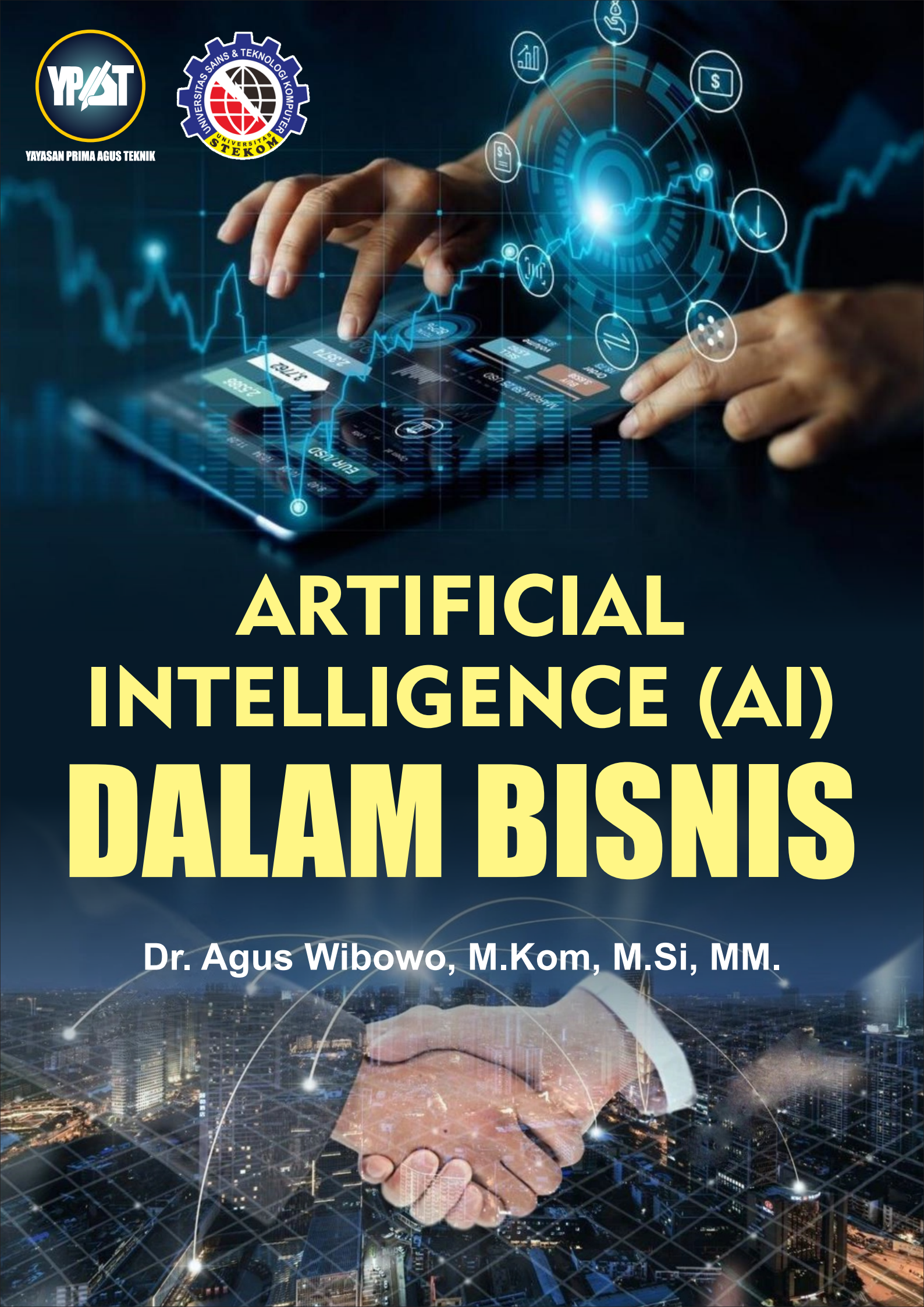




YAYASAN PRIMA AGUS TEKNIK

The background of the entire page is a dark blue, high-tech image. It features a pair of hands interacting with a futuristic, glowing blue interface. The interface includes various icons such as a bar chart, a dollar sign, a downward arrow, and a circular arrow, along with a line graph. The hands are positioned as if they are touching or manipulating these digital elements. The overall aesthetic is futuristic and digital, with a grid-like pattern visible in the background.

ARTIFICIAL INTELLIGENCE (AI) DALAM BISNIS

Dr. Agus Wibowo, M.Kom, M.Si, MM.

The bottom section of the page features a background image of a city skyline at night, with various buildings and lights. Overlaid on this is a large, semi-transparent image of two hands shaking in a firm grip. The handshake is positioned in the center, with the hands appearing to be in a business deal. The city lights and the handshake are both rendered in a slightly desaturated, blue-toned style, creating a cohesive visual theme.



ARTIFICIAL INTELLIGENCE (AI) DALAM BISNIS

Dr. Agus Wibowo, M.Kom, M.Si, MM.



YAYASAN PRIMA AGUS TEKNIK

PENERBIT :
YAYASAN PRIMA AGUS TEKNIK
Jl. Majapahit No. 605 Semarang
Telp. (024) 6723456. Fax. 024-6710144
Email : penerbit_ypat@stekom.ac.id

ISBN 978-634-7227-31-7 (PDF)



9

786347

227317

ARTIFICIAL INTELLIGENCE (AI) DALAM BISNIS

Penulis :

Dr. Agus Wibowo, M.Kom, M.Si, MM.

ISBN : 978-634-7227-31-7

Editor :

Dr. Joseph Teguh Santoso, S.Kom., M.Kom.

Penyunting :

Dr. Mars Caroline Wibowo. S.T., M.Mm.Tech

Desain Sampul dan Tata Letak :

Irdha Yuniarto, S.Ds., M.Kom

Penebit :

Yayasan Prima Agus Teknik Bekerja sama dengan
Universitas Sains & Teknologi Komputer (Universitas STEKOM)

Anggota IKAPI No: 279 / ALB / JTE / 2023

Redaksi :

Jl. Majapahit no 605 Semarang

Telp. 08122925000

Fax. 024-6710144

Email : penerbit_ypat@stekom.ac.id

Distributor Tunggal :**Universitas STEKOM**

Jl. Majapahit no 605 Semarang

Telp. 08122925000

Fax. 024-6710144

Email : info@stekom.ac.id

Hak cipta dilindungi undang-undang

Dilarang memperbanyak karya tulis ini dalam bentuk dan dengan cara
apapun tanpa ijin dari penulis

KATA PENGANTAR

Puji syukur penulis panjatkan atas terselesaikannya buku berjudul ***“Artificial Intelligence (AI) dalam Bisnis”*** yang hadir sebagai sumber referensi komprehensif mengenai pemanfaatan kecerdasan buatan dalam dunia bisnis modern. Buku ini disusun dengan tujuan memberikan pemahaman mendalam tentang konsep, teknologi, serta aplikasi AI yang kian berkembang pesat dan menjadi pendorong utama inovasi serta efisiensi di berbagai sektor industri.

Buku ini terdiri dari enam bab utama yang disusun secara sistematis dan mendalam. Bab pertama mengupas pengantar kecerdasan buatan, mencakup dasar-dasar AI, konsep fundamental, serta peran big data dan AI dalam mendorong inovasi bisnis. Bab kedua menjelaskan alur kerja pembelajaran mesin beserta algoritma-algoritma yang menjadi fondasi teknologi AI untuk mendukung pengambilan keputusan otomatis dan analisis data. Pada bab ketiga, dibahas teknik-teknik pembelajaran mendalam, analisis big data, serta aplikasinya dalam bisnis dengan contoh nyata dari industri e-commerce. Bab keempat menguraikan berbagai teknik sistem mesin rekomendasi, penerapannya dalam dunia bisnis, dan disertai studi kasus sebagai ilustrasi pemanfaatan yang efektif. Bab kelima fokus pada pemrosesan bahasa alami (*Natural Language Processing/NLP*), menjelaskan teknologi terkait, penerapannya dalam bisnis, serta bagaimana analisis sentimen dapat memperkaya interaksi dengan pelanggan. Terakhir, bab keenam menyoroti lanskap analisis AI, strategi integrasi AI ke dalam proses bisnis, potensi pertumbuhan yang ditimbulkan oleh AI, dan gambaran terkini tentang bisnis AI di Indonesia.

Semoga buku ini dapat menjadi panduan yang berguna bagi pelaku bisnis, akademisi, serta praktisi yang ingin memahami sekaligus mengaplikasikan kecerdasan buatan secara efektif dalam kegiatan bisnisnya. Penulis berharap karya ini mampu menginspirasi pemanfaatan AI yang tidak hanya mengoptimalkan keuntungan, tetapi juga membawa kemajuan yang berdampak positif bagi masyarakat luas.

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dan kontribusi sehingga buku ini dapat terwujud. Semoga bermanfaat dan dapat menjadi sumber ilmu yang dapat memperkaya wawasan serta membuka peluang baru dalam era digital yang penuh dinamika ini.

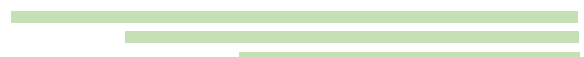
Semarang, 2025

Penulis

Dr. Agus Wibowo, M.Kom, M.Si, MM.

DAFTAR ISI

Halaman Judul	i
Kata Pengantar	ii
DAFTAR ISI.....	iii
BAB 1 PENGANTAR KECERDASAN BUATAN	1
1.1 Dasar Kecerdasan Buatan	1
1.2 Konsep Dasar Kecerdasan Buatan	3
1.3 Inovasi Bisnis Dengan Big Data Dan Kecerdasan Buatan.....	11
BAB 2 PEMBELAJARAN MESIN	20
2.1 Pendahuluan.....	20
2.2 Alur Kerja Pembelajaran Mesin	22
2.3 Algoritma Pembelajaran	24
BAB 3 PEMBELAJARAN MENDALAM	35
3.1 Pendahuluan.....	35
3.2 Menganalisis <i>Big Data</i>	36
3.3 Berbagai Model Pembelajaran Mendalam.....	42
3.4 Aplikasi Pembelajaran Mendalam Dalam Bisnis	46
3.5 Contoh Penggunaan Bisnis Untuk <i>E-Commerce</i>	46
BAB 4 SISTEM MESIN REKOMENDASI	50
4.1 Pendahuluan.....	50
4.2 Teknik Sistem Rekomendasi	54
4.3 Penerapan Mesin Rekomendasi Dalam Bisnis.....	58
4.4 Kasus Penggunaan Bisnis.....	61
BAB 5 PEMROSESAN BAHASA ALAMI	64
5.1 Pendahuluan.....	64
5.2 Contoh Penggunaan NLP	67
5.3 Aplikasi NLP Dalam Bisnis	70
5.4 Teknologi Sentimen Dalam Bisnis.....	73
BAB 6 MEMANFAATKAN AI DALAM BISNIS.....	76
6.1 Lanskap Analisis.....	76
6.2 Menanamkan AI Ke Dalam Proses Bisnis	83
6.3 Kecerdasan Buatan Untuk Pertumbuhan	85
6.4 Bisnis AI Di Indonesia	90
DAFTAR PUSTAKA	94



BAB 1

PENGANTAR KECERDASAN BUATAN

1.1 DASAR KECERDASAN BUATAN

Data

Hal-hal faktual, diskrit, statis, dan dinamis serta observasi mentah dari area minat tertentu dikenal sebagai data. Informasi dapat dihasilkan setelah pemrosesan data tersebut secara sistematis. Data sering diidentifikasi sebagai nilai numerik dalam lingkungan. Data juga dapat diamati sebagai catatan fisik transaksional dari aktivitas perusahaan, yang dianggap sebagai blok pembangun dasar dari setiap sistem informasi. Kita perlu memprosesnya sebelum menggunakannya. Data dapat didefinisikan sebagai:

Data adalah simbol yang merepresentasikan properti objek, peristiwa, dan lingkungannya. Data merupakan produk observasi. Mengamati berarti merasakan. Teknologi instrumentasi penginderaan, tentu saja, harus dikembangkan secara intensif.

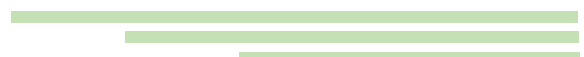
Data adalah hal-hal yang diberikan kepada analis, peneliti, atau pemecah masalah; data dapat berupa angka, kata, kalimat, catatan, dan asumsi apa pun yang diberikan, apa pun bentuk dan asalnya. Hal ini dulunya sudah diketahui oleh para akademisi di sebagian besar bidang: beberapa menginginkan kata data merujuk pada fakta, terutama pembacaan instrumen. Bagi yang lain yang berurusan dengan hipotesis, data adalah asumsi.

Meskipun data merupakan bukti untuk sesuatu, data tersebut tidak harus selalu benar; namun, terdapat kesulitan dalam "mengetahui" apakah data tersebut benar atau tidak. Hal ini menyebabkan pemrosesan lebih lanjut untuk menghasilkan informasi dan pengetahuan dari data yang tersedia. Misalnya, suhu pada waktu tertentu pada hari tertentu merupakan atom data tunggal dan diperlakukan sebagai fakta tertentu. Mungkin terdapat beberapa atom seperti itu, dan atom-atom ini dapat digabungkan dengan berbagai cara menggunakan operasi logika standar. Namun, terdapat juga pernyataan universal, seperti "Setiap hari suhu maksimum di atas 30 derajat". Namun, dari sudut pandang logika, pernyataan universal tersebut lebih kuat daripada atom atau senyawa atom, sehingga lebih sulit untuk memastikan kebenarannya. Data tersebut juga perlu disaring lebih lanjut untuk menghasilkan informasi yang benar dan diperlukan. Yang terpenting, data mungkin berupa data empiris. Sangat sulit untuk menetapkan nilai kebenaran pada data non-empiris yang fiktif.

Informasi

Ketika data diproses, diorganisasikan, distrukturkan, atau disajikan dalam konteks tertentu agar bermanfaat, data tersebut disebut informasi. Namun, ada informasi yang bukan data. Informasi khusus tersebut dapat dianggap sebagai data olahan yang mempermudah pengambilan keputusan. Pemrosesan melibatkan agregasi data, perhitungan data, koreksi data, dll. sedemikian rupa sehingga menghasilkan aliran pesan. Informasi biasanya memiliki makna dan tujuan tertentu. Artinya, data dalam suatu konteks dapat dianggap sebagai informasi. Nilai tambah pada data dapat diperoleh dengan beberapa cara:

- *Dikontekstualisasikan*: memberi tahu kita tujuan pengumpulan data



- *Dikategorikan*: memberi tahu kita unit analisis atau komponen kunci data
- *Dihitung*: memberi tahu kita apakah data dianalisis secara matematis atau statistik
- *Dikoreksi*: memberi tahu kita apakah kesalahan telah dihilangkan dari data
- *Diringkas*: memberi tahu kita apakah data diringkas dalam bentuk yang lebih ringkas

Lebih lanjut, informasi dapat diproses, diakses, dihasilkan, ditransmisikan, disimpan, dikirim, didistribusikan, diproduksi dan dikonsumsi, dicari, digunakan, dikompresi, dan diduplikasi. Informasi juga dapat beragam jenisnya dengan atribut yang berbeda. Informasi tersebut dapat berupa informasi sensitif, informasi kualitatif, atau informasi kuantitatif.

Pengetahuan

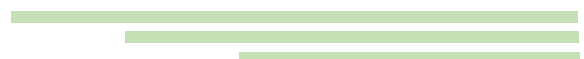
Pengetahuan dianggap sebagai pemahaman manusia tentang suatu subjek yang diperoleh melalui studi dan pengalaman yang tepat. Informasi dan data dapat dikaitkan dengan sekelompok manusia dan dianggap sebagai massa kolektif, sedangkan pengetahuan biasanya didasarkan pada pembelajaran, pemikiran, dan pemahaman yang tepat tentang area masalah oleh seorang individu. Pengetahuan diperoleh dari informasi dengan cara yang sama seperti informasi diperoleh dari data. Pengetahuan dapat dianggap sebagai sintesis dan integrasi proses persepsi manusia yang membantu mereka menarik kesimpulan yang bermakna.

Pengetahuan adalah "keyakinan sejati yang dibenarkan" yang terkait dengan tindakan manusia dan diciptakan dari aliran pesan. Pengetahuan umumnya bersifat pribadi, subjektif, dan secara inheren lokal pengetahuan ditemukan "di dalam kepala karyawan" alih-alih ada secara objektif. Selain itu, pengetahuan dapat dimiliki di luar pikiran manusia dan disarankan bahwa agen mampu memanipulasi keyakinan dan penilaian. Dia menggambarkan pengetahuan sebagai "kebenaran dan keyakinan, perspektif dan konsep, penilaian dan harapan, metodologi dan pengetahuan dan dimiliki oleh manusia atau agen lain". Informasi adalah data yang menceritakan tentang bisnisnya dan bagaimana fungsinya.

Langkah tambahan diterapkan pada informasi untuk mengubahnya menjadi pengetahuan, dengan mengidentifikasi tiga "I" dalam bisnis sebagai berikut:

- Dampak: Dampak bisnis terhadap kelompok pengguna dan pasar sasaran
- Interaksi: Bagaimana sistem bisnis berinteraksi dengan pengguna dan sistem lain di lingkungan
- Dipengaruhi: Bagaimana bisnis dipengaruhi oleh pesaing dan tren pasar

Dalam bidang manajemen pengetahuan, terdapat dua jenis pengetahuan yang cukup berbeda dan diterima secara luas: tacit dan eksplisit. Pengetahuan tacit, sebagaimana diidentifikasi oleh Polanyi, adalah pengetahuan yang sulit dikodekan dan dikomunikasikan. Pengetahuan ini bersifat sementara dan sementara, serta "tidak dapat diuraikan menjadi informasi atau dirinci dengan cara yang menjadi ciri khas informasi". Lebih lanjut, pengetahuan tacit bersifat personal, spesifik konteks, dan sulit diformalkan. Di sisi lain, pengetahuan eksplisit adalah jenis pengetahuan yang dapat dikodekan dan ditransmisikan dalam bahasa. Pengetahuan eksplisit inilah yang coba, dan memang dapat, ditangkap, diperoleh, diciptakan, dimanfaatkan, dipertahankan, dikodifikasi, disimpan, ditransfer, dan dibagikan oleh sebagian besar praktik manajemen pengetahuan saat ini.



Data dan informasi merupakan aspek yang sangat penting dari pengetahuan. Data dan informasi memerlukan pemrosesan yang tepat untuk menghasilkan informasi terstruktur yang bermakna guna membantu pengambilan keputusan dan memperoleh keahlian untuk pemecahan masalah. Artinya, tingkat pemrosesan inilah yang membuat konten bermakna dan dapat diterapkan. Dengan pemrosesan yang tepat, kita dapat menghasilkan laporan yang membantu pengambilan keputusan, konsep untuk pembelajaran, dan model untuk pemecahan masalah.

Kecerdasan

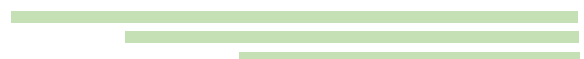
Pengetahuan tentang konsep dan model mengarah pada tingkat pengetahuan yang lebih tinggi yang disebut kebijaksanaan. Seseorang perlu menerapkan moral, prinsip, dan keahlian untuk mendapatkan dan memanfaatkan kebijaksanaan. Hal ini membutuhkan waktu dan kedewasaan yang seiring bertambahnya usia dan pengalaman.

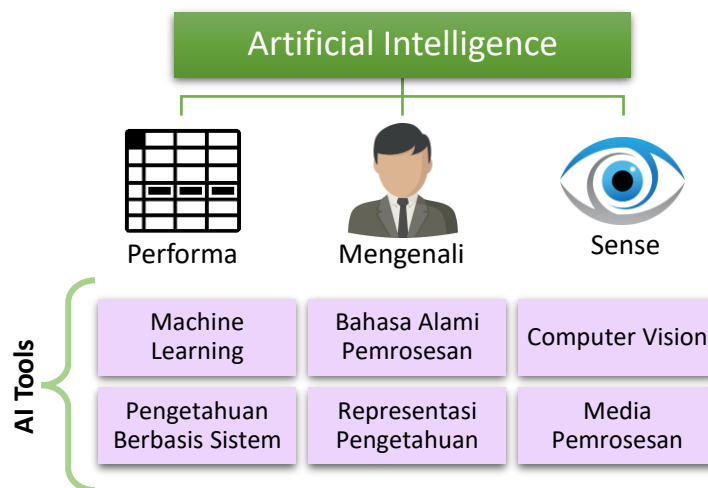
Konsep kebijaksanaan telah dibahas oleh para filsuf Yunani kuno, seperti Plato dan Aristoteles; meskipun belum menjadi topik diskusi yang populer akhir-akhir ini. Tampaknya ada beberapa aliran berbeda dalam kebijaksanaan. Seseorang mungkin memiliki pengetahuan ensiklopedis tentang fakta dan angka yang berkaitan dengan negara-negara di dunia; tetapi pengetahuan itu sendiri tidak akan membuat orang tersebut bijaksana. Sebaliknya, seseorang menjadi bijaksana dengan menerapkan pengetahuan pada masalah-masalah kompleks yang bersifat etis dan praktis serta mencari solusi potensial. Peningkatan lebih lanjut dari kebijaksanaan adalah kecerdasan. Kecerdasan adalah tujuan suatu entitas untuk menjadi entitas yang sepenuhnya memiliki kecerdasan buatan.

1.2 KONSEP DASAR KECERDASAN BUATAN

Kecerdasan buatan (AI) telah ada selama bertahun-tahun; namun, di mana ia dapat dikembangkan masih menjadi bahan diskusi. Dengan perkembangan teknologi, saat ini terdapat tuntutan besar akan pembelajaran manusia yang komprehensif dalam aspek komputasional yang mampu mengubah keyakinan perilakunya sendiri. Memiliki kemampuan untuk memutuskan, belajar, dan menanamkan dirinya berdasarkan peristiwa sebelumnya dan menindaklanjutinya dengan sangat tekun.

AI mengacu pada beragam alat dan teknologi yang dapat dikombinasikan dengan berbagai cara untuk merasakan, mengenali, dan bertindak dengan kemampuan untuk belajar dari pengalaman dan beradaptasi seiring waktu, seperti yang diilustrasikan pada Gambar 1.1.





Gambar 1.1 Apa itu AI?

Secara umum, kecerdasan adalah kemampuan seseorang untuk memahami dunia objektif dan menerapkan pengetahuan untuk memecahkan masalah. Kecerdasan individu terdiri dari berbagai kemampuan, seperti: kemampuan untuk memahami dan menghayati hal-hal objektif, dunia objektif, dan diri sendiri; kemampuan untuk memperoleh pengalaman dan memperoleh pengetahuan melalui pembelajaran; kemampuan untuk memahami pengetahuan dan menerapkan pengetahuan serta pengalaman untuk analisis dan pemecahan masalah; kemampuan berasosiasi, bernalar, menilai, dan mengambil keputusan; kemampuan abstraksi dan generalisasi linguistik; kemampuan penemuan, invensi, kreativitas, dan inovasi; kemampuan untuk mengatasi lingkungan yang kompleks dengan tepat, cepat, dan wajar; serta kemampuan untuk memprediksi dan memberikan wawasan tentang perkembangan dan perubahan berbagai hal.

AI bukanlah konsep baru, faktanya banyak landasan teoretis dan teknologinya telah dikembangkan selama 62 tahun terakhir. Sebagai catatan, permulaan resmi AI dianggap sebagai "konferensi Dartmouth" pada tahun 1956. Dan sampai batas tertentu, uji Turing bahkan mendahuluinya dan menawarkan pemikiran tentang cara mengenali "mesin cerdas". Namun, perjalanan AI cukup bergejolak. Menengok ke belakang, terdapat kemajuan substansial di hampir semua bidang yang terutama dianggap sebagai bagian dari AI. Mari kita lihat beberapa perkembangan yang merangsang dalam hal signifikansi praktis.

Sistem berbasis pengetahuan mungkin merupakan cabang praktis AI yang paling sukses. Ada beberapa aplikasi yang diterapkan di berbagai organisasi di seluruh dunia. Ratusan alat, yang umumnya disebut kerangka sistem pakar, telah dikembangkan. Sistem semacam itu mencapai kemegahan yang cukup untuk menjadi disiplin ilmu yang independen, hingga memiliki mata kuliah akademis yang terpisah.

Seiring dengan keberhasilan praktis, bidang ini juga berkontribusi pada pertumbuhan AI itu sendiri. Konsep representasi pengetahuan berbasis aturan, penekanan pada penalaran dengan ketidakpastian, isu-isu verifikasi pengetahuan domain, pembelajaran mesin dalam konteks akuisisi pengetahuan otomatis, dll. adalah beberapa bidang pertumbuhan akademis. Bidang kemajuan lainnya adalah pemrosesan bahasa alami. Sistem penerjemahan yang

memadai kini tersedia untuk digunakan dalam konteks terbatas, terutama efektif jika sedikit panduan manusia dapat diberikan kepada sistem. Systran adalah contoh relevan, yang memberikan solusi bahasa waktu nyata untuk kolaborasi internal, pencarian, eDiscovery, manajemen konten, dukungan pelanggan daring, dan e-Commerce. Bidang ini juga berkontribusi pada pengembangan bidang pencarian informasi. World Wide Web adalah salah satu alasan utama minat di bidang ini, dengan informasi yang tersedia jauh melampaui batas imajinasi manusia. Tanpa analisis dan penyaringan otomatis, mengidentifikasi dan mengambil item yang menarik dari tambang besar ini merupakan tugas yang menantang. Web semantik, analisis konten dan tautan halaman web, penambahan teks, ekstraksi informasi spesifik dari dokumen, klasifikasi otomatis, dan agen personalisasi yang mencari informasi yang menarik bagi individu tertentu adalah beberapa bidang aktif saat ini.

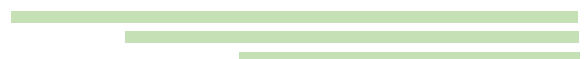
Pemrosesan ucapan telah menghasilkan alat-alat yang bernilai fungsional. Saat ini, tersedia perangkat lunak yang dapat mengubah teks lisan Anda menjadi teks yang dapat diproses mesin seperti dokumen Word. Alat-alat ini memang memerlukan pelatihan dan belum terlalu efektif dalam beradaptasi dengan banyak pembicara. Alat-alat semacam itu berguna bagi orang-orang yang tidak memiliki kecepatan mengetik yang baik, dan yang lebih penting lagi bagi mereka yang memiliki disabilitas untuk berinteraksi dengan komputer.

Robotika juga sedang berada di jalur momentum yang tinggi. Ada inisiatif besar di Jepang, yang bertujuan untuk mengembangkan robot humanoid guna membantu para lansia dalam pekerjaan rutin mereka. Inisiatif semacam ini saat ini sedang mendorong perkembangan robotika di Jepang dan Amerika Serikat. Honda dan Sony dari Jepang telah membangun robot yang dapat berjalan, melambaikan tangan, melakukan beberapa langkah tarian dasar, dll. Hewan peliharaan robotik telah mencapai status komersial dengan beberapa perusahaan memasarkan anjing peliharaan yang canggih.

Apa yang telah kami catat hanyalah sebagian dari keberhasilan AI. Dari awal yang sederhana sekitar setengah abad yang lalu, AI telah berkembang dalam banyak dimensi. Sementara beberapa praktisi AI mengejar tujuan awal untuk mencapai kecerdasan mesin, sebagian besar penelitian AI saat ini difokuskan pada pemecahan masalah praktis yang kompleks. Meskipun AI telah menjadi bagian dari kehidupan kita sehari-hari selama beberapa waktu, teknologi ini berada di titik balik, terutama karena kemajuan penting terbaru dalam aplikasi pembelajaran mendalam.

Pembelajaran mendalam memanfaatkan jaringan yang mampu melakukan pembelajaran tanpa pengawasan dari data yang tidak terstruktur atau tidak berlabel. Jaringan saraf yang mendasari kemampuan pembelajaran mendalam menjadi lebih efisien dan akurat karena dua kemajuan teknologi terkini yang signifikan: akses yang belum pernah terjadi sebelumnya ke big data dan peningkatan daya komputasi. Efektivitas jaringan saraf berkorelasi dengan jumlah data yang tersedia.

Pembelajaran mesin (ML), salah satu bidang AI yang paling menarik, melibatkan pengembangan pendekatan komputasional untuk memahami data secara otomatis teknologi ini memanfaatkan wawasan bahwa pembelajaran adalah proses dinamis, yang dimungkinkan melalui contoh dan pengalaman, alih-alih aturan yang telah ditentukan sebelumnya. Seperti



manusia, mesin dapat menyimpan informasi dan menjadi lebih pintar seiring waktu. Berbeda dengan manusia, mesin tidak cenderung mengalami kurang tidur, gangguan, kelebihan informasi, dan kehilangan memori jangka pendek di situlah teknologi berpengaruh ini menjadi menarik.

Dengan aplikasi di hampir setiap industri, AI menjanjikan transformasi signifikan terhadap model bisnis yang ada sekaligus menciptakan model bisnis baru. Dalam layanan keuangan, misalnya, terdapat manfaat yang jelas dari peningkatan akurasi dan kecepatan dalam sistem deteksi penipuan yang dioptimalkan AI, yang diperkirakan akan menjadi pasar senilai Rp 30 triliun pada tahun 2020.

Keunggulan utama AI dibandingkan kecerdasan manusia adalah skalabilitas, umur panjang, dan kemampuan peningkatan berkelanjutan. Atribut-atribut tersebut diantisipasi akan meningkatkan produktivitas, menurunkan biaya, dan mengurangi kesalahan manusia secara signifikan. Meskipun masih dalam tahap yang menjanjikan, teknologi AI diharapkan dapat memperkenalkan standar baru bagi produktivitas perusahaan, keunggulan kompetitif, dan pada akhirnya, pertumbuhan ekonomi.

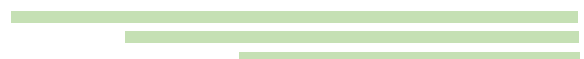
Jumlah data yang sangat besar yang dikumpulkan dalam basis data saat ini sangat terbatas manfaatnya jika hanya mekanisme pengambilan data yang biasa diterapkan. Mengajukan pertanyaan yang tepat dan menghubungkan data dengan cara yang sesuai dapat menghasilkan informasi yang relevan. Kini, tersedia berbagai metode untuk melakukan hal ini, misalnya, gudang data, metode statistik klasik, jaringan saraf tiruan, dan algoritma pembelajaran mesin. AI telah memberikan kontribusi substansial terhadap teknologi data yang digunakan secara komersial selama beberapa tahun dan sering disebut penambangan data.

Manfaat AI

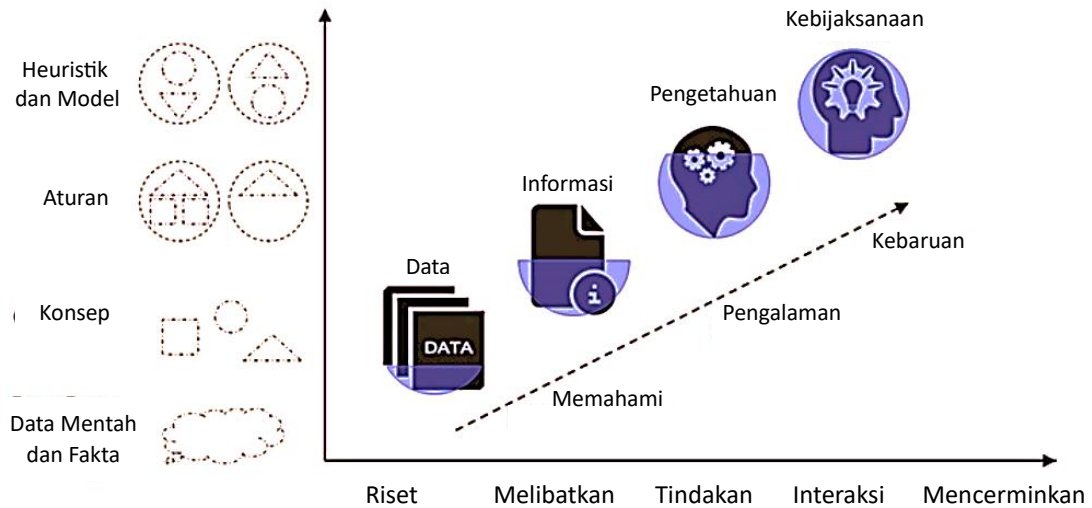
- Menemukan pola yang bermakna dan bermanfaat dalam data bervolume besar, apa pun jenisnya, termasuk teks, gambar, video, dan data tidak terstruktur lainnya.
- Model pembelajaran mandiri memungkinkan Anda beradaptasi dengan cepat terhadap perubahan pola data dan kondisi bisnis yang mendasarinya.
- Membuat keputusan yang lebih baik dan lebih cepat dengan memaksimalkan nilai semua data Anda dan beralih dari analitik prediktif ke analitik preskriptif.
- Mengenali wawasan unik tentang arketipe dalam data yang memungkinkan segmentasi dan penanganan pelanggan yang optimal.
- Memudahkan pengguna bisnis dan ilmuwan data untuk menafsirkan, merespons, dan menggunakan data melalui visualisasi dan transparansi yang lebih baik.
- Menawarkan model bisnis baru dan penciptaan nilai dengan mempercepat inovasi melalui penemuan pola baru dalam data dan memanfaatkan aset pengetahuan secara penuh.

Piramida Data

Sistem AI menggunakan teknik AI, yang memungkinkan mereka mencapai kompetensi tingkat ahli dalam memecahkan masalah di area tertentu. Sistem yang menggunakan satu atau lebih pengetahuan pakar untuk memecahkan masalah dalam domain tertentu disebut sistem



berbasis pengetahuan atau sistem pakar. Sistem informasi tradisional bekerja berdasarkan data dan/atau informasi.



Gambar 1.2 Konvergensi data dan kecerdasan



Gambar 1.3 Piramida data: perspektif pengambilan keputusan

Gambar 1.2 dan 1.3 merepresentasikan piramida data yang menyatakan hubungan antara data, informasi, pengetahuan, dan kecerdasan. Gambar 1.2 menyatakan konvergensi data menjadi pengetahuan dengan menerapkan aktivitas seperti meneliti, terlibat, bertindak, berinteraksi, dan berefleksi. Dalam proses ini, manusia biasanya memperoleh pemahaman dan pengalaman serta dapat menghasilkan ide-ide inovatif. Aktivitas-aktivitas ini ditunjukkan pada sumbu X. Sumbu Y menyajikan bentuk-bentuk konvergensi, yaitu observasi mentah, konsep, aturan, model, dan heuristik.

Gambar 1.3 menunjukkan piramida data melalui perspektif manajemen. Staf tingkat operasional umumnya bekerja dengan lingkungan terstruktur dan menggunakan prosedur yang telah ditentukan sebelumnya untuk menjalankan transaksi rutin bisnis, yang merupakan operasi dasar bisnis. Untuk menjalankan transaksi rutin bisnis, staf operasional menggunakan sistem seperti sistem pemrosesan transaksi (TPS). Dengan lingkungan yang sepenuhnya terstruktur dan serangkaian prosedur yang telah ditentukan sebelumnya, pengembangan dan otomatisasi sistem tersebut (TPS) menjadi mudah. TPS semacam itu mempertimbangkan observasi lapangan mentah dan memprosesnya untuk menghasilkan informasi yang bermakna. Inilah tingkat data piramida.

Informasi yang dihasilkan melalui transaksi bisnis dianalisis untuk membentuk laporan rutin dan luar biasa, yang membantu para manajer dan eksekutif untuk mengambil keputusan. Sistem yang melakukan ini disebut sistem informasi manajemen (MIS). TPS dan MIS bekerja pada lingkungan terstruktur yang bekerja dengan data dan/atau informasi.

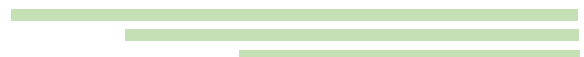
Manajemen juga perlu mengambil keputusan dengan mempertimbangkan rasio biaya-manfaat dari berbagai solusi alternatif yang tersedia untuk secara efektif memanfaatkan sumber daya yang langka dan kendala lingkungan. Kategori sistem yang dimaksudkan untuk itu adalah sistem pendukung keputusan (DSS). Tidak seperti TPS yang hanya menggunakan basis data dan bekerja di lingkungan terstruktur, DSS biasanya bekerja pada lingkungan terstruktur hingga semi-terstruktur dan memanfaatkan basis model dan basis data untuk pemanfaatan sumber daya yang optimal.

Sistem seperti TPS, MIS, dan DSS melakukan transaksi rutin bisnis, menyediakan analisis terperinci tentang informasi yang dihasilkan, dan mendukung proses pengambilan keputusan bisnis. Namun, sistem ini tidak mengambil keputusan sendiri atau membenarkannya dengan penjelasan dan penalaran yang tepat karena mereka tidak memiliki pengetahuan. Manajemen tingkat tinggi membutuhkan pengetahuan dan kebijaksanaan untuk pengambilan kebijakan dan strategi, sehingga diperlukan sistem berbasis pengetahuan dan kebijaksanaan (KBS dan WBS). Dengan menerapkan etika, prinsip, dan penilaian pada keputusan yang diambil dan setelah mencapai tingkat kematangan (pengalaman), informasi dapat digeneralisasi dan diubah menjadi pengetahuan.

Sifat Otonomi

Dalam Ilmu Komputer, sistem terdistribusi telah dikembangkan selama beberapa tahun dan terbukti menjadi cara yang ampuh untuk meningkatkan efisiensi pemecahan masalah, tetapi juga membuka ranah aplikasi baru bagi teknologi komputer. Kontribusi AI di bidang sistem terdistribusi tidak hanya mengembangkan algoritma baru untuk sistem tersebut, tetapi juga membekali komponen sistem terdistribusi dengan tingkat otonomi tertentu.

Hingga dekade terakhir, sistem komputer sebagian besar dipandang sebagai mesin komputasi dan penyimpanan data massal serta digunakan untuk mendukung keputusan. Manusia ragu untuk menerima keputusan yang dibuat oleh sistem komputer. Namun, dalam beberapa ranah teknis, komputer telah mengendalikan instalasi besar, yang dalam beberapa hal mencakup pengambilan keputusan, meskipun manusia masih berperan sebagai pengawas. Dengan demikian, sistem komputer telah mulai menjadi otonom.



Dalam sistem multi-agen, yang merupakan kontribusi AI terhadap keterdistribusian, otonomi merupakan isu kunci. Suatu komponen sistem dapat dipandang atau dimodelkan sebagai agen hanya jika memiliki tingkat otonomi tertentu, jika tidak, ia dianggap sebagai komponen pasif. Otonomi dapat dicirikan oleh beberapa fitur. Otonomi berarti kemampuan untuk memilih suatu tindakan dari serangkaian tindakan yang mungkin dalam situasi tertentu, termasuk keputusan antara tetap tidak aktif atau menjadi aktif ketika diperlukan. Kemampuan ini disebut proaktivitas.

Fitur penting lainnya adalah kemampuan untuk mempertahankan tujuan. Agen tidak hanya dituntut memiliki tujuan yang memandu perencanaan dan aktivitasnya, tetapi juga dapat mengubahnya, meninggalkannya, atau mengadopsi tujuan baru. Fitur ketiga adalah kemampuan komunikasi dan kerja sama aktif. Kata "aktif" penting di sini karena pertukaran pesan dalam arti panggilan fungsional antara komponen pasif biasanya juga disebut komunikasi, tetapi bukan itu yang dimaksud di sini.

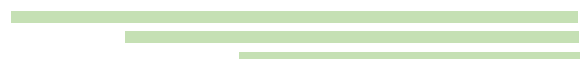
Komunikasi dan kerja sama antar agen juga merupakan perilaku proaktif; seorang agen dapat memulai komunikasi atau mencari kerja sama dengan orang lain kapan pun ia merasa perlu. Properti otonomi menciptakan hubungan baru antara agen apakah mereka diwujudkan sebagai robot atau sebagai sistem perangkat lunak dan manusia yang memiliki kualitas kemitraan. Agen otonom tidak bisa lagi dianggap sebagai mesin belaka yang dimulai saat dibutuhkan, melakukan pekerjaan mereka, dan dihentikan.

Seseorang dapat memandang mereka sebagai pelayan, tetapi pelayan memiliki kemauan mereka sendiri dan kemampuan kognitif yang kompleks, dan ini harus diakui kepada agen juga. Kita harus berharap untuk hidup dalam masyarakat yang jauh lebih kompleks di masa depan daripada yang kita lakukan saat ini bersama dengan agen sebagai semacam entitas "hidup". Mereka akan hadir dalam kehidupan kita sehari-hari, misalnya, sebagai asisten pribadi dalam sistem komputer yang dapat dikenakan kita, sebagai pengemudi mobil kita atau sebagai manajer rumah tangga kita. Peneliti AI harus bersama dengan sosiolog menangani masalah yang mungkin disebabkan oleh perspektif ini.

Kesadaran Situasi

Secara umum, semua sistem komputer bersifat situasional, tetapi sistem tradisional berada dalam situasi yang sangat terbatas dan terdefinisi dengan baik, yang sepenuhnya ditentukan oleh manusia. Kesituasian baru menjadi topik penelitian penting dengan munculnya agen dalam sistem multi-agen dan robot bergerak yang adaptif. Jelas, kesituasian berkaitan erat dengan isu otonomi. Hanya sistem otonom yang harus menemukan dan mengorientasikan diri serta bertindak dalam situasi. Situasional di sini berarti serangkaian pengaruh dari lingkungan yang setidaknya sebagian tidak dapat diramalkan tetapi sangat penting bagi sistem, sehingga harus bereaksi dengan cara yang tepat.

Sistem yang tersituasi harus memecahkan dua masalah utama, yaitu bagaimana merasakan situasi dan bagaimana memilih reaksi yang tepat. Merasakan situasi mungkin merupakan tugas yang cukup sederhana bagi agen perangkat lunak biasa dalam lingkungan yang terdefinisi dengan baik. Namun, jika kita memikirkan agen yang bertindak, misalnya, di Internet, segalanya menjadi jauh lebih rumit karena lingkungan ini sangat tidak terstruktur dan



dinamis. Penginderaan dengan robot bahkan lebih kompleks. Di sini, sinyal fisik pertama-tama harus diubah menjadi data, sebuah tugas yang dapat mereka delegasikan kepada fisikawan dan insinyur. Namun, pada langkah selanjutnya, sinyal dari berbagai sumber harus digabungkan untuk menghasilkan deskripsi situasi yang memungkinkan sistem yang ada untuk bereaksi dengan tepat.

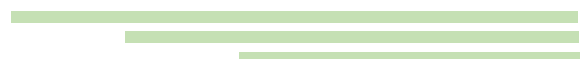
Tugas ini disebut fusi sensor, dan di sinilah metode AI berperan. Dari Ilmu Kognitif, kita belajar bahwa memahami suatu situasi adalah proses yang sangat kompleks yang membutuhkan banyak pengetahuan latar belakang dan banyak aktivitas saraf. Otak membangun situasi dari berbagai masukan. Kita masih sedikit mengetahui tentang proses ini dan untuk alasan yang jelas, sulit untuk merekonstruksinya. Simulasi dengan metode AI dapat membantu untuk mendapatkan wawasan tentangnya. Jika simulasi memungkinkan, simulasi juga dapat digunakan oleh sistem situasional buatan untuk membangun deskripsi situasi untuk tujuan mereka sendiri.

Bagi sistem situasional, tujuan utama membangun deskripsi situasi adalah menggunakannya untuk aktivitas mereka sendiri. Tugas dasar lokalisasi dan orientasi diri serta tugas turunan tindakan hanya dapat dilakukan berdasarkan deskripsi situasi. Ini berarti, di sisi lain, konstruksi deskripsi situasi selalu bertujuan untuk mendukung sistem situasional dalam memenuhi tugasnya, bukan tujuan akhir.

Kriteria seperti kelengkapan atau konsistensi tidak diprioritaskan, melainkan deskripsi memenuhi kebutuhan sistem situasional jika membantu memilih tindakan yang tepat. Basis pengetahuan yang tersituasi diperlukan, terdiri dari pengetahuan latar belakang yang luas dan potongan-potongan pengetahuan relevan yang tidak perlu konsisten satu sama lain maupun dengan pengetahuan latar belakang. Pengetahuan dalam potongan-potongan ini bahkan dapat direpresentasikan dalam berbagai bentuk dan tingkat detail yang berbeda. Metode untuk memilih potongan pengetahuan yang tepat, menggabungkannya, dan mentransformasikan pengetahuan dari satu bentuk ke bentuk lainnya harus dikembangkan.

Tentu saja, sistem yang tersituasi harus memiliki kemampuan perencanaan untuk memilih urutan tindakan. Sistem tersebut juga harus memiliki kemampuan belajar karena, sebagaimana disebutkan di atas, pengaruh yang membentuk suatu situasi sebagian tidak dapat diramalkan. Namun, situasi baru tidak sepenuhnya berbeda satu sama lain, melainkan di semua lingkungan yang wajar terdapat kesamaan di antara mereka, sehingga pembelajaran, yaitu mendeteksi dan mengklasifikasikan kasus serupa, menjadi masuk akal. Pembelajaran dapat meningkatkan perilaku sistem yang tersituasi.

Penciptaan informasi, otonomi, dan kesusuaian dapat dianggap sebagai fokus penelitian dan pengembangan AI di masa depan. Untuk menghadapi tantangan ini, banyak metode tunggal harus diintegrasikan ke dalam sistem yang lebih besar. Jadi, arah umum penelitian dan pengembangan AI dapat dicirikan oleh pengembangan sistem kompleks yang mengintegrasikan berbagai metode dan memenuhi ketiga persyaratan.



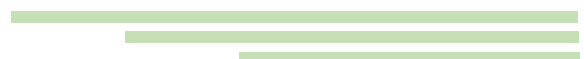
1.3 INOVASI BISNIS DENGAN BIG DATA DAN KECERDASAN BUATAN

Permintaan data telah meningkat selama beberapa tahun terakhir. Bisnis berlomba-lomba mengadopsi gudang data internal dan perangkat lunak analitik bisnis, serta beralih ke basis data publik dan privat untuk mencari data guna memacu strategi AI mereka. Karena meningkatnya permintaan, data menjadi komoditas yang berharga dan bisnis mulai bersaing untuk mendapatkan cadangan yang paling menguntungkan.

Hingga baru-baru ini, bisnis tidak menyadari bahwa mereka memiliki gudang data yang sangat besar dan tidak tahu apa yang harus dilakukan dengannya. Dengan kemajuan inovatif dalam penambangan data dan AI, bisnis kini dapat memanfaatkan data yang dihasilkan oleh konsumen dan pengguna. Misalnya, Moz menggunakan AI untuk memprediksi churn pelanggan menggunakan jaringan saraf pembelajaran mendalam yang menganalisis tindakan pengguna dan dapat memprediksi perilaku pengguna. Karena tindakan yang akan dilakukan pelanggan dalam sistem disebabkan oleh beberapa faktor dari masa lalu, hal ini memungkinkan untuk menggali beberapa wawasan bisnis yang berharga dan mengurangi churn pelanggan yang sudah ada, yang berdampak besar pada pertumbuhan perusahaan secara keseluruhan. Belakangan ini, aktivitas konsumen daring seperti kueri penelusuran, klik, atau pembelian menjadi sumber data utama bagi perusahaan-perusahaan besar. Namun, ternyata, data berlimpah di lingkungan fisik dan juga pengalaman luring kita. Perusahaan besar seperti Amazon telah menetapkan strategi pengawasan korporat di toko-toko swalayan. Sensor dan aktuator baru yang dipasang di toko-toko dapat mengumpulkan data tentang preferensi dan perilaku konsumen. Drone, asisten pribadi AI, dan bahkan *Internet of Things* (IoT) adalah alat yang dapat mengubah setiap momen kehidupan manusia menjadi data berharga. Data ini menjadi penggerak algoritma penetapan harga yang bereaksi terhadap perubahan permintaan konsumen. Uber telah mulai menggunakan model ini dalam mekanisme harganya. Bisnis-bisnis yang berada di ujung tombak inovasi semacam itu akan memiliki prospek terbaik untuk mengekstrak nilai dari perilaku konsumen.

Salah satu jalur yang paling menjanjikan adalah analisis sentimen yang menggunakan teknik NLP untuk memahami dinamika emosi dan umpan balik pengguna. Dengan analisis sentimen, seseorang juga dapat mengidentifikasi ulasan positif dan negatif terhadap produk mereka di platform e-commerce seperti Amazon. Selain itu, mengetahui sentimen yang terkait dengan pesaing Anda dapat membantu perusahaan menilai kinerja mereka sendiri dan menemukan cara untuk meningkatkannya. Salah satu manfaat analisis sentimen untuk mengelola reputasi online adalah otomatisasi, karena memproses banyak umpan balik pengguna secara manual bisa jadi sulit. Mengubah umpan balik menjadi data yang akan disalurkan ke perangkat lunak intelijen bisnis Anda adalah salah satu solusi paling efisien yang akan membedakan Anda dari pesaing.

Dari chatbot dan generator narasi cerdas hingga alat analitik bisnis, AI menjadi keunggulan kompetitif yang nyata bagi bisnis yang mempromosikan otomatisasi, pengurangan biaya, dan pengambilan keputusan yang cerdas. Namun, untuk mengembangkan strategi AI dan melatih model pembelajaran mesin mereka, bisnis membutuhkan data berkualitas tinggi. Facebook dan Google telah memecahkan masalah ini dengan memanfaatkan model



pengguna-dalam-lingkaran (user-in-the-loop) di mana pengguna menghasilkan data untuk mereka melalui postingan, komentar, atau kueri pencarian. Beberapa bisnis mendapatkan akses ke data dengan menjangkau basis data publik dan komersial, layanan pengumpulan dan klasifikasi data crowdsourcing, berkolaborasi dengan bisnis berbasis data, dll. Apa pun pendekatan yang paling sesuai dengan model bisnis Anda, Anda perlu memperkenalkan strategi akuisisi data yang efektif untuk memanfaatkan kekuatan AI.

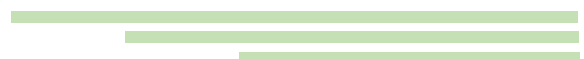
Tumpang Tindih Kecerdasan Buatan dengan Bidang Lain

Kecerdasan buatan adalah bidang yang memungkinkan mesin menjadi cerdas dan mengambil keputusan dengan justifikasi. Bidang ini juga menggunakan data dan membuat pembelajaran mesin. Kecerdasan buatan adalah bidang yang diterapkan secara luas di sebagian besar bidang lain dan dapat berkontribusi di domain apa pun. Kecerdasan buatan memiliki kemampuan untuk belajar dari sejumlah besar data, kemampuan untuk mensimulasikan perilaku yang terinspirasi oleh alam di samping model dan algoritma cerdas yang umum. Hal ini membuat kecerdasan buatan dapat diterapkan secara universal di mana model formal yang umum gagal.

Mungkin, perbedaan yang paling signifikan adalah kekuatan komputasi dan jumlah data yang dapat kita kumpulkan dan analisis dibandingkan dengan dekade-dekade sebelumnya. Ponsel pintar yang muat di telapak tangan saat ini dapat menyimpan dan memproses lebih banyak data daripada komputer mainframe tahun 1960an, yang menempati beberapa ruangan. Alih-alih mengandalkan kumpulan data yang dikurasi secara menyeluruh dan kecil, kita dapat menggunakan data yang besar dan tidak terorganisir dengan ribuan parameter untuk melatih algoritma dan membuat prediksi. Jumlah dan kualitas data inilah yang juga membedakan teknik pembelajaran mesin modern dari statistik. Meskipun statistik biasanya bergantung pada beberapa variabel untuk menangkap suatu pola, pembelajaran mesin dapat dimanfaatkan secara efektif dengan ribuan karakteristik data.

Pembelajaran mesin dianggap sebagai komponen integral ilmu komputer dan bidang yang terkait dengan TIK. Dalam bidang pembelajaran mesin, penekanan diberikan pada berbagai algoritma dan teknik untuk membuat mesin belajar secara otomatis dari data. Hasil ini kemudian digunakan dalam interpretasi dan penerapan data untuk pemecahan masalah. Namun, bidang pembelajaran mesin menerapkan beberapa teknik statistik dan matematika.

Istilah ilmu data dicetuskan kembali pada tahun 1960an. Seiring berkembangnya ilmu data dan mendapatkan "instrumen" baru dari waktu ke waktu, tujuan bisnis inti tetap berfokus pada menemukan pola yang bermanfaat dan menghasilkan wawasan berharga dari data. Saat ini, ilmu data digunakan di berbagai industri dan membantu dalam berbagai masalah analitis. Misalnya, dalam pemasaran, mengeksplorasi usia, jenis kelamin, lokasi, dan perilaku pelanggan memungkinkan pembuatan kampanye yang sangat tertarget, mengevaluasi seberapa besar kemungkinan pelanggan melakukan pembelian atau pergi. Dalam perbankan, menemukan tindakan klien yang tidak biasa membantu dalam mendeteksi penipuan. Dalam bidang kesehatan, analisis rekam medis pasien dapat menunjukkan kemungkinan menderita penyakit, dll.



Penambangan data juga berkaitan erat dengan pembelajaran mesin dan AI. Istilah "penambangan data" kurang tepat dan terdengar berbeda dari maknanya. Alih-alih menambang data itu sendiri, disiplin ilmu ini berfokus pada pembuatan algoritma untuk mengekstrak wawasan berharga dari data yang besar dan mungkin tidak terstruktur.

Masalah dasar penambangan data adalah memetakan data yang tersedia dan mengubahnya menjadi pola yang mudah dipahami. Penambangan data dianggap sebagai bagian dari proses yang lebih luas yang disebut penemuan pengetahuan dalam basis data (KDD) yang diperkenalkan pada tahun 1984 oleh Gregory Piatetsky-Shapiro. Beberapa teknik yang umum digunakan meliputi pengenalan pola, klasifikasi, partisi, dan pengelompokan, serta beberapa model statistik. Artinya, penambangan data juga memiliki beberapa tumpang tindih dengan statistik.

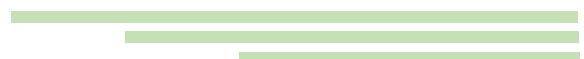
Di era yang didominasi oleh media sosial, personalisasi pelanggan menjadi salah satu sumber utama keunggulan kompetitif bagi perusahaan yang menawarkan produk dan layanan mereka secara daring. Alat analisis konsumen dan perangkat lunak AI canggih untuk mesin rekomendasi merupakan pengubah permainan utama yang memungkinkan personalisasi yang efisien dalam bisnis. Data tentang preferensi, minat, dan perilaku pengguna saat ini dan di masa lalu kini dapat dengan mudah dikumpulkan, disimpan, dan dianalisis menggunakan perangkat analitik bisnis dan algoritma AI.

Sebagai contoh, wawasan dari data ini memungkinkan pemasar untuk menyajikan konten yang relevan kepada pengunjung situs web, desainer gim video untuk menyesuaikan tingkat kesulitan dan fitur gim bagi pemain, atau mesin rekomendasi untuk menyarankan musik, video, atau produk yang mungkin disukai konsumen. Personalisasi yang didukung oleh data ini menjadi alat yang ampuh untuk mempertahankan konsumen dan menawarkan produk, layanan, dan fitur yang benar-benar mereka cari.

Masalah Etika dan Privasi

Sebagian besar aplikasi AI membutuhkan data dalam jumlah besar untuk belajar dan membuat keputusan yang cerdas. AI menjadi prioritas di sebagian besar sektor karena potensinya untuk meningkatkan layanan secara radikal, terobosan komersial, dan keuntungan finansial. Di masa depan, kita akan menghadapi berbagai dilema hukum dan etika dalam mencari keseimbangan antara kemajuan sosial yang signifikan atas nama AI dan hak privasi fundamental. Data dan algoritma yang membentuk AI tidak hanya harus akurat dan berkinerja tinggi; keduanya juga harus memenuhi persyaratan privasi dan memenuhi persyaratan regulasi.

Masalah data dapat terlihat jelas dalam industri yang sangat diatur seperti asuransi, yang bergeser dari model historis berdasarkan pengumpulan risiko menuju pendekatan yang menggabungkan elemen-elemen yang memprediksi risiko spesifik. Namun, beberapa atribut bersifat eksklusif. Misalnya, meskipun faktor jenis kelamin dan agama dapat digunakan untuk memprediksi beberapa risiko, faktor-faktor tersebut tidak dapat diterima oleh regulator dalam beberapa aplikasi dan yurisdiksi. Seiring perkembangan teknologi yang melampaui ekspektasi dan preferensi konsumen, bisnis melangkah di garis yang semakin tipis antara inisiatif AI, perlindungan privasi, dan layanan pelanggan. Misalnya, penyedia layanan keuangan



menggunakan teknologi pengenalan suara untuk mengidentifikasi pelanggan melalui telepon guna menghemat waktu verifikasi identitas. Pelanggan menyambut baik, alih-alih menolak, pengalaman ini, sebagian karena mereka menghargai layanan dan mempercayai perusahaan untuk tidak menyalahgunakan kapabilitas atau data yang mendukungnya.

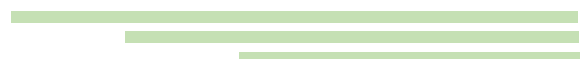
Peraturan perlindungan data Uni Eropa yang baru, yang mulai berlaku pada Mei 2018, akan memperkuat hak privasi kita, sekaligus memperketat persyaratan bagi mereka yang memproses data tersebut. Organisasi akan memikul lebih banyak tanggung jawab untuk memproses data pribadi sesuai dengan peraturan, dan persyaratan transparansi akan lebih ketat. Seiring dengan meningkatnya persyaratan, permintaan akan data juga meningkat.

Sistem berbasis AI hanya dapat menjadi cerdas jika memiliki cukup data relevan untuk dipelajari. Chatbot cerdas menganalisis semua informasi yang diberikan, kombinasi pertanyaan yang diajukan oleh pelanggan dan tanggapan yang dikomunikasikan oleh layanan pelanggan. Dari analisisnya, chatbot dapat "memahami" pertanyaan pelanggan dan karenanya mampu memberikan jawaban yang bermakna. Semakin banyak informasi yang dapat dijadikan dasar analisis chatbot, semakin baik dan akurat jawaban yang diberikannya.

Ketentuan Peraturan Perlindungan Data Umum (GDPR) mengatur tugas pengendali data dan hak subjek data ketika informasi pribadi diproses. Oleh karena itu, GDPR berlaku ketika kecerdasan buatan sedang dikembangkan dengan bantuan data pribadi dan juga ketika digunakan untuk menganalisis atau mencapai keputusan tentang individu. Aturan yang mengatur pemrosesan data pribadi memiliki dasar dalam beberapa prinsip dasar. Pasal 5 GDPR mencantumkan prinsip-prinsip yang berlaku untuk semua pemrosesan data pribadi. Inti dari prinsip-prinsip ini adalah bahwa informasi pribadi harus digunakan dengan cara yang melindungi privasi subjek data dengan cara sebaik mungkin, dan bahwa setiap orang berhak untuk memutuskan bagaimana data pribadinya digunakan. Penggunaan data pribadi dalam pengembangan kecerdasan buatan menantang beberapa prinsip ini. Singkatnya, prinsip-prinsip ini mensyaratkan bahwa data pribadi:

- ❖ Diproses secara sah, adil, dan transparan (prinsip legalitas, keadilan, dan transparansi)
- ❖ Dikumpulkan untuk tujuan tertentu yang dinyatakan secara tegas dan dapat dibenarkan, serta tidak diperlakukan dengan cara baru yang tidak sesuai dengan tujuan tersebut (prinsip pembatasan tujuan)
- ❖ Memadai, relevan, dan terbatas pada apa yang diperlukan untuk memenuhi tujuan pemrosesannya (prinsip minimisasi data)
- ❖ Benar dan jika perlu, diperbarui (prinsip akurasi)
- ❖ Tidak disimpan dalam bentuk yang dapat diidentifikasi untuk jangka waktu lebih lama dari yang diperlukan untuk tujuan tersebut (prinsip terkait periode penyimpanan data)
- ❖ Diproses dengan cara yang menjamin perlindungan data pribadi yang memadai (prinsip integritas dan kerahasiaan)

Kecerdasan buatan adalah teknologi yang berkembang pesat. Hal yang sama berlaku untuk perangkat dan metode yang dapat membantu memenuhi tantangan perlindungan data yang ditimbulkan oleh AI. Kami telah mengumpulkan beberapa contoh untuk mengilustrasikan



beberapa opsi yang tersedia. Metode-metode ini belum dievaluasi dalam praktiknya, tetapi dinilai berdasarkan potensinya. Artinya, secara teknis metode-metode ini mungkin tidak cocok saat ini, tetapi konsepnya menarik, dan memiliki potensi untuk penelitian lebih lanjut dan penggunaan di masa mendatang.

AI dan Analisis Prediktif

Analisis prediktif dan AI adalah dua hal yang berbeda. Ketika digabungkan, keduanya menghasilkan yang terbaik. AI memberdayakan analisis prediktif agar lebih cepat, lebih cerdas, dan lebih mudah ditindaklanjuti daripada sebelumnya. Ketika bisnis ingin membuat prediksi berbasis data tentang peristiwa masa depan, mereka mengandalkan analisis prediktif. Di era big data, analisis prediktif dengan cepat menjadi bagian penting dari banyak bisnis dan fungsi. Analisis prediktif adalah tentang penggunaan data historis untuk membuat prediksi. Contoh terbaik dari analisis prediktif adalah skor kredit.

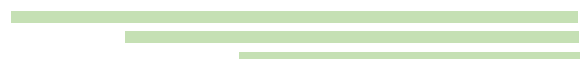
Skor tersebut didasarkan pada riwayat kredit Anda di masa lalu dan digunakan untuk memprediksi seberapa besar kemungkinan Anda melunasi utang. Meskipun analisis prediktif telah digunakan selama beberapa dekade dalam layanan keuangan, baru-baru ini menjadi alat penting dalam bisnis lain. Kemajuan teknologi pengumpulan dan pemrosesan data telah memungkinkan penerapan analitik prediktif pada hampir setiap aspek bisnis, mulai dari logistik, penjualan, hingga sumber daya manusia. Inti dari analitik prediktif adalah modelnya. Meskipun teknik statistik yang digunakan untuk membuat model bergantung pada tugas spesifik, teknik-teknik tersebut terbagi dalam dua jenis besar.

Pertama, model regresi, yang digunakan untuk mengukur korelasi antara variabel dan hasil tertentu. Koefisien yang dihasilkan memberikan ukuran kuantitatif dari hubungan tersebut, yaitu seberapa besar kemungkinan suatu hasil tertentu didasarkan pada serangkaian variabel. Jenis model lainnya adalah model klasifikasi. Jika model regresi menetapkan kemungkinan pada suatu peristiwa, model klasifikasi memprediksi apakah peristiwa tersebut termasuk dalam satu kategori atau lainnya.

Pemodelan dan analitik prediktif telah ada sejak lama. Namun, model ini kekurangan tiga hal penting untuk mendorong nilai pemasaran yang nyata: skala, kecepatan, dan penerapan. Di sinilah AI berperan. Dengan AI, model prediktif dapat memperhitungkan volume informasi waktu nyata yang luar biasa. Model semacam itu dapat mempertimbangkan lebih banyak informasi daripada sebelumnya, sehingga keluarannya lebih tepat dan dapat ditindaklanjuti. Lebih lanjut, AI dapat mengevaluasi miliaran variabel secara real-time dan dapat membuat keputusan simultan untuk menganalisis sejumlah besar peluang pemasaran per detik. Tanpa AI, model prediktif tidak dapat memahami volume data tersebut dengan cepat, dan model prediktif juga tidak memiliki kemampuan "kognitif" untuk mengambil tindakan.

Area Aplikasi

Manajemen hubungan pelanggan (CRM). Dengan menggunakan kombinasi analisis regresi dan teknik pengelompokan, perangkat CRM dapat memisahkan pelanggan perusahaan ke dalam kelompok-kelompok berdasarkan demografi dan posisi mereka dalam siklus hidup



pelanggan, sehingga memungkinkan Anda menargetkan upaya pemasaran dengan cara yang paling mungkin efektif.

Mendeteksi outlier dan penipuan. Jika sebagian besar aplikasi analitik prediktif mencari pola yang mendasarinya, deteksi anomali mencari item yang menonjol. Layanan keuangan telah menggunakannya untuk mendeteksi penipuan selama bertahun-tahun, tetapi teknik statistik yang sama juga berguna untuk aplikasi lain, termasuk penelitian medis dan farmasi.

Mengantisipasi permintaan. Tugas penting namun menantang bagi setiap bisnis adalah memprediksi permintaan produk dan layanan baru. Sebelumnya, prediksi semacam ini dibuat menggunakan data deret waktu untuk membuat prakiraan umum, tetapi kini peritel dapat menganonimkan data pencarian untuk memprediksi penjualan produk tertentu hingga ke tingkat regional.

Meningkatkan proses. Bagi produsen, produsen energi, dan bisnis lain yang bergantung pada mesin yang kompleks dan sensitif, analitik prediktif dapat meningkatkan efisiensi dengan mengantisipasi mesin dan komponen apa yang mungkin memerlukan perawatan. Dengan menggunakan data kinerja historis dan data sensor waktu nyata, model prediktif ini dapat meningkatkan kinerja dan mengurangi waktu henti sekaligus membantu mencegah penghentian kerja besar yang dapat terjadi ketika sistem utama tiba-tiba gagal.

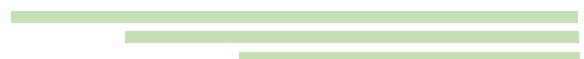
Membangun mesin rekomendasi. Rekomendasi yang dipersonalisasi diandalkan oleh layanan streaming, pengecer online, layanan kencan, dan lainnya untuk meningkatkan loyalitas dan keterlibatan pengguna. Teknik penyaringan kolaboratif menggunakan kombinasi perilaku masa lalu dan kesamaan dengan pengguna lain untuk menghasilkan rekomendasi, sementara penyaringan berbasis konten menetapkan karakteristik pada item dan merekomendasikan item terbaru berdasarkan kesamaannya dengan item sebelumnya.

Meningkatkan waktu perekrutan dan retensi. Bisnis dapat menggunakan data dari sistem sumber daya manusia untuk mengoptimalkan proses perekrutan mereka dan mengidentifikasi kandidat yang berhasil yang mungkin terlewatkan oleh penyaring manusia. Selain itu, beberapa departemen menggunakan campuran data kinerja dan profil kepribadian untuk mengidentifikasi kapan karyawan kemungkinan akan keluar atau mengantisipasi potensi konflik sehingga dapat diselesaikan secara proaktif.

Pengelompokan atau Segmentasi

Pengelompokan adalah proses pengorganisasian objek ke dalam kelompok-kelompok yang anggotanya serupa dalam beberapa hal. Sedangkan, segmentasi pelanggan adalah praktik membagi basis pelanggan menjadi kelompok-kelompok individu yang serupa dalam hal-hal tertentu yang relevan dengan pemasaran, seperti usia, jenis kelamin, minat, kebiasaan belanja, dan sebagainya. Segmentasi atau pengelompokan pelanggan bermanfaat dalam banyak hal. Ini dapat digunakan untuk pemasaran yang tertarget. Terkadang, ketika membangun model prediktif, cukup efektif untuk mengelompokkan data dan membangun model prediktif terpisah untuk setiap kluster.

Pengelompokan adalah teknik penambangan data tak terarah. Ini berarti dapat digunakan untuk mengidentifikasi pola dan struktur tersembunyi dalam data tanpa merumuskan hipotesis tertentu. Tidak ada variabel target dalam pengelompokan. Misalnya,



pengecer bahan makanan tidak secara aktif mencoba mengidentifikasi pecinta makanan segar pada awal analisis. Penelitian ini hanya mencoba memahami berbagai perilaku pembelian basis pelanggannya.

Pengelompokan dilakukan untuk mengidentifikasi kesamaan terkait perilaku atau dimensi tertentu. Misalnya, kami ingin mengidentifikasi segmen pelanggan dengan perilaku pembelian yang serupa. Oleh karena itu, pengelompokan dilakukan menggunakan variabel yang mewakili pola pembelian pelanggan.

Analisis kluster dapat digunakan untuk menemukan struktur dalam data tanpa memberikan penjelasan atau interpretasi. Analisis kluster hanya menemukan pola dalam data tanpa menjelaskan mengapa pola tersebut ada. Kluster yang dihasilkan tidak memiliki makna tersendiri. Kluster-kluster tersebut perlu diprofilkan secara ekstensif untuk membangun identitasnya, yaitu untuk memahami apa yang mereka wakili dan bagaimana mereka berbeda dari populasi induk.

Pengelompokan terutama digunakan untuk melakukan segmentasi, baik itu pelanggan, produk, maupun toko. Misalnya, produk dapat dikelompokkan ke dalam kelompok hierarkis berdasarkan atributnya seperti penggunaan, ukuran, merek, rasa, dll.; toko dengan karakteristik serupa penjualan, ukuran, basis pelanggan, dll, dapat dikelompokkan bersama.

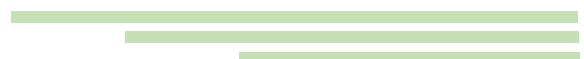
Prosedur pengelompokan dapat bersifat hierarkis di mana pengelompokan dicirikan oleh pengembangan hierarki atau struktur seperti pohon.

- Pengelompokan aglomeratif dimulai dengan setiap objek dalam kluster terpisah dan kluster dibentuk dengan mengelompokkan objek ke dalam kluster yang lebih besar.
- Pengelompokan divisif di sisi lain dimulai dengan semua objek dikelompokkan ke dalam satu kluster dan kluster kemudian dibagi atau dipecah hingga setiap objek berada dalam kluster terpisah.
- Pengelompokan K -means adalah pengelompokan non-hierarkis dan merupakan prosedur yang pertama-tama menetapkan atau menentukan pusat kluster dan kemudian mengelompokkan semua objek dalam nilai ambang batas yang telah ditentukan bersama-sama mulai dari pusat.

Penentuan jumlah kluster didasarkan pada pertimbangan teoretis atau praktis. Dalam pengelompokan hierarkis, jarak penggabungan kluster dapat digunakan sebagai kriteria. Dalam pengelompokan non-hierarkis, rasio total varians dalam grup terhadap varians antar grup dapat diplot terhadap jumlah kluster.

Penafsiran dan pembuatan profil kluster melibatkan pemeriksaan sentroid kluster. Senroid tersebut mewakili nilai rata-rata objek yang terdapat dalam kluster pada setiap variabel. Senroid dapat diberi nama atau label. Untuk menilai reliabilitas dan validitas, seseorang harus melakukan analisis kluster pada data yang sama menggunakan ukuran jarak yang berbeda dan membandingkan hasilnya untuk menentukan stabilitas solusi.

Membagi data secara acak menjadi dua bagian dan melakukan pengelompokan secara terpisah pada setiap bagian dan membandingkan sentroid kluster di dua sub-sampel adalah salah satu cara favorit saya. Dalam pengelompokan hierarkis, solusinya mungkin bergantung pada urutan kasus dalam set data. Untuk mencapai hasil terbaik, lakukan beberapa kali



percobaan menggunakan urutan kasus yang berbeda hingga solusinya stabil. Pengelompokan juga dapat digunakan untuk deteksi anomali, misalnya, mengidentifikasi transaksi penipuan.

Metode deteksi klaster dapat digunakan pada sampel yang hanya berisi transaksi valid untuk menentukan bentuk dan ukuran klaster "normal". Ketika sebuah transaksi muncul dan berada di luar klaster karena alasan apa pun, transaksi tersebut patut dicurigai. Pendekatan ini telah digunakan dalam bidang kedokteran untuk mendeteksi keberadaan sel abnormal dalam sampel jaringan dan dalam telekomunikasi untuk mendeteksi pola panggilan yang mengindikasikan penipuan. Pengelompokan sering digunakan untuk memecah kumpulan data besar menjadi kelompok-kelompok yang lebih kecil yang lebih mudah diterapkan dengan teknik lain. Misalnya, hasil regresi logistik dapat ditingkatkan dengan melakukannya secara terpisah pada klaster yang lebih kecil yang berperilaku berbeda dan mungkin mengikuti distribusi yang sedikit berbeda.

Singkatnya, pengelompokan adalah teknik yang ampuh untuk mengeksplorasi struktur pola dalam data dan memiliki aplikasi yang luas dalam analitik bisnis. Ada berbagai metode untuk pengelompokan. Seorang analis harus terbiasa dengan berbagai algoritma pengelompokan dan harus mampu menerapkan teknik yang paling relevan sesuai kebutuhan bisnis.

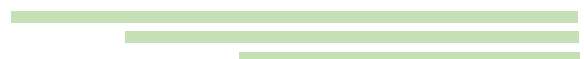
Persona Psikografis

Psikografis adalah indikator minat, perilaku, sikap, dan opini seseorang yang membantu memahami alasan mengapa seseorang mungkin/tidak mungkin membeli suatu produk. Data psikografis, jika digabungkan dengan data demografis, dapat memberikan gambaran yang hampir lengkap tentang persona tersebut dan membantu Anda memilih jenis produk yang sesuai dengan persona tersebut.

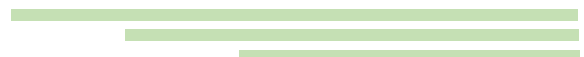
Parameter penargetan psikografis untuk suatu persona ditentukan oleh kecenderungan psikologis sekelompok orang untuk berperilaku dengan cara tertentu atau tertarik pada hal-hal serupa. Jadi, bagi seorang ibu muda, parameter psikografis mencakup minat untuk menjelajahi sumber daya yang memberinya pengetahuan tentang cara merawat bayinya. Di dunia maya, indikator untuk menentukan jiwa seorang persona mencakup aktivitas penelusuran sebelumnya, aktivitas di situs web, riwayat pembelian sebelumnya, klaim minat di halaman jejaring sosial, dan data lainnya. Data psikografis, yang dikumpulkan dan disatukan, dapat memberikan wawasan yang sangat baik tentang jenis produk apa yang mungkin diminati atau dapat dibeli oleh seorang persona.

Segmentasi pasar adalah proses pemisahan pasar menjadi segmen atau kelompok konsumen yang serupa, tetapi berbeda dari konsumen di kelompok lain. Segmentasi membagi pasar menjadi subkelompok. Pemasaran target melibatkan penentuan segmen mana yang paling menguntungkan. Lebih lanjut, penentuan posisi melibatkan penciptaan citra produk yang menarik bagi satu atau beberapa pasar target.

Segmentasi psikografis membantu membangun produk atau memposisikannya sedemikian rupa sehingga membuatnya lebih menarik daripada pesaing. Membuat peta persepsi membantu Anda memahami bagaimana konsumen melihat merek Anda dan memungkinkan Anda memposisikan merek Anda untuk manfaat maksimal. AI mengumpulkan pelanggan ke dalam kelompok audiens berdasarkan titik kontak dan analisis sentimen, yang membantu



pemasar memahami bagaimana berbagai segmen pelanggan mungkin bereaksi terhadap postingan sosial, papan reklame, atau blog. Dengan mempertimbangkan cara pelanggan berbicara satu sama lain, AI dapat menyarankan frasa dan suasana hati yang paling sesuai dengan setiap segmen audiens.



BAB 2

PEMBELAJARAN MESIN

2.1 PENDAHULUAN

Konsep pembelajaran mesin pertama kali diperkenalkan pada tahun 1950an, yang luar biasa sebagai era pelopor AI. Pada tahun 1950, Alan Turing menerbitkan makalah "*Computing Machinery and Intelligence*" yang mengusulkan uji evaluasi AI terkenal yang kita kenal sekarang sebagai Uji Turing. Pada tahun 1959, Arthur Lee Samuel menciptakan istilah "pembelajaran mesin". Pembelajaran mesin (ML) dapat didefinisikan secara luas sebagai metode komputasi yang menggunakan pengalaman untuk meningkatkan kinerja atau membuat prediksi yang akurat.

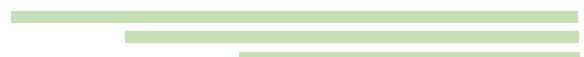
Kami mendefinisikan pembelajaran mesin sebagai serangkaian manipulasi matematika yang dilakukan pada data penting untuk mendapatkan wawasan berharga. Ini adalah studi tentang algoritma yang belajar dari contoh dan pengalaman, alih-alih aturan yang kaku. Umumnya, ada tiga jenis utama permasalahan pembelajaran mesin: terawasi, tak terawasi, dan penguatan.

- ✓ Permasalahan pembelajaran mesin terawasi adalah permasalahan di mana kita ingin membuat prediksi berdasarkan serangkaian contoh.
- ✓ Masalah pembelajaran mesin tanpa pengawasan adalah masalah di mana data kita tidak memiliki serangkaian kategori yang ditentukan, tetapi kita mencari algoritma pembelajaran mesin untuk membantu kita mengatur data tersebut.

Artinya, masalah pembelajaran mesin tanpa pengawasan memiliki serangkaian titik data historis yang ingin kita gunakan untuk memprediksi masa depan, sedangkan masalah pembelajaran mesin tanpa pengawasan memiliki serangkaian data yang kita cari bantuan pembelajaran mesin untuk mengatur atau memahaminya.

Penguatan mencakup tugas atau tujuan spesifik yang harus diselesaikan sistem. Sepanjang proses, sistem menerima umpan balik untuk mempelajari perilaku yang diinginkan. Misalnya, sistem mengalami kesalahan saat melakukan tindakan atau mendapatkan hadiah karena mencapai hasil yang paling menguntungkan. Dengan demikian, program dapat mempelajari pendekatan yang paling efektif melalui sinyal penguatan. Meskipun penambangan data dan penemuan pengetahuan dalam basis data (KDD) tampaknya hanya mengatasi masalah utama ilmu data, pembelajaran mesin justru meningkatkan efisiensi bisnis. Teknik pembelajaran mesin secara garis besar dapat dibagi menjadi empat area berbeda: klasifikasi, pengelompokan, pembelajaran asosiasi, dan prediksi numerik. Klasifikasi yang diterapkan pada teks merupakan subjek kategorisasi teks, yaitu tugas mengurutkan sekumpulan dokumen secara otomatis ke dalam kategori (atau kelas, atau topik) dari sekumpulan dokumen yang telah ditentukan sebelumnya.

Klasifikasi dokumen yang sederhana digunakan dalam pengindeksan dokumen untuk sistem temu kembali informasi, penyaringan teks (termasuk perlindungan dari spam email), kategorisasi halaman web, dan banyak aplikasi lainnya. Klasifikasi juga dapat digunakan pada



bagian teks yang lebih kecil (paragraf, kalimat, kata) tergantung pada aplikasi konkretnya, seperti segmentasi dokumen, pelacakan topik, atau disambiguasi makna kata. Dalam pendekatan pembelajaran mesin, algoritma klasifikasi (pengklasifikasi) dilatih terlebih dahulu pada data berlabel yang telah diurutkan sebelumnya, sebelum diterapkan untuk mengurutkan teks yang belum terlihat.

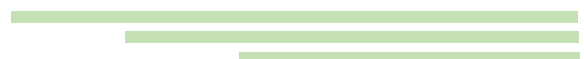
Penggunaan teknik pengelompokan dengan teks dapat dicapai pada dua tingkat. Menganalisis kumpulan dokumen dengan mengidentifikasi kluster yang serupa hanya memerlukan pemanfaatan algoritma pengelompokan yang telah dikenal dan pengukuran kesamaan dokumen. Pengelompokan dalam dokumen bisa sedikit lebih menantang, karena memerlukan prapemrosesan teks dan isolasi objek untuk pengelompokan kalimat, kata, atau beberapa konstruksi yang memerlukan derivasi.

Pembelajaran asosiasi pada dasarnya merupakan generalisasi dari klasifikasi, yang bertujuan untuk menangkap hubungan antara fitur arbitrer (juga disebut atribut) dari contoh dalam suatu set data. Dalam hal ini, klasifikasi hanya menangkap hubungan semua fitur dengan satu fitur yang menentukan kelasnya. Penerapan pembelajaran asosiasi secara langsung pada teks tidak terlalu memungkinkan karena tingginya dimensionalitas representasi dokumen, yaitu banyaknya fitur (yang banyak di antaranya mungkin tidak terlalu informatif). Memanfaatkan pembelajaran asosiasi pada informasi yang diekstraksi dari teks (misalnya, menggunakan klasifikasi dan/atau pengelompokan) adalah hal yang berbeda dan dapat menghasilkan banyak wawasan yang bermanfaat.

Prediksi numerik (juga disebut regresi, dalam arti kata yang lebih luas) adalah generalisasi lain dari klasifikasi, di mana fitur kelas tidak diskrit melainkan kontinu. Pergeseran kecil dalam definisi ini menghasilkan perbedaan besar dalam internal algoritma klasifikasi dan regresi. Namun, dengan membagi fitur numerik yang diprediksi menjadi sejumlah interval terbatas, setiap algoritma regresi juga dapat digunakan untuk klasifikasi. Kebalikannya biasanya tidak memungkinkan. Sekali lagi, seperti halnya pembelajaran asosiasi, penerapan regresi sederhana pada teks tidak terlalu berguna, kecuali untuk klasifikasi (terutama ketika ukuran keyakinan diperlukan, tetapi ini juga dapat dicapai dengan sebagian besar algoritma klasifikasi).

Ada perbedaan antara penambahan data dan pembelajaran mesin yang sangat populer. Namun, pembelajaran mesin adalah tentang menciptakan algoritma untuk mengekstrak wawasan berharga, sangat berfokus pada penggunaan berkelanjutan dalam lingkungan yang berubah secara dinamis dan menekankan pada penyesuaian, pelatihan ulang, dan pembaruan algoritma berdasarkan pengalaman sebelumnya. Tujuan pembelajaran mesin adalah untuk terus beradaptasi dengan data baru dan menemukan pola atau aturan baru di dalamnya. Terkadang hal itu dapat diwujudkan tanpa panduan manusia dan pemrograman ulang yang eksplisit.

Pembelajaran mesin merupakan bidang ilmu data yang paling pesat perkembangannya saat ini berkat sejumlah terobosan teoretis dan teknologi terkini. Terobosan-terobosan tersebut menghasilkan pemrosesan bahasa alami, pengenalan gambar, atau bahkan



pembuatan gambar, musik, dan teks baru oleh mesin. Pembelajaran mesin tetap menjadi "instrumen" utama dalam membangun kecerdasan buatan.

Untuk menggunakan pembelajaran mesin dalam aplikasi atau bahkan mempelajarinya, ada dua cara. Pertama, mempelajari cara menggunakan pustaka yang bertindak sebagai kotak hitam, yaitu menyediakan fungsionalitas yang berbeda. Kedua, mempelajari cara menulis algoritma dan menemukan koefisien, menyesuaikan model, menemukan titik optimasi, dan banyak lagi sehingga Anda dapat mengkurasi aplikasi sesuai kebutuhan. Namun, jika Anda hanya ingin bereksperimen, ada beberapa pustaka dan antarmuka pemrograman aplikasi yang dapat membantu Anda.

Berbagai bisnis menggunakan teknologi pembelajaran mesin untuk menganalisis riwayat pembelian pelanggan mereka dan membuat rekomendasi produk yang dipersonalisasi untuk pembelian berikutnya. Kemampuan untuk menangkap, menganalisis, dan menggunakan data pelanggan untuk memberikan pengalaman berbelanja yang dipersonalisasi ini adalah masa depan penjualan dan pemasaran. Di sektor transportasi, berdasarkan riwayat perjalanan dan pola perjalanan di berbagai rute, pembelajaran mesin dapat membantu perusahaan transportasi memprediksi potensi masalah yang mungkin timbul pada rute tertentu dan dengan demikian menyarankan pelanggan mereka untuk memilih rute yang berbeda. Perusahaan transportasi dan logistik secara bertahap menggunakan teknologi pembelajaran mesin untuk melakukan analisis data dan pemodelan data guna membuat keputusan yang tepat dan membantu pelanggan mereka membuat keputusan cerdas saat bepergian.

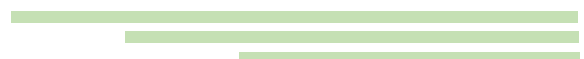
2.2 ALUR KERJA PEMBELAJARAN MESIN

Perbedaan utama antara pembelajaran mesin dan algoritma terprogram tradisional adalah kemampuannya untuk memproses data tanpa diprogram secara eksplisit. Ini berarti bahwa seorang insinyur tidak diharuskan memberikan instruksi yang rumit kepada mesin tentang cara menangani setiap jenis rekaman data. Sebaliknya, mesin mendefinisikan aturan-aturan ini sendiri berdasarkan data masukan.

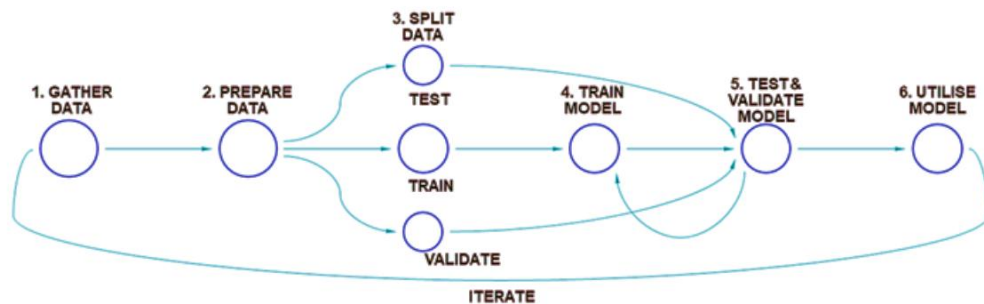
Terlepas dari aplikasi pembelajaran mesin yang digunakan, alur kerja inti tetap sama dan berulang secara iteratif setelah hasilnya menjadi usang atau membutuhkan akurasi yang lebih tinggi. Bagian ini difokuskan pada pengenalan konsep-konsep dasar yang membentuk alur kerja pembelajaran mesin seperti yang diilustrasikan pada Gambar 2.1.

Alur kerja mengikuti langkah-langkah berikut:

1. *Mengumpulkan data.* Gunakan infrastruktur TI Anda untuk mengumpulkan sebanyak mungkin data yang sesuai dan menyatukannya menjadi satu set data.
2. *Siapkan data.* Siapkan data Anda untuk diproses dengan cara yang paling praktis. Prosedur pra-pemrosesan dan pembersihan data bisa jadi cukup rumit, tetapi umumnya bertujuan untuk mengisi nilai yang hilang dan memperbaiki kekurangan lain dalam data, seperti representasi yang berbeda dari nilai yang sama dalam satu kolom.
3. *Pisahkan data.* Pisahkan subset data untuk melatih model dan mengevaluasi lebih lanjut kinerjanya terhadap data baru.



4. *Latih model.* Gunakan subset data historis agar algoritme dapat mengenali pola di dalamnya.
5. *Uji dan validasi model.* Evaluasi performa model menggunakan subset pengujian dan validasi data historis dan pahami seberapa akurat prediksinya.
6. *Manfaatkan model.* Tanamkan model yang telah diuji ke dalam konteks pengambilan keputusan Anda sebagai bagian dari solusi analitik.
7. *Iterasi.* Kumpulkan data baru setelah menggunakan model untuk meningkatkannya secara bertahap.



Gambar 2.1 Alur kerja pembelajaran mesin

Artefak penting dari setiap eksekusi pembelajaran mesin adalah model matematika, yang menjelaskan bagaimana algoritme memproses data baru setelah dilatih dengan subset data historis. Tujuan pelatihan adalah untuk mengembangkan model yang mampu merumuskan nilai target (atribut), suatu nilai yang tidak diketahui dari setiap objek data.

Misalnya, Anda perlu memprediksi apakah pelanggan toko e-commerce Anda akan melakukan pembelian atau pergi. Prediksi ini membeli atau meninggalkan atribut target yang kita cari. Untuk melatih model dalam melakukan prediksi jenis ini, Anda "memberi makan" algoritma dengan kumpulan data yang menyimpan berbagai catatan perilaku pelanggan dan hasilnya, seperti apakah pelanggan meninggalkan atau menyelesaikan pembelian. Dengan mempelajari data historis ini, model akan dapat membuat prediksi untuk data mendatang.

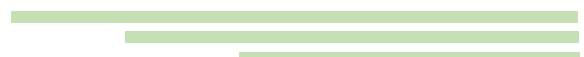
Teknik Pembelajaran Mesin

Setelah memahami konsep dasar machine learning, berikut ini dijelaskan dua teknik pembelajaran mesin yang penting untuk diketahui:

- **Supervised Learning**

Secara sederhana, supervised learning adalah teknik pembelajaran mesin yang menggunakan data berlabel sebagai panduan. Dengan teknik ini, mesin belajar untuk memetakan input ke output yang diinginkan berdasarkan pengalaman dari data sebelumnya. Misalnya, Anda memiliki berbagai laporan keuangan yang sudah diklasifikasikan dalam kategori tertentu, seperti neraca dan laporan tahunan. Ketika laporan keuangan baru dibuat, mesin dapat mengidentifikasi isi laporan tersebut dan secara otomatis mengelompokkan dokumen ke dalam kategori yang tepat.

- **Unsupervised Learning**



Teknik unsupervised learning digunakan ketika data yang tersedia tidak memiliki label atau informasi pembimbing. Teknik ini membantu dalam menemukan pola atau struktur tersembunyi dalam data yang tidak terorganisir. Berbeda dengan supervised learning, pada teknik ini Anda tidak memiliki contoh data sebelumnya. Contohnya, jika Anda belum pernah memiliki data investor dan kemudian mendapatkan investor baru, Anda dapat mengelompokkan investor tersebut berdasarkan jumlah investasi yang mereka berikan dengan mengidentifikasi kategori yang sesuai sebelum menyimpan data tersebut.

2.3 ALGORITMA PEMBELAJARAN

Regresi Linier

Model linier menggunakan rumus sederhana untuk menemukan garis "paling sesuai" melalui serangkaian titik data. Anda menemukan variabel yang ingin Anda prediksi melalui persamaan variabel yang Anda ketahui. Untuk menemukan prediksi, kita memasukkan variabel yang kita ketahui untuk mendapatkan jawaban kita. Dengan kata lain, untuk mengetahui berapa lama waktu yang dibutuhkan kue untuk dipanggang, kita cukup memasukkan bahan-bahannya. Ada berbagai bentuk algoritma model linier.

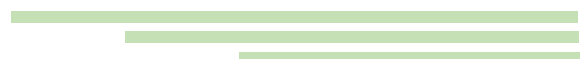
Regresi linier, juga dikenal sebagai "regresi kuadrat terkecil", adalah bentuk model linier yang paling standar. Untuk masalah regresi (variabel yang kita coba prediksi bersifat numerik), regresi linier adalah model linier yang paling sederhana.

Algoritma regresi umumnya digunakan untuk analisis statistik dan merupakan algoritma kunci dalam pembelajaran mesin. Algoritma regresi membantu analisis memodelkan hubungan antar titik data. Algoritma regresi mendefinisikan nilai target numerik, alih-alih kelas. Dengan memperkirakan variabel numerik, algoritma ini ampuh dalam memprediksi permintaan produk, angka penjualan, laba pemasaran, dll. Misalnya:

- *Berapa banyak barang dari produk ini yang dapat kita jual tahun ini?*
- *Berapa biaya perjalanan untuk kota ini?*
- *Berapa kecepatan maksimum mobil untuk mempertahankan masa pakainya?*

Algoritma regresi dapat mengukur kekuatan korelasi antar variabel dalam suatu set data. Selain itu, analisis regresi dapat berguna untuk memprediksi nilai data di masa mendatang berdasarkan nilai historis. Namun, penting untuk diingat bahwa analisis regresi mengasumsikan bahwa korelasi berkaitan dengan sebab-akibat. Tanpa memahami konteks seputar data, analisis regresi dapat menghasilkan prediksi yang tidak akurat. Regresi logistik hanyalah adaptasi regresi linier untuk masalah klasifikasi (variabel yang kita coba prediksi adalah jawaban "Ya/Tidak"). Regresi logistik sangat baik untuk masalah klasifikasi karena bentuknya.

Baik regresi linier maupun regresi logistik memiliki kelemahan yang sama. Keduanya memiliki kecenderungan untuk "overfit", yang berarti model beradaptasi terlalu tepat terhadap data sehingga mengorbankan kemampuan untuk menggeneralisasi ke data yang sebelumnya tidak terlihat. Dengan demikian, kedua model tersebut seringkali "terregularisasi", yang berarti mereka memiliki penalti tertentu untuk mencegah overfit.

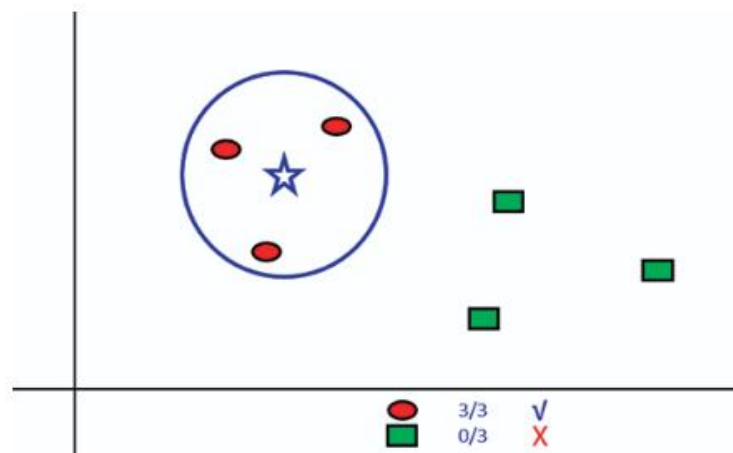


Kelemahan lain dari model linier adalah, karena sederhana, mereka cenderung kesulitan memprediksi perilaku yang lebih kompleks.

***k*-Nearest Neighbor**

Algoritma *k*-nearest neighbor adalah metode untuk mengklasifikasikan objek berdasarkan contoh pelatihan terdekat dalam ruang fitur. Algoritma ini memeriksa ruang fitur dan dapat dengan yakin memberikan prediksi berdasarkan tetangga terdekat. Algoritma ini bekerja dengan fakta bahwa objek yang dekat akan memiliki nilai prediksi yang serupa, dan setelah kita mengetahui nilai prediksi suatu objek, mudah untuk memprediksi tetangga terdekatnya.

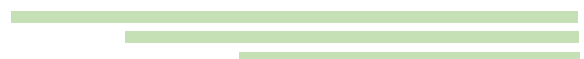
Algoritma *k*-tetangga terdekat (*k*-nearest neighbor) adalah salah satu algoritma pembelajaran mesin yang paling sederhana dan sering disebut sebagai lazy learner karena hanya bergantung pada prediksi dari pilihan instans tertentu yang paling mirip dengan instans set pengujian. Sampel pelatihan dijelaskan oleh atribut numerik *n*-dimensi. Setiap sampel merepresentasikan sebuah titik dalam ruang *n*-dimensi; semua sampel pelatihan disimpan dalam ruang pola *n*-dimensi. Ketika terdapat sampel yang tidak diketahui, algoritma mencari *k* sampel pelatihan yang paling dekat dengan sampel yang tidak diketahui di ruang pola. *K* sampel pelatihan tersebut merupakan "*k* nearest neighbor" dari sampel yang tidak diketahui, seperti yang ditunjukkan pada Gambar 2.2.



Gambar 2.2 *k*-nearest neighbor

k-nearest neighbor terbukti menjadi metode yang sangat efektif untuk data pelatihan yang bising. *K*-nearest neighbor berkinerja baik dalam hal otomatisasi, karena banyak algoritma yang robust dan memberikan prediksi yang baik ketika kita memiliki kumpulan data dengan data yang hilang. *K*-nearest neighbor dapat ditingkatkan jika kita memiliki versi pohon keputusan untuk melakukan praproses pada kumpulan data tersebut.

Praproses dapat diterapkan dalam beberapa cara pada kumpulan data kita dengan mempertimbangkan sifat kumpulan data tersebut; jika kumpulan data terdiri dari atribut numerik, algoritma diskritisasi akan sangat efektif dengan mengurangi jumlah tupel dari kumpulan data ke interval yang jauh lebih sedikit. Keluaran setelah diskritisasi akan



dimasukkan ke dalam k -nearest neighbor. Dengan ini, prosesnya akan jauh lebih cepat karena lebih sedikit tupel yang dipertimbangkan.

Pohon Keputusan

Pohon keputusan bekerja dalam bentuk struktur pohon dari atas yang disebut sebagai simpul akar hingga ke daun; Setiap cabang mewakili hasil pengujian, dan simpul daun mewakili kelas. Untuk mengklasifikasikan sampel yang tidak diketahui, kami menguji atribut sampel terhadap pohon keputusan. Sebuah jalur dilacak dari akar, yaitu puncak pohon, ke simpul daun yang menyimpan prediksi kelas untuk sampel tersebut. Pohon keputusan rentan terhadap banyak gangguan, dan teknik standar untuk menangani hal ini adalah memangkas pohon. Pemangkasan melibatkan penghapusan kondisi apa pun dalam antesedennya yang tidak meningkatkan akurasi estimasi aturan. Proses ini dimaksudkan untuk meningkatkan akurasi klasifikasi pada data yang tidak terlihat.

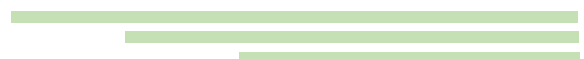
Untuk membuat atau melatih pohon keputusan, kami mengambil data yang kami gunakan untuk melatih model dan menemukan atribut mana yang paling baik membagi set pelatihan sehubungan dengan target. Misalnya, pohon keputusan dapat digunakan dalam deteksi penipuan kartu kredit. Kami akan menemukan atribut yang paling baik memprediksi risiko penipuan adalah jumlah pembelian (misalnya, seseorang dengan kartu kredit telah melakukan pembelian yang sangat besar). Ini bisa menjadi pemisahan pertama (atau percabangan), kartu-kartu yang memiliki nilai pembelian luar biasa tinggi dan yang tidak. Kemudian, kita menggunakan atribut terbaik kedua (misalnya, kartu kredit tersebut sering digunakan) untuk membuat pemisahan berikutnya. Kita kemudian dapat melanjutkan hingga kita memiliki atribut yang cukup untuk memenuhi kebutuhan kita.

Algoritma klasifikasi menentukan kategori objek dari kumpulan data. Dengan demikian, kategori biasanya dikaitkan sebagai kelas. Dengan memecahkan masalah klasifikasi, Anda dapat menjawab berbagai pertanyaan:

- *Apakah email ini spam atau bukan?*
- *Apakah transaksi ini duplikasi atau bukan?*
- *Jenis produk apa yang kemungkinan besar akan dibeli oleh pembeli ini: sofa, meja makan, atau kursi taman?*

Skalabilitas merupakan masalah signifikan dengan pohon keputusan karena tidak berskala dengan baik pada kumpulan data besar, dan dalam penambahan data, set pelatihan umumnya terdiri dari jutaan sampel. Masalah skalabilitas muncul karena set pelatihan disimpan di memori utama. Batasan ini membatasi skalabilitas algoritma tersebut, di mana konstruksi pohon keputusan dapat menjadi tidak efisien karena pertukaran sampel pelatihan masuk dan keluar dari memori utama dan cache.

Salah satu pilihannya adalah mendiskritisasi atribut kontinu dan melakukan pengambilan sampel di setiap node. Namun, hal ini juga memiliki inefisiensi tersendiri. Pilihan lain adalah mempartisi pohon keputusan yang besar menjadi subset dan membangun pohon keputusan dari subset tersebut. Karena kita hanya bekerja pada subset, akurasi hasil kita tidak sebaik jika kita menggunakan seluruh data.



Kelemahan pohon keputusan adalah sifat serakah dari algoritmanya. Sifat serakah ini berarti algoritma membuat komitmen terhadap pilihan tertentu terlalu dini sehingga mencegah algoritma menemukan solusi keseluruhan terbaik di kemudian hari. Pohon keputusan sangat cepat dan akurasi klasifikasi secara alami tinggi untuk data yang pemetaan kelasnya terdiri dari wilayah panjang dan tipis dalam ruang konsep.

Penyempurnaan teknik pembelajaran ini dapat dilakukan dengan memodifikasi algoritma untuk menangani atribut bernilai kontinu. Pohon keputusan memiliki atribut yang tangguh terhadap banyak jenis prediktor. Hal ini membuatnya cocok sebagai metode praproses yang baik untuk algoritma lain. Contohnya adalah melakukan praproses data untuk jaringan saraf; karena kecepatannya, pohon keputusan akan dengan mudah melakukan first pass pada data yang akan menghasilkan subset prediktor yang akan dimasukkan ke dalam jaringan saraf atau k -nearest neighbor. Hal ini tentunya akan mengurangi kandungan derau yang harus ditangani oleh jaringan saraf tiruan, dan ini tentunya akan meningkatkan kinerja jaringan saraf tiruan.

Tugas klasifikasi lain yang sangat presisi adalah deteksi anomali. Hal ini biasanya dikenal sebagai klasifikasi satu kelas karena tujuan deteksi anomali adalah untuk menemukan outlier, objek yang tidak biasa dalam data yang tidak muncul dalam distribusi normalnya. Masalah apa saja yang dapat dipecahkannya:

- *Apakah ada pembeli dengan kualitas yang berbeda dalam kumpulan data kita?*
- *Dapatkah kita menemukan perilaku yang tidak biasa di antara nasabah asuransi kita?*

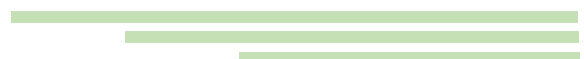
Konstruksi Fitur dan Reduksi Data

Peran representasi telah diakui sebagai isu krusial dalam AI dan ML. Dalam paradigma pembelajaran dari contoh dan representasi atribut-nilai data masukan, representasi asli merupakan vektor atribut (fitur, variabel) yang mendeskripsikan contoh (instans, objek). Proses transformasi atribut masukan, yang digunakan dalam konstruksi fitur, dapat dirumuskan sebagai berikut: dengan vektor asli fitur dan set pelatihan, buatlah representasi turunan yang lebih baik dengan beberapa kriteria (misalnya akurasi prediktif, ukuran representasi).

Atribut baru yang telah ditransformasi dapat menggantikan atribut asli atau dapat ditambahkan ke deskripsi contoh. Contoh transformasi atribut adalah penghitungan, pengelompokan, konstruksi/diskretisasi interval, penskalaan, perataan, normalisasi (nilai numerik), pengelompokan, analisis komponen utama, dll. Banyak transformasi yang dimungkinkan dengan menerapkan berbagai macam rumus matematika, tetapi dalam praktiknya, hanya sejumlah transformasi terbatas yang efektif.

Hutan Acak

Hutan acak adalah rata-rata dari beberapa pohon keputusan, yang masing-masing dilatih dengan sampel data acak. Setiap pohon tunggal di hutan lebih lemah daripada pohon keputusan penuh, tetapi dengan menggabungkan semuanya, kita mendapatkan kinerja keseluruhan yang lebih baik berkat keragaman. Hutan acak adalah algoritma yang sangat umum dalam pembelajaran mesin saat ini. Sangat mudah dilatih, dan cenderung berkinerja

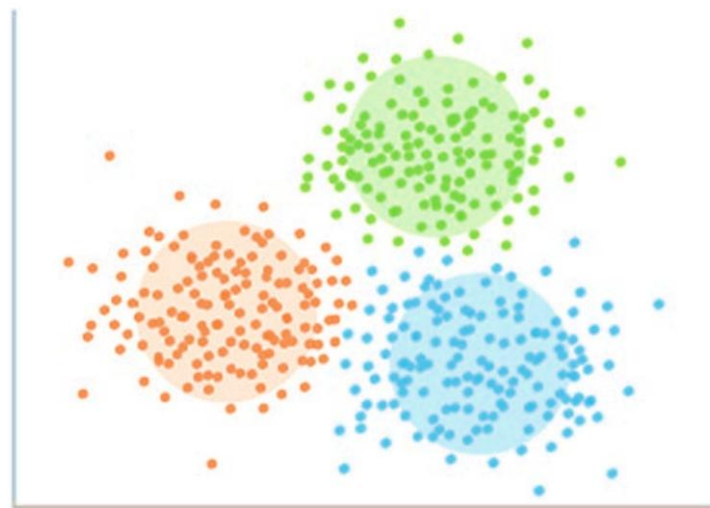


baik. Kelemahannya adalah lambatnya dalam menghasilkan prediksi dibandingkan dengan algoritma lain, sehingga Anda mungkin tidak menggunakannya saat membutuhkan prediksi secepat kilat. Hutan acak memberikan prediksi yang jauh lebih akurat dibandingkan dengan model regresi dalam banyak skenario. Kasus-kasus ini umumnya memiliki jumlah variabel prediktif yang tinggi dan ukuran sampel yang besar. Hal ini karena hutan acak menangkap varians dari beberapa variabel input secara bersamaan dan memungkinkan sejumlah besar observasi untuk berpartisipasi dalam prediksi.

Algoritma *k*-Means

k-means adalah jenis algoritma tanpa pengawasan yang memecahkan masalah pengelompokan. Perbedaan utama antara klasifikasi reguler dan pengelompokan adalah bahwa algoritma ditantang untuk mengelompokkan item dalam kluster tanpa kelas yang telah ditentukan sebelumnya. Artinya, algoritma harus memutuskan prinsip-prinsip pembagiannya sendiri tanpa bimbingan manusia. Analisis kluster biasanya dilakukan dalam gaya pembelajaran tanpa pengawasan. Pengelompokan dapat memecahkan masalah-masalah berikut:

- *Apa saja segmen utama pelanggan yang kita miliki berdasarkan demografi dan perilaku mereka?*
- *Adakah hubungan antara risiko gagal bayar beberapa nasabah bank dengan perilaku mereka?*
- *Bagaimana kita dapat mengklasifikasikan kata kunci yang digunakan orang untuk mengakses situs web kita?*

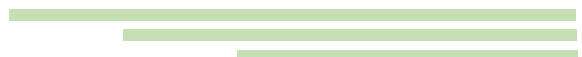


Gambar 2.3 Tiga kluster

Prosedurnya mengikuti cara yang sederhana dan mudah untuk mengklasifikasikan kumpulan data tertentu melalui sejumlah kluster (asumsikan *k* kluster), lihat Gambar 2.3. Titik-titik data di dalam kluster bersifat homogen dan heterogen terhadap kelompok sejawat.

Bagaimana *k*-means membentuk kluster:

1. *k*-means memilih *k* titik untuk setiap kluster yang dikenal sebagai sentroid.
2. Setiap titik data membentuk kluster dengan sentroid terdekat, yaitu *k* kluster.



3. Menemukan sentroid setiap klaster berdasarkan anggota klaster yang ada. Di sini kita memiliki sentroid baru.
4. Karena kita memiliki sentroid baru, ulangi langkah 2 dan 3. Temukan jarak terdekat untuk setiap titik data dari sentroid baru dan hubungkan dengan k -klaster baru. Ulangi proses ini hingga terjadi konvergensi, yaitu, sentroid tidak berubah.

Dalam k -means, kita memiliki klaster dan setiap klaster memiliki sentroidnya sendiri. Jumlah kuadrat selisih antara sentroid dan titik data dalam suatu klaster merupakan nilai kuadrat total untuk klaster tersebut. Selain itu, ketika jumlah kuadrat total untuk semua klaster dijumlahkan, hasilnya menjadi total kuadrat total untuk solusi klaster.

Kita tahu bahwa seiring bertambahnya jumlah klaster, nilai ini terus menurun, tetapi jika Anda memplot hasilnya, Anda akan melihat bahwa jumlah kuadrat jarak menurun tajam hingga mencapai nilai k tertentu, dan kemudian jauh lebih lambat setelahnya. Di sinilah, kita dapat menemukan jumlah klaster yang optimal.

Reduksi Dimensionalitas

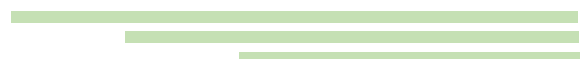
Reduksi dimensionalitas membantu sistem menghilangkan data yang tidak berguna untuk analisis. Kelompok algoritma ini digunakan untuk menghilangkan data redundan, outlier, dan data lain yang tidak berguna. Reduksi dimensionalitas dapat membantu saat menganalisis data dari sensor dan kasus penggunaan *Internet of Things* (IoT) lainnya. Dalam sistem IoT, mungkin terdapat ribuan titik data yang hanya memberi tahu Anda bahwa sensor menyala. Menyimpan dan menganalisis data yang "menyala" tersebut tidak membantu dan akan menghabiskan ruang penyimpanan yang penting. Selain itu, dengan menghilangkan data redundan ini, kinerja sistem pembelajaran mesin akan meningkat. Terakhir, reduksi dimensionalitas juga akan membantu analisis memvisualisasikan data.

Pembelajaran Penguatan

Pembelajaran penguatan adalah pendekatan komputasional untuk memahami dan mengotomatiskan pembelajaran dan pengambilan keputusan yang berorientasi pada tujuan. Pembelajaran ini mempelajari apa yang harus dilakukan dan bagaimana memetakan situasi ke tindakan. Hasilnya adalah memaksimalkan sinyal imbalan numerik. Pembelajar tidak diberi tahu tindakan mana yang harus diambil, tetapi harus menemukan tindakan mana yang akan menghasilkan imbalan maksimum.

Pembelajaran penguatan didefinisikan bukan dengan mengkarakterisasi algoritma pembelajaran, melainkan dengan mengkarakterisasi masalah pembelajaran. Algoritma apa pun yang cocok untuk memecahkan masalah tersebut kami anggap sebagai algoritma pembelajaran penguatan. Pembelajaran penguatan berbeda dengan pembelajaran terbimbing, jenis pembelajaran yang dipelajari.

Pembelajaran terbimbing adalah pembelajaran dari contoh-contoh yang diberikan oleh supervisor eksternal yang berpengetahuan. Ini merupakan jenis pembelajaran yang penting, tetapi pembelajaran ini saja tidak memadai untuk belajar dari interaksi. Dalam masalah interaktif, seringkali tidak praktis untuk mendapatkan contoh perilaku yang diinginkan yang benar dan representatif dari semua situasi di mana agen harus bertindak. Di wilayah yang



belum dipetakan, di mana seseorang mengharapkan pembelajaran paling bermanfaat, seorang agen harus mampu belajar dari pengalamannya sendiri.

Salah satu tantangan yang muncul dalam pembelajaran penguatan, dan bukan dalam jenis pembelajaran lainnya, adalah trade-off antara eksplorasi dan eksploitasi. Untuk mendapatkan banyak imbalan, agen pembelajaran penguatan harus lebih memilih tindakan yang telah dicoba di masa lalu dan terbukti efektif dalam menghasilkan imbalan. Namun, untuk menemukan tindakan tersebut, ia harus mencoba tindakan yang belum pernah dipilih sebelumnya. Agen harus mengeksploitasi apa yang sudah diketahuinya untuk mendapatkan imbalan, tetapi juga harus melakukan eksplorasi agar dapat membuat pilihan tindakan yang lebih baik di masa mendatang.

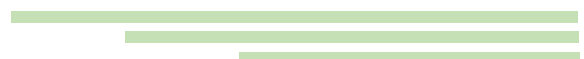
Dilemanya adalah bahwa baik eksploitasi maupun eksplorasi tidak dapat dilakukan secara eksklusif tanpa gagal dalam tugas tersebut. Agen harus mencoba berbagai tindakan dan secara progresif mengutamakan tindakan yang tampaknya terbaik. Pada tugas stokastik, setiap tindakan harus dicoba berkali-kali untuk memperkirakan imbalan yang diharapkan secara andal. Dilema eksplorasi-eksploitasi telah dipelajari secara intensif oleh para matematikawan selama beberapa dekade. Fitur kunci lain dari pembelajaran penguatan adalah ia secara eksplisit mempertimbangkan keseluruhan masalah agen yang diarahkan pada tujuan yang berinteraksi dengan lingkungan yang tidak pasti.

Hal ini berbeda dengan banyak pendekatan yang membahas sub-masalah tanpa membahas bagaimana sub-masalah tersebut dapat dimasukkan ke dalam gambaran yang lebih besar. Misalnya, kami telah menyebutkan bahwa sebagian besar penelitian pembelajaran mesin berkaitan dengan pembelajaran terawasi tanpa secara eksplisit menentukan bagaimana kemampuan tersebut pada akhirnya akan berguna. Peneliti lain telah mengembangkan teori perencanaan dengan tujuan umum, tetapi tanpa mempertimbangkan peran perencanaan dalam pengambilan keputusan waktu nyata, atau pertanyaan tentang dari mana model prediktif yang diperlukan untuk perencanaan akan berasal.

Meskipun pendekatan-pendekatan ini telah menghasilkan banyak hasil yang bermanfaat, fokus mereka pada sub-masalah yang terisolasi merupakan batasan yang signifikan. Pembelajaran penguatan mengambil pendekatan yang berlawanan, dengan memulai dengan agen yang lengkap, interaktif, dan mencari tujuan. Semua agen pembelajaran penguatan memiliki tujuan yang eksplisit, dapat merasakan aspek-aspek lingkungan mereka dan dapat memilih tindakan untuk memengaruhi lingkungan mereka.

Selain itu, biasanya diasumsikan sejak awal bahwa agen harus beroperasi meskipun ada ketidakpastian yang signifikan tentang lingkungan yang dihadapinya. Ketika pembelajaran penguatan melibatkan perencanaan, ia harus mengatasi interaksi antara perencanaan dan pemilihan tindakan waktu nyata, serta pertanyaan tentang bagaimana model lingkungan diperoleh dan ditingkatkan. Ketika pembelajaran penguatan melibatkan pembelajaran yang diawasi, ia melakukannya untuk alasan yang sangat spesifik yang menentukan kemampuan mana yang penting dan mana yang tidak.

Agar penelitian pembelajaran dapat mengalami kemajuan, sub-masalah yang penting harus diisolasi dan dipelajari, tetapi sub-masalah tersebut haruslah sub-masalah yang



dimotivasi oleh peran yang jelas dalam agen-agen yang lengkap, interaktif, dan mencari tujuan, bahkan jika semua rincian agen yang lengkap belum dapat diisi.

Peningkatan Gradien

Peningkatan gradien, seperti halnya hutan acak, juga dibuat dari pohon keputusan yang "lemah". Perbedaan signifikannya adalah dalam peningkatan gradien, pohon-pohon dilatih satu demi satu. Setiap pohon berikutnya dilatih terutama dengan data yang telah diidentifikasi secara salah oleh pohon sebelumnya. Hal ini memungkinkan peningkatan gradien untuk lebih berfokus pada kasus-kasus yang sulit daripada kasus-kasus yang mudah diprediksi.

Ensemble hanyalah kumpulan prediktor yang merupakan rata-rata dari semua prediksi untuk menghasilkan prediksi akhir. Alasan kami menggunakan ensemble adalah karena banyak prediktor berbeda yang mencoba memprediksi variabel target yang sama akan berkinerja lebih baik daripada satu prediktor saja. Teknik ensembling diklasifikasikan lebih lanjut menjadi Bagging dan Boosting.

- Boosting adalah teknik ensemble di mana prediktor tidak dibuat secara independen, tetapi secara berurutan.
- Bagging adalah teknik ensembling sederhana di mana kami membangun banyak prediktor/model/pelajar independen dan menggabungkannya menggunakan beberapa teknik perata-rataan model.

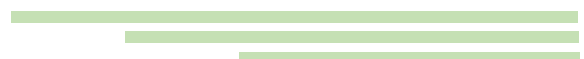
Peningkatan gradien adalah contoh algoritma peningkatan. Algoritma ini cepat dilatih dan berkinerja sangat baik. Namun, perubahan kecil pada set data pelatihan dapat menyebabkan perubahan radikal pada model, sehingga mungkin tidak menghasilkan hasil yang paling dapat dijelaskan.

Jaringan Saraf Tiruan

Jaringan saraf dalam biologi adalah neuron yang saling terhubung yang bertukar pesan satu sama lain. Ide ini kini telah diadaptasi ke dalam pembelajaran mesin dan disebut jaringan saraf tiruan (JST). Pembelajaran mendalam, yang merupakan konsep baru, hanyalah beberapa lapisan jaringan saraf tiruan yang disusun satu demi satu.

Jaringan saraf tiruan adalah sekumpulan unit masukan-keluaran yang terhubung di mana setiap koneksi memiliki bobot yang terkait dengannya. Jaringan saraf tiruan terdiri dari neuron formal yang terhubung sedemikian rupa sehingga setiap keluaran neuron selanjutnya berfungsi sebagai masukan bagi neuron yang lebih banyak, sama seperti terminal akson neuron biologis yang terhubung melalui ikatan sinaptik dengan dendrit neuron lain. Jumlah neuron dan cara mereka saling terhubung menentukan arsitektur (topologi) jaringan saraf tiruan. Berdasarkan tujuannya, neuron masukan, neuron kerja (lapisan tersembunyi, mediasi), dan neuron keluaran dapat dibedakan dalam jaringan.

Neuron masukan dan keluaran masing-masing mewakili reseptor dan efektor, dan neuron kerja yang terhubung menciptakan kanal yang sesuai di antara keduanya untuk menyebarkan sinyal masing-masing. Kanal-kanal ini disebut lintasan dalam model matematika. Perambatan sinyal dan pemrosesan informasi di sepanjang lintasan jaringan diwujudkan dengan mengubah status neuron pada lintasan ini. Status semua neuron dalam jaringan



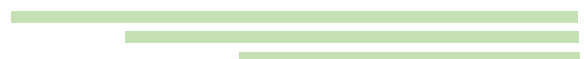
membentuk status jaringan saraf, dan bobot sinaptik yang terkait dengan semua koneksi mewakili konfigurasi jaringan saraf.

Jaringan saraf berkembang secara bertahap seiring waktu, interkoneksi serta status neuron berubah, dan bobotnya diadaptasi. Dalam konteks pembaruan atribut jaringan ini seiring waktu, penting untuk membagi dinamika global jaringan saraf menjadi tiga dinamika dan mempertimbangkan tiga fase operasi jaringan:

1. Arsitektur (perubahan topologi): menentukan topologi jaringan dan kemungkinan perubahannya. Pembaruan arsitektur biasanya diterapkan dalam kerangka mode adaptif sedemikian rupa sehingga jaringan disuplai dengan neuron dan koneksi tambahan saat dibutuhkan. Namun, dalam kebanyakan kasus, dinamika arsitektur mengasumsikan topologi jaringan saraf tiruan tetap, yang tidak lagi berubah.
2. Komputasional (perubahan status): menentukan status awal jaringan dan aturan untuk pembaruannya seiring waktu, dengan syarat topologi dan konfigurasi jaringan tetap. Pada awal mode komputasi, status neuron masukan ditetapkan ke masukan jaringan dan neuron yang tersisa berada dalam status awal. Semua masukan dan status jaringan potensial masing-masing membentuk ruang masukan dan ruang status jaringan saraf tiruan. Setelah menginisialisasi status jaringan, komputasi yang tepat dilakukan.
3. Adaptif (perubahan konfigurasi): menentukan konfigurasi awal jaringan dan cara bobot dalam jaringan diadaptasi seiring waktu. Semua konfigurasi jaringan potensial membentuk ruang bobot jaringan saraf tiruan. Pada awal mode adaptif, bobot semua koneksi jaringan ditetapkan ke konfigurasi awal. Setelah menginisialisasi konfigurasi jaringan, adaptasi yang tepat dilakukan. Demikian pula, untuk dinamika komputasi, model dengan evolusi waktu bobot jaringan saraf tiruan yang kontinu ketika konfigurasinya merupakan fungsi waktu kontinu yang biasanya dijelaskan oleh persamaan diferensial umumnya dapat dipertimbangkan. Namun, dalam kebanyakan kasus, waktu adaptasi diskrit diasumsikan.

Klasifikasi ini tidak sesuai dengan realitas neurofisiologis karena dalam sistem saraf semua perubahan yang terkait berlangsung secara simultan. Dinamika jaringan saraf tiruan yang diperkenalkan di atas biasanya ditentukan oleh kondisi awal dan oleh persamaan atau aturan matematika yang menentukan perkembangan karakteristik jaringan tertentu (topologi, status, konfigurasi) seiring waktu. Pembaruan yang dikendalikan oleh aturan-aturan ini dilakukan dalam mode operasional jaringan saraf tiruan yang sesuai.

Melalui konkretisasi dinamika yang diperkenalkan, diperoleh berbagai model jaringan saraf tiruan yang sesuai untuk menyelesaikan tugas-tugas spesifik. Ini berarti bahwa untuk menentukan model jaringan saraf tiruan tertentu, cukup dengan mendefinisikan dinamika arsitektur, komputasi, dan adaptifnya. Jaringan saraf tiruan memiliki kemampuan untuk mengklasifikasikan pola pada data yang belum dilatih dan juga, memiliki toleransi yang lebih baik daripada pengklasifikasi lain dalam menangani data yang bising. Kelemahan jaringan saraf tiruan adalah kesulitan dalam menginterpretasikan keluarannya yang berupa pola simbolik. Pengklasifikasi lain lebih baik dalam hal ini. Jaringan saraf tiruan memungkinkan interaksi



variabel tingkat tinggi karena konektivitasnya yang meningkat dan cenderung menangani data berkorelasi lebih baik daripada teknik pembelajaran lainnya.

Jaringan saraf tiruan memiliki penerapan yang luas untuk permasalahan bisnis dunia nyata. Bahkan, jaringan saraf tiruan telah berhasil diterapkan di banyak industri. Karena jaringan saraf paling baik dalam mengidentifikasi pola atau tren dalam data, jaringan saraf sangat cocok untuk kebutuhan prediksi atau peramalan, termasuk:

- Peramalan penjualan
- Pengendalian proses industri
- Riset pelanggan
- Validasi data
- Manajemen risiko
- Pemasaran target

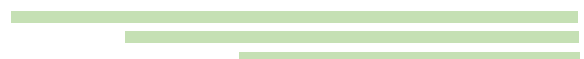
Ada berbagai manfaat yang dirasakan analis dari penggunaan jaringan saraf dalam pekerjaannya.

Pengenalan pola merupakan teknik yang ampuh untuk memanfaatkan informasi dalam data dan menggeneralisasikannya. Jaringan saraf belajar mengenali pola yang ada dalam kumpulan data. Sistem ini dikembangkan melalui pembelajaran, alih-alih pemrograman. Pemrograman jauh lebih memakan waktu bagi analis dan mengharuskan analis untuk menentukan perilaku model yang tepat. Jaringan saraf mempelajari sendiri pola-pola dalam data, sehingga analis dapat melakukan pekerjaan yang lebih menarik.

Jaringan saraf fleksibel dalam lingkungan yang terus berubah. Sistem berbasis aturan atau sistem terprogram terbatas pada situasi yang dirancang untuknya, ketika kondisi berubah, sistem tersebut tidak lagi valid. Meskipun jaringan saraf mungkin membutuhkan waktu untuk mempelajari perubahan drastis yang tiba-tiba, jaringan saraf sangat baik dalam beradaptasi dengan informasi yang terus berubah.

Jaringan saraf dapat membangun model informatif di mana pendekatan yang lebih konvensional gagal. Karena jaringan saraf dapat menangani interaksi yang sangat kompleks, jaringan saraf dapat dengan mudah memodelkan data yang terlalu sulit untuk dimodelkan dengan pendekatan tradisional seperti statistik inferensial atau logika pemrograman. Kinerja jaringan saraf setidaknya sama baiknya dengan pemodelan statistik klasik dan lebih baik untuk sebagian besar permasalahan. Jaringan saraf membangun model yang lebih mencerminkan struktur data dalam waktu yang jauh lebih singkat. Jaringan saraf sekarang beroperasi dengan baik dengan perangkat keras komputer yang sederhana. Meskipun jaringan saraf membutuhkan komputasi yang intensif, rutinitasnya telah dioptimalkan hingga dapat berjalan dalam waktu yang wajar di komputer pribadi. Jaringan saraf tidak memerlukan komputer super seperti pada masa-masa awal penelitian jaringan saraf. Ada beberapa keterbatasan dalam komputasi saraf.

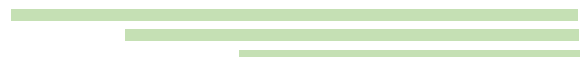
Keterbatasan utamanya adalah ketidakmampuan jaringan saraf untuk menjelaskan model yang telah dibangunnya dengan cara yang bermanfaat. Analis sering kali ingin tahu mengapa model berperilaku seperti itu. Jaringan saraf mendapatkan jawaban yang lebih baik, tetapi kesulitan menjelaskan bagaimana mereka mendapatkannya. Inilah alasan mengapa



koneksi saraf menyediakan begitu banyak alat untuk mengeksplorasi keluaran dan begitu banyak pilihan operasional pada alat yang membangun model. Dengan bereksperimen dengan berbagai parameter, dan mengeksplorasi hasil secara menyeluruh, baik dalam teks maupun grafik, pengguna koneksi saraf dapat memperoleh pemahaman yang lebih baik tentang perilaku model dan keyakinan yang lebih besar terhadap hasilnya.

Ada beberapa keterbatasan lain yang perlu dipahami. Pertama, sulit untuk mengekstrak aturan dari jaringan saraf. Hal ini terkadang penting bagi orang-orang yang harus menjelaskan jawaban mereka kepada orang lain dan bagi mereka yang telah terlibat dengan kecerdasan buatan, khususnya sistem pakar yang berbasis aturan. Dalam sebagian besar metode analisis, Anda tidak bisa begitu saja memasukkan data ke jaringan saraf dan mendapatkan jawaban yang konklusif, tetapi Anda harus meluangkan waktu untuk memahami masalah atau hasil yang ingin Anda prediksi. Dan, Anda harus memastikan bahwa data yang digunakan untuk melatih sistem sudah tepat (yaitu mencerminkan faktor-faktor yang terlibat) dan diukur dengan cara yang mencerminkan perilaku faktor-faktor tersebut. Jika data tidak mewakili masalah, komputasi neural tidak akan menghasilkan hasil yang baik. Ini adalah situasi klasik di mana "sampah masuk" pasti akan menghasilkan "sampah keluar".

Terakhir, melatih model dari kumpulan data yang sangat kompleks dapat memakan waktu. Teknik neural membutuhkan komputer yang intensif dan akan lambat pada PC kelas bawah atau mesin tanpa ko-prosesor matematika. Namun, penting untuk diingat bahwa waktu keseluruhan untuk mendapatkan hasil masih bisa lebih cepat daripada pendekatan analisis data lainnya, bahkan ketika sistem membutuhkan waktu lebih lama untuk dilatih. Kecepatan pemrosesan saja bukan satu-satunya faktor dalam kinerja, dan jaringan neural tidak memerlukan waktu pemrograman dan debugging atau pengujian asumsi seperti yang dilakukan oleh pendekatan analitis lainnya.



BAB 3

PEMBELAJARAN MENDALAM

3.1 PENDAHULUAN

Pembelajaran mendalam terdiri dari banyak lapisan hierarkis untuk memproses informasi secara nonlinier, di mana beberapa konsep tingkat rendah membantu mendefinisikan konsep tingkat tinggi.

Pembelajaran mendalam didefinisikan sebagai pembelajaran mendalam adalah kelas teknik pembelajaran mesin yang memanfaatkan banyak lapisan pemrosesan informasi nonlinier untuk ekstraksi dan transformasi fitur terawasi atau tak terawasi, serta untuk analisis dan klasifikasi pola.

Jaringan saraf tiruan dangkal tidak mampu menangani data kompleks dalam jumlah besar, yang jelas terlihat dalam banyak aplikasi rutin seperti ucapan alami, gambar, pengambilan informasi, dan aplikasi pemrosesan informasi mirip manusia lainnya. Pembelajaran mendalam disarankan untuk aplikasi semacam itu. Dengan pembelajaran mendalam, dimungkinkan untuk mengenali, mengklasifikasikan, dan mengkategorikan pola dalam data untuk mesin dengan upaya yang relatif lebih sedikit. Google adalah pelopor dalam eksperimen pembelajaran mendalam, yang diprakarsai oleh Andrew Ng.

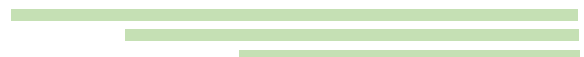
Pembelajaran mendalam menawarkan pemrosesan berlapis-lapis mirip manusia dibandingkan dengan arsitektur dangkal. Ide dasar pembelajaran mendalam adalah menggunakan pemrosesan hierarkis menggunakan banyak lapisan arsitektur. Lapisan-lapisan arsitektur disusun secara hierarkis. Masukan dari setiap lapisan diberikan ke lapisan di sebelahnya setelah beberapa pra-pelatihan.

Seringkali, pra-pelatihan lapisan tertentu dilakukan tanpa pengawasan. Pembelajaran mendalam mengikuti pendekatan terdistribusi untuk mengelola big data. Pendekatan ini mengasumsikan bahwa data dihasilkan dengan mempertimbangkan berbagai faktor, waktu, dan tingkatan yang berbeda. Pembelajaran mendalam memfasilitasi pengaturan dan pemrosesan data ke dalam berbagai lapisan berdasarkan waktu (kemunculannya), tingkatan, atau sifatnya. Pembelajaran mendalam sering dikaitkan dengan jaringan saraf tiruan.

Ada tiga kategori arsitektur pembelajaran mendalam:

- (i) Generatif
- (ii) Diskriminatif
- (iii) Arsitektur pembelajaran mendalam hibrida

Arsitektur yang termasuk dalam kategori generatif berfokus pada pra-pelatihan suatu lapisan tanpa pengawasan. Pendekatan ini menghilangkan kesulitan dalam melatih arsitektur tingkat bawah, yang bergantung pada lapisan sebelumnya. Setiap lapisan dapat dilatih terlebih dahulu dan kemudian dimasukkan ke dalam model untuk penyetelan dan pembelajaran umum lebih lanjut. Hal ini menyelesaikan masalah pelatihan arsitektur jaringan saraf tiruan dengan beberapa lapisan dan memungkinkan pembelajaran mendalam. Arsitektur jaringan saraf tiruan dapat memiliki kemampuan pemrosesan diskriminatif dengan menumpuk keluaran



setiap lapisan dengan data asli atau dengan berbagai kombinasi informasi, sehingga membentuk arsitektur pembelajaran mendalam.

Model deskriptif sering kali menganggap keluaran jaringan saraf tiruan sebagai distribusi bersyarat atas semua kemungkinan urutan label untuk urutan masukan yang diberikan, yang akan dioptimalkan lebih lanjut melalui fungsi objektif. Arsitektur hibrida menggabungkan properti arsitektur generatif dan diskriminatif. Pembelajaran mendalam yang umum dapat dilakukan sebagai berikut.

- ✚ Bangun jaringan yang terdiri dari lapisan masukan dan lapisan tersembunyi dengan simpul-simpul yang diperlukan
- ✚ Latih jaringan
- ✚ Tambahkan lapisan tersembunyi lain di atas jaringan yang telah dipelajari sebelumnya untuk menghasilkan jaringan baru
- ✚ Latih ulang jaringan
- ✚ Ulangi penambahan lapisan dan setelah setiap penambahan, latih ulang jaringan

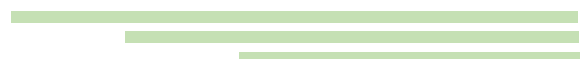
3.2 MENGANALISIS BIG DATA

Analisis big data diperlukan untuk mengelola data dalam jumlah besar secara efisien. Aspek utama yang dapat dipertimbangkan saat menangani big data adalah optimasi skala besar, penanganan data berdimensi tinggi, dan penanganan data dinamis. Optimasi berkaitan dengan menemukan solusi paling efektif untuk masalah menggunakan prosedur dan model yang terdefinisi dengan baik. Setiap orang menghadapi masalah optimasi, baik secara langsung (sistematis) maupun tidak langsung (informal). Permasalahan yang memerlukan optimasi meliputi masalah tenaga penjual keliling, pemilihan mata kuliah dari mata kuliah yang tersedia di bawah aliran (misalnya aliran sains) dan jenjang (sarjana), aktivitas e-commerce (memilih ponsel terbaik melalui berbagai situs belanja online), dll.

Optimasi membantu menemukan alternatif yang hemat biaya untuk menyelesaikan tugas. Biasanya, optimasi dianggap sebagai maksimalisasi atau minimalisasi fungsi sumber daya. Beberapa contohnya adalah maksimalisasi keuntungan, minimalisasi biaya, dan kesalahan. Untuk domain dengan dimensi terbatas, berbagai model untuk permasalahan tersebut tersedia; namun, ketika menyangkut data besar, tugas optimasi menjadi menantang. Dalam kasus data besar, tidak hanya ukuran transaksi yang besar, tetapi jumlah variabel dan kendala juga tinggi.

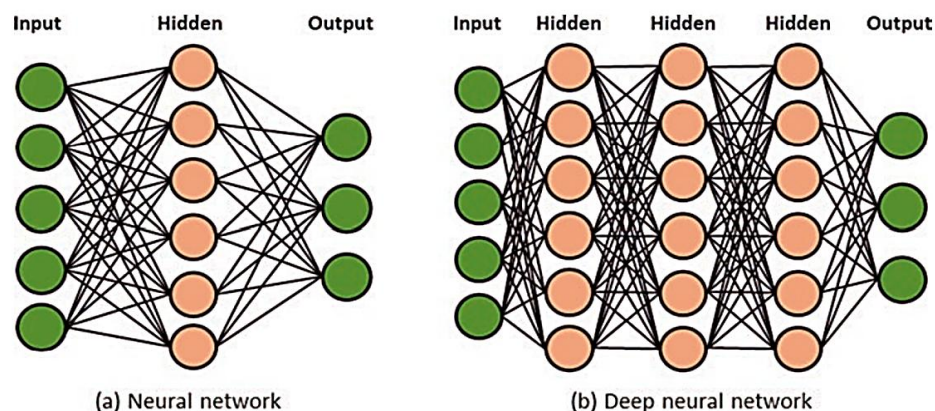
Sebaliknya, terkadang data dan kendala sedang tetapi strukturnya kompleks, sehingga tidak dapat ditangani dengan metode saat ini. Misalnya, pembelajaran fitur dari repositori besar citra medis secara optimal akan sulit dilakukan dengan dukungan metode tradisional. Lebih lanjut, hal ini memerlukan penyetelan manual beberapa parameter. Selain pendekatan tradisional, pembelajaran mesin dan metode optimasi paralel juga semakin populer. Ketika dimensi dalam masalah meningkat, kompleksitasnya pun meningkat.

Dimensi masalah yang terbatas membuat masalah mudah dipecahkan; namun, solusinya tidak kuat dan tidak memberikan pengetahuan tingkat tinggi. Peningkatan jumlah dimensi mengakibatkan pertumbuhan data yang luar biasa besarnya sehingga sulit untuk



ditangani, divisualisasikan, dan dipecahkan. Dikatakan bahwa karena pertumbuhan eksponensial jumlah nilai yang mungkin dengan setiap dimensi, enumerasi lengkap semua subruang menjadi sulit dengan meningkatnya dimensionalitas, yang dikenal sebagai kutukan dimensionalitas. Pembelajaran mendalam akan membantu dalam mengelola data berdimensi tinggi tersebut dan membantu dalam pengelompokan, pemrosesan, dan visualisasi data tersebut.

Bioinformatika, web sensor, visi, dan pengenalan suara adalah bidang-bidang di mana seseorang dapat menemukan data berdimensi tinggi tersebut. Selain volume dan struktur, waktu merupakan faktor utama lain yang meningkatkan kompleksitas data dan karenanya membuat pekerjaan pengelolaan data menjadi lebih sulit. Data dinamis bervariasi dalam hal ukuran, volume, dan struktur yang mendasarinya. Data berskala besar dan dinamis dihasilkan dan dimanipulasi di berbagai bidang seperti dinamika fluida, ilmu material, dinamika modular, dan sistem yang terinspirasi oleh biologi. Contohnya adalah pembangkitan ucapan manusia. Pembangkitan ucapan manusia mengikuti struktur hierarkis. Pembelajaran mendalam akan berguna dalam pemodelan ucapan terstruktur.



Gambar 3.1 ANN untuk pembelajaran mendalam

Perbedaan Antara Data Tradisional dan Big Data

Data tradisional dan big data memiliki perbedaan utama pada jenis data yang dikelola, volume data, serta alat yang digunakan untuk menganalisisnya.

- Data tradisional biasanya terdiri dari data terstruktur yang tersimpan dalam basis data relasional. Basis data ini mengorganisir data dalam bentuk tabel yang telah terdefinisi dengan jelas, sehingga memudahkan dalam melakukan kueri menggunakan alat standar seperti SQL. Metode analisis data tradisional umumnya menggunakan pendekatan statistik yang sesuai untuk dataset dengan format yang teratur dan ukuran relatif kecil.
- Sementara itu, big data mencakup kumpulan data yang sangat besar dengan berbagai format, termasuk data terstruktur, semi-terstruktur, dan tidak terstruktur. Kompleksitas dan keragaman data ini mengharuskan penggunaan teknik analitik yang lebih maju, seperti machine learning, penggalian data (data mining), serta visualisasi data untuk

mendapatkan wawasan yang bermakna. Selain itu, skala besar big data memerlukan sistem pemrosesan terdistribusi agar dapat dikelola dan dianalisis secara efisien.

Lima Karakteristik Big Data

"V dari Big Data" mencakup lima sifat utama yang membedakan big data dari jenis data lainnya. Karakteristik ini menjelaskan perbedaan fundamental antara big data dan data tradisional serta kebutuhan pengelolaannya yang khusus.

❖ **Volume**

Big data disebut “besar” bukan tanpa alasan. Saat ini, jumlah data yang dihasilkan sangat besar dan terus bertambah dari berbagai sumber seperti aplikasi web, perangkat *Internet of Things* (IoT), serta catatan transaksi. Organisasi sering menghadapi tantangan dalam menyimpan dan mengelola data dalam skala ini, karena sistem tradisional sulit mengikuti pertumbuhan data tersebut. Solusi big data modern, termasuk penyimpanan berbasis cloud, memungkinkan organisasi mengelola kumpulan data yang sangat besar dengan efisien dan memastikan data penting tidak hilang.

❖ **Kecepatan**

Kecepatan menunjukkan seberapa cepat data mengalir ke dalam sistem. Saat ini, data masuk dengan sangat cepat, mulai dari pembaruan media sosial secara langsung hingga transaksi saham berfrekuensi tinggi. Data yang diproses dengan cepat ini memungkinkan organisasi mendapatkan wawasan secara waktu nyata yang menunjang pengambilan keputusan yang cepat dan tepat. Untuk mengatasi hal ini, alat seperti kerangka kerja pemrosesan aliran dan sistem penyimpanan dalam memori digunakan agar data bisa dianalisis dan ditindaklanjuti hampir secara langsung.

❖ **Variasi**

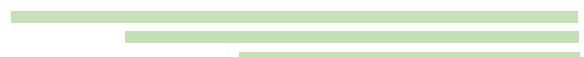
Variasi mengacu pada banyaknya format dan jenis data dalam big data. Selain data terstruktur tradisional, big data juga mencakup data tidak terstruktur seperti teks bebas, gambar, dan video, serta data semi-terstruktur seperti file JSON dan XML yang memiliki struktur longgar. Untuk mengelola keberagaman ini, diperlukan solusi fleksibel seperti basis data NoSQL dan data lake dengan kerangka kerja schema-on-read, yang mampu menyimpan dan mengintegrasikan beragam format data guna mendukung analisis yang mendalam.

❖ **Kebenaran (Veracity)**

Kebenaran merujuk pada tingkat keakuratan dan keandalan data. Karena big data berasal dari berbagai sumber dan jumlahnya sangat besar, ada risiko data mengandung kesalahan (noise) yang dapat mengganggu kualitas pengambilan keputusan. Oleh karena itu, organisasi harus menerapkan proses pengelolaan data yang ketat, termasuk pembersihan, validasi, dan verifikasi data, guna meningkatkan kualitas dan keakuratan analisis yang dihasilkan.

❖ **Nilai**

Nilai mengacu pada manfaat konkret yang dapat diperoleh dari big data oleh organisasi. Manfaat tersebut bisa berupa optimalisasi operasional bisnis, identifikasi



peluang bisnis baru, hingga pengembangan strategi pemasaran yang lebih efektif. Proses analisis big data, yang sering didukung oleh teknik canggih seperti machine learning dan kecerdasan buatan, memainkan peran penting dalam mengubah data mentah menjadi wawasan berharga yang dapat segera ditindaklanjuti.

Manajemen Big Data

Manajemen big data adalah proses sistematis yang mencakup pengumpulan, pemrosesan, dan analisis data untuk mengubah data mentah menjadi wawasan yang berguna bagi organisasi. Inti dari proses ini adalah rekayasa data, yang memastikan jalur data, sistem penyimpanan, dan integrasi bekerja secara efisien pada skala besar.

I. Pengumpulan Big Data

Pada tahap ini, data dalam jumlah besar dikumpulkan dari berbagai sumber penyusun big data. Untuk menangani kecepatan dan variasi data, organisasi sering menggunakan teknologi khusus seperti Apache Kafka untuk streaming data secara real-time dan Apache NiFi untuk mengotomatisasi aliran data. Alat ini membantu menjaga data tetap akurat dan konsisten saat berpindah melalui jalur data. Selain itu, alat integrasi data menggabungkan data dari berbagai sumber menjadi satu tampilan terpadu yang mendukung analisis. Metadata tentang asal, format, dan karakteristik data juga dikumpulkan untuk memudahkan pengelolaan di masa depan. Menjaga kualitas data menjadi sangat penting, dengan penerapan validasi skema, deduplikasi, dan pembersihan untuk mengatasi kesalahan dan inkonsistensi.

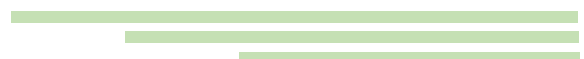
II. Penyimpanan Big Data

Setelah terkumpul, data disimpan dalam solusi penyimpanan yang sesuai, yaitu data lake, data warehouse, dan data lakehouse.

- Data Lake adalah sistem penyimpanan berbiaya rendah yang mampu menampung data mentah terstruktur maupun tidak terstruktur dalam format aslinya tanpa proses pembersihan ketat. Ideal untuk volume dan variasi data besar dan digunakan untuk pelatihan AI serta machine learning.
- Data Warehouse mengintegrasikan dan membersihkan data dari berbagai sumber ke dalam format yang terstruktur dan konsisten, mendukung analisis bisnis dan intelijen. Biaya penyimpanan lebih tinggi karena skema ketat, biasanya digunakan untuk subset data tertentu.
- Data Lakehouse menggabungkan fleksibilitas data lake dengan struktur dan kemampuan query data warehouse, menciptakan platform terpadu yang memanfaatkan keunggulan keduanya, meski lebih kompleks dalam pengelolaan.

III. Analisis Big Data

Analisis big data melibatkan penggunaan teknik machine learning, penambangan data, dan statistik untuk menemukan pola, korelasi, dan tren dalam data yang besar dan kompleks. Dengan analisis ini, bisnis dapat berpindah dari pelaporan tradisional ke insight prediktif dan preskriptif, memungkinkan pengambilan keputusan proaktif, seperti rekomendasi produk yang dipersonalisasi atau solusi medis yang disesuaikan. Hal ini berdampak pada peningkatan kepuasan pelanggan, pendapatan, dan inovasi.



Manfaat Big Data

Big data telah merubah cara organisasi dalam menggali wawasan dan mengambil keputusan strategis. Sebuah studi dari Harvard Business Review menunjukkan bahwa perusahaan yang berorientasi pada data cenderung lebih menguntungkan dan lebih inovatif dibandingkan pesaingnya. Organisasi yang menerapkan big data dan kecerdasan buatan secara efektif melaporkan keunggulan signifikan dalam berbagai indikator bisnis utama, seperti efisiensi operasional (81% dibanding 58%), pertumbuhan pendapatan (77% dibanding 61%), dan peningkatan pengalaman pelanggan (77% dibanding 45%). Berikut ini beberapa manfaat utama big data beserta contohnya:

- **Pengambilan Keputusan Lebih Tepat:**

Analisis data dalam jumlah besar memungkinkan organisasi untuk mengidentifikasi pola dan tren yang mendukung keputusan yang lebih akurat. Contohnya, rantai toko bahan makanan dapat menggabungkan data penjualan dan prediksi cuaca untuk memperkirakan permintaan produk musiman, sehingga stok di toko dapat diatur dengan lebih efisien dan limbah dapat diminimalkan.

- **Peningkatan Pengalaman Pelanggan:**

Big data membantu perusahaan memahami perilaku pelanggan secara mendalam, memungkinkan interaksi yang lebih personal. Misalnya, analisis data dapat mengenali pelanggan yang sering membeli produk perawatan kulit dari merek tertentu, sehingga merek tersebut dapat menargetkan kampanye promosi khusus yang relevan, seperti penawaran waktu terbatas atau diskon produk serupa.

- **Efisiensi Operasional yang Lebih Baik:**

Data real-time memungkinkan organisasi mengoptimalkan operasi dan meminimalisir pemborosan. Dalam industri manufaktur, data sensor waktu nyata dapat digunakan untuk memprediksi kerusakan mesin sebelum terjadi (pemeliharaan prediktif), yang membantu mencegah downtime dan mengurangi biaya perawatan.

- **Pengembangan Produk yang Responsif:**

Insight dari big data membantu perusahaan menyesuaikan produk dengan kebutuhan pelanggan. Contohnya, jika banyak pengguna mengeluhkan fitur tertentu pada smartphone yang cepat menguras baterai, pengembang dapat mengutamakan perbaikan fitur tersebut dalam pembaruan perangkat lunak berikutnya.

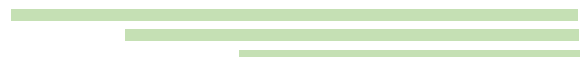
- **Optimasi Strategi Penetapan Harga:**

Big data memungkinkan organisasi menyesuaikan harga secara dinamis sesuai kondisi pasar. Contoh nyatanya adalah maskapai penerbangan yang memanfaatkan data untuk mengubah harga tiket secara real-time berdasarkan permintaan dan persaingan pasar.

- **Peningkatan Manajemen Risiko dan Deteksi Penipuan:**

Dengan big data, organisasi bisa lebih proaktif dalam mengidentifikasi dan memantau risiko. Bank, misalnya, menganalisis pola transaksi untuk mendeteksi potensi penipuan. Jika terjadi pembelian besar yang tidak biasa dengan kartu kredit di luar negeri, transaksi itu akan ditandai dan nasabah akan dikonfirmasi keasliannya.

- **Inovasi dalam Layanan Kesehatan:**



Penyedia layanan kesehatan menggunakan big data dari catatan pasien, data genetik, dan perangkat wearable untuk meningkatkan perawatan. Contohnya, monitor glukosa kontinu pada pasien diabetes dapat melacak kadar gula darah secara real-time, memungkinkan deteksi cepat lonjakan atau penurunan berbahaya dan penyesuaian rencana perawatan sesuai kebutuhan.

Tantangan dalam Big Data

Meski big data memiliki potensi besar, penggunaannya juga menghadapi berbagai tantangan signifikan, khususnya terkait dengan skala dan kecepatan data.

Beberapa tantangan utama dalam mengelola big data antara lain:

- **Kualitas dan Manajemen Data**

Menghubungkan berbagai titik data dan menjaga keakuratan informasi menjadi tugas yang kompleks, terutama dengan aliran data yang terus-menerus dari media sosial, perangkat IoT, dan sumber lainnya. Sebagai contoh, perusahaan logistik dapat mengalami kesulitan mengintegrasikan data GPS armada mereka dengan masukan pelanggan serta data inventaris gudang sehingga sulit memperoleh gambaran lengkap tentang kinerja pengiriman.

- **Skalabilitas**

Seiring dengan pertumbuhan volume data, organisasi perlu memperbesar kapasitas penyimpanan dan kemampuan pemrosesan untuk mengimbangnya. Misalnya, platform streaming yang memproses jutaan interaksi pengguna setiap hari harus terus meningkatkan kapasitas penyimpanan dan komputasi mereka untuk memenuhi permintaan. Meski layanan cloud menawarkan solusi yang lebih fleksibel dan mudah diskalakan dibandingkan sistem lokal, tantangan dalam mengelola volume dan kecepatan data yang sangat besar tetap ada.

- **Privasi dan Keamanan**

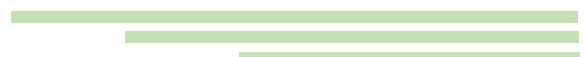
Peraturan seperti GDPR dan HIPAA mengharuskan penerapan standar keamanan dan privasi yang ketat, termasuk pengendalian akses dan enkripsi data guna mencegah akses tidak sah, terutama pada data sensitif seperti informasi pasien. Mematuhi regulasi ini menjadi tantangan tersendiri saat berhadapan dengan kumpulan data besar yang terus berkembang.

- **Kompleksitas Integrasi**

Menggabungkan berbagai jenis data dari sumber yang berbeda secara teknis seringkali sulit. Contohnya, sebuah rantai ritel mungkin mengalami kesulitan dalam mengonsolidasikan data penjualan terstruktur, ulasan pelanggan yang tidak terstruktur, dan data pemasok semi-terstruktur untuk mendapatkan pemahaman menyeluruh tentang kinerja produk.

- **Ketersediaan Tenaga Kerja Terampil**

Pengelolaan big data membutuhkan keahlian khusus dalam bidang ilmu data, teknik, dan analisis. Banyak organisasi menghadapi kesulitan dalam mencari profesional yang mampu mengelola dan menganalisis data besar. Sebagai contoh, institusi keuangan



mungkin kekurangan ilmuwan data yang memiliki kemampuan machine learning dan pemodelan keuangan untuk menganalisis data transaksi serta meramalkan tren pasar.

Peran Big Data dalam Machine Learning dan Kecerdasan Buatan (AI)

Sebanyak 72% CEO dari perusahaan berkinerja terbaik sepakat bahwa keunggulan kompetitif di era saat ini sangat bergantung pada penerapan AI generatif yang paling canggih. Untuk mencapai tingkat kecanggihan tersebut, hal yang paling krusial adalah ketersediaan sejumlah besar data berkualitas tinggi.

Sistem AI modern dan model machine learning yang kompleks, seperti model bahasa besar (*Large Language Models/LLM*), bergantung pada metode pembelajaran mendalam (*deep learning*) sebagai fondasi utama mereka. Pembelajaran mendalam memanfaatkan kumpulan data yang sangat besar dan sebagian besar tidak berlabel untuk melatih model dalam menyelesaikan berbagai tugas kompleks, misalnya pengenalan gambar, pemahaman ucapan, dan pemrosesan bahasa alami.

Dengan fondasi big data yang kuat ini, algoritma machine learning mampu mengenali pola tersembunyi dalam data, mengembangkan wawasan mendalam, serta memberikan prediksi yang akurat. Kemampuan ini memungkinkan organisasi untuk mendorong inovasi, meningkatkan kualitas pengalaman pelanggan, dan mempertahankan posisi kompetitif dalam pasar yang dinamis dan penuh tantangan.

3.3 BERBAGAI MODEL PEMBELAJARAN MENDALAM

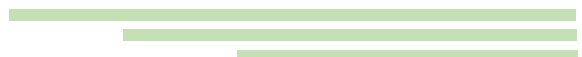
Autoencoder

Autoencoder adalah jaringan saraf tiruan yang mampu mempelajari berbagai pola pengkodean. Bentuk sederhana autoencoder sama seperti perceptron multilayer yang terdiri dari lapisan masukan, satu atau lebih lapisan tersembunyi, dan lapisan keluaran. Perbedaan utama antara perceptron multilayer, jaringan saraf tiruan feedforward, dan autoencoder terletak pada jumlah simpul pada lapisan keluaran.

Dalam kasus autoencoder, lapisan keluaran berisi jumlah simpul yang sama dengan lapisan masukan. Alih-alih memprediksi nilai target sesuai vektor keluaran, autoencoder harus memprediksi masukannya sendiri. Garis besar mekanisme pembelajarannya adalah sebagai berikut.

```
For each input  $x$ ,  
- Do a feedforward pass to compute activation functions provided  
  at all the hidden layers and output layers  
- Find the deviation between the calculated values with the inputs  
  using appropriate error function  
- Backpropagate the error in order to update weights  
Repeat the task till satisfactory output.
```

Jika jumlah simpul pada lapisan tersembunyi lebih sedikit daripada simpul masukan/keluaran, maka aktivasi lapisan tersembunyi terakhir dianggap sebagai representasi masukan yang terkompresi. Ketika simpul lapisan tersembunyi lebih banyak daripada lapisan masukan,



autoencoder berpotensi mempelajari fungsi identitas dan menjadi tidak berguna dalam sebagian besar kasus.

Jaringan Kepercayaan Dalam

Jaringan kepercayaan dalam adalah solusi untuk masalah penanganan fungsi objektif non-konveks dan minimum lokal saat menggunakan perceptron multilapis yang umum. Ini adalah jenis pembelajaran mendalam alternatif yang terdiri dari beberapa lapisan variabel laten dengan koneksi antar lapisan. Jaringan kepercayaan dalam dapat dipandang sebagai mesin Boltzmann terbatas (RBM) di mana setiap lapisan tersembunyi subjaringan bertindak sebagai lapisan masukan yang terlihat untuk lapisan jaringan yang berdekatan.

Hal ini menjadikan lapisan terlihat terendah sebagai set pelatihan untuk lapisan jaringan yang berdekatan. Dengan cara ini, setiap lapisan jaringan dilatih secara independen dan rakus. Variabel tersembunyi digunakan sebagai variabel teramati untuk melatih setiap lapisan struktur dalam. Algoritma pelatihan untuk jaringan keyakinan mendalam tersebut disediakan sebagai berikut:

- ✓ Pertimbangkan sebuah vektor masukan
- ✓ Latih mesin Boltzmann terbatas menggunakan vektor masukan dan dapatkan matriks bobot
- ✓ Latih dua lapisan bawah jaringan menggunakan matriks bobot ini
- ✓ Hasilkan vektor masukan baru menggunakan jaringan (RBM) melalui pengambilan sampel atau aktivasi rata-rata unit tersembunyi
- ✓ Ulangi prosedur hingga dua lapisan teratas jaringan tercapai

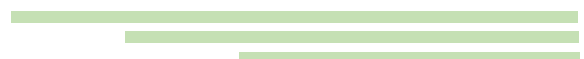
Penyetelan halus jaringan keyakinan mendalam sangat mirip dengan perceptron multilapis. Jaringan keyakinan mendalam semacam itu berguna dalam pemodelan akustik.

Jaringan Saraf Tiruan Konvolusional

Jaringan saraf tiruan konvolusional (CNN) adalah varian lain dari perceptron multilapis umpan maju. Ini adalah jenis jaringan saraf tiruan umpan maju, di mana neuron-neuron individual disusun sedemikian rupa sehingga mereka merespons daerah yang tumpang tindih dalam bidang visual. CNN dalam bekerja dengan memodelkan potongan-potongan kecil informasi secara berurutan dan menggabungkannya lebih dalam di jaringan. Salah satu cara untuk memahaminya adalah bahwa lapisan pertama akan mencoba mendeteksi tepi dan membentuk templat untuk deteksi tepi.

Kemudian, lapisan-lapisan berikutnya akan mencoba menggabungkannya menjadi bentuk yang lebih sederhana dan akhirnya menjadi templat dengan posisi objek, iluminasi, skala, dll. yang berbeda. Lapisan terakhir akan mencocokkan gambar masukan dengan semua templat dan prediksi akhir seperti jumlah tertimbang dari semuanya. Jadi, CNN dalam dapat memodelkan variasi dan perilaku kompleks yang menghasilkan prediksi yang sangat akurat.

CNN terdiri dari setidaknya tiga jenis lapisan utama yang saling terkait, yaitu lapisan konvolusi, lapisan penyatuan, dan lapisan terhubung penuh (*fully connected/FC*). Dalam aplikasi yang kompleks, sebuah CNN dapat memiliki ribuan lapisan, di mana setiap lapisan dibangun berdasarkan hasil proses dari lapisan sebelumnya sehingga membentuk hierarki pemrosesan data yang bertingkat.



Lapisan konvolusi merupakan inti dari CNN. Proses konvolusi melibatkan pengolahan berulang dari input asli untuk mengekstrak pola-pola rinci, seperti tepi, tekstur, dan bagian kecil dari gambar. Pada lapisan-lapisan awal, CNN fokus mengidentifikasi fitur-fitur dasar seperti warna, garis, dan tepi. Ketika data gambar diproses secara berantai melalui lapisan berikutnya, kompleksitas informasi yang dikenali oleh CNN meningkat, sehingga dapat mengidentifikasi bagian yang lebih besar dan lebih kompleks dari gambar, seperti objek atau bentuk yang lebih utuh.

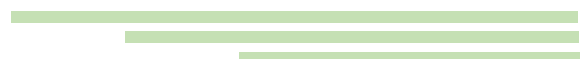
Lapisan penyatuan berfungsi merangkum informasi dari lapisan sebelumnya dengan mengurangi dimensi data, biasanya dengan metode seperti max pooling atau average pooling. Meskipun proses ini dapat menyebabkan hilangnya beberapa detail informasi, manfaatnya jauh lebih besar karena membantu mengurangi kompleksitas representasi, meningkatkan efisiensi pengolahan, dan mengurangi risiko overfitting yang muncul akibat kelebihan parameter.

Lapisan terhubung penuh berada di akhir jaringan, menghubungkan semua neuron dalam lapisan tersebut dengan neuron pada lapisan sebelumnya. Fungsi utamanya adalah menggabungkan fitur-fitur yang telah diekstrak ke dalam bentuk yang dapat diklasifikasikan secara efektif, seperti menentukan kategori objek dalam gambar.

CNN dikenal unggul dalam menangani jenis data seperti gambar, suara, dan sinyal audio dibandingkan dengan jaringan saraf lainnya. Sebelum kemunculan CNN, proses ekstraksi fitur dilakukan secara manual, yang sangat memakan waktu dan kurang efisien. CNN menawarkan pendekatan otomatis dan dapat diskalakan untuk klasifikasi dan pengenalan objek pada data berdimensi tinggi. Selain itu, CNN memungkinkan pertukaran informasi antar lapisan secara efisien, yang mempercepat proses pembelajaran dan pemrosesan data. Namun, CNN juga memiliki sejumlah keterbatasan. Proses pelatihan dan inferensinya membutuhkan sumber daya komputasi yang besar, khususnya kartu grafis (GPU) dengan kemampuan tinggi. Selain itu, penerapan CNN memerlukan tenaga ahli yang memahami penyesuaian arsitektur jaringan, tuning hiperparameter, serta pengujian konfigurasi secara mendetail. Keahlian lintas domain juga penting untuk memahami dengan baik konteks data dan aplikasi CNN.

Jaringan semacam itu mengikuti mekanisme visual organisme hidup. Sel-sel di korteks visual sensitif terhadap subwilayah kecil dari bidang visual, yang disebut bidang reseptif. Subwilayah tersebut disusun untuk menutupi seluruh bidang visual, dan sel-sel bertindak sebagai filter lokal di atas ruang masukan. Algoritma backpropagation digunakan untuk melatih parameter setiap kernel konvolusi.

Selanjutnya, setiap kernel direplikasi di seluruh gambar dengan parameter yang sama. Ada operator konvolusional yang mengekstraksi fitur unik dari masukan. Selain lapisan konvolusional, jaringan ini juga mengandung lapisan unit linier terkoreksi, lapisan pengumpulan untuk menghitung nilai maksimum atau rata-rata suatu fitur pada suatu wilayah gambar, dan lapisan kerugian yang terdiri dari fungsi kerugian spesifik aplikasi. Pengenalan gambar dan analisis video serta pemrosesan bahasa alami merupakan aplikasi utama dari jaringan saraf tiruan tersebut. Bidang visi komputer telah mengalami kemajuan pesat dalam beberapa tahun terakhir. Salah satu kemajuan yang paling menonjol adalah CNN. Kini, CNN



mendalam membentuk inti dari sebagian besar aplikasi visi komputer canggih, seperti mobil tanpa pengemudi, penandaan otomatis teman di foto Facebook kita, pengenalan gestur, fitur keamanan wajah, dan pengenalan plat nomor otomatis.

Jaringan Saraf Tiruan Rekursif

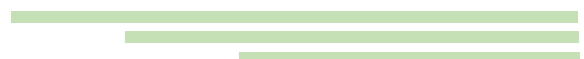
Model konvolusional bekerja pada jumlah masukan tetap, menghasilkan vektor berukuran tetap sebagai keluaran dengan jumlah langkah yang telah ditentukan sebelumnya. Jaringan saraf tiruan rekursif memungkinkan kita untuk beroperasi pada rangkaian vektor masukan dan keluaran. Dalam kasus jaringan saraf tiruan rekursif, hubungan antar unit membentuk siklus terarah. Tidak seperti jaringan saraf tiruan tradisional, masukan dan keluaran jaringan saraf tiruan rekursif tidak independen, melainkan saling terkait. Lebih lanjut, jaringan saraf tiruan rekursif berbagi parameter yang sama di setiap lapisan. Kita dapat melatih jaringan saraf tiruan rekursif dengan cara yang mirip dengan jaringan saraf tiruan tradisional menggunakan metode propagasi balik.

Di sini, perhitungan gradien tidak bergantung pada langkah saat ini, tetapi juga pada langkah sebelumnya. Varian yang disebut jaringan saraf tiruan rekursif dwiarah juga digunakan untuk banyak aplikasi. Jaringan saraf tiruan dwiarah tidak hanya mempertimbangkan keluaran sebelumnya tetapi juga keluaran yang diharapkan di masa mendatang. Dalam jaringan saraf tiruan dwiarah dan sederhana, pembelajaran mendalam dapat dicapai dengan memperkenalkan beberapa lapisan tersembunyi. Jaringan mendalam semacam itu menyediakan kapasitas pembelajaran yang lebih tinggi dengan banyak data pembelajaran. Ucapan, pemrosesan gambar, dan pemrosesan bahasa alami adalah beberapa area kandidat di mana jaringan saraf rekuren dapat digunakan.

Pembelajaran Penguatan ke Jaringan Saraf Tiruan

Pembelajaran penguatan adalah semacam hibridisasi pemrograman dinamis dan pembelajaran terawasi. Komponen umum dari pendekatan ini adalah fungsi lingkungan, agen, tindakan, kebijakan, dan biaya. Agen bertindak sebagai pengontrol sistem; kebijakan menentukan tindakan yang akan diambil; dan fungsi imbalan menentukan tujuan keseluruhan dari permasalahan pembelajaran penguatan.

Agen yang menerima imbalan semaksimal mungkin dapat dipandang melakukan tindakan terbaik untuk suatu keadaan tertentu. Agen di sini mengacu pada entitas abstrak yang dapat berupa objek atau subjek apa pun yang melakukan tindakan: mobil otonom, robot, manusia, bot obrolan dukungan pelanggan, dll. Keadaan agen mengacu pada posisi dan keadaan agen dalam lingkungan abstraknya; misalnya, posisi tertentu dalam dunia realitas virtual, gedung, papan catur, atau posisi dan kecepatan di arena pacuan kuda. Pembelajaran penguatan mendalam menjanjikan prosedur pembelajaran yang sangat umum yang dapat mempelajari perilaku bermanfaat dengan umpan balik yang sangat minim. Ini merupakan area yang menarik sekaligus menantang, yang tentunya akan menjadi bagian penting dari lanskap AI di masa depan.



3.4 APLIKASI PEMBELAJARAN MENDALAM DALAM BISNIS

Analisis telah mengubah laba bisnis selama beberapa waktu. Kini, semakin banyak perusahaan yang menguasai penggunaan analitik, mereka menggali data lebih dalam untuk meningkatkan efisiensi, mendapatkan keunggulan kompetitif yang lebih besar, dan meningkatkan laba mereka lebih jauh lagi. Itulah sebabnya perusahaan ingin menerapkan pembelajaran mesin dan kecerdasan buatan; mereka menginginkan strategi analitik yang lebih komprehensif untuk mencapai tujuan bisnis ini.

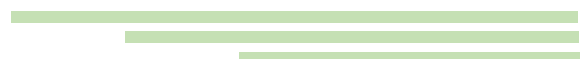
Mempelajari cara menggabungkan teknik pembelajaran mesin modern ke dalam infrastruktur data mereka adalah langkah pertama. Untuk itu, banyak yang mencari perusahaan yang telah memulai proses implementasi dengan sukses. Tepatnya, bisnis di bidang keterlibatan pelanggan memanfaatkan AI dan pembelajaran mesin untuk menganalisis percakapan, baik yang berakhir dengan penjualan maupun yang tidak, dan untuk secara otomatis mengidentifikasi bahasa yang secara alami mengarah pada penjualan atau yang memprediksi kapan penjualan akan terjadi. Permasalahan penting adalah apakah akan memanfaatkan seluruh korpus masukan big data yang tersedia saat menganalisis data dengan algoritma pembelajaran mendalam. Fokus umumnya adalah menerapkan algoritma pembelajaran mendalam untuk melatih pola representasi data tingkat tinggi berdasarkan sebagian korpus masukan yang tersedia, lalu memanfaatkan korpus masukan yang tersisa dengan pola yang telah dipelajari untuk mengekstraksi abstraksi dan representasi data. Dalam konteks permasalahan ini, pertanyaan yang perlu dikaji adalah berapa volume data masukan yang umumnya diperlukan untuk melatih representasi data yang bermanfaat (baik) dengan algoritma pembelajaran mendalam yang kemudian dapat digeneralisasi untuk data baru dalam domain aplikasi big data tertentu.

Salah satu aplikasi bisnis pembelajaran mendalam yang paling banyak dibahas saat ini adalah mobil self-driving, sebuah konsep yang sedang digarap oleh setiap pemain besar, mulai dari Volkswagen hingga Google. Sistem ini menggunakan sensor dan jaringan saraf untuk memproses sejumlah besar data. Mobil belajar mengenali rintangan dan bereaksi dengan tepat, meningkatkan pengetahuannya melalui penggunaan di luar pemrograman pabriknya. Sistem mobil self-driving menggunakan sensor dan jaringan saraf untuk memproses sejumlah besar data. Akhirnya, dengan data yang cukup, mesin belajar cara mengemudi lebih baik daripada manusia.

Penggunaan umum lainnya adalah deteksi gambar dan klasifikasi objek, seperti yang terlihat saat ini dengan Facebook. Perusahaan ini memiliki lebih dari cukup data gambar untuk diolah, sehingga pembelajaran mendalam untuk deteksi gambar sangat mudah diakses. Saat ini, Facebook dapat mengklasifikasikan berbagai objek dalam sebuah gambar dengan akurasi yang sangat tinggi.

3.5 CONTOH PENGGUNAAN BISNIS UNTUK E-COMMERCE

Sejak awal tahun 2010an, pembelajaran mendalam (*deep learning*) mulai diadopsi secara luas dalam dunia bisnis, didorong oleh kemajuan teknologi komputasi berbasis GPU dan melimpahnya data dalam jumlah besar (*big data*). Di Indonesia, penerapan deep learning



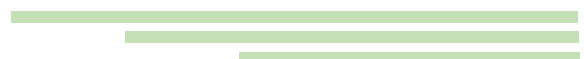
mulai menjadi nyata terutama sejak sekitar tahun 2015, khususnya di sektor e-commerce, layanan finansial, dan telekomunikasi. Perkembangan e-commerce seluler semakin menambah ketersediaan data besar yang sangat beragam dan terus berkembang, memberikan peluang emas bagi perusahaan teknologi dan startup untuk memanfaatkan deep learning. Mereka menggunakan teknologi ini untuk menganalisis perilaku pengguna secara lebih mendalam, mendeteksi aktivitas penipuan secara real-time, serta mengoptimalkan layanan pelanggan melalui chatbot berbasis kecerdasan buatan, sehingga mampu memberikan pengalaman yang lebih personal dan responsif. Melimpahnya data yang berasal dari transaksi dan interaksi di platform e-commerce seluler menjadi fondasi penting dalam pengembangan model deep learning yang akurat dan efektif.

Data dalam jumlah besar tersedia melalui e-commerce seluler. Data ini memungkinkan algoritma pembelajaran mendalam untuk melacak perjalanan pembeli, dan dengan demikian, mesin kini memiliki gambaran yang jelas tentang jenis informasi produk yang dicari pembeli ketika mereka membuat keputusan pembelian untuk berbagai barang. Karena mesin dapat menyimpan dan merujuk sejumlah besar informasi sekaligus sambil menerapkan pengetahuan yang dipelajari, mereka dapat memprediksi jenis pembelian yang kemungkinan dilakukan konsumen sebelum mereka membuat keputusan.

Artinya, dengan mengikuti aktivitas pengguna, situs e-commerce Anda dapat menawarkan opsi pembelian yang menarik bagi pembeli berdasarkan spesifikasi barang lain yang telah mereka lihat, berapa lama mereka menghabiskan waktu untuk melihat setiap barang, dan apa yang mereka lakukan sebelum dan sesudah melihat setiap barang. Situs ini bahkan akan mempertimbangkan waktu dan lokasi pengunjung Anda, apakah mereka pria atau wanita, dan konteks relevan lainnya yang tersedia bagi mesin. Misalnya, Rita sedang mengunjungi toko online dan melihat sepatu lari wanita Blue. Sepatu yang dilihatnya cantik, tetapi tidak persis seperti yang diinginkannya. Ia menginginkan sepatu yang memiliki penyangga lengkung kaki yang lebih baik. Untungnya bagi Rita, situs tersebut menggunakan algoritma pembelajaran mendalam. Rita dapat mengeklik gambar sepatu tersebut dan ia akan ditawarkan produk-produk yang secara visual sebanding untuk dijelajahi. Produk-produk yang direkomendasikan untuknya dipilih dengan menganalisis gambar yang diklik Rita, beserta aktivitas lainnya sejak ia masuk ke situs tersebut. Algoritma tersebut telah menyarankan sepasang sepatu lain dengan karakteristik warna dan bentuk yang serupa.

Sistem ini bekerja dengan sistem masukan-keluaran di mana tindakan dan preferensi Rita bertindak sebagai masukan, dan sepatu yang direkomendasikan kepadanya bertindak sebagai keluaran. Kumpulan piksel dalam gambar yang diklik Rita berfungsi sebagai masukan, memaksa sistem untuk mencoba dan menemukan kumpulan data yang serupa.

Jelas bagaimana ini merupakan cara yang ampuh tidak hanya untuk memasarkan produk Anda tetapi juga menawarkan layanan pelanggan tanpa harus hadir. Pelanggan menemukan apa yang mereka inginkan dengan mudah, mereka memiliki banyak pilihan yang disajikan kepada mereka dan semua informasi yang mereka butuhkan tersedia dengan mudah. Pada dasarnya, situs e-commerce Anda menjadi seperti asisten toko virtual. Pembelajaran mendalam berarti lebih banyak peluang untuk meningkatkan rasio konversi bisnis dan



meningkatkan citra merek melalui pengalaman pelanggan yang positif. Pelanggan Anda tidak hanya akan disugahi konten personalisasi yang ditargetkan untuk pembelian yang mereka lakukan, tetapi juga ditawarkan produk terkait yang kemungkinan besar akan mereka minati, berdasarkan perilaku individu mereka.

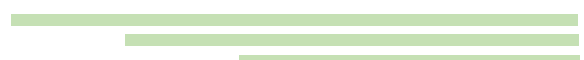
Meskipun pembelajaran mendalam telah diterapkan secara efektif pada tugas-tugas inferensi pola yang menantang, tujuan sebenarnya dari bidang ini jauh melampaui aplikasi spesifik tugas. Cakupan ini dapat membuat perbandingan berbagai metodologi semakin kompleks dan kemungkinan akan membutuhkan upaya bersama oleh komunitas riset untuk mengatasinya. Selain itu, terlepas dari prospek tak terbatas yang ditawarkan oleh teknologi pembelajaran mendalam, beberapa tugas spesifik domain mungkin tidak secara langsung ditingkatkan oleh skema tersebut.

Platform pembelajaran mendalam juga dapat memanfaatkan fitur-fitur rekayasa sambil mempelajari representasi yang lebih kompleks yang biasanya tidak dimiliki oleh sistem rekayasa. Terlepas dari banyaknya pertanyaan penelitian terbuka dan fakta bahwa bidang ini masih dalam tahap awal, kemajuan yang dibuat sehubungan dengan pengembangan sistem pembelajaran mendalam akan membentuk masa depan pembelajaran mesin dan sistem kecerdasan buatan secara luas.

Studi Kasus: Penerapan Deep Learning dalam Industri E-Commerce

Salah satu contoh nyata penerapan deep learning dapat ditemukan pada platform e-commerce besar di Indonesia. Melalui pemanfaatan deep learning, perusahaan berhasil mengatasi berbagai tantangan bisnis sebagai berikut:

Tantangan yang Dihadapi	Solusi Berbasis Deep Learning	Hasil yang Dicapai
Rekomendasi produk yang kurang relevan	Menggunakan jaringan neural untuk personalisasi konten	Meningkatkan tingkat konversi penjualan hingga 25%
Sulit mendeteksi transaksi penipuan	Penerapan deep learning untuk mendeteksi pola anomali	Mengurangi kasus penipuan sebesar 30%
Lambatnya respons layanan pelanggan	Pemanfaatan chatbot berbasis deep learning	Meningkatkan kepuasan pelanggan secara signifikan



Manfaat Utama Deep Learning bagi Bisnis:

🎨 **Prediksi dan Analitik Data:**

Deep learning memungkinkan prediksi yang lebih tepat dengan menganalisis pola dalam data historis perusahaan.

🎨 **Otomatisasi Proses:**

Banyak proses manual seperti klasifikasi dokumen dan inspeksi kualitas di bidang manufaktur sekarang dapat diotomatisasi.

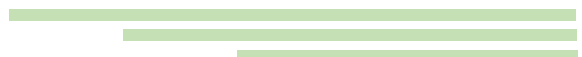
🎨 **Pemahaman Bahasa Alami (NLP):**

Teknologi ini mendukung pengembangan chatbot, pencarian dokumen, serta analisis sentimen di media sosial.

🎨 **Visi Komputer (*Computer Vision*):**

Deep learning digunakan dalam pengenalan wajah, deteksi objek otomatis, hingga pemrosesan gambar medis untuk diagnosa yang lebih akurat.

Dengan integrasi deep learning, platform e-commerce dapat meningkatkan efisiensi operasional, memperkaya pengalaman pengguna secara personal, dan membangun keunggulan kompetitif yang sulit disaingi. Teknologi ini juga memungkinkan penyesuaian strategi bisnis secara cepat berdasarkan data real-time dan wawasan mendalam, menjadikan bisnis lebih responsif terhadap perubahan pasar.



BAB 4

SISTEM MESIN REKOMENDASI

4.1 PENDAHULUAN

Sistem rekomendasi telah muncul sebagai bidang penelitian yang signifikan sejak pertengahan 1990an. Minat terhadap bidang penelitian ini tetap tinggi karena, di satu sisi, aplikasi semacam itu membantu pengguna menemukan item yang relevan, dan di sisi lain, sistem ini berguna untuk mengambil item yang tidak dapat diakses karena pengguna tidak mengetahui keberadaannya. Contoh aplikasi semacam itu antara lain rekomendasi film, berita, buku, dan produk lainnya di *Amazon.com*.

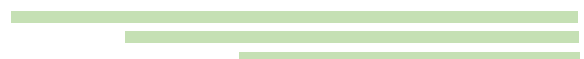
Sistem rekomendasi awalnya dikembangkan untuk menyarankan barang atau layanan yang disesuaikan dengan preferensi masing-masing pengguna. Namun, terdapat situasi di mana sekelompok pengguna secara kolektif berpartisipasi dalam satu aktivitas, seperti menonton film atau bertamasya di kota. Untuk kasus seperti itu, kami memerlukan teknik yang mengatasi masalah dalam mengidentifikasi rekomendasi untuk sekelompok pengguna dan memuaskan hampir semua individu dalam kelompok tersebut semaksimal mungkin.

Salah satu motivasi untuk sistem rekomendasi kelompok; Bidang-bidang yang berkaitan dengan pilihan sosial (juga disebut pengambilan keputusan kelompok) dalam ekonomi, politik, dan sosiologi, di mana kita ingin memutuskan apa yang terbaik bagi suatu kelompok mengingat perbedaan pendapat individu-individu penyusunnya, juga menggambarkan pentingnya rekomendasi kelompok.

Permasalahan rekomendasi kelompok juga telah dipelajari dalam bidang metasearch, middleware basis data, penyaringan kolaboratif, dan sistem multi-agen. Misalnya, dalam metasearch, daftar peringkat yang dihasilkan oleh beberapa mesin pencari harus digabungkan menjadi satu daftar; permasalahan ini dikenal sebagai permasalahan agregasi peringkat. Middleware basis data melibatkan pengurutan objek ketika setiap objek diberi nilai numerik yang terkait dengan beberapa bidang. Dalam sistem multi-agen, agen harus memanfaatkan keputusan yang rasional, tidak hanya dari sudut pandang individu tetapi juga dari sudut pandang sosial.

Oleh karena itu, dapat dikatakan bahwa rekomendasi kelompok berupaya mengidentifikasi hal-hal yang akan diterima oleh suatu kelompok secara keseluruhan, alih-alih hanya oleh anggota kelompok secara individu. Kelompok yang diminati dapat bervariasi dari kelompok yang mapan hingga kelompok acak yang hanya memerlukan rekomendasi sesekali. Sebagian besar penelitian tentang rekomendasi kelompok berfokus pada penyelidikan algoritma inti yang digunakan untuk pembuatan rekomendasi. Sejauh ini, dua strategi utama telah digunakan untuk menghasilkan rekomendasi kelompok. Teknik-teknik ini melibatkan agregasi prediksi individu menjadi prediksi kelompok (prediksi agregat) atau agregasi model individu menjadi model kelompok (model agregat).

Strategi-strategi ini berbeda dalam pengaturan waktu langkah agregasi data. Dalam strategi prediksi agregat, prediksi individual dihasilkan berdasarkan model preferensi



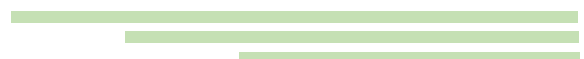
individual dan kemudian diagregasi menjadi prediksi kelompok. Strategi alternatifnya adalah membangun model preferensi kelompok sebelum melakukan rekomendasi item. Dengan kata lain, dalam strategi model agregat, model pengguna individual digabungkan menjadi model berbasis kelompok, dan rekomendasi kemudian dihasilkan menggunakan model kelompok agregat. Namun, banyak penelitian telah menunjukkan bahwa rekomendasi kelompok merupakan tugas yang jauh lebih kompleks daripada rekomendasi individual, dan berbagai tantangan muncul ketika kelompok dipertimbangkan yang mencegah teknik rekomendasi tradisional bekerja secara efisien untuk merekomendasikan item kepada kelompok.

Sistem rekomendasi adalah alat dan teknik yang memberikan saran item yang akan digunakan oleh pengguna. Sistem rekomendasi umumnya diarahkan untuk membantu pengguna menemukan item yang kemungkinan besar mereka minati di antara sejumlah besar item; sistem semacam itu mencoba memprediksi produk atau layanan yang paling sesuai berdasarkan preferensi dan batasan pengguna. Klasifikasi umum data yang digunakan oleh sistem rekomendasi menggunakan tiga kategori: item, pengguna, dan transaksi, yaitu hubungan antara pengguna dan item.

1. Item adalah objek yang direkomendasikan. Item dapat dikarakterisasi berdasarkan kompleksitas dan nilai atau kegunaannya.
2. Pengguna sistem rekomendasi dapat memiliki tujuan dan karakteristik yang sangat beragam. Untuk mempersonalisasi rekomendasinya, sistem rekomendasi mencoba memanfaatkan informasi tentang penggunanya. Informasi ini dapat disusun dengan berbagai cara; pemilihan informasi yang akan dimodelkan bergantung pada teknik rekomendasi yang digunakan.
3. Transaksi adalah interaksi yang terekam antara pengguna dan sistem rekomendasi. Transaksi adalah data yang berguna bagi algoritma pembangkitan rekomendasi. Peringkat adalah bentuk data transaksi paling populer yang dikumpulkan oleh sistem rekomendasi. Jenis data transaksi lainnya dapat mencakup riwayat pembacaan atau seberapa sering pengguna memanfaatkan rekomendasi yang diberikan. Peringkat dapat berupa berbagai bentuk, seperti peringkat numerik, peringkat ordinal, dan peringkat biner.

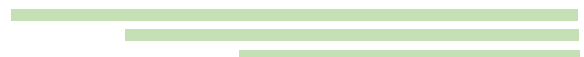
Sistem rekomendasi dapat digunakan untuk berbagai tujuan, seperti yang dijelaskan saat mengidentifikasi tugas-tugas umum yang dapat dibantu oleh sistem rekomendasi, seperti topik-topik yang tercantum di bawah ini.

- ❖ Menemukan item yang diinginkan: Tugas ini melibatkan pencarian daftar item yang diperingkat untuk pengguna dan memprediksi seberapa besar pengguna akan menyukai item tersebut (misalnya peringkat bintang).
- ❖ Menemukan semua item yang diinginkan: Tugas ini melibatkan pencarian semua, bukan hanya beberapa, item yang akan diminati pengguna. Terkadang perlu menemukan semua item yang memungkinkan bagi pengguna, bukan hanya beberapa saja. Contohnya termasuk situasi medis, keuangan, dan hukum, di mana hasil positif palsu berpotensi memiliki konsekuensi besar jika ada.



- ❖ Menyediakan anotasi dalam konteks: Tugas ini melibatkan pencarian item yang berguna dalam konteks pengguna saat ini. Misalnya, sistem dapat memberikan rekomendasi seperti "Pelanggan yang membeli item ini dalam konteks spesifik ini (Natal) juga membeli ini" atau "Film serupa" saat pengguna mencari film.
- ❖ Menemukan urutan yang direkomendasikan: Tugas ini melibatkan pencarian urutan item yang "menyenangkan secara keseluruhan" bagi pengguna, alih-alih hanya mencocokkan preferensi pengguna berdasarkan komponen individual. Contoh sistem rekomendasi tersebut antara lain pembuatan daftar putar lagu dan fungsi radio di Spotify.
- ❖ Penjelajahan Sederhana: Tugas ini melibatkan bantuan kepada pengguna yang ingin menelusuri serangkaian item dengan memberikan saran yang relevan bagi mereka. Rekomendasi semacam itu ditawarkan oleh berbagai layanan, termasuk *Amazon.com* dan eBay.
- ❖ Menemukan Pemberi Rekomendasi yang Kredibel: Beberapa pengguna cenderung tidak mempercayai sistem rekomendasi. Mereka dapat bereksperimen dengan sistem dan mengamati bagaimana sistem tersebut berperilaku dalam berbagai pengaturan, misalnya dengan mengubah profil mereka. Sistem rekomendasi dapat membantu menegaskan kredibilitas mereka dengan menyertakan alasan mengapa suatu item dalam hasil direkomendasikan, misalnya, "lagu ini direkomendasikan karena Anda mendengarkan *xy*".
- ❖ Meningkatkan profil dengan memberi peringkat item: Pengguna dapat membantu meningkatkan hasil algoritma rekomendasi dengan memungkinkan algoritma tersebut memahami jenis item yang disukai pengguna dengan lebih baik.
- ❖ Mengekspresikan diri: Pengguna juga dapat memberi peringkat item karena mereka senang melakukannya, karena mereka suka mengekspresikan diri dan merasa puas. Di sini, tujuan utamanya bukanlah untuk meningkatkan hasil sistem rekomendasi, melainkan untuk mengungkapkan perasaan mereka tentang suatu item.
- ❖ Membantu orang lain: Tujuan utama memberi peringkat suatu item mungkin juga untuk membantu orang lain. Jika seorang pengguna, misalnya, memiliki pengalaman buruk dengan perusahaan tur, memposting tanggapan dan peringkat di situs web seperti *tripadvisor.com* untuk memperingatkan orang lain mungkin menjadi motivasi bagi pengguna untuk menggunakan sistem rekomendasi.
- ❖ Mempengaruhi orang lain: Sistem rekomendasi dapat digunakan untuk mempromosikan suatu bisnis. Situs seperti *hotels.com* dan *tripadvisor.com* dapat digunakan untuk memberikan peringkat kepada hotel, di antara fasilitas dan layanan lainnya. Pendekatan yang tidak etis, seperti mendorong individu (misalnya karyawan) untuk memberikan peringkat tinggi dalam jumlah besar kepada suatu bisnis sementara memberikan peringkat yang lebih rendah kepada pesaingnya, dapat digunakan untuk mempromosikan suatu bisnis.

Cara Kerja Sistem Rekomendasi pada E-Commerce



Sistem rekomendasi di platform e-commerce berfungsi dengan mengumpulkan data perilaku pelanggan dan memanfaatkannya untuk menyajikan rekomendasi produk yang relevan. Proses kerja sistem ini umumnya melibatkan beberapa tahapan penting berikut:

1. **Pengumpulan Data**

Sistem mengumpulkan informasi dari berbagai sumber, seperti riwayat pembelian, pencarian produk, penilaian atau rating, serta preferensi lain yang disampaikan pelanggan. Data ini juga meliputi atribut produk seperti kategori, harga, merek, dan karakteristik lainnya.

2. **Pengolahan dan Analisis Data**

Data yang terkumpul diolah dan dianalisis untuk mengenali pola serta preferensi pelanggan. Proses ini menggunakan metode seperti analisis data statistik, machine learning, dan algoritma khusus untuk pemodelan rekomendasi.

3. **Pemilihan Algoritma Rekomendasi**

Sistem memilih dan menerapkan algoritma yang paling sesuai untuk menghasilkan rekomendasi, antara lain:

- *Collaborative Filtering*: Algoritma yang menganalisis kesamaan pola preferensi antar pengguna. Jika dua pengguna memiliki selera yang mirip, produk yang disukai oleh satu pengguna akan direkomendasikan kepada pengguna lainnya.
- *Content-Based Filtering*: Algoritma yang merekomendasikan produk serupa berdasarkan atribut dan karakteristik produk yang sebelumnya diminati oleh pelanggan.
- *Hybrid Methods*: Menggabungkan berbagai algoritma untuk menghasilkan rekomendasi yang lebih akurat dan personal.

4. **Penyajian Rekomendasi**

Rekomendasi produk ditampilkan kepada pelanggan dalam beragam format, seperti daftar produk di halaman utama situs, email promosi yang dipersonalisasi, atau notifikasi dalam aplikasi mobile.

5. **Pengumpulan Umpan Balik dan Pembaruan Model**

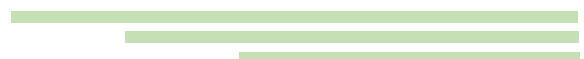
Sistem terus menerus memantau interaksi pelanggan dengan rekomendasi dan mengumpulkan umpan balik untuk memperbaiki kualitas rekomendasi. Pembaruan model dilakukan secara dinamis agar rekomendasi semakin relevan seiring waktu.

6. **Evaluasi Kinerja Sistem**

Kinerja sistem rekomendasi dinilai secara berkala menggunakan metrik seperti tingkat konversi (berapa banyak rekomendasi yang berujung pada pembelian), akurasi prediksi, dan tingkat kepuasan pengguna. Evaluasi ini penting untuk memastikan sistem terus memberikan nilai tambah bagi pengguna dan bisnis.

Dengan mekanisme tersebut, sistem rekomendasi membantu pelanggan menemukan produk yang sesuai dengan preferensi pribadi mereka, memperkaya pengalaman berbelanja, sekaligus meningkatkan efektivitas penjualan bagi pelaku bisnis e-commerce.

Manfaat Sistem Rekomendasi bagi Bisnis E-Commerce



Sistem rekomendasi memberikan dampak signifikan dalam dunia e-commerce. Dengan penggunaan yang optimal, platform dapat membuat pengguna betah berlama-lama menjelajahi situs dan melakukan pembelian lebih banyak. Berikut adalah beberapa kisah sukses dari perusahaan e-commerce yang mampu meningkatkan performa bisnis mereka melalui integrasi sistem rekomendasi:

- Meningkatkan tingkat konversi penjualan secara signifikan
- Meningkatkan loyalitas serta retensi pelanggan
- Mempercepat proses pengambilan keputusan pembelian pelanggan
- Mengoptimalkan penawaran produk yang relevan sesuai kebutuhan pengguna

Sistem rekomendasi telah menjadi elemen yang sangat penting. Teknologi ini memungkinkan situs-situs e-commerce untuk memberikan rekomendasi produk kepada pelanggan berdasarkan preferensi mereka. Sistem rekomendasi memanfaatkan data pelanggan, seperti riwayat pembelian, penelusuran, peringkat produk, dan preferensi lainnya, untuk memberikan rekomendasi yang sesuai.

Proses kerja sistem rekomendasi melibatkan beberapa tahap, dimulai dari pengumpulan data yang mencakup berbagai informasi yang relevan, seperti kategori produk, harga, merek, dan sebagainya. Data ini kemudian diproses dan dianalisis menggunakan teknik seperti analisis data, machine learning, dan algoritma rekomendasi. Berbagai jenis algoritma rekomendasi digunakan, termasuk Collaborative Filtering, Content-Based Filtering, dan Hybrid Methods, untuk menghasilkan rekomendasi yang sesuai.

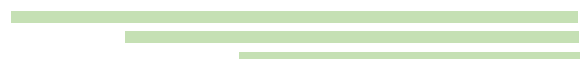
Sistem rekomendasi memberikan rekomendasi kepada pelanggan dalam berbagai bentuk, seperti daftar produk yang direkomendasikan di situs web, rekomendasi dalam email, atau pemberitahuan di aplikasi seluler. Selain itu, sistem ini terus memantau perilaku pelanggan dan menerima umpan balik untuk meningkatkan kualitas rekomendasi seiring berjalannya waktu.

Manfaat dari sistem rekomendasi dalam e-commerce sangat signifikan. Mereka meningkatkan pengalaman pelanggan dengan membantu mereka menemukan produk yang sesuai dengan preferensi mereka. Ini, pada gilirannya, berkontribusi pada peningkatan penjualan dan pertumbuhan bisnis e-commerce. Beberapa perusahaan terkenal, seperti Amazon dan Moviefinder.com, telah berhasil mengimplementasikan sistem rekomendasi untuk memberikan layanan yang lebih baik kepada pelanggan mereka.

Dalam rangka memastikan keberhasilan dan efektivitas sistem rekomendasi, evaluasi kinerja secara berkala dilakukan dengan mempertimbangkan berbagai metrik, seperti tingkat konversi, akurasi rekomendasi, dan kepuasan pelanggan. Dengan cara ini, sistem rekomendasi telah membawa dampak positif yang besar pada industri e-commerce, menciptakan pengalaman belanja yang lebih personal dan memadukan preferensi pelanggan dengan penawaran produk yang tersedia.

4.2 TEKNIK SISTEM REKOMENDASI

Teknik rekomendasi dapat menggunakan beberapa klasifikasi berbeda. Klasifikasi ini didasarkan pada data yang digunakan sebagai masukan untuk rekomendasi. Pada bagian



berikut, kami akan menyajikan klasifikasi yang diusulkan dalam literatur dan akan memberikan survei tentang beragam jenis sistem rekomendasi. Sistem rekomendasi biasanya diklasifikasikan ke dalam kategori berikut, berdasarkan cara rekomendasi dibuat:

Rekomendasi Berbasis Konten

Pendekatan berbasis konten berlandaskan pada pencarian informasi dan didasarkan pada gagasan bahwa pengguna akan lebih menyukai item yang serupa dengan item yang sebelumnya mereka sukai. Sistem rekomendasi berbasis konten menganalisis deskripsi item untuk mengidentifikasi item yang secara khusus menarik bagi pengguna, dan berfokus pada fitur item tersebut.

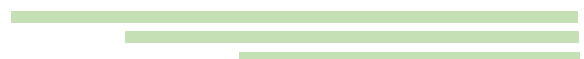
Kesamaan item ditentukan dengan mengukur kesamaan fiturnya. Sistem rekomendasi memang berusaha merekomendasikan item yang serupa dengan item yang pernah disukai pengguna sebelumnya. Sistem semacam itu menganalisis serangkaian fitur item yang sebelumnya telah dinilai oleh pengguna dan membangun model atau profil minat pengguna berdasarkan fitur item yang dinilai oleh pengguna tersebut. Misalnya, model semacam itu dapat digunakan untuk memfilter hasil pencarian dengan menentukan apakah pengguna tertarik pada halaman web tertentu. Contoh rekomendasi semacam itu adalah rekomendasi hotel tertentu berdasarkan jumlah bintang, tingkat harga, atau apakah hotel tersebut menawarkan layanan luar ruangan. Algoritma penyaringan berbasis konten untuk rekomendasi hotel secara konseptual dapat dijelaskan sebagai berikut:

Sistem mengumpulkan serangkaian fitur yang dikombinasikan dengan probabilitas bahwa setiap pengguna akan menyukai setiap fitur. Untuk menentukan hotel yang kemungkinan besar akan disukai pengguna, sistem melakukan perbandingan antara rangkaian fitur ini dan berbagai hotel dalam sistem. Perbandingan tersebut dapat mencakup heuristik atau sekadar perhitungan kesamaan seperti perhitungan berbasis kosinus. Sistem kemudian mengembalikan hotel yang paling sesuai dengan preferensi pengguna. Pemfilteran berbasis konten dapat sangat berguna ketika sejumlah besar data deskriptif tersedia untuk item-item tersebut. Data ini dapat dikodifikasi dan direpresentasikan secara internal sebagai ontologi fitur, sehingga memberikan deskripsi detail domain yang berlaku untuk perhitungan kesamaan terhadap preferensi pengguna. Namun demikian, agar bermanfaat, skema berbasis konten harus memiliki ontologi tersebut selain kemampuan untuk mengklasifikasikan item menurut ontologi tersebut. Oleh karena itu, pemfilteran berbasis konten dapat menjadi masalah ketika informasi yang tersedia untuk setiap item jarang atau sulit diinterpretasikan menggunakan algoritma komputer. Lebih lanjut, rekomendasi berbasis konten hanya dapat didasarkan pada informasi eksplisit yang tersedia mengenai item tertentu.

Beberapa komponen, seperti representasi item, profil pengguna, dan kemampuan untuk mempelajari model pengguna, hadir dalam arsitektur tingkat tinggi pemfilteran berbasis konten; komponen-komponen ini akan dijelaskan secara rinci di bawah ini.

Representasi Item

Item yang dapat direkomendasikan kepada pengguna direpresentasikan oleh serangkaian fitur. Misalnya, setiap film dikaitkan dengan fitur-fitur tertentu, seperti genre, aktor, subjek, dan tanggal. Ketika setiap item dideskripsikan oleh serangkaian fitur dan nilai



fitur-fitur ini diketahui, item-item tersebut dapat direpresentasikan oleh data terstruktur. Ketika data tidak memiliki struktur, seperti halnya teks, langkah prapemrosesan diperlukan untuk mengekstrak informasi terstruktur yang relevan.

Tugas utama komponen tersebut adalah merepresentasikan konten item, misalnya, dokumen, halaman web, item berita, dan deskripsi produk, dalam bentuk yang sesuai untuk digunakan dalam langkah-langkah pemrosesan selanjutnya. Item data dianalisis menggunakan teknik ekstraksi fitur untuk menggeser representasi item dari ruang informasi asli ke ruang target. Misalnya, halaman web dapat direpresentasikan sebagai vektor kata kunci.

Profil Pengguna

Komponen ini membuat model berdasarkan minat pengguna yang disimpulkan dari suka dan tidak suka item sebelumnya. Biasanya, strategi generalisasi diimplementasikan menggunakan teknik pembelajaran mesin, yang mampu menyimpulkan model minat pengguna dari item yang disukai atau tidak disukai sebelumnya. Misalnya, untuk rekomendasi berita berdasarkan riwayat membaca pengguna, berbagai properti artikel berita, seperti konten berita, pola akses, dan entitas bernama, dapat digabungkan untuk membangun profil pengguna.

Profil ini kemudian digunakan sebagai representasi item untuk mengidentifikasi item yang sesuai untuk pengguna target. Representasi ini biasanya terdiri dari vektor fitur berbobot: semakin besar bobotnya, semakin pengguna menyukai fitur tersebut. Ukuran vektor sesuai dengan jumlah fitur. Terdapat beberapa pendekatan yang dapat digunakan untuk membangun profil pengguna dari profil item. Dimungkinkan untuk menghitung setiap profil pengguna sebagai vektor rata-rata berdasarkan profil item yang dievaluasi oleh pengguna.

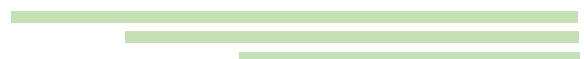
Pembelajaran Model Pengguna

Modul ini memanfaatkan profil pengguna untuk menyarankan item yang relevan dengan mencocokkan representasi profil dengan item yang akan direkomendasikan. Hasilnya mungkin berdasarkan pada penilaian relevansi biner atau berkelanjutan menggunakan berbagai metrik kesamaan.

Rekomendasi Kolaboratif

Dalam pendekatan ini, alih-alih menggunakan fitur item untuk menentukan kesamaannya, fokusnya ditempatkan pada kesamaan peringkat pengguna untuk item tersebut. Misalnya, dalam rekomendasi film, untuk memprediksi peringkat pengguna u untuk film i , teknik penyaringan kolaboratif mencoba mengidentifikasi pengguna yang serupa dengan u , yaitu pengguna lain yang memiliki selera film yang sama (yaitu, mereka menilai film yang sama dengan cara yang sama). Kemudian, teknik ini menggabungkan peringkat yang diungkapkan oleh pengguna dengan pola peringkat yang serupa dan menghitung peringkat yang diprediksi untuk pengguna tertentu.

Pemfilteran kolaboratif sejauh ini merupakan teknik rekomendasi yang paling luas karena kekuatan dan kesederhanaannya. Kekuatan terbesar dari teknik kolaboratif adalah bahwa teknik ini sepenuhnya independen dari representasi objek yang direkomendasikan yang dapat dibaca mesin dan dapat diterapkan secara efektif pada objek yang kompleks, seperti musik dan film, yang mungkin disukai atau tidak disukai pengguna karena berbagai alasan.



Misalnya, pengguna mungkin menyukai genre film tertentu, sutradara tertentu, aktor atau aktris tertentu, dll.

Penyaringan kolaboratif sering disebut sebagai penyaringan sosial karena teknik ini mengenali kesamaan di antara pengguna berdasarkan peringkat mereka dan menghasilkan rekomendasi baru berdasarkan perbandingan antar-pengguna. Proses ini meniru proses sosial merekomendasikan item kepada teman karena sistem memberikan rekomendasi kepada pengguna berdasarkan item yang disukai oleh orang-orang yang serupa dengan pengguna dalam hal peringkat di masa lalu.

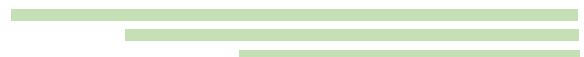
Algoritma yang digunakan dalam sistem penyaringan kolaboratif dapat dikelompokkan menjadi dua kelas umum: algoritma berbasis memori (atau berbasis heuristik) dan algoritma berbasis model. Algoritma berbasis memori memberikan prediksi peringkat untuk pengguna berdasarkan peringkat mereka sebelumnya. Pendekatan berbasis memori menerapkan heuristik yang menggunakan semua atau sampel data pengguna-item untuk menghasilkan prediksi. Dua pendekatan umum untuk penyaringan kolaboratif berbasis memori (CF) digunakan: penyaringan kolaboratif berbasis pengguna dan penyaringan kolaboratif berbasis item.

Penyaringan kolaboratif berbasis pengguna menghasilkan prediksi untuk pengguna tertentu dengan terlebih dahulu mengidentifikasi pengguna yang serupa (tetangga terdekat) dengan pengguna tersebut, lalu menghitung item yang paling sering diberi peringkat yang belum pernah dilihat oleh pengguna tersebut. Misalnya, sistem rekomendasi penyaringan kolaboratif untuk buku dapat menyarankan buku yang belum pernah dibaca oleh pengguna tertentu, tetapi telah dibaca oleh pengguna dengan selera yang sama. Penyaringan kolaboratif berbasis item menghasilkan prediksi untuk pengguna tertentu dengan terlebih dahulu menghitung item yang paling mirip dengan setiap item berdasarkan kemiripannya.

Algoritme berbasis model memodelkan pengguna berdasarkan peringkat mereka sebelumnya, dan sistem rekomendasi kemudian menggunakan model ini untuk memprediksi peringkat item yang belum diberi peringkat. Dalam metode ini, setelah algoritma mempelajari model, algoritma tersebut tidak lagi memerlukan akses ke setiap peringkat dan dapat menggunakan parameter model untuk menghitung prediksinya. Pendekatan berbasis model (seperti algoritma pembelajaran mesin dan penambangan data) digunakan untuk merancang dan mengembangkan model yang memungkinkan sistem mempelajari pola kompleks berdasarkan data pelatihan, lalu memberikan prediksi berdasarkan model yang dipelajari.

Keuntungan pendekatan ini adalah setelah model dipelajari, sistem tidak lagi memerlukan akses ke setiap peringkat, dan parameter model dapat digunakan untuk menghitung prediksi; namun, fase pembelajaran dapat memakan waktu. Sebaliknya, pendekatan berbasis memori lebih lambat selama fase prediksi karena harus mengeksplorasi rangkaian peringkat secara heuristik, sedangkan pendekatan berbasis model cukup cepat karena hanya menggunakan parameter model untuk menghitung peringkat yang diprediksi.

Berdasarkan analisis terhadap perilaku pengguna, seperti penilaian (rating) dan interaksi mereka dengan berbagai item, sistem rekomendasi kolaboratif dibagi menjadi dua kategori utama:



- **User-Based Collaborative Filtering:** Teknik ini merekomendasikan item kepada pengguna dengan cara memanfaatkan preferensi pengguna lain yang memiliki selera atau minat yang serupa. Dengan kata lain, rekomendasi berasal dari kesamaan pola penggunaan antar pengguna.
- **Item-Based Collaborative Filtering:** Pendekatan ini memberikan rekomendasi dengan mencari item yang memiliki kemiripan atau karakteristik yang serupa dengan item yang sebelumnya telah dinikmati atau disukai oleh pengguna tersebut. Jadi, fokusnya adalah pada hubungan antar item, bukan antar pengguna.

Pendekatan Hibrida

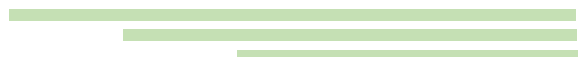
Metode populer lainnya untuk merumuskan rekomendasi adalah penggunaan pendekatan hibrida. Seseorang dapat mengimplementasikan sistem kolaboratif dan berbasis konten secara terpisah dan menggabungkan hasilnya, sehingga menggabungkan karakteristik dari satu metode ke metode lainnya atau membuat model pemersatu yang memiliki karakteristik dari kedua pendekatan. Penggunaan teknik semacam itu membantu mengurangi keterbatasan tertentu dari sistem berbasis konten dan kolaboratif.

- Menerapkan metode kolaboratif dan berbasis konten secara terpisah dan menggabungkan prediksinya
- Menggabungkan beberapa karakteristik berbasis konten ke dalam pendekatan kolaboratif
- Menggabungkan beberapa karakteristik kolaboratif ke dalam pendekatan berbasis konten
- Membangun model pemersatu umum yang menggabungkan karakteristik berbasis konten dan kolaboratif

4.3 PENERAPAN MESIN REKOMENDASI DALAM BISNIS

Jika Anda pernah menggunakan layanan streaming Netflix, membeli sesuatu dari Amazon, atau terhubung dengan "orang yang mungkin Anda kenal" di LinkedIn atau Facebook, maka Anda telah menggunakan mesin rekomendasi. Mesin rekomendasi mencocokkan fitur-fitur tentang Anda dengan hal-hal yang mungkin Anda minati. Misalnya, sebuah film memiliki tahun rilis, genre, aktor, dan hasil box office. Anda memiliki fitur-fitur. Anda memiliki preferensi, usia, dan Anda mungkin telah menyelesaikan survei yang mengungkapkan beberapa sikap Anda terhadap film-film tertentu. Anda mungkin telah memberi peringkat pada beberapa film yang Anda tonton. Dengan menentukan film mana yang akan ditampilkan kepada Anda, dan respons Anda terhadap rekomendasi tersebut, mesin tersebut belajar seiring waktu untuk memberikan saran yang lebih baik.

Jika Anda menonton beberapa film fiksi ilmiah, dan Anda memberi peringkat tinggi, maka mesin akan belajar untuk menampilkan lebih banyak film fiksi ilmiah kepada Anda, dan, untuk variasi, film-film yang mungkin juga disukai oleh orang lain seperti Anda, yang menyukai film fiksi ilmiah. Sistem rekomendasi telah diterapkan di berbagai industri dan konteks,



terutama sebagai bagian dari situs belanja online, jelas Thompson. Anda biasanya akan melihatnya digunakan di:

- Supermarket: Kasir yang menghasilkan kupon khusus untuk pembelian berikutnya melihat pembelian sebelumnya dan menyesuaikan kupon agar mencakup barang-barang yang mungkin Anda minati.
- Toko buku dan musik: Perusahaan dapat mengirimkan email kepada pelanggan setelah pembelian dan merekomendasikan buku atau album baru atau seperti Amazon, Pandora, dan situs serupa, memberikan rekomendasi produk secara real-time berdasarkan apa yang sedang dilihat atau didengarkan pelanggan.
- Perusahaan investasi: Sistem rekomendasi dapat menganalisis saham mana yang mungkin Anda minati berdasarkan pilihan pelanggan serupa.
- Layanan TV dan film: Perusahaan seperti Netflix menganalisis pilihan konten sebelumnya setiap pelanggan dan membuat rekomendasi berdasarkan pilihan tersebut, serta secara proaktif merekomendasikan barang secara real-time berdasarkan riwayat penelusuran mereka.
- Situs jejaring sosial: Situs seperti LinkedIn dan Facebook menggunakan sistem rekomendasi untuk menyarankan koneksi atau teman tambahan berdasarkan jaringan yang dimiliki seseorang.

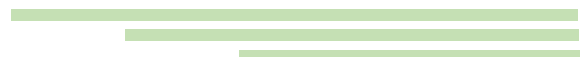
Banyak algoritma yang digunakan dalam mesin rekomendasi dan pembelajaran mesin bukanlah hal yang baru. Regresi, pohon keputusan, k -nearest neighbor, support vector machines (SVM), jaringan saraf tiruan, dan Naive Bayes adalah metode-metode yang telah mapan dengan batasan-batasan yang umum diketahui dan penggunaannya yang tepat. Banyak dari metode ini telah lama digunakan untuk mendukung pengambilan keputusan bisnis berbasis data. Mesin rekomendasi pada umumnya memproses data melalui empat fase berikut, yaitu pengumpulan, penyimpanan, analisis, dan pemfilteran.

Pengumpulan Data

Langkah pertama dalam membuat mesin rekomendasi adalah mengumpulkan data. Data dapat berupa data eksplisit maupun implisit. Data eksplisit terdiri dari data yang dimasukkan oleh pengguna seperti peringkat dan komentar produk. Data implisit meliputi riwayat pesanan/riwayat pengembalian, peristiwa Keranjang, Tampilan Halaman, Klik-tayang, dan log pencarian. Kumpulan data ini akan dibuat untuk setiap pengguna yang mengunjungi situs.

Data perilaku mudah dikumpulkan karena Anda dapat menyimpan log aktivitas pengguna di situs Anda. Pengumpulan data ini juga mudah karena tidak memerlukan tindakan tambahan dari pengguna; mereka sudah menggunakan aplikasi. Kelemahan dari pendekatan ini adalah lebih sulit untuk menganalisis data. Misalnya, memfilter log yang dibutuhkan dari yang kurang dibutuhkan bisa jadi rumit.

Karena setiap pengguna pasti memiliki kesukaan atau ketidaksukaan yang berbeda terhadap suatu produk, kumpulan data mereka akan berbeda. Seiring waktu, saat Anda "memberi" lebih banyak data ke mesin, mesin tersebut akan semakin pintar dalam memberikan rekomendasi sehingga pelanggan dan pelanggan email Anda lebih mungkin untuk



berinteraksi, mengklik, dan membeli. Sama seperti cara kerja mesin rekomendasi Amazon dengan tab "Sering dibeli bersama" dan "Direkomendasikan untuk Anda".

Menyimpan Data

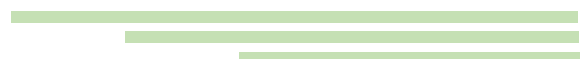
Semakin banyak data yang dapat Anda sediakan untuk algoritma Anda, semakin baik pula rekomendasi yang dihasilkan. Ini berarti setiap proyek rekomendasi dapat dengan cepat berubah menjadi proyek big data. Jenis data yang Anda gunakan untuk membuat rekomendasi dapat membantu Anda menentukan jenis penyimpanan yang sebaiknya digunakan. Anda dapat memilih untuk menggunakan database NoSQL, database SQL standar, atau bahkan penyimpanan objek. Masing-masing opsi ini dapat diterapkan tergantung pada apakah Anda menyimpan masukan atau perilaku pengguna dan faktor-faktor seperti kemudahan implementasi, jumlah data yang dapat dikelola penyimpanan, integrasi dengan lingkungan lainnya, dan portabilitas. Saat menyimpan peringkat atau komentar pengguna, database yang skalabel dan terkelola meminimalkan jumlah tugas yang diperlukan dan membantu untuk fokus pada rekomendasi.

Menganalisis Data

Untuk menemukan item yang memiliki data keterlibatan pengguna yang serupa, kami memfilter data dengan menggunakan metode analisis yang berbeda. Jika Anda ingin memberikan rekomendasi langsung kepada pengguna saat mereka melihat produk, maka Anda memerlukan jenis analisis yang lebih gesit. Beberapa cara untuk menganalisis data adalah: Sistem waktu nyata (*real-time*) dapat memproses data saat data tersebut dibuat. Sistem jenis ini biasanya melibatkan alat yang dapat memproses dan menganalisis aliran peristiwa. Sistem waktu nyata diperlukan untuk memberikan rekomendasi langsung.

Analisis batch mengharuskan Anda untuk memproses data secara berkala. Pendekatan ini menyiratkan bahwa data yang cukup perlu dibuat agar analisis relevan, seperti volume penjualan harian. Sistem batch mungkin berfungsi dengan baik untuk mengirim email di kemudian hari. Analisis mendekati waktu nyata (*near-real-time*) memungkinkan Anda mengumpulkan data dengan cepat, sehingga Anda dapat memperbarui analitik setiap beberapa menit atau detik. Sistem mendekati waktu nyata (*near-real-time*) berfungsi paling baik untuk memberikan rekomendasi selama sesi penelusuran yang sama.

Algoritma Rekomendasi Produk



Filtering the data: To get the relevant data necessary to provide recommendations to the user. We have to choose an algorithm that would better suit the recommendation engine.

Content-based: A popular, recommended product has similar characteristics to what a user views or likes.

Cluster: Recommended products go well together, no matter what other users have done.

Collaborative: Other users, who like the same products as another user views or likes, will also like a recommended product. Collaborative filtering enables you to make product attributes theoretical and make predictions based on user tastes.

The output of this filtering is based on the assumption that two users who liked the same products in the past will probably like the same ones now or in the future.

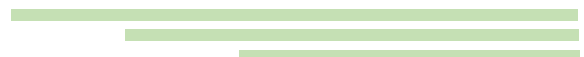
Anda dapat merepresentasikan data tentang peringkat atau interaksi sebagai sekumpulan matriks, dengan produk dan pengguna sebagai dimensinya. Asumsikan bahwa dua matriks berikut serupa, tetapi kemudian kita kurangi yang kedua dari yang pertama dengan mengganti peringkat yang ada dengan angka satu dan peringkat yang hilang dengan angka nol. Matriks yang dihasilkan adalah tabel kebenaran di mana angka satu merepresentasikan interaksi pengguna dengan suatu produk.

Pada akhirnya, hasil yang diperoleh setelah penyaringan dan penggunaan algoritma, rekomendasi diberikan kepada pengguna berdasarkan ketepatan waktu jenis rekomendasi, baik rekomendasi waktu nyata maupun pengiriman email setelah beberapa waktu. Karena mesin rekomendasi produk sebagian besar berjalan pada data, bisnis mungkin tidak memiliki kapasitas penyimpanan untuk menyimpan sejumlah besar data dari pengunjung situs Anda. Anda dapat menggunakan kerangka kerja daring seperti Hadoop dan Spark yang memungkinkan Anda menyimpan data di beberapa perangkat untuk mengurangi ketergantungan pada satu mesin. Hadoop menggunakan HDFS untuk membagi berkas menjadi blok-blok besar dan mendistribusikannya ke seluruh node dalam sebuah kluster. Hal ini memungkinkan kumpulan data diproses lebih cepat dan lebih efisien dibandingkan dengan arsitektur superkomputer konvensional yang mengandalkan sistem berkas paralel di mana komputasi dan data didistribusikan melalui jaringan berkecepatan tinggi.

Terakhir, kami memproses kumpulan data besar menggunakan model pemrograman MapReduce. Dengan ini, kami dapat menjalankan algoritme dalam sistem berkas terdistribusi secara bersamaan dan memilih kluster yang paling mirip. Dengan demikian, setiap bisnis dapat mengembangkan arsitektur mesin rekomendasinya sendiri menggunakan perangkat lunak sumber terbuka, dan kami dapat membantu mereka dalam mengimplementasikan mesin tersebut dengan memanfaatkan keahlian teknis kami.

4.4 KASUS PENGGUNAAN BISNIS

Mesin rekomendasi memainkan peran penting dalam keterlibatan dan retensi pelanggan di industri media dan hiburan daring. Dengan volume data media yang



eksponensial, mesin rekomendasi dengan big data menunjukkan pendekatan penyampaian media yang modern dan berpusat pada pengguna melalui pemrosesan data yang efisien, pembelajaran mesin, dan analitik prediktif.

Mesin rekomendasi media dapat dibuat untuk film atau video musik, buku, atau produk apa pun. Kasus penggunaan data besar (big data) untuk video film tersebut berfokus pada arsitektur mesin rekomendasi bagi konsumen yang menggunakan dekoder (STB), yang:

- Menggunakan arsitektur Hadoop sumber terbuka sebagai fondasi big data
- Mengumpulkan data mentah pengguna dari video sesuai permintaan, log aktivitas dekoder, rekaman terjadwal, dan beberapa katalog media
- Memproses dan menganalisis data log pengguna dalam kerangka big data Hadoop
- Memasukkan hasilnya ke mesin pencari yang kemudian memberikan rekomendasi unik melalui antarmuka peramban yang dapat diakses pengguna

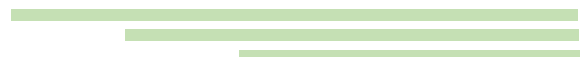
Pada tingkat yang lebih rinci, perilaku pengguna individu, seperti video yang ditonton, katalog yang diklik, program yang dijadwalkan untuk direkam, rata-rata waktu menonton video, dianalisis secara menyeluruh mengikuti metodologi analitik log big data. Pada tingkat yang lebih tinggi, data yang sangat besar ini dapat menggambarkan "apa yang sedang tren" bagi pengguna tertentu atau di antara kelompok pengguna dengan selera yang sama.

Bagi analisis data, arsitektur mesin rekomendasi ini lebih dari sekadar menjalankan kueri SQL terhadap gudang data untuk memprediksi tren dan preferensi. Big data memungkinkan mereka menjadi lebih efisien dengan memproses data pengguna dalam jumlah besar dalam waktu yang jauh lebih singkat dibandingkan dengan SQL tradisional. Hasil yang dipersonalisasi dimuat ke mesin pencari dan ditampilkan di aplikasi web bawaan.

Bagi pemasar, ini merupakan alat yang sangat baik untuk menggali persona pengguna dan memberikan apa yang diinginkan pengguna. Bagi pengguna akhir, rekomendasi yang dipersonalisasi menghemat pekerjaan manual mereka dalam menelusuri basis data video yang sangat besar. Pembelajaran mesin dan analitik prediktif memungkinkan mesin rekomendasi menjadi lebih akurat dalam memprediksi preferensi pengguna, sehingga meningkatkan kepuasan dan retensi pengguna.

Big data, dengan skalabilitas dan kemampuannya untuk memproses data terstruktur dalam jumlah besar (misalnya, judul video yang dicari pengguna, genre musik yang mereka sukai) dan tidak terstruktur (misalnya, pola menonton/mendengarkan pengguna), memungkinkan perusahaan untuk menganalisis miliaran klik dan data tayangan dari Anda dan pengguna lain seperti Anda untuk mendapatkan rekomendasi terbaik. Meskipun mesin rekomendasi telah diadopsi dan dipelajari secara luas selama lebih dari satu dekade, beberapa tantangan utama tetap ada.

Skalabilitas: Jumlah data yang digunakan sebagai input ke mesin rekomendasi tumbuh pesat seiring bertambahnya pengguna dan item. Misalnya, untuk situs web populer, ukuran data perilaku pengguna yang tersimpan dapat dengan mudah mencapai terabyte per hari. Meskipun jumlah datanya besar, sebagian besar mesin rekomendasi berupaya untuk merespons secara interaktif dalam waktu kurang dari satu detik agar pengguna tetap terlibat.



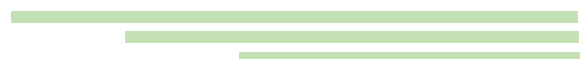
Tantangan utama di sini adalah merancang algoritma pembelajaran yang efisien yang dapat menangani kumpulan data berskala besar tersebut.

Privasi: Dengan memahami nilai data pengguna, sebagian besar situs web mengumpulkan data pengguna sebanyak mungkin. Pendekatan ini menimbulkan kekhawatiran privasi karena data tersebut mungkin berisi informasi sensitif yang ingin dirahasiakan oleh pengguna, misalnya alamat dan riwayat pembayaran pengguna. Meskipun pengguna diberikan kebijakan privasi terkait penggunaan data, mereka biasanya tidak memiliki kendali eksplisit atas data tersebut.

Kelangkaan: Kelangkaan adalah masalah kurangnya informasi. Di toko online yang memiliki banyak pengguna dan item, hampir selalu ada pengguna yang hanya menilai beberapa item. Dengan menggunakan pendekatan kolaboratif dan pendekatan lainnya, mesin rekomendasi pada dasarnya membuat lingkungan pengguna menggunakan profil mereka. Jika pengguna hanya mengevaluasi beberapa item, maka cukup sulit untuk menentukan seleranya dan pengguna tersebut mungkin berada di lingkungan yang salah.

Rekomendasi terstruktur: Mesin rekomendasi saat ini memprediksi item individual yang mungkin diinginkan pengguna. Ekstensi yang menarik adalah memprediksi preferensi untuk set item. Misalnya, jika sistem mengetahui bahwa pengguna akan melakukan olahraga musim dingin untuk pertama kalinya, sistem dapat merekomendasikan sepasang sepatu bot, helm, kacamata, dan pakaian ski dengan warna dan harga yang serasi. Dengan cara ini, pengguna dapat memperoleh semua yang dibutuhkan untuk bermain ski hanya dengan satu pembelian. Ada dua tantangan dalam rekomendasi terstruktur semacam ini. Pertama, jumlah set yang memungkinkan bertambah secara eksponensial seiring dengan ukuran grup. Mengingat jumlah item sudah sangat besar, efisiensi algoritma pembelajaran dapat menjadi masalah. Kedua, tidak seperti item individual, tidak jelas bagaimana memilih fungsi skor yang tepat untuk set.

Kepercayaan: Suara orang-orang dengan riwayat singkat mungkin tidak serelevan suara mereka yang memiliki riwayat panjang di profil mereka. Masalah kepercayaan muncul dalam hal evaluasi terhadap pelanggan tertentu. Masalah ini dapat diatasi dengan mendistribusikan prioritas kepada pengguna.



BAB 5

PEMROSESAN BAHASA ALAMI

5.1 PENDAHULUAN

Ketika kita berkomunikasi satu sama lain, kita menggunakan proses yang sangat kompleks namun masih sedikit dipahami, hampir tanpa usaha. Sangat sulit untuk membangun komputer yang dapat menghasilkan dan memahami bahkan fragmen bahasa alami. Hal ini karena bahasa telah berkembang sebagai media komunikasi yang efektif antarmakhluk cerdas.

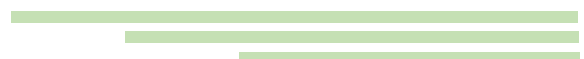
Bahasa dianggap mentransmisikan sedikit "struktur mental" dari satu otak ke otak lainnya ketika masing-masing otak memiliki struktur mental yang besar dan sangat mirip yang berfungsi sebagai konteks umum. Kesamaan dalam konteks ini membantu dalam menghasilkan dan memahami pesan yang sangat padat. Dengan demikian, pemahaman bahasa alami merupakan masalah pengkodean dan dekode yang sangat kompleks.

Salah satu tujuan jangka panjang AI adalah menciptakan program yang mampu memahami dan menghasilkan bahasa manusia. Kemampuan untuk menggunakan dan memahami bahasa alami tidak hanya tampaknya menjadi aspek fundamental dari kecerdasan manusia, tetapi juga keberhasilan otomatisasinya akan berdampak luar biasa pada kegunaan dan efektivitas komputer itu sendiri. Banyak upaya telah dilakukan untuk menulis program yang memahami bahasa alami. Meskipun program-program ini telah mencapai keberhasilan dalam konteks terbatas, sistem yang dapat menggunakan bahasa alami dengan fleksibilitas dan generalitas yang menjadi ciri khas ucapan manusia berada di luar metodologi yang ada saat ini.

Memahami bahasa alami melibatkan lebih dari sekadar mengurai kalimat menjadi bagian-bagian ujarannya masing-masing dan mencari kata-kata tersebut di kamus. Pemahaman yang sesungguhnya bergantung pada pengetahuan latar belakang yang luas tentang ranah wacana dan idiom yang digunakan dalam ranah tersebut serta kemampuan untuk menerapkan pengetahuan kontekstual umum guna mengatasi kelalaian dan ambiguitas yang merupakan bagian normal dari ucapan manusia.

Oleh karena itu, untuk membangun sistem komputer yang dapat memahami bahasa alami, diperlukan pengetahuan kontekstual dan proses untuk membuat inferensi yang efektif. Pemrosesan bahasa alami (NLP) telah mentransformasi kecerdasan bisnis (BI), dengan cara yang jauh melampaui sekadar mempermudah antarmuka. Pemrosesan bahasa alami adalah disiplin ilmu yang berkaitan dengan membuat bahasa alami dapat diakses oleh mesin. NLP menangani tugas-tugas seperti mengidentifikasi batasan kalimat dalam dokumen, mengekstraksi hubungan dari dokumen, serta mencari dan mengambil dokumen, dan lain-lain. NLP merupakan sarana yang diperlukan untuk memfasilitasi analisis teks dengan membangun struktur dalam teks yang tidak terstruktur untuk memungkinkan analisis lebih lanjut.

Pemahaman sederhananya adalah bahwa kalimat-kalimat dalam sebuah teks pertama-tama dianalisis berdasarkan sintaksisnya; hal ini memberikan tatanan dan struktur yang lebih mudah dianalisis berdasarkan semantik, atau makna literal; dan ini diikuti oleh tahap analisis



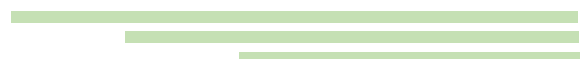
pragmatik yang menentukan makna ujaran atau teks dalam konteksnya. Tahap terakhir ini sering dianggap berkaitan dengan wacana, sedangkan dua tahap sebelumnya umumnya berkaitan dengan hal-hal yang berkaitan dengan kalimat.

Upaya korelasi antara perbedaan stratifikasional (sintaksis, semantik, dan pragmatik) dan perbedaan dalam hal granularitas (kalimat versus wacana) terkadang menyebabkan kebingungan dalam memikirkan isu-isu yang terlibat dalam pemrosesan bahasa alami, dan secara luas diakui bahwa dalam praktiknya tidaklah mudah untuk memisahkan pemrosesan bahasa secara rapi ke dalam kotak-kotak yang sesuai dengan masing-masing strata. Namun, pemisahan semacam itu berfungsi sebagai alat bantu pedagogis yang bermanfaat dan juga merupakan dasar bagi model arsitektur yang membuat tugas analisis bahasa alami lebih mudah dikelola dari sudut pandang rekayasa perangkat lunak. Dua masalah menyulitkan pemrosesan bahasa alami dan menyebabkan penggunaan teknik yang berbeda dengan teknik yang digunakan untuk membangun kompilar, dll., untuk memproses bahasa buatan. Masalah-masalah ini adalah tingkat ambiguitas yang terdapat dalam bahasa alami dan kompleksitas informasi semantik yang terkandung bahkan dalam kalimat sederhana.

Tugas-tugas NLP dasar meliputi tokenisasi dan penguraian, lemmatisasi/stemming, penandaan part-of-speech, deteksi bahasa, dan identifikasi hubungan semantik. Tugas-tugas NLP memecah bahasa menjadi bagian-bagian yang lebih pendek dan mendasar, mencoba memahami hubungan antar bagian, dan mengeksplorasi bagaimana bagian-bagian tersebut bekerja sama untuk menciptakan makna. Tugas-tugas mendasar ini sering digunakan dalam kapabilitas NLP tingkat tinggi, seperti:

- Kategorisasi konten: Ringkasan dokumen berbasis linguistik, termasuk pencarian dan pengindeksan, peringatan konten, dan deteksi duplikasi
- Penemuan dan pemodelan topik: Menangkap makna dan tema secara akurat dalam kumpulan teks dan menerapkan analitik tingkat lanjut pada teks, seperti optimasi dan peramalan
- Ekstraksi kontekstual: Secara otomatis menarik informasi terstruktur dari sumber berbasis teks
- Analisis sentimen: Mengidentifikasi suasana hati atau opini subjektif dalam sejumlah besar teks, termasuk sentimen rata-rata dan penggalan opini
- Konversi ucapan ke teks dan teks ke ucapan: Mengubah perintah suara menjadi teks tertulis dan sebaliknya
- Peringkasan dokumen: Secara otomatis menghasilkan sinopsis dari kumpulan teks yang besar
- Terjemahan mesin: Terjemahan otomatis teks atau ucapan dari satu bahasa ke bahasa lain

Dalam semua kasus ini, tujuan utamanya adalah mengambil masukan bahasa mentah dan menggunakan linguistik dan algoritma untuk mengubah atau memperkaya teks sedemikian rupa sehingga memberikan nilai yang lebih besar. Pandangan dasar NLP menyoroti empat tahap berbeda: pemrosesan morfologi, analisis sintaksis (penguraian), analisis semantik, dan analisis pragmatik.



Pemrosesan Morfologi

Tahap awal yang dilakukan sebelum analisis sintaksis adalah pemrosesan morfologis. Tujuan tahap pemrosesan bahasa ini adalah memecah rangkaian masukan bahasa menjadi kumpulan token yang sesuai dengan kata-kata diskrit, subkata, dan bentuk tanda baca. Misalnya, kata seperti "discontentedly" dapat dipecah menjadi tiga token subkata, yaitu: dis-contented -ly.

Morfologi terutama berkaitan dengan pengenalan bagaimana kata dasar telah dimodifikasi untuk membentuk kata lain dengan makna yang serupa tetapi seringkali dengan kategori sintaksis yang berbeda. Modifikasi biasanya terjadi dengan penambahan awalan dan/atau akhiran, tetapi perubahan tekstual lainnya juga dapat terjadi.

Sintaks dan Semantik

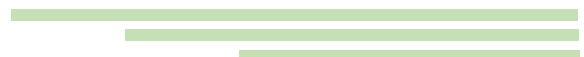
Pemroses bahasa harus menjalankan beberapa fungsi berbeda yang terutama didasarkan pada analisis sintaksis dan analisis semantik. Tujuan analisis sintaksis ada dua: untuk memeriksa apakah rangkaian kata (kalimat) terbentuk dengan baik dan memecahnya menjadi struktur yang menunjukkan hubungan sintaksis antara kata-kata yang berbeda. Penganalisis sintaksis (atau parser) melakukan hal ini menggunakan kamus definisi kata (leksikon) dan seperangkat aturan sintaksis (tata bahasa). Leksikon sederhana hanya memuat kategori sintaksis setiap kata; tata bahasa sederhana menjelaskan aturan yang hanya menunjukkan bagaimana kategori sintaksis dapat digabungkan untuk membentuk frasa dengan tipe yang berbeda.

Semantik dan Pragmatik

Setelah analisis semantik, tahap pemrosesan selanjutnya berkaitan dengan pragmatik. Sayangnya, tidak ada perbedaan yang disepakati secara universal antara semantik dan pragmatik. Analisis semantik mengaitkan makna dengan ujaran/kalimat yang terisolasi; analisis pragmatik menafsirkan hasil analisis semantik dari perspektif konteks tertentu. Ini berarti bahwa dengan kalimat seperti "Kucing besar mengejar tikus", analisis semantik dapat menghasilkan ekspresi yang berarti kucing besar tetapi tidak dapat melakukan langkah inferensi lebih lanjut yang diperlukan untuk mengidentifikasi kucing besar sebagai Felix. Hal ini akan diserahkan kepada analisis pragmatik.

Dalam beberapa kasus, seperti contoh yang baru saja dijelaskan, analisis pragmatik hanya mencocokkan objek/peristiwa aktual yang ada dalam konteks tertentu dengan referensi objek yang diperoleh selama analisis semantik. Dalam kasus lain, analisis pragmatik dapat menghilangkan ambiguitas kalimat yang tidak dapat sepenuhnya dihilangkan ambiguitasnya selama fase sintaksis dan analisis semantik.

Saat ini, NLP cenderung didasarkan pada pengubahan bahasa alami menjadi bahasa mesin. Namun seiring perkembangan teknologi, khususnya komponen AI, komputer akan semakin baik dalam "memahami" pertanyaan dan mulai memberikan jawaban, alih-alih hasil pencarian. Ini selangkah lebih maju dari mengajukan pertanyaan dalam bahasa alami. Komputer juga menerimanya dengan cara itu. Namun, setelah mempelajari hubungan semantik dan inferensi dari pertanyaan tersebut, komputer akan dapat secara otomatis



melakukan penyaringan dan proses bisnis yang diperlukan untuk memberikan jawaban yang dapat dipahami, alih-alih hanya menampilkan data kepada Anda.

5.2 CONTOH PENGGUNAAN NLP

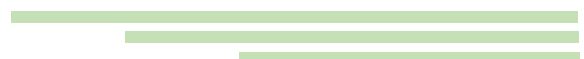
Chatbot telah beredar di pasaran selama beberapa tahun, tetapi versi yang lebih baru memiliki pemahaman bahasa yang lebih baik dan lebih interaktif. Beberapa bisnis menggunakan chatbot untuk menjawab pertanyaan rutin di meja bantuan. Beberapa lainnya menggunakan bot untuk membantu mengarahkan pertanyaan di meja bantuan. Di sini, berdasarkan siapa Anda dan apa yang Anda minta, Anda akan diarahkan ke petugas pusat panggilan yang tepat untuk menjawab pertanyaan spesifik Anda. Bisnis lain menggunakannya untuk belanja personalisasi yang melibatkan pemahaman tentang apa yang Anda dan orang-orang seperti Anda beli, selain apa yang Anda cari. Contoh penggunaan ini membutuhkan pencarian cerdas berbasis NLP serta pembelajaran mesin.

Kebanyakan orang familiar dengan aplikasi pencarian interaktif seperti Siri atau Alexa. Namun, lebih sering, jenis aplikasi lain menggunakan semacam antarmuka bahasa alami. Ini termasuk aplikasi BI di mana pengguna dapat mengajukan pertanyaan, baik melalui suara maupun teks, dalam bahasa alami dan mendapatkan jawaban dari aplikasi. Selain itu, pembelajaran mesin digunakan dalam manajemen data dan aplikasi BI untuk membantu segala hal, mulai dari integrasi data, persiapan data, hingga analisis analitik yang sebenarnya. Beberapa aplikasi memberi pengguna wawasan yang mungkin tidak terpikirkan oleh mereka sendiri. Beberapa sponsor laporan ini menyediakan fungsi semacam ini.

Dengan menggunakan pembelajaran mendalam, suatu sistem dapat dilatih untuk mengenali gambar dan suara. Sistem belajar dari contoh-contoh yang diberi label agar akurat dalam mengklasifikasikan gambar atau suara baru. Misalnya, komputer dapat dilatih untuk mengidentifikasi suara-suara tertentu yang menunjukkan bahwa sebuah motor sedang rusak. Aplikasi semacam ini digunakan dalam mobil dan penerbangan. Meskipun kita berpikir tentang penandaan gambar otomatis di jejaring sosial, teknologi semacam ini juga digunakan untuk mengklasifikasikan foto dalam bisnis untuk penjualan mobil daring atau untuk mengidentifikasi produk lainnya.

Analisis Teks

Analisis teks mempelajari nilai nominal kata-kata, termasuk tata bahasa dan hubungan antarkata. Sederhananya, analisis teks memberi Anda makna. Analisis sentimen memberikan wawasan tentang emosi di balik kata-kata. Analisis teks mengacu pada ekstraksi informasi bermanfaat dari sumber teks. Ini adalah istilah luas yang menggambarkan tugas-tugas mulai dari memberi anotasi pada sumber teks dengan meta-informasi seperti orang dan tempat yang disebutkan dalam teks hingga berbagai model tentang dokumen (misalnya analisis sentimen, pengelompokan dan kategorisasi teks). Untuk lebih jelasnya, istilah dokumen adalah gagasan abstrak yang dapat mewakili setiap bagian teks yang koheren dalam koleksi yang lebih besar seperti satu postingan blog dalam kumpulan postingan WordPress, artikel surat kabar, halaman di Wikipedia, dan sebagainya.



Selama aktivitas analisis teks, seorang peneliti dapat mengembangkan fitur-fitur yang menggambarkan beberapa aspek dokumen, misalnya:

- Dokumen tersebut tentang "kesehatan" atau "perjalanan".
- Panggilan telepon tersebut mengandung banyak bahasa negatif.
- Situs web tersebut menunjukkan produk tertentu.
- Tweet tersebut menggambarkan hubungan antara suatu produk dan masalah yang terkait dengan produk tersebut.
- Penulis postingan blog tersebut mungkin seorang fisioterapis.
- Email tersebut melanggar kepatuhan karena mengungkapkan informasi pribadi.

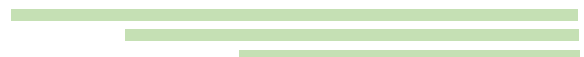
Aplikasi analisis teks yang umum di industri keuangan berfokus pada kepatuhan dan pencegahan penipuan. Tujuan pemrosesan bahasa alami dalam kasus penggunaan ini adalah untuk memahami isi utas komunikasi melalui interpretasi semantik dan untuk mengidentifikasi hubungan dan entitas di seluruh utas. Namun, analisis teks bertanggung jawab untuk menentukan apakah suatu pesan, atau serangkaian pesan, melanggar kepatuhan. Departemen kepatuhan mendapatkan manfaat dari penggabungan data terstruktur, seperti perdagangan dan transaksi, dengan informasi yang diekstrak dari email dan pesan instan. Dengan kedua jenis aset data tersebut, dimungkinkan untuk menyimpulkan maksud di balik suatu transaksi.

Organisasi keuangan menghadapi masalah kepatuhan mendasar lainnya, anti-pencucian uang. Lembaga keuangan diwajibkan untuk menyaring semua transaksi di seluruh unit bisnis mereka untuk mencegah transaksi antara pihak-pihak yang masuk daftar hitam. Tugas ini melibatkan analisis teks bebas yang terdapat dalam transaksi dan pencocokan nama serta entitas dengan daftar pantauan dari Kantor Pengawasan Aset Luar Negeri (OFAC) dan lembaga pemerintah lainnya. Salah satu tugas penting adalah mencocokkan nama yang ditransliterasikan dengan "satu" representasi dalam sebuah daftar. Misalnya, nama Rajendra dapat ditransliterasikan menjadi Rajinder, Raj, atau Raja. Pencocokan dengan beberapa daftar harus sangat presisi karena analisis hanya dapat meninjau sebagian kecil peringatan secara manual.

Dalam bisnis asuransi, perusahaan memiliki koleksi besar data teks pusat panggilan, klaim, penagihan, dan nota penyesuaian yang tidak terstruktur. Untuk mendapatkan pemahaman yang lebih baik tentang pemegang polis, perusahaan-perusahaan ini dapat memanfaatkan analisis sentimen untuk mengukur apakah pelanggan mereka puas atau tidak puas dengan produk, layanan, dan proses mereka. Analisis teks dapat mengidentifikasi area masalah pada produk dan prosedur, dan dapat memberikan panduan untuk meningkatkan layanan atau mengembangkan produk baru.

Analisis Sentimen

Analisis sentimen adalah penggunaan pemrosesan bahasa alami, statistik, dan analisis teks untuk mengekstrak dan mengidentifikasi sentimen teks ke dalam kategori positif, negatif, atau netral. Kita sering melihat analisis sentimen digunakan untuk sampai pada keputusan biner: seseorang mendukung atau menentang sesuatu, pengguna menyukai atau tidak menyukai sesuatu, atau produk tersebut baik atau buruk. Ada manfaat strategis dalam mengetahui sentimen konsumen terkait pesaing Anda. Analisis sentimen dapat membantu



memprediksi tren pelanggan, sehingga memantau opini publik tentang bisnis lain di industri Anda menyediakan kelompok kontrol untuk membandingkan skor Anda.

Analisis sentimen juga disebut penambangan opini karena mencakup identifikasi sikap, emosi, dan opini konsumen terhadap produk, merek, atau layanan perusahaan. Analisis sentimen memberikan wawasan tentang setiap perubahan opini publik terkait merek Anda yang akan mendukung atau menentang arah bisnis Anda. Skor sentimen yang tinggi atau rendah membantu Anda mengidentifikasi cara untuk merestrukturisasi tim atau mengembangkan strategi kreatif baru.

Kasus Penggunaan Analisis Sentimen

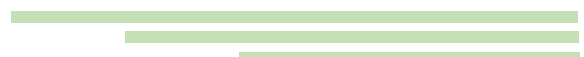
Analisis sentimen sering diterapkan pada ulasan dan media sosial untuk membantu tim pemasaran dan layanan pelanggan mengidentifikasi perasaan konsumen. Dalam media, seperti ulasan produk, analisis sentimen dapat digunakan untuk mengungkapkan apakah konsumen puas atau tidak puas dengan suatu produk. Demikian pula, perusahaan dapat menggunakan analisis sentimen untuk mengukur dampak produk baru, kampanye iklan, atau respons konsumen terhadap berita perusahaan terbaru di media sosial.

Seorang agen layanan pelanggan di sebuah perusahaan dapat menggunakan analisis sentimen untuk secara otomatis mengurutkan email pengguna yang masuk ke dalam kategori "mendesak" atau "tidak mendesak" berdasarkan sentimen email tersebut, secara proaktif mengidentifikasi pengguna yang frustrasi. Agen tersebut kemudian dapat mengalokasikan waktu mereka untuk menyelesaikan kebutuhan pengguna yang paling mendesak terlebih dahulu.

Analisis sentimen sering digunakan dalam intelijen bisnis untuk memahami alasan subjektif mengapa konsumen merespons atau tidak merespons sesuatu (misalnya, Mengapa konsumen membeli suatu produk? Apa pendapat mereka tentang pengalaman pengguna? Apakah dukungan layanan pelanggan memenuhi harapan mereka?). Analisis sentimen juga dapat digunakan dalam bidang ilmu politik, sosiologi, dan psikologi untuk menganalisis tren, bias ideologis, opini, mengukur reaksi, dll.

Tantangan Analisis Sentimen

Orang-orang mengungkapkan pendapat dengan cara yang kompleks, sehingga sulit untuk memahami subjek opini manusia. Ada beberapa elemen yang terdefinisi dalam sebuah teks yang menjadi faktor dalam analisis sentimen: objek, atribut, pemegang opini, orientasi opini, dan kekuatan opini.



Object: The product, service, individual, business, event or topic being analysed
Example: iPhone
Attributes: The specific components and properties of the object
Component examples: Battery, touch screen, headphone jack
Property examples: Size, weight, processing speed
Opinion holder: The person or business who is expressing the sentiment
Example: The person who purchased the iPhone
Opinion orientation (polarity): The general position of the opinion
Examples: Positive, negative or neutral
Opinion strength: The level, scale or intensity of the opinion
Examples: Thrilled > jubilant > happy > satisfied

Untuk mendapatkan informasi yang lengkap, akurat, dan dapat ditindaklanjuti dari sebuah teks, penting untuk tidak hanya mengidentifikasi kelima elemen ini secara individual, tetapi juga memahami bagaimana elemen-elemen tersebut bekerja bersama untuk memberikan konteks dan sentimen yang utuh. Karena pemrosesan kata kunci hanya mengidentifikasi sentimen yang tercermin dalam kata tertentu, pemrosesan ini gagal menyediakan semua elemen yang diperlukan untuk memahami konteks lengkap dari keseluruhan teks.

5.3 APLIKASI NLP DALAM BISNIS

Kami mengeksplorasi beberapa contoh bagaimana aplikasi bisnis NLP dapat diterapkan dalam skala besar untuk menjawab pertanyaan bisnis.

Layanan Pelanggan

NLP digunakan oleh komputer untuk memanipulasi bahasa manusia, baik untuk mengekstrak makna, menghasilkan teks, atau untuk tujuan lainnya. Interaksi antara bahasa komputer dikategorikan berdasarkan tugas yang perlu diselesaikan: meringkas dokumen yang panjang, menerjemahkan antara dua bahasa manusia, atau mendeteksi email spam adalah contoh tugas yang saat ini dapat diselesaikan dengan baik oleh mesin.

Sebagian besar NLP saat ini didasarkan pada pembelajaran mesin, yaitu metode statistik yang mampu mensimulasikan apa yang akan dilakukan manusia dalam situasi serupa. NLP banyak digunakan dalam layanan pelanggan. Tugas NLP yang relevan dalam layanan pelanggan meliputi:

- Pengenalan ucapan, yang mengubah bahasa lisan menjadi teks
- Menjawab pertanyaan, yang melibatkan hal tersebut. Menjawab pertanyaan yang diajukan oleh manusia dalam bahasa alami

Pemantauan Reputasi

Seiring dengan penurunan biaya komputasi dan peningkatan algoritma, bisnis mulai mengadopsi perangkat yang memungkinkan mereka untuk melihat lebih jauh dari sekadar basis data. Data jenis ini umumnya disebut sebagai data eksternal, data publik, atau intelijen sumber terbuka (OSINT). Meskipun sebagian data ini terstruktur dan siap dianalisis (misalnya data sensus, harga saham), sebagian besar lainnya masih tersimpan dalam teks buatan

manusia yang tidak terstruktur seperti berita, postingan blog, forum, paten, lowongan pekerjaan, laporan, media sosial, dan situs web perusahaan. Sumber-sumber ini berisi banyak informasi berharga tentang bagaimana pesaing, pelanggan, dan pasar secara keseluruhan berkembang.

Saat ini, seiring konsumen mulai menyuarkan keluhan mereka di Twitter dan Facebook, pemantauan dan pengelolaan reputasi telah menjadi prioritas utama bagi bisnis. Perusahaan kini dapat memindai seluruh web untuk mencari penyebutan merek dan produk mereka serta mengenali kapan mereka harus bertindak. Tugas NLP yang relevan untuk aplikasi ini meliputi:

- Analisis sentimen, yang menentukan sikap, keadaan emosi, penilaian, atau niat penulis. Hal ini dilakukan dengan menetapkan polaritas pada teks (positif, netral, atau negatif) atau mencoba mengenali suasana hati yang mendasarinya (senang, sedih, tenang, marah, dll.).
- Resolusi koreferensi, yang menghubungkan kata ganti dengan objek yang tepat. Ini merupakan tugas yang menantang, tetapi sangat penting untuk menafsirkan teks dengan benar. Misalnya, jika seorang pelanggan menulis: "Seorang pria dari penjual menelepon untuk menanyakan apakah saya menyukai apartemen baru saya. Baiklah, tidak, bung, apartemennya jelek?", penting untuk menyadari bahwa "itu" mengacu pada apartemen dan bukan orangnya. Singkatnya, pelanggan mengeluh tentang produk, bukan layanannya.

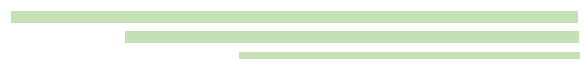
Penempatan Iklan Pembelian media biasanya merupakan lini utama dalam anggaran periklanan perusahaan, sehingga penargetan apa pun yang dapat dilakukan untuk memastikan bahwa iklan disajikan kepada orang yang tepat sangatlah penting. Email, media sosial, e-commerce, dan perilaku penjelajahan kita mengandung banyak informasi tentang apa yang kita minati. Tugas NLP yang relevan untuk aplikasi ini meliputi:

- Pencocokan kata kunci, yang memeriksa apakah kata-kata yang menarik disertakan dalam suatu teks. Meskipun perkiraan awal ini seringkali cukup baik, kurangnya kecanggihannya dapat menghasilkan hasil yang kurang tepat.
- Disambiguasi makna atau identifikasi makna kata yang digunakan dalam sebuah kalimat. Ini adalah salah satu masalah utama yang masih terbuka dalam NLP.

Kecerdasan Pasar

Mengetahui status suatu industri sangat penting untuk mengembangkan strategi yang efektif, tetapi saluran distribusi konten saat ini (umpan RSS, media sosial, email) menghasilkan begitu banyak informasi sehingga sulit untuk mengikutinya. Tugas NLP yang relevan untuk aplikasi ini meliputi:

- ❖ Ekstraksi peristiwa: yang mengenali apa yang terjadi pada suatu entitas. Basis data terstruktur berisi peristiwa tentang perusahaan, pemerintah, dan individu merupakan alat yang sangat ampuh untuk menganalisis ekosistem bisnis.
- ❖ Klasifikasi kalimat: sering digunakan sebagai langkah awal untuk mengekstrak konten yang relevan dari repositori besar.



- ❖ Kepatuhan regulasi: Contoh penting kepatuhan adalah studi yang dilakukan setelah suatu obat dipasarkan untuk mengumpulkan informasi tentang efek sampingnya. Tugas NLP yang relevan untuk aplikasi ini meliputi:
- ❖ Pengenalan entitas bernama: mengekstrak nama obat, penyakit, pasien, dan perusahaan farmasi menggunakan metode berbasis aturan atau statistik.
- ❖ Deteksi relasi: digunakan untuk mengidentifikasi konteks di mana kejadian efek samping obat disebutkan. Hal ini sering dilakukan dengan kerangka atau pola kata yang sesuai dengan suatu konsep.

Kedua tugas ini mendapatkan manfaat dari penggunaan ontologi, yaitu pengetahuan domain terstruktur yang menyediakan kamus objek dan hubungan antar objek.

Natural Language Processing (NLP) Di Dunia Bisnis

Natural Language Processing (NLP) merupakan teknologi kecerdasan buatan yang memungkinkan komputer untuk memahami, mengolah, dan menganalisis bahasa manusia secara alami. Dalam konteks bisnis, NLP berperan penting dalam mengubah data teks yang tidak terstruktur—seperti ulasan pelanggan, email, dan percakapan chat—menjadi wawasan yang dapat mendukung pengambilan keputusan yang lebih tepat dan strategis.

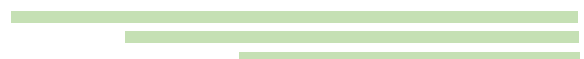
Berdasarkan berbagai penelitian terbaru, NLP digunakan secara luas untuk melakukan analisis sentimen pelanggan dengan memanfaatkan model machine learning seperti Naïve Bayes dan Support Vector Machine. Melalui teknologi ini, opini pelanggan dapat diklasifikasikan menjadi sentimen positif, negatif, maupun netral dengan tingkat akurasi yang tinggi. Hal ini memungkinkan perusahaan untuk memahami kebutuhan dan persepsi pelanggan secara cepat tanpa harus membaca setiap feedback secara manual.

Selain itu, NLP juga mendukung pengembangan chatbot cerdas yang mampu memberikan respons cepat dan relevan berdasarkan konteks percakapan, sehingga meningkatkan kualitas layanan pelanggan sekaligus efisiensi operasional. Proses NLP sendiri dimulai dari pengumpulan data teks, pembersihan dan persiapan data, ekstraksi fitur penting, hingga analisis menggunakan model yang sesuai dengan kebutuhan bisnis.

Manfaat yang diperoleh bisnis dari penerapan NLP sangat beragam, mulai dari pengoptimalan layanan pelanggan, peningkatan efektivitas pemasaran melalui analisis tren dan sentimen, hingga percepatan proses rekrutmen dengan memanfaatkan analisis otomatis terhadap dokumen seperti CV. Dengan demikian, NLP telah menjadi elemen kunci dalam transformasi digital bisnis yang berfokus pada pemanfaatan data dan pengalaman pelanggan, memberikan keunggulan kompetitif yang signifikan di era modern.

Penggunaan AI di Berbagai Sektor Manufaktur

Kecerdasan buatan (AI) dalam bidang manufaktur memanfaatkan algoritma pembelajaran mesin (machine learning) dan pembelajaran mendalam (deep learning) untuk menganalisis berbagai kumpulan data besar guna menemukan pola tertentu. Berdasarkan hasil analisis tersebut, AI mampu mengambil tindakan seperti menyelesaikan tugas, mengotomatiskan proses, atau memberikan wawasan yang berguna bagi produsen untuk meningkatkan kinerja bisnis mereka.



Meskipun AI memiliki beragam penerapan, beberapa penggunaan utamanya dalam manufaktur meliputi otomatisasi pabrik, seperti penjadwalan dan pengelolaan sumber daya, pengoperasian dan manajemen cerdas, pemantauan kualitas dan proses produksi, optimalisasi rantai pasokan, serta pengambilan keputusan yang berbasis data.

5.4 TEKNOLOGI SENTIMEN DALAM BISNIS

Untuk mengatasi variasi, teknologi sentimen paling berguna ketika dapat diadaptasi untuk mengkaji isu-isu kontemporer, memungkinkan Anda menjawab pertanyaan yang dibingkai dalam beberapa cara yang beragam atau mengajukan pertanyaan yang sama sekali berbeda dari data yang sama.

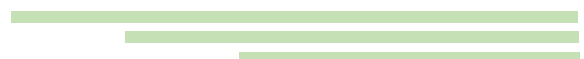
Ketika Anda dapat mengkaji pertanyaan dari berbagai perspektif, kemungkinan besar Anda melayani berbagai kebutuhan, dan sering kali Anda ingin berbagi hasil analitik dalam berbagai cara, misalnya dengan mengintegrasikannya dengan proses peringatan, alur kerja, dan laporan pemantauan Anda saat ini. Aplikasi analitik memiliki fleksibilitas untuk memeriksa semua informasi yang tersedia dengan berbagai cara, menggunakan berbagai teknik untuk memecahkan masalah, dan metode alternatif untuk memberikan wawasan yang diperoleh.

Untuk masalah analitik seperti sentimen, Anda memerlukan teknologi yang tidak membatasi Anda untuk memeriksa pertanyaan hanya dengan satu atau dua cara, dengan satu definisi atau hanya dengan beberapa data. Dengan pendekatan analitik terintegrasi, Anda memiliki kekuatan untuk mendapatkan jawaban dari data, jawaban atas pertanyaan yang sudah Anda miliki, serta pertanyaan yang belum terpikirkan untuk Anda ajukan. Jika Anda menggunakan metode yang sangat terspesialisasi misalnya, berlangganan laporan media sosial yang memantau tren untuk Anda. Anda tidak mendapatkan diferensiasi kompetitif apa pun. Ketika teknologi terintegrasi sehingga Anda dapat menggunakan berbagai metode dan pendekatan, lingkungan analitik Anda mendukung inovasi.

Memiliki beragam kemampuan analisis di ujung jari Anda adalah cara terbaik untuk menyelesaikan pertanyaan sentimen dan, dalam hal ini, pertanyaan analitik lainnya. Misalnya, dengan memasukkan kategori yang telah ditentukan sebelumnya yang mengidentifikasi segmen tertentu, Anda dapat menjelaskan di mana perhatian difokuskan, berdasarkan aspek bisnis yang dapat Anda pengaruhi. Menilai bagaimana kata-kata digunakan untuk menggambarkan konsep-konsep ini dapat memberikan wawasan tentang tingkat perhatian. Ekstraksi semacam ini memberikan detail tentang fitur-fitur spesifik, menawarkan analisis yang jauh lebih detail. Dengan wawasan yang lebih detail, Anda dapat lebih spesifik dalam menentukan prioritas dan tindakan terkait.

Misalnya, mempertimbangkan tenses suatu kata kerja dapat membantu mengidentifikasi prioritas atau niat seseorang. Penggunaan kata-kata emosional, termasuk frekuensinya, dapat mengidentifikasi tingkat kepedulian seseorang.

Beberapa orang menggambarkan hal ini sebagai penggunaan aspek seni dan sains atau manusia dan mesin dalam analitik. Hal yang menarik adalah bahwa dengan sistem terintegrasi, Anda memiliki keduanya. Anda dapat menggunakan entitas kustom Anda (yang ditentukan manusia) dalam penambangan teks Anda (dipelajari mesin). Jika sistem terintegrasi



sepenuhnya, Anda dapat menggabungkan analisis terstruktur dan tidak terstruktur hanya dengan menyeret dan melepaskan algoritma model prediktif pada topik yang ditemukan untuk melihat apakah topik tersebut memengaruhi kemungkinan churn, misalnya.

Integrasi memberi Anda pilihan dalam cara mengembangkan analitik dan cara Anda menerapkan hasilnya. Fleksibilitas semacam ini sangat penting bagi operasional Anda karena Anda menjadi lebih responsif terhadap pasar Anda, audiens Anda yang beragam, dan topik investigasi Anda yang beragam. Untuk merasakan berbagai manfaat dari analisis media sosial, Anda memerlukan teknologi analitik yang terbuka, fleksibel, dan terintegrasi yang dapat memberi tahu Anda bagaimana sentimen penting dan dalam konteks apa. Organisasi yang memanfaatkan sepenuhnya manfaat analisis sentimen meningkatkan hasil dengan manfaat kinerja yang jauh melampaui manajemen reputasi hingga penghematan biaya dan peningkatan penjualan.

Penerapan AI dalam analisis sentimen menunjukkan bahwa penggunaan teknologi Natural Language Processing (NLP) dan machine learning memungkinkan perusahaan untuk secara otomatis memantau dan menganalisis sentimen pelanggan yang terdapat dalam ulasan produk, media sosial, atau survei pelanggan. Dengan pemahaman sentimen yang lebih baik, perusahaan dapat menyusun strategi pemasaran yang lebih efektif, meningkatkan kualitas produk dan layanan, serta membangun hubungan yang lebih kuat dengan pelanggan.

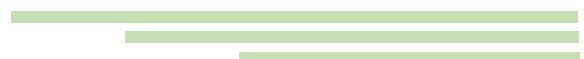
Analisis sentimen ini membantu bisnis dalam mengidentifikasi pola perilaku dan preferensi konsumen, sehingga memungkinkan prediksi kebutuhan pelanggan di masa mendatang. Dalam konteks pemasaran, AI yang digunakan untuk analisis sentimen mampu menyajikan wawasan yang meningkatkan pengalaman pelanggan (customer experience) dengan cara mengoptimalkan personalisasi pesan pemasaran dan pengembangan produk.

Perusahaan seperti Kata.ai di Indonesia telah berhasil mengaplikasikan solusi berbasis AI untuk analisis sentimen dan pengambilan keputusan bisnis, yang membantu lebih dari 140 perusahaan meningkatkan efisiensi operasional dan aktivitas pemasaran mereka melalui pengolahan sentimen yang akurat dari berbagai platform digital.

Meskipun demikian, penelitian juga menunjukkan adanya kesenjangan dalam pemahaman dan penerapan analisis sentimen AI dari skala global hingga lokal. Hal ini menjadi tantangan sekaligus peluang untuk pengembangan lebih lanjut agar AI dapat lebih optimal dalam mendukung strategi bisnis secara komprehensif.

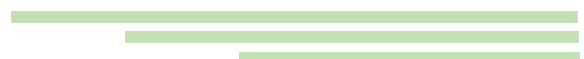
Secara keseluruhan, analisis sentimen berbasis AI memberikan kontribusi signifikan dalam meningkatkan kualitas pengambilan keputusan bisnis, memperbaiki strategi pemasaran, dan meningkatkan kepuasan serta loyalitas pelanggan dengan memanfaatkan data teks yang sangat besar dan beragam secara cepat dan akurat.

Penggunaan kecerdasan buatan (AI) dalam analisis sentimen bisnis membawa berbagai dampak positif yang signifikan. Dengan kemampuan AI dalam mengolah dan menganalisis data dari beragam sumber seperti media sosial, ulasan pelanggan, dan survei, perusahaan bisa mendapatkan pemahaman yang lebih mendalam mengenai persepsi dan opini konsumen. Hal ini memungkinkan bisnis untuk merespons kebutuhan dan harapan pelanggan dengan lebih cepat dan akurat, sehingga meningkatkan kualitas pengalaman pelanggan sekaligus



memperkuat loyalitas mereka. Selain itu, AI memungkinkan pemantauan reputasi perusahaan secara real-time, sehingga perusahaan dapat segera menindaklanjuti masukan negatif dan menjaga citra positif di mata publik. Otomatisasi dalam analisis sentimen juga meningkatkan efisiensi operasional dengan mengurangi ketergantungan pada analisis manual, sekaligus mendukung pengambilan keputusan strategis yang lebih tepat berbasis data. AI juga membantu bisnis mengidentifikasi tren dan peluang pasar dengan lebih cepat, sehingga mendorong inovasi dan kemampuan beradaptasi terhadap perubahan kebutuhan konsumen. Namun demikian, penerapan AI dalam analisis sentimen tidak lepas dari tantangan dan dampak negatif yang perlu diperhatikan. Salah satu masalah utama adalah potensi bias algoritma yang muncul jika data pelatihan tidak lengkap atau kurang mewakili kondisi sebenarnya, sehingga analisis yang dihasilkan kurang objektif. Selain itu, AI masih memiliki keterbatasan dalam memahami nuansa bahasa manusia, termasuk ironi, sindiran, atau konteks budaya tertentu, yang berisiko menimbulkan kesalahan dalam interpretasi sentimen. Kualitas dan kelengkapan data menjadi faktor penting karena data yang kurang memadai dapat menghasilkan analisis yang menyesatkan. Aspek privasi dan etika juga menjadi perhatian utama, mengingat penggunaan data pelanggan harus sesuai dengan regulasi agar terhindar dari penyalahgunaan dan pelanggaran privasi. Di sisi lain, otomatisasi ini berpotensi menggeser peran tenaga kerja manusia, menimbulkan kekhawatiran terkait pengurangan lapangan pekerjaan. Selain itu, ketergantungan berlebihan pada AI tanpa pengawasan manusia yang memadai juga dapat menyebabkan kesalahan dalam pengambilan keputusan bisnis yang berdampak negatif.

Dengan demikian, meskipun AI dalam analisis sentimen memberikan peluang besar untuk meningkatkan efektivitas bisnis, keberhasilannya sangat bergantung pada pengelolaan risiko secara cermat serta keseimbangan antara pemanfaatan teknologi yang canggih dengan peran manusia secara bijak dan etis, agar manfaatnya dapat dirasakan secara maksimal dan berkelanjutan.



BAB 6

MEMANFAATKAN AI DALAM BISNIS

6.1 LANSKAP ANALISIS

Saat ini, sejumlah besar data tambahan akan tersedia, termasuk data terstruktur, seperti dari sensor, dan data tidak terstruktur, seperti dari kamera, media sosial, dan sentimen dari jejaring sosial. Faktor utama di balik pertumbuhan AI berbasis data adalah ketersediaan data dalam jumlah besar. Ini mencakup data yang dihasilkan mesin dari sumber-sumber baru seperti sensor atau yang disebut Internet of Things serta data yang dihasilkan manusia, baik dari dalam maupun luar perusahaan.

Analisis bisnis digunakan dalam industri teknologi informasi untuk merujuk pada penggunaan komputasi guna mendapatkan wawasan dari data. Data tersebut dapat diperoleh dari sumber internal perusahaan, seperti aplikasi perencanaan sumber daya perusahaan (ERP), gudang data/mart, dari penyedia data pihak ketiga, atau dari sumber publik. Perusahaan berupaya memanfaatkan data digital dari sistem transaksi dan proses bisnis otomatis untuk mendukung pengambilan keputusan "berbasis fakta".

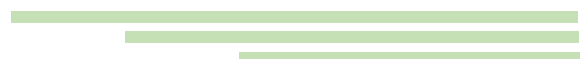
Dengan demikian, analitik bisnis merupakan kategori komputasi, bukan metode, aplikasi, atau produk tertentu. Analitik merupakan bagian dari evolusi yang dapat menghasilkan sistem AI yang sukses. Contohnya, model pembelajaran mesin dilatih pada set data yang sangat besar. Dalam organisasi yang sadar analitik, yang menangani penemuan data, big data, dan tugas-tugas seperti pengolahan data, persiapan, dan integrasi data, AI merupakan perkembangan alami.

Sistem analitik yang matang akan mendukung kesuksesan AI. Area penting lainnya yang perlu dipertimbangkan adalah bahwa sistem AI akan matang seiring waktu seiring dengan semakin banyaknya data yang diberikan dan data yang tepat dan berkualitas. Oleh karena itu, bisnis berinvestasi dalam penyimpanan data dan gudang data; ini merupakan bagian dari proses penyelarasan aset untuk implementasi AI.

Area Aplikasi

Banyak fungsi dalam bisnis yang dapat memperoleh manfaat dari analitik. Kategori fungsional yang paling umum meliputi:

- (a) Analisis Pelanggan: Kategori ini mencakup aplikasi untuk pemasaran (profil pelanggan, segmentasi, analisis jejaring sosial, analisis reputasi merek, optimasi bauran pemasaran) dan pengalaman pelanggan. Teknologi SNAzzy (Analisis Jejaring Sosial di Telekomunikasi) dan VOCA (Analisis Suara Pelanggan) termasuk dalam kategori ini.
- (b) Analisis Rantai Pasokan: Ini mencakup peramalan permintaan, dan optimasi inventaris, harga, penjadwalan, transportasi, dan penyimpanan, sekaligus memitigasi risiko. Optimalisasi rantai pasokan internal IBM telah menghasilkan hasil yang mengesankan. Subbidang yang dikenal sebagai Analisis Modal Manusia atau Analisis Tenaga Kerja berlaku untuk industri jasa di mana sumber daya manusia merupakan alat produksi utama.



- (c) Analisis Penipuan dan Risiko: Ini mencakup penilaian beberapa jenis risiko (pasar, operasional, kredit) khususnya di sektor Keuangan.
- (d) Analisis dalam ranah publik: Didorong oleh keterbatasan sumber daya alam, pemerintah menggunakan analisis untuk tugas-tugas seperti mendeteksi kebocoran air dalam sistem distribusi, membuat jaringan energi dan sistem lalu lintas lebih cerdas, dan meningkatkan keselamatan publik.

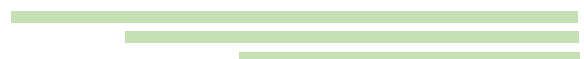
Kompleksitas Analisis

Kompleksitas analisis dapat dipecah menjadi tiga lapisan: Analisis deskriptif dapat diimplementasikan menggunakan spreadsheet atau kekuatan industri. Analisis prediktif membahas tentang apa yang akan terjadi selanjutnya, dan analisis preskriptif membahas tentang cara mencapai hasil terbaik. Analisis bisnis berfokus pada lima area utama kebutuhan pelanggan:

- Akses informasi: Segmen pertama ini merupakan fondasi bagi analisis bisnis. Intinya adalah mendorong pengambilan keputusan yang terinformasi/kolaboratif di seluruh bisnis, memastikan bahwa para pengambil keputusan dapat memahami kinerja area bisnis mereka sehingga mereka dapat membuat keputusan yang tepat.
- Wawasan: Memperoleh pemahaman yang lebih mendalam tentang mengapa suatu hal terjadi, misalnya, mendapatkan gambaran lengkap tentang pelanggan Anda (riwayat transaksi, segmentasi, sentimen dan opini, dll.) untuk membuat keputusan yang lebih baik dan memungkinkan pertumbuhan yang menguntungkan.
- Foresight: Memanfaatkan masa lalu untuk memprediksi potensi hasil di masa depan sehingga tindakan dan keputusan diperhitungkan untuk memenuhi tujuan dan persyaratan bisnis.
- Kelincahan bisnis: Mendorong optimalisasi keputusan secara real-time, baik dalam proses yang berpusat pada manusia maupun yang berpusat pada proses/otomatisasi.
- Penyelarasan strategis: Segmen pasar ini berfokus pada penyelarasan strategis semua pihak dalam bisnis, mulai dari strategi hingga eksekusi. Ini tentang memungkinkan visibilitas perusahaan dan operasional. Ini tentang mendokumentasikan preferensi, prioritas, tujuan, dan persyaratan yang mendorong pengambilan keputusan.

Organisasi yang melakukan perjalanan ke dalam penerapan analitik bisnis harus memulai dengan agenda manajemen informasi yang memperlakukan data dan informasi sebagai aset strategis. Setelah informasi diperlakukan sebagai aset, analitik deskriptif, prediktif, dan preskriptif dapat diterapkan. Secara klasik, bisnis atau organisasi memulai perjalanan ini dengan memeriksa data yang dihasilkan dari sistem otomasinya: perencanaan sumber daya perusahaan, manajemen hubungan pelanggan, waktu dan kehadiran, e-commerce, manajemen garansi, dan sejenisnya.

Bisnis mungkin juga memiliki data tidak terstruktur, seperti kontrak, keluhan pelanggan, surel internal, dan, yang semakin meningkat, data gambar dari sistem pemantauan fasilitas, serta data tidak terstruktur dari berbagai sumber web, seperti Facebook, Twitter, dan blog. Gabungan data terstruktur dan tidak terstruktur menghasilkan pandangan informasi 360



derajat untuk meningkatkan pengambilan keputusan. Banyak teknik analitik bisnis yang digunakan untuk data terstruktur juga dapat diterapkan pada data tidak terstruktur.

Analisis Deskriptif

Sebagian besar bisnis memulai dengan analisis deskriptif, penggunaan data untuk mengetahui apa yang terjadi di masa lalu. Analisis deskriptif mempersiapkan dan menganalisis data historis serta mengidentifikasi pola dari sampel untuk pelaporan tren. Teknik seperti pemodelan data, visualisasi, dan analisis regresi sebagian besar berada di bidang ini. Ini adalah serangkaian teknologi dan proses yang menggunakan data untuk memahami dan menganalisis kinerja bisnis.

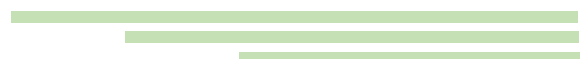
Analisis deskriptif dapat diklasifikasikan menjadi tiga area yang menjawab beberapa jenis pertanyaan:

- Pelaporan dan dasbor standar: Apa yang terjadi? Bagaimana perbandingannya dengan rencana kita? Apa yang terjadi sekarang?
- Pelaporan ad hoc: Berapa banyak? Seberapa sering? Di mana?
- Analisis/kueri/drill-down: Apa sebenarnya masalahnya? Mengapa itu terjadi?

Analisis deskriptif adalah jenis analisis yang paling umum digunakan dan paling dipahami. Analisis deskriptif mengkategorikan, mengkarakterisasi, mengkonsolidasikan, dan mengklasifikasikan data. Analisis deskriptif mencakup dasbor, laporan (misalnya anggaran, penjualan, pendapatan, dan biaya), dan beragam jenis kueri. Alat untuk analisis deskriptif dapat menyediakan mekanisme untuk berinteraksi dengan sumber data perusahaan. Alat-alat ini biasanya mencakup pembuatan laporan, kemampuan distribusi, dan fasilitas visualisasi data. Teknik analisis deskriptif paling umum diterapkan pada data terstruktur, meskipun telah banyak upaya untuk memperluas jangkauannya ke data tidak terstruktur, seringkali melalui pembuatan metadata dan indeks terstruktur.

Analisis deskriptif membantu memberikan pemahaman tentang masa lalu serta peristiwa yang terjadi secara real-time. Banyak aplikasi analisis deskriptif diimplementasikan melalui solusi perangkat lunak intelijen bisnis atau alat spreadsheet yang siap pakai; namun, kesulitan kontrol versi dapat timbul akibat maraknya penggunaan spreadsheet. Keuntungan dari platform perangkat lunak analisis deskriptif (intelijen bisnis dan perangkat lunak manajemen informasi/data) adalah konektivitas yang disediakan ke sistem manajemen informasi tepercaya yang mendasarinya, serta kemampuan untuk bekerja dengan data di berbagai dimensi untuk mendapatkan wawasan. Wawasan tentang apa yang terjadi sekarang atau telah terjadi di masa lalu dapat berguna dalam pengambilan keputusan tentang masa depan, tetapi analisis deskriptif bergantung pada peninjauan data oleh manusia dan tidak memiliki teknik yang kuat yang memfasilitasi pemahaman tentang apa yang mungkin terjadi di masa depan, juga tidak menyediakan alat untuk menyarankan keputusan tentang apa yang harus dilakukan selanjutnya.

Analisis deskriptif memang memberikan wawasan yang signifikan tentang kinerja bisnis dan memungkinkan pengguna untuk memantau dan mengelola proses bisnis mereka dengan lebih baik. Selain itu, analisis deskriptif sering kali berfungsi sebagai langkah pertama dalam keberhasilan penerapan analisis prediktif atau preskriptif. Organisasi yang secara efektif



menggunakan analitik deskriptif biasanya memiliki satu pandangan tentang masa lalu dan dapat memfokuskan perhatian mereka pada masa kini, alih-alih menyelaraskan berbagai pandangan tentang masa lalu.

Analisis Prediktif

Analisis prediktif menggunakan data untuk mengetahui apa yang mungkin terjadi di masa depan. Tentu saja, ini merupakan penggunaan analitik yang lebih halus dan tingkat tinggi. Analisis prediktif memprediksi probabilitas dan tren masa depan dan menemukan hubungan dalam data yang tidak jelas dengan analisis gaya lama. Teknik seperti penambangan data dan pemodelan prediktif berada di bidang ini. Secara khusus, analisis prediktif, suatu kategori analisis data terstruktur, menggunakan data dan teknik matematika untuk mengungkap model eksplanatori dan prediktif kinerja bisnis yang mewakili hubungan bawaan antara masukan dan keluaran data.

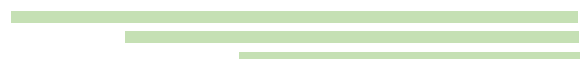
Analisis prediktif menggunakan pemahaman masa lalu untuk membuat "prediksi" tentang masa depan. Analisis prediktif diterapkan baik secara waktu nyata (real-time) untuk memengaruhi proses operasional (tindakan retensi waktu nyata melalui pesan obrolan atau identifikasi transaksi mencurigakan secara waktu nyata) maupun secara batch (menargetkan pelanggan baru di situs web atau surat langsung untuk mendorong penjualan silang/peningkatan penjualan, memprediksi churn, dll.). Prediksi ini dibuat dengan memeriksa data masa lalu, mendeteksi pola atau hubungan dalam data ini, dan kemudian mengekstrapolasi hubungan ini ke masa depan. Misalnya, klaim asuransi yang termasuk dalam kategori (pola) yang terbukti mengkhawatirkan di masa lalu mungkin ditandai untuk penyelidikan lebih lanjut.

Analisis prediktif dapat diklasifikasikan menjadi enam kategori:

- Penambangan data: Data apa yang berkorelasi dengan data lain?
- Pengenalan pola dan peringatan: Kapan saya harus mengambil tindakan untuk memperbaiki atau menyesuaikan suatu proses atau peralatan?
- Simulasi Monte Carlo: Apa yang mungkin terjadi?
- Peramalan: Bagaimana jika tren ini berlanjut?
- Analisis akar penyebab: Mengapa sesuatu terjadi?
- Pemodelan prediktif: Apa yang akan terjadi selanjutnya?

Analisis deskriptif mungkin dimulai dengan memberikan pandangan statis tentang masa lalu, tetapi seiring bertambahnya contoh yang terakumulasi dalam sumber data yang mendokumentasikan pengalaman masa lalu, langkah-langkah evaluasi, klasifikasi, dan kategorisasi dapat dilakukan berulang kali oleh algoritma yang cepat, sehingga memberikan keseluruhan proses kerja ukuran adaptabilitas. Ketika analisis deskriptif mencapai tahap yang mendukung tindakan antisipatif, ambang batas diteruskan ke ranah analisis prediktif.

Analisis prediktif menerapkan teknik-teknik canggih untuk memeriksa skenario dan membantu mendeteksi pola tersembunyi dalam data dalam jumlah besar untuk memproyeksikan peristiwa masa depan. Analisis ini menggunakan teknik yang mengelompokkan dan mengelompokkan data (transaksi, individu, peristiwa, dll.) ke dalam kumpulan yang koheren untuk memprediksi perilaku dan mendeteksi tren. Analisis ini

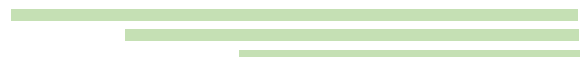


memanfaatkan teknik-teknik seperti pengelompokan, aturan pakar, pohon keputusan, dan jaringan saraf. Analisis prediktif paling umum digunakan untuk menghitung potensi perilaku dengan cara yang memungkinkan seseorang untuk:

- Memeriksa deret waktu, mengevaluasi data dan tren masa lalu untuk memprediksi permintaan masa depan (level, tren, musiman). Metode lanjutan meliputi identifikasi pola siklus, isolasi dampak peristiwa eksternal (misalnya, cuaca), karakterisasi variabilitas inheren, dan deteksi tren.
- Menentukan hubungan "kausalitas" antara dua atau lebih deret waktu, misalnya, peramalan permintaan suku cadang pengganti di fasilitas perawatan bus kota dengan mempertimbangkan tingkat penggunaan historis dan perubahan permintaan penumpang yang diketahui, diprediksi, atau musiman.
- Mengekstrak pola dari data dalam jumlah besar melalui penambangan data, untuk memprediksi perilaku nonlinier yang tidak mudah diidentifikasi melalui pendekatan lain. Perilaku yang diprediksi ini dapat digunakan untuk membuat kebijakan yang mengotomatiskan tindakan yang akan diambil di masa mendatang; misalnya, dengan mengklasifikasikan klaim asuransi masa lalu, klaim di masa mendatang dapat ditandai untuk diselidiki jika memiliki probabilitas tinggi merupakan penipuan. Dalam istilah operasional, analitik prediktif dapat diterapkan sebagai panduan untuk menjawab pertanyaan seperti:
 - Siapa pelanggan terbaik saya dan apa cara terbaik untuk menargetkan mereka?
 - Pasien mana yang paling mungkin merespons pengobatan tertentu?
 - Apakah aplikasi asuransi ini kemungkinan besar akan ditolak?
 - Apakah ini transaksi mencurigakan yang mungkin merupakan penipuan?

Pada tingkat inilah istilah "analitik lanjutan" lebih tepat diterapkan. Termasuk di dalamnya adalah teknik untuk pemodelan prediktif dan simulasi serta peramalan. Dalam simulasi, sebuah model sistem dibuat; estimasi atau prediksi tentang perilaku sistem di masa mendatang dibuat dengan menguji model tersebut dalam berbagai skenario. Simulasi membutuhkan kemampuan untuk membangun algoritma atau konstruksi matematika yang memberikan representasi yang cukup akurat dari perilaku sistem yang dapat diamati. Hal ini pada gilirannya dapat digunakan untuk mengevaluasi usulan perubahan pada sistem sebelum diimplementasikan, sehingga meminimalkan biaya dan risiko.

Sebagian besar pemodelan proses bisnis termasuk dalam kategori ini. Peramalan, yang merupakan bagian dari analitik prediktif, dapat diterapkan dalam berbagai cara, salah satunya adalah memprediksi beban kerja, yang seringkali diterjemahkan menjadi sumber daya yang dibutuhkan, termasuk sumber daya manusia. Aktivitas peramalan menetapkan kondisi akhir yang diinginkan, dan detailnya kemudian diterjemahkan ke dalam rencana operasional yang disepakati (masuk ke aktivitas perencanaan perusahaan). Bersama-sama, aktivitas deskriptif, analitik lanjutan, perencanaan perusahaan, dan penutupan/konsolidasi/kepatuhan tahap akhir membentuk sistem manajemen kinerja loop tertutup yang berulang-ulang dalam suatu bisnis.



Teknik pemodelan prediktif juga dapat digunakan untuk memeriksa data guna mengevaluasi hipotesis. Jika setiap titik data (atau observasi) terdiri dari beberapa atribut, maka mungkin bermanfaat untuk memahami apakah beberapa kombinasi dari subset atribut dapat memprediksi kombinasi atribut lainnya. Misalnya, seseorang dapat memeriksa klaim asuransi untuk memvalidasi hipotesis bahwa usia, jenis kelamin, dan kode pos dapat memprediksi kemungkinan klaim asuransi mobil. Alat pemodelan prediktif dapat membantu dalam memvalidasi dan menghasilkan hipotesis. Hal ini khususnya berguna ketika beberapa atribut merupakan tindakan yang ditentukan oleh para pengambil keputusan bisnis.

Data merupakan inti dari analitik prediktif, dan untuk menghasilkan tampilan yang lengkap, data digabungkan dari data deskriptif (atribut, karakteristik, geografis/demografis), data perilaku (pesanan, transaksi, riwayat pembayaran), data interaksi (email, transkrip obrolan, aliran klik web), dan data sikap (opini, preferensi, kebutuhan, dan keinginan). Dengan tampilan yang lengkap, pelanggan dapat mencapai kinerja yang lebih tinggi seperti menurunkan biaya klaim secara drastis, memerangi penipuan dan memaksimalkan pengembalian modal, mengubah pusat panggilan menjadi pusat laba, melayani pelanggan lebih cepat, dan secara efektif mengurangi biaya.

Selain menangkap data, mengakses data terpercaya dan sosial di dalam dan di luar bisnis, serta memodelkan dan menerapkan algoritma prediktif, penerapan model ini sama pentingnya untuk memaksimalkan dampak analitik dalam operasi waktu nyata (real-time).

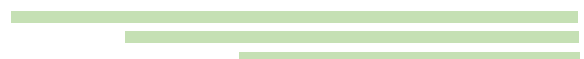
Aplikasi analitik prediktif. Analitik prediktif dapat digunakan dalam banyak aplikasi. Berikut ini beberapa contoh yang telah memberikan dampak positif.

Sistem pendukung keputusan medis. Para ahli menggunakan analitik prediktif dalam layanan kesehatan terutama untuk menentukan pasien mana yang berisiko terkena kondisi tertentu seperti diabetes, asma, penyakit jantung, dan penyakit seumur hidup lainnya.

Deteksi penipuan. Penipuan tersebar luas di berbagai industri. Kasus penipuan muncul di berbagai bidang seperti aktivasi kartu kredit, faktur, SPT, aktivitas daring, klaim asuransi, dan aktivitas panggilan telekomunikasi. Semua industri ini tertarik untuk mendeteksi penipuan dan menindak mereka yang bertanggung jawab, serta mencegah dan memantau penipuan dengan biaya yang wajar. Pemodelan prediktif dapat membantu mereka mencapai tujuan ini. Pemodelan ini juga dapat digunakan untuk mendeteksi penipuan laporan keuangan di perusahaan.

Asuransi. Seperti halnya penipuan, klaim yang tinggi dan mencurigakan secara tak terduga merupakan momok bagi perusahaan asuransi. Mereka ingin menghindari pembayaran klaim tersebut. Meskipun tujuannya cukup sederhana, pemodelan prediktif hanya berhasil sebagian dalam menghilangkan sumber kerugian tinggi ini bagi perusahaan. Ini merupakan area yang menjanjikan untuk penelitian lebih lanjut.

Kesehatan. Meskipun aplikasi sistematis pemodelan prediktif dalam layanan kesehatan relatif baru, aplikasi fundamentalnya serupa dengan yang ada di bidang lain. Bagaimanapun, meminimalkan risiko pelanggan adalah tujuannya. Dalam layanan kesehatan, ini adalah risiko readmisi, yang dapat dikurangi dengan mengidentifikasi pasien berisiko tinggi dan memantaunya.



Prediksi Keuangan: Analisis prediktif berguna dalam prediksi keuangan.

Manajemen Hubungan Pelanggan Analitis (CRM): Manajemen hubungan pelanggan analitis adalah aplikasi komersial yang sering digunakan dari analisis prediktif. CRM menggunakan analisis prediktif dalam aplikasi untuk kampanye pemasaran, penjualan, dan layanan pelanggan, sebagai contoh.

Retensi Pelanggan: Dengan pemeriksaan berkala terhadap penggunaan layanan, kinerja, pengeluaran, dan pola perilaku pelanggan lainnya di masa lalu, model prediktif dapat menentukan kemungkinan pelanggan ingin segera menghentikan layanan.

Berbagai macam teknik dapat digunakan untuk membangun model analitik prediktif:

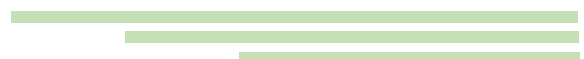
- Jaringan saraf tiruan dapat menilai seberapa besar kemungkinan pemegang kartu melakukan transaksi kartu kredit dengan mengevaluasi seberapa dekat transaksi ini dengan pola yang diprediksi oleh perilaku masa lalu orang tersebut.
- Model regresi dapat menentukan karakteristik pelanggan mana yang membuatnya lebih mungkin untuk berhenti berlangganan, atau mengalami attrite, sehingga memungkinkan perhitungan risiko churn di masa mendatang.
- Model respons dapat memprediksi seberapa besar kemungkinan seseorang akan merespons tawaran pemasaran tertentu, berdasarkan keberhasilan atau kegagalan tawaran yang dibuat di masa lalu.
- Kartu skor prediktif dapat menentukan kemungkinan seseorang akan gagal membayar pinjamannya di tahun mendatang.

Prediksi risiko, penipuan, dan peluang pelanggan ini semuanya dibuat dari sejumlah besar data historis. Untuk membangun model-model ini, Anda tidak hanya membutuhkan data, tetapi data dari waktu ke waktu. Data yang berubah seiring waktu mengungkapkan perilaku dan pola. Seperti disebutkan di atas, kami memiliki data terbanyak untuk operasi kami. Oleh karena itu, lingkungan operasional kami adalah tempat kami memiliki data yang kami perlukan untuk membangun model analitik prediktif.

Analitik Preskriptif

Analitik preskriptif menggunakan data untuk menentukan tindakan terbaik guna meningkatkan peluang mencapai hasil terbaik. Analitik preskriptif mengevaluasi dan menentukan cara-cara baru untuk beroperasi, menargetkan tujuan bisnis, dan menyeimbangkan semua kendala. Teknik seperti optimasi dan simulasi berada di bidang ini. Bisnis, seiring upaya mereka untuk menjadi lebih matang secara analitis, telah menunjukkan tujuan untuk naik dalam hierarki analitik guna mengoptimalkan proses bisnis atau operasional mereka.

Mereka memandang penggunaan analitik preskriptif sebagai faktor pembeda bagi bisnis mereka yang akan memungkinkan mereka untuk melepaskan diri dari persaingan. Jelas, analitik mengarah pada optimasi. Namun, jelas juga bahwa optimasi bergantung pada proses analitik. Optimasi adalah serangkaian teknik matematika yang secara komputasional menentukan serangkaian tindakan atau keputusan alternatif bernilai tinggi dengan serangkaian tujuan, persyaratan, dan kendala yang kompleks, dengan tujuan meningkatkan



kinerja bisnis. Analisis preskriptif, yang merupakan bagian dari "analisis lanjutan", didasarkan pada konsep optimasi, yang dapat dibagi menjadi dua area:

- Optimasi: Bagaimana kita bisa mencapai hasil terbaik?
- Optimasi stokastik: Bagaimana kita bisa mencapai hasil terbaik dan mengatasi ketidakpastian dalam data untuk membuat keputusan yang lebih baik?

Setelah masa lalu dipahami dan prediksi dapat dibuat tentang apa yang mungkin terjadi di masa depan, saatnya untuk memikirkan respons atau tindakan terbaik, mengingat keterbatasan sumber daya perusahaan. Inilah area analisis preskriptif. Banyak masalah melibatkan terlalu banyak pilihan atau alternatif bagi pengambil keputusan manusia untuk mempertimbangkan, menimbang, dan mengambil keputusan secara efektif misalnya, masalah penjadwalan atau perencanaan kerja. Di masa lalu, masalah ini hanya dapat diselesaikan menggunakan komputer yang menjalankan algoritma pada kumpulan data tertentu selama berjam-jam atau bahkan sehari-hari.

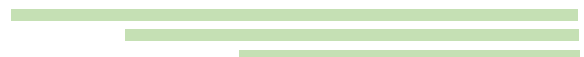
Menanamkan kemampuan pemecahan masalah seperti itu ke dalam sistem pendukung keputusan tidaklah bermanfaat karena tidak dapat memberikan hasil yang tepat waktu. Namun, kini, dengan peningkatan kecepatan dan ukuran memori komputer, serta kemajuan signifikan dalam kinerja algoritma matematika yang mendasarinya, komputasi serupa dapat dilakukan dalam hitungan menit. Meskipun informasi semacam ini mungkin telah digunakan di masa lalu untuk membentuk kebijakan dan menawarkan panduan tindakan dalam berbagai situasi, penilaian kini dapat diselesaikan secara real-time untuk mendukung keputusan dalam memodifikasi tindakan, mengalokasikan sumber daya, dan sebagainya.

6.2 MENANAMKAN AI KE DALAM PROSES BISNIS

Penggunaan AI dalam bisnis sedang meningkat, tetapi AI dan pembelajaran mesin belum efektif untuk setiap operasi bisnis. Meskipun ada banyak aspek bisnis yang dapat diotomatisasi, dan seharusnya diotomatisasi, tugas-tugas yang memerlukan penilaian, prioritas, dan pertimbangan yang matang tetap membutuhkan kecerdasan manusia.

Menanamkan AI ke dalam proses bisnis spesifik yang dapat menghasilkan hasil positif adalah langkah awal yang baik. Seiring waktu, integrasi AI ke dalam setiap inisiatif bisnis baru akan menjadi hal yang alami. Pada akhirnya, kemahiran AI di seluruh bisnis Anda harus mencakup:

- ✚ *Menentukan siapa yang akan memimpin dan bertanggung jawab atas inisiatif AI*
- ✚ *Mengidentifikasi orang-orang di bisnis Anda yang memiliki keahlian AI atau merekrut talenta dari luar bisnis Anda*
- ✚ *Menentukan di mana AI dapat memberikan nilai tambah atau mengatasi permasalahan*
- ✚ *Menggunakan AI untuk meningkatkan kapabilitas yang paling krusial bagi kesuksesan bisnis*
- ✚ *Meningkatkan pengambilan keputusan dan proses lainnya dengan kecerdasan otomatis*



Tindakan yang memerlukan komputasi atau pengorganisasian data dalam jumlah besar dapat ditangani dengan lebih baik oleh mesin, dan menerapkan solusi otomatis yang menggunakan AI akan menghapus model bisnis lama lainnya. Bisnis mengumpulkan data dalam jumlah besar dari operasi bisnis, interaksi sosial, dan sensor. Memanfaatkan kekuatan big data untuk membangun solusi otomatis dan memberikan wawasan akan menghasilkan keunggulan kompetitif yang berkelanjutan.

Wajar jika perbedaan tersebut menjadi kabur, terutama seiring kemajuan teknologi AI. Untuk menentukan strategi terbaik, bisnis harus mempertanyakan apakah tugas yang dihadapi memungkinkan pengulangan, volume tinggi, pola, dan biaya kesalahan yang rendah. Tugas yang memenuhi kriteria spesifik ini merupakan titik awal untuk implementasi AI. AI dapat digunakan untuk memberikan respons yang lebih baik terhadap permintaan transaksi dengan mengumpulkan data obrolan historis secara efisien. Proses di pinggiran bisnis, dengan biaya kesalahan yang rendah, juga merupakan peluang yang tepat untuk menerapkan AI.

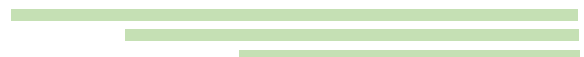
Setelah bisnis mengatur tugas mana yang dapat diotomatisasi, bisnis dapat menetapkan bagaimana AI dapat diimplementasikan dengan sukses. Untuk mencapai hal ini, bisnis harus merujuk pada daftar periksa berikut:

- *Kenali masalahnya*
 - Mulailah dengan mengenali masalah bisnis dan tanyakan di mana penggunaan AI dapat meningkatkan efisiensi.
- *Identifikasi sumber data*
 - Setelah masalah diidentifikasi, bedakan sumber data dan kumpulkan data dari titik kontak pelanggan yang relevan.
- *Kembangkan solusi berbasis AI*
 - Kembangkan solusi berbasis AI untuk membantu pengambilan keputusan algoritmik dan memanfaatkan jaringan saraf tiruan dan NLP.

Implementasi dan Tindakan

Selanjutnya, solusi AI dapat diimplementasikan secara berulang. Pelatihan harus diikuti agar staf dapat bekerja secara efektif dengan AI dan memahami penggunaan serta aplikasinya dalam situasi yang tepat, serta menyediakan intervensi manusia bila diperlukan. Dengan menjelaskan proses implementasi, bisnis dapat menetapkan ekspektasi yang realistis tentang bagaimana AI akan mulai memengaruhi bisnis mereka, sekaligus memanfaatkan apa yang dapat ditawarkan oleh teknologi tersebut.

Terakhir, ingatlah bahwa tujuan dari setiap produk, layanan, atau pengalaman baru adalah untuk menyediakan solusi terbaik bagi pelanggan Anda. Hindari pendekatan yang mengutamakan teknologi yang mengutamakan teknologi daripada kebutuhan mereka dan gunakan teknologi modern untuk memenuhi kebutuhan yang telah Anda identifikasi sebelumnya. Teknologi bekerja paling baik ketika tidak terlihat dan bekerja di latar belakang. Kecerdasan buatan pun demikian. Orang-orang seharusnya tidak pernah tahu bahwa mereka sedang berhadapan dengan AI. Mereka seharusnya merasa bahwa mereka dilayani dengan baik oleh perusahaan yang telah mereka pilih untuk menyediakan produk, layanan, atau pengalaman tersebut.



Seperti teknologi lainnya, AI terapan sebenarnya tidak lebih dari sekadar alat yang digunakan bisnis untuk menyelesaikan suatu tugas. Nilainya terletak pada seberapa baik ia melakukannya. Pada akhirnya, ini berarti nilainya diukur dari seberapa baik perusahaan menggunakannya untuk melayani pelanggan mereka dengan baik. Setelah inovasi kecerdasan buatan memudar, hanya inilah yang tersisa. Jadi, untuk membuka nilai AI, yang terbaik adalah meninggalkan teknologi ini sepenuhnya dan melayani pelanggan Anda dengan baik.

6.3 KECERDASAN BUATAN UNTUK PERTUMBUHAN

AI untuk Layanan Pelanggan

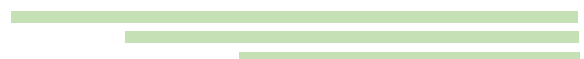
Saat ini, pelanggan terikat dengan perangkat mereka sepanjang waktu dan beralih di antara perangkat tersebut dengan sempurna. Terdapat kesenjangan yang semakin besar antara jumlah data pelanggan yang dihasilkan dan kapasitas teknik pemasaran tradisional yang sebagian besar didukung oleh analisis manusia untuk memproses data ini. Kemajuan dalam teknologi big data seperti Hadoop telah memudahkan pengambilan data mentah dalam beragam format dan menyimpannya di beberapa penyimpanan data berbeda yang biasanya disebut danau data yang mencakup sistem SQL, sistem NoSQL, file datar, dan lembar Excel.

Ini adalah tambang emas mentah yang Anda gunakan dan Anda harus memprioritaskan pengambilan data dalam format apa pun daripada memaksakannya ke penyimpanan data atau skema. Perangkat AI yang Anda investasikan harus beradaptasi dengan campuran data terstruktur dan tidak terstruktur ini. Waktu paruh niat konsumen semakin pendek setiap tahunnya, dan pelanggan mengharapkan pengalaman "sesuai permintaan" yang relevan secara kontekstual dan dipersonalisasi untuk mereka di setiap perangkat. Pemasar pertumbuhan harus memprioritaskan algoritma dan proses AI yang lebih sederhana yang dapat beradaptasi dengan baik terhadap data waktu nyata daripada solusi mode batch yang lebih kompleks yang mungkin memerlukan waktu beberapa jam atau hari untuk dieksekusi. Perhatikan dengan saksama waktu pelatihan yang dibutuhkan untuk membangun dan menerapkan model AI dan seberapa cepat mereka dapat mengintegrasikan data baru.

Meskipun idealnya setiap atribut dan preferensi diketahui tentang semua pengguna, Anda akan berakhir dengan bidang data yang tidak lengkap atau hanya diketahui sebagian meskipun Anda telah berusaha sebaik mungkin. Pemasar pertumbuhan B2C khususnya harus mengantisipasi hal ini sejak hari pertama dan berinvestasi pada alat dan solusi yang dapat beradaptasi dengan baik terhadap data yang tidak lengkap.

Ambil contoh lokasi pengguna, mungkin terdapat campuran data lokasi yang diberikan pengguna, dengan lintang/bujur perangkat, IP ke geografis, inferensi dari konten yang dilihat atau pencarian yang dilakukan, dan banyak lagi. Sebagai pemasar pertumbuhan, Anda sebaiknya memilih alat AI yang dapat beradaptasi dengan baik terhadap campuran semua data ini dan menghasilkan jawaban terbaik untuk basis pengguna terluas daripada hanya pada beberapa pengguna dengan data yang lengkap dan bersih.

Beberapa algoritma AI mengharapkan data pelatihan untuk diumpankan kepada mereka dan ukuran serta ketersediaan data pelatihan merupakan kendala besar yang harus diatasi untuk menggunakannya secara efektif. Kelas algoritma AI tertentu seperti Boosted



Random Forests lebih baik dalam beradaptasi dengan ukuran data pelatihan daripada jaringan saraf konvolusional, sejenis pembelajaran mendalam. Pemasar pertumbuhan sebaiknya memilih algoritma yang dapat bekerja dengan data pelatihan terbatas dan memiliki teknik pengambilan sampel bawaan untuk menangani ukuran kelas yang tidak proporsional.

Pemanfaatan AI yang efektif dapat mempersempit tantangan kognitif manusia ini dengan memproses data dalam jumlah besar secara cepat dan menggunakan pembelajaran mesin untuk mengenali pola dalam data serta memprediksi langkah selanjutnya berdasarkan perilaku audiens serupa di masa lalu. Meskipun terdengar logis dan lugas, hal ini tidak semudah membeli sistem AI kotak hitam dan langsung menggunakannya.

Keputusan untuk menerapkan sistem AI guna meningkatkan kinerja pemasaran merupakan keputusan besar yang harus dilakukan dengan tingkat perencanaan dan persiapan yang sama seperti saat menerapkan sistem atau sistem otomatisasi pemasaran Anda. Sistem bertenaga AI akan menjadi sistem kecerdasan Anda yang harus bekerja sama erat dengan sistem-sistem lain ini untuk meningkatkan hasil.

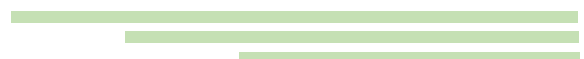
Menerapkan AI untuk Pemasaran

Dalam pemasaran, penyatuan data merupakan tantangan terbesar mereka ketika ingin memaksimalkan data pelanggan. Adanya silo data merupakan salah satu alasan utama mengapa pemasar tidak dapat sepenuhnya memanfaatkan semua data pelanggan mereka. Namun, dengan AI, pemasar dapat berhasil mendobrak silo data dan secara efektif menghasilkan serta mengorkestrasi wawasan pelanggan dan intelijen yang dapat ditindaklanjuti.

Banyak bisnis bercita-cita menjadi berpusat pada pelanggan dan didorong oleh data, namun hanya sedikit yang dapat mengubah data menjadi tindakan yang menguntungkan. AI dapat menjembatani kesenjangan ini dan memungkinkan pemasar untuk menjadi berorientasi pelanggan, didorong oleh wawasan, cepat, dan terhubung. AI akan memberi mereka visibilitas yang lebih besar terhadap perilaku pelanggan, membuat penawaran yang tepat dan kontekstual, serta memberikan pengalaman yang personal dan terpadu di semua saluran. Dan berdasarkan wawasan yang dihasilkan, AI memberdayakan pemasar untuk lebih mengoptimalkan penawaran dan strategi keseluruhan mereka. Adopsi AI tidak hanya berdampak pada proses pemasaran tetapi juga seluruh bisnis.

Tentukan kesenjangan organisasi yang harus ditutup dan atasi faktor-faktor serta kesalahan yang dapat menghambat bisnis dalam menerapkan AI. Penting untuk menguraikan perjalanan pelanggan dan mencakup semua langkah dalam proses tersebut, melampaui apa yang terjadi dalam pemasaran. Pahami di mana hambatannya dan atasi bagian-bagian penting yang ingin Anda bantu selesaikan dengan AI. Lebih lanjut, penting untuk memastikan bahwa setiap orang dalam bisnis memiliki pemahaman yang benar tentang AI. Misalnya, mereka harus menyadari bahwa AI saja tidak dapat menyampaikan pengetahuan. AI membutuhkan manusia dan teknologi. Hal ini membawa kita ke langkah berikutnya.

Pemasar harus tahu cara menerapkan AI secara efektif dan menetapkan kontrol serta sistem pemantauan yang tepat. Haruskah mereka mengaktifkan model dan membiarkannya menghasilkan hasil untuk ditinjau orang? Atau, haruskah mereka mengintegrasikannya ke



dalam aplikasi yang mengotomatiskan proses seperti mempersonalisasi konten dan mengoptimalkan kampanye email? Bagaimana mereka akan menerapkan kontrol dan pemantauan yang tepat untuk memastikan bahwa model berfungsi dengan baik dan memberikan hasil? Memberdayakan pemasaran yang didukung AI dapat menghasilkan perjalanan pelanggan yang optimal, efisiensi yang lebih tinggi, keputusan yang lebih cerdas, peningkatan kecepatan, dan peningkatan kinerja yang berkelanjutan.

Bisnis yang mengadopsi AI berharap melihat peningkatan produktivitas. Kompetensi baru dapat diperoleh dari perampingan tugas yang sebelumnya membutuhkan waktu sehari-hari untuk diselesaikan manusia dan peningkatan proses kerja dengan memadukan manusia dan mesin dengan cara-cara baru. Banyak bisnis juga percaya AI akan mempercepat dan meningkatkan inovasi, yang selanjutnya menciptakan lapangan kerja baru. Beberapa berharap melihat peningkatan permintaan konsumen dari produk dan layanan yang disempurnakan dengan AI yang lebih personal dan berkualitas tinggi.

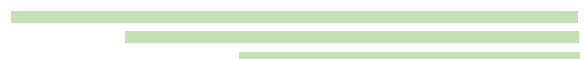
Di saat yang sama, beberapa khawatir bahwa AI akan mengakibatkan hilangnya lapangan kerja. Namun untuk memanfaatkan teknologi ini, bisnis perlu mempekerjakan orang dengan pengalaman AI atau keterampilan untuk menganalisis dan menggunakan data. Bisnis juga membutuhkan daya komputasi dan infrastruktur sistem untuk mendukung produk dan layanan yang mendukung AI, dan mereka membutuhkan platform untuk mengatur dan mengintegrasikan data mereka.

Menyiapkan infrastruktur ini bisa mahal. Data yang dikumpulkan oleh AI menawarkan tantangan besar lainnya. Bagaimana bisnis menjamin validitasnya? Batasan apa yang perlu mereka terapkan pada penggunaannya? Dapatkah perlindungan memastikan bahwa mesin menjalankan perintah manusia sebagaimana mestinya? Ada kekhawatiran bahwa bias yang tidak disengaja dapat masuk ke dalam algoritma AI atau model pengambilan keputusan. Bisnis perlu membangun kontrol yang kuat untuk mencegah hal ini terjadi dan memantau sistem yang belajar melalui AI. Memberi tahu para pemangku kepentingan tentang pengawasan bisnis dapat membantu membangun kepercayaan dengan para pemangku kepentingan bahwa bisnis menggunakan AI dengan andal.

Teknologi **AI generatif** telah menjadi pilar utama dalam transformasi digital saat ini. Dengan kemampuannya untuk menciptakan konten, menyederhanakan proses, dan mempersonalisasi pengalaman pengguna, teknologi ini menawarkan peluang besar bagi berbagai sektor untuk tetap kompetitif dan relevan dalam pasar yang terus berubah. Artikel ini mengulas pengertian AI generatif dan menggambarkan 10 cara utama teknologi ini mendukung bisnis pada tahun 2025.

Apa Itu AI Generatif?

AI generatif adalah teknologi kecerdasan buatan yang mampu menghasilkan konten baru, mulai dari teks, gambar, musik, kode, hingga desain produk. Berbeda dengan AI tradisional yang biasanya hanya mengenali pola atau mengambil keputusan berdasarkan data, AI generatif menciptakan output kreatif yang meniru kemampuan berpikir manusia. Contohnya adalah model AI seperti **GPT-4** (OpenAI), **Claude** (Anthropic), dan **Gemini** (Google), yang



dilatih menggunakan data dalam jumlah besar untuk menjawab pertanyaan serta menghasilkan ide inovatif.

Dalam lingkungan perusahaan, konsep seperti **Retrieval-Augmented Generation (RAG)** memungkinkan integrasi data internal secara aman, sehingga AI dapat memberikan respons yang lebih relevan dan personal. Karena itu, AI generatif bukan hanya alat teknologi, melainkan juga pendorong utama inovasi strategis.

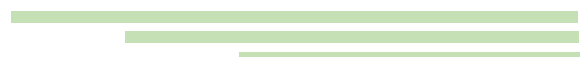
Pentingnya AI Generatif dalam Dunia Bisnis

Di tengah percepatan digitalisasi, AI generatif muncul sebagai kunci untuk mendorong pertumbuhan dan kesinambungan bisnis. Teknologi ini menyediakan kemampuan untuk menghasilkan solusi baru dan ide kreatif secara otomatis. Berikut adalah alasan mengapa AI generatif sangat penting bagi bisnis:

- **Menghemat Biaya melalui Otomatisasi**
AI generatif mengurangi biaya operasional dengan mengotomatisasi tugas-tugas yang sebelumnya memerlukan banyak sumber daya manusia, seperti pembuatan konten pemasaran dan simulasi desain produk, sehingga memungkinkan pengalokasian dana ke area yang lebih strategis.
- **Meningkatkan Produktivitas dengan Efisiensi Waktu**
AI generatif mempercepat proses bisnis, misalnya dalam pembuatan laporan analisis data atau pembuatan konsep kreatif, membebaskan waktu tim untuk fokus pada pengambilan keputusan penting.
- **Mengasah Kreativitas dengan Ide-ide Baru**
Teknologi ini mampu menghasilkan inovasi mulai dari konsep produk baru hingga kampanye pemasaran segar, membuka jalan bagi eksplorasi ide yang sebelumnya sulit dijangkau.
- **Memberikan Pengalaman Personal kepada Pelanggan**
Dengan memahami preferensi individual pelanggan, AI generatif memungkinkan bisnis untuk memberikan rekomendasi dan layanan yang disesuaikan secara personal, meningkatkan kepuasan dan loyalitas pelanggan.

10 Penerapan AI Generatif dalam Bisnis Tahun 2025

1. **Deteksi Penipuan dan Keamanan**
AI generatif menciptakan data sintetis untuk melatih sistem deteksi penipuan dan mensimulasikan serangan siber, memperkuat kemampuan pertahanan perusahaan. Contohnya, Bank HSBC menggunakan AI untuk mendeteksi aktivitas mencurigakan secara real-time.
2. **Percepatan Pengembangan Perangkat Lunak**
AI membantu penulisan kode, deteksi bug, dan pembuatan dokumentasi otomatis, mempercepat siklus pengembangan aplikasi seperti yang diterapkan Microsoft.
3. **Otomatisasi Pembuatan Konten**
AI menghasilkan artikel, deskripsi produk, dan email pemasaran dengan cepat, seperti yang dilakukan Amazon untuk pasar globalnya.
4. **Kampanye Pemasaran yang Dipersonalisasi**



AI menganalisis data pelanggan untuk menyusun kampanye yang relevan, contohnya Spotify yang membuat playlist yang disesuaikan secara otomatis.

5. **Desain Produk Adaptif**

AI menggabungkan preferensi pelanggan dan tren pasar untuk mempercepat pengembangan produk, seperti Tesla dalam mengoptimalkan desain baterai dan aerodinamika.

6. **Optimalisasi Rantai Pasokan**

AI memprediksi permintaan dan mengelola inventaris untuk meningkatkan efisiensi, contohnya Walmart dalam penyesuaian stok selama musim liburan.

7. **Chatbot Layanan Pelanggan**

Chatbot AI menyediakan layanan 24/7 dengan respons cepat dan solusi yang relevan, seperti yang digunakan Delta Airlines.

8. **Otomatisasi Penjualan**

AI membuat pesan personal dan mengotomatiskan tindak lanjut prospek, membantu meningkatkan konversi seperti yang diimplementasikan Salesforce.

9. **Pemodelan Keuangan dan Peramalan**

AI menghasilkan proyeksi pendapatan dan laporan keuangan otomatis, digunakan oleh perusahaan besar seperti Procter & Gamble.

10. **Perekrutan dan Otomasi HR**

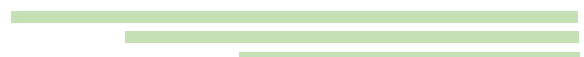
AI menyaring kandidat dan menyusun modul pelatihan personal, mempercepat proses perekrutan seperti yang dijalankan oleh Unilever dengan jumlah aplikasi besar.

Dengan berbagai aplikasi ini, AI generatif tidak hanya menjadi alat bantu, tetapi juga menjadi motor utama inovasi dan efisiensi dalam bisnis modern tahun 2025. Transformasi melalui AI generatif memungkinkan perusahaan untuk merespons perubahan dengan cepat, meningkatkan produktivitas, dan menciptakan nilai kompetitif yang berkelanjutan.

Contoh nyata implementasi AIoT (Artificial Intelligence of Things) untuk predictive maintenance di pabrik dapat dilihat pada beberapa perusahaan manufaktur modern yang telah mengintegrasikan sensor IoT dengan teknologi AI untuk memantau kondisi mesin secara real-time.

Misalnya, dalam sebuah proyek di industri manufaktur, sensor-sensor yang dipasang pada mesin-mesin produksi secara terus-menerus mengumpulkan data tentang parameter kritis seperti suhu, getaran, tekanan, dan kecepatan operasi mesin. Data ini kemudian dikirim ke platform cloud yang menggunakan algoritma machine learning untuk menganalisis data tersebut dan memprediksi potensi kegagalan atau kebutuhan perawatan sebelum kerusakan terjadi. Dengan demikian, perusahaan dapat melakukan pemeliharaan secara terjadwal dan efektif, menghindari downtime tidak terduga yang merugikan produksi.

Salah satu contoh nyata adalah perusahaan manufaktur otomotif "The AutomobileTech" yang mengimplementasikan sistem predictive maintenance berbasis AIoT. Mereka melengkapi mesin produksinya dengan sensor pintar dan platform analitik data. Sistem ini mampu memantau kesehatan mesin secara terus-menerus dan memberikan peringatan dini jika terdapat indikasi kerusakan, sehingga perusahaan dapat mengurangi



downtime tak terduga, mengoptimalkan jadwal perawatan, dan menekan biaya operasional akibat kerusakan tiba-tiba.

Selain itu, General Electric (GE) menggunakan teknologi digital twin dan AI dalam operasinya untuk memprediksi kegagalan pada mesin pabrik mereka secara global. Sistem ini menganalisis data dari ribuan sensor yang terhubung, memungkinkan GE menghemat miliaran dolar melalui pengurangan waktu henti dan perbaikan yang lebih terencana.

Implementasi ini tidak hanya meningkatkan efisiensi operasional dan memperpanjang umur mesin, tetapi juga membantu perusahaan menjadi lebih responsif dan adaptif terhadap kebutuhan produksi, menjaga kelancaran proses manufaktur dengan biaya yang lebih terkendali.

6.4 BISNIS AI DI INDONESIA

Kata.ai merupakan sebuah perusahaan teknologi kecerdasan buatan yang berasal dari Indonesia, didirikan pada tahun 2015. Perusahaan ini fokus mengembangkan teknologi pemrosesan bahasa alami (Natural Language Processing) yang memungkinkan chatbot mereka memahami dan merespons percakapan dalam bahasa Indonesia dengan tingkat akurasi sekitar 87 persen.

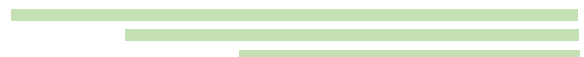
Awalnya, Kata.ai beroperasi sebagai layanan asisten virtual bernama YesBoss, yang melayani pengguna lewat SMS dan aplikasi mobile. Pada 2016, perusahaan bertransformasi menjadi startup B2B dengan nama Kata.ai, menyediakan platform mesin dialog berbasis bahasa alami untuk berbagai merek dan perusahaan membangun chatbot mereka sendiri. Chatbot ini dapat menjalankan fungsi seperti pemasaran, penjualan, dan pengumpulan data perilaku pelanggan melalui platform pesan populer seperti WhatsApp, Facebook Messenger, dan LINE.

Kata.ai telah melayani lebih dari 100 perusahaan dari beragam sektor seperti telekomunikasi, keuangan, consumer goods, ritel, dan perbankan. Mereka juga menawarkan solusi untuk bisnis kecil dan menengah melalui produk Kata Omnichat, yang mengintegrasikan WhatsApp Business dan Instagram dalam satu dashboard.

Pada 2021, perusahaan mengantongi sertifikasi ISO 27001:2013 sebagai tanda pengelolaan keamanan data yang sesuai standar. Kata.ai juga telah menerima investasi besar dari berbagai pihak seperti Trans-Pacific Technology Fund dan MDI Ventures, yang mendukung pengembangan dan ekspansi mereka.

Selain itu, Kata.ai terus berinovasi dengan memperluas jangkauan, termasuk menjalin kerja sama dengan startup AI internasional seperti Kanari AI dari Qatar untuk memperkuat layanan percakapan multibahasa di wilayah Asia Tenggara dan Timur Tengah.

Transformasi dari YesBoss ke Kata.ai bermula sejak pendirian YesBoss pada tahun 2015, yang pada awalnya berfokus pada layanan asisten virtual melalui SMS dan aplikasi mobile dengan berbagai fitur seperti pemesanan tiket, makanan, dan reservasi. YesBoss mengadopsi pendekatan Human Assisted Artificial Intelligence (HAAI) untuk memberikan layanan yang personal kepada para penggunanya.



Pada tahun 2016, para pendiri menyadari bahwa segmen bisnis B2B (business to business) memiliki potensi pasar yang jauh lebih besar dibandingkan layanan langsung ke konsumen (B2C). Oleh karena itu, mereka memutuskan untuk mengubah model bisnis menjadi pengembangan platform chatbot dan kecerdasan buatan yang dapat dimanfaatkan perusahaan untuk meningkatkan interaksi dengan pelanggan. Bersamaan dengan perubahan ini, nama perusahaan pun berubah menjadi Kata.ai.

Transformasi tersebut disertai dengan pengembangan teknologi Natural Language Processing (NLP) yang memungkinkan chatbot untuk memahami bahasa Indonesia sehari-hari, termasuk bahasa gaul dan singkatan, sehingga percakapan menjadi lebih alami dan efektif. Berkat teknologi ini, Kata.ai berhasil menjalin kemitraan dengan perusahaan besar seperti Telkomsel (melalui chatbot Veronika), Bank Rakyat Indonesia (chatbot Sabrina), dan Unilever (chatbot Jemma).

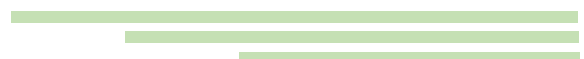
Perubahan fokus bisnis dan kemajuan teknologi ini membuat Kata.ai mendapatkan pendanaan besar, termasuk Seri A senilai US\$3,5 juta pada tahun 2017, yang semakin menguatkan posisi mereka sebagai pionir AI percakapan di Indonesia. Saat ini, Kata.ai terus memperluas jangkauan layanan chatbotnya serta menambah jumlah klien di berbagai platform pesan instan populer.

Singkatnya, perjalanan YesBoss ke Kata.ai merupakan pergeseran dari layanan asisten virtual berbasis konsumen menjadi penyedia platform chatbot AI B2B yang inovatif dan berfokus pada pengembangan teknologi percakapan berbasis bahasa alami untuk pasar Indonesia dan Asia Tenggara. Kata.ai memainkan peran krusial dalam mendukung perkembangan sektor e-commerce dan fintech di Indonesia melalui pemanfaatan teknologi kecerdasan buatan, khususnya chatbot dan analisis data yang bertujuan untuk meningkatkan pelayanan dan efisiensi bisnis.

Di bidang e-commerce, Kata.ai menghadirkan chatbot AI yang memberikan layanan pelanggan yang cepat dan personal di platform populer seperti WhatsApp dan media sosial lainnya. Chatbot ini mampu menjawab pertanyaan konsumen secara otomatis, memberikan rekomendasi produk sesuai selera pelanggan, serta membantu proses pembelian. Selain itu, dengan analisis data pelanggan secara real-time, Kata.ai membantu pelaku e-commerce memahami tren pasar, mengelola stok, dan mengoptimalkan strategi pemasaran, sehingga mempercepat penjualan dan meningkatkan kepuasan pelanggan. Inovasi ini mendorong pertumbuhan e-commerce dan memperluas inklusi digital di Indonesia.

Sementara di sektor fintech, Kata.ai menyediakan solusi AI yang meningkatkan layanan pelanggan, seperti chatbot cerdas yang dapat melakukan verifikasi otomatis, mendeteksi potensi penipuan, dan menawarkan produk keuangan yang disesuaikan. Teknologi AI ini juga membantu perusahaan fintech dalam mengotomatisasi proses administratif dan menganalisis data keuangan untuk mendukung pengelolaan risiko, pemberian pinjaman, dan investasi. Hal ini mempercepat akses ke layanan keuangan yang aman dan mudah, terutama bagi masyarakat yang belum terlayani oleh perbankan konvensional.

Secara keseluruhan, Kata.ai berkontribusi signifikan dalam mempercepat transformasi digital di kedua sektor tersebut dengan mengintegrasikan AI untuk memperlancar komunikasi



dengan pelanggan, meningkatkan produktivitas, dan membuka peluang bisnis yang lebih luas di era digital Indonesia.

Kata.ai didirikan pada tahun 2015 di Indonesia oleh Irzan Raditya dan tiga rekannya, yang semuanya memiliki latar belakang teknologi dan pengalaman pendidikan di Jerman. Pada awalnya, perusahaan ini menghadirkan layanan asisten virtual bernama YesBoss yang beroperasi melalui SMS dan aplikasi mobile, menawarkan layanan pemesanan tiket pesawat, makanan, reservasi restoran, serta kendaraan. YesBoss berhasil mengumpulkan sejumlah besar data pesan yang kemudian dimanfaatkan untuk mengembangkan teknologi pemrosesan bahasa alami (Natural Language Processing/NLP) khusus untuk bahasa Indonesia.

Melihat peluang bisnis yang lebih besar di segmen B2B (business to business), pada tahun 2016 mereka mengubah model bisnis dan meluncurkan platform Kata.ai. Platform ini memungkinkan perusahaan untuk menciptakan chatbot cerdas berbasis AI yang digunakan untuk layanan pelanggan, pemasaran, serta transaksi jual-beli. Irzan Raditya, sebagai CEO, berhasil meraih dukungan dari berbagai investor internasional dari Taiwan, Amerika Serikat, Korea Selatan, dan Indonesia.

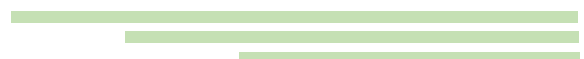
Kata.ai kini menjadi pionir dalam teknologi chatbot berbasis kecerdasan buatan di Indonesia, dengan kemampuan NLP yang dapat memahami bahasa sehari-hari, termasuk singkatan dan bahasa gaul. Hingga saat ini, platform ini telah memproses ratusan juta pesan dan digunakan oleh banyak perusahaan besar untuk meningkatkan interaksi dengan pelanggan melalui berbagai aplikasi pesan populer seperti WhatsApp, Facebook Messenger, dan LINE.

Melihat peluang bisnis yang lebih besar di segmen B2B (business to business), pada tahun 2016 mereka mengubah model bisnis dan meluncurkan platform Kata.ai. Platform ini memungkinkan perusahaan untuk menciptakan chatbot cerdas berbasis AI yang digunakan untuk layanan pelanggan, pemasaran, serta transaksi jual-beli. Irzan Raditya, sebagai CEO, berhasil meraih dukungan dari berbagai investor internasional dari Taiwan, Amerika Serikat, Korea Selatan, dan Indonesia.

Kata.ai kini menjadi pionir dalam teknologi chatbot berbasis kecerdasan buatan di Indonesia, dengan kemampuan NLP yang dapat memahami bahasa sehari-hari, termasuk singkatan dan bahasa gaul. Hingga saat ini, platform ini telah memproses ratusan juta pesan dan digunakan oleh banyak perusahaan besar untuk meningkatkan interaksi dengan pelanggan melalui berbagai aplikasi pesan populer seperti WhatsApp, Facebook Messenger, dan LINE.

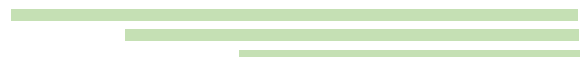
AI untuk Bisnis dalam tahun 2025 bukan lagi konsep masa depan, tetapi kenyataan yang semakin meluas. Seiring berkembangnya teknologi, Artificial Intelligence semakin menjadi kunci transformasi di berbagai sektor industri.

Tahun 2025 akan menjadi masa penting di mana AI tidak hanya mempercepat otomatisasi, tetapi juga mendorong strategi bisnis yang lebih cerdas, hemat biaya, dan berkelanjutan. Dari manufaktur hingga layanan kesehatan, potensi AI dalam industri semakin luas dan menjanjikan.

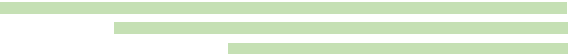


Namun, agar implementasi AI berhasil, perusahaan perlu memahami kebutuhan bisnisnya, memilih solusi yang tepat, dan membekali tenaga kerja dengan keterampilan digital yang relevan. Dengan pendekatan yang tepat, AI bukan hanya alat bantu, tetapi aset strategis yang membawa industri ke level berikutnya.

Siapkah Kamu Menghadapi Era Bisnis Berbasis AI?



DAFTAR PUSTAKA

- Abrokwah-Larbi, K., & Awuku-Larbi, Y. (2024). The impact of artificial intelligence in marketing on the performance of business organizations: evidence from SMEs in an emerging economy. *Journal of Entrepreneurship in Emerging Economies*, 16(4), 1090-1117.
- Agarwal, P., Swami, S., & Malhotra, S. K. (2024). Artificial intelligence adoption in the post COVID-19 new-normal and role of smart technologies in transforming business: a review. *Journal of Science and Technology Policy Management*, 15(3), 506-529.
- Alam, A., & Mohanty, A. (2022, November). Business models, business strategies, and innovations in EdTech companies: integration of learning analytics and artificial intelligence in higher education. In 2022 IEEE 6th Conference on Information and Communication Technology (CICT) (pp. 1-6). IEEE.
- Alghizzawi, M., Ezmigna, I., Jado, K., Alzeaiden, K., & Kanan, M. (2024). Theoretical perspective of artificial intelligence in digital marketing for business organizations. In *Business Analytical Capabilities and Artificial Intelligence-Enabled Analytics: Applications and Challenges in the Digital Era, Volume 1* (pp. 167-177). Cham: Springer Nature Switzerland.
- Badghish, S., & Soomro, Y. A. (2024). Artificial intelligence adoption by SMEs to achieve sustainable business performance: application of technology–organization–environment framework. *Sustainability*, 16(5), 1864.
- Bharadiya, J. (2023). The impact of artificial intelligence on business processes. *European journal of technology*, 7(2), 15-25.
- Bharadiya, J. P., Thomas, R. K., & Ahmed, F. (2023). Rise of artificial intelligence in business and industry. *Journal of Engineering Research and Reports*, 25(3), 85-103.
- Buntak, K., Kovačić, M., & Mutavdžija, M. (2021). Application of Artificial Intelligence in the business. *International journal for quality research*, 15(2), 403.
- Chen, L., Jiang, M., Jia, F., & Liu, G. (2022). Artificial intelligence adoption in business-to-business marketing: toward a conceptual framework. *Journal of Business & Industrial Marketing*, 37(5), 1025-1044.
- Drydakis, N. (2022). Artificial Intelligence and reduced SMEs' business risks. A dynamic capabilities analysis during the COVID-19 pandemic. *Information Systems Frontiers*, 24(4), 1223-1247.
- Enholm, I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2022). Artificial intelligence and business value: A literature review. *Information systems frontiers*, 24(5), 1709-1734.
- 

- Gellert, R. M. (2021). The role of the risk-based approach in the General data protection Regulation and in the European Commission's proposed Artificial Intelligence Act. Business as usual?.
- Getchell, K. M., Carradini, S., Cardon, P. W., Fleischmann, C., Ma, H., Aritz, J., & Stapp, J. (2022). Artificial intelligence in business communication: The changing landscape of research and teaching. *Business and Professional Communication Quarterly*, 85(1), 7-33.
- Gurjar, K., Jangra, A., Baber, H., Islam, M., & Sheikh, S. A. (2024). An analytical review on the impact of artificial intelligence on the business industry: Applications, trends, and challenges. *IEEE Engineering Management Review*, 52(2), 84-102.
- Han, R., Lam, H. K., Zhan, Y., Wang, Y., Dwivedi, Y. K., & Tan, K. H. (2021). Artificial intelligence in business-to-business marketing: a bibliometric analysis of current research status, development and future directions. *Industrial Management & Data Systems*, 121(12), 2467-2497.
- Kar, A. K., & Kushwaha, A. K. (2023). Facilitators and barriers of artificial intelligence adoption in business—insights from opinions using big data analytics. *Information Systems Frontiers*, 25(4), 1351-1374.
- Kitsios, F., & Kamariotou, M. (2021). Artificial intelligence and business strategy towards digital transformation: A research agenda. *Sustainability*, 13(4), 2025.
- Kubatko, O. V., Ozims, S. C., & Voronenko, V. I. (2024). Influence of artificial intelligence on business decision-making.
- Kulkov, I. (2021). The role of artificial intelligence in business transformation: A case of pharmaceutical companies. *Technology in Society*, 66, 101629.
- Kumar, S., Lim, W. M., Sivarajah, U., & Kaur, J. (2023). Artificial intelligence and blockchain integration in business: trends from a bibliometric-content analysis. *Information systems frontiers*, 25(2), 871-896.
- Limna, P. (2023). Artificial Intelligence (AI) in the hospitality industry: A review article. *International Journal of Computing Sciences Research*, 7, 1306-1317.
- Loureiro, S. M. C., Guerreiro, J., & Tussyadiah, I. (2021). Artificial intelligence in business: State of the art and future research agenda. *Journal of business research*, 129, 911-926.
- Mah, P. M., Skalna, I., & Muzam, J. (2022). Natural language processing and artificial intelligence for enterprise management in the era of industry 4.0. *Applied Sciences*, 12(18), 9207.
- Maslak, O. I., Maslak, M. V., Grishko, N. Y., Hlazunova, O. O., Pererva, P. G., & Yakovenko, Y. Y. (2021, September). Artificial intelligence as a key driver of business operations

- transformation in the conditions of the digital economy. In 2021 IEEE International Conference on Modern Electrical and Energy Systems (MEES) (pp. 1-5). IEEE.
- Nafizah, U. Y., Roper, S., & Mole, K. (2024). Estimating the innovation benefits of first-mover and second-mover strategies when micro-businesses adopt artificial intelligence and machine learning. *Small Business Economics*, 62(1), 411-434.
- Naim, A. (2022). Role of artificial intelligence in business risk management. *American Journal of Business Management, Economics, and Banking*, 1, 55-66.
- Olan, F., Arakpogun, E. O., Suklan, J., Nakpodia, F., Damij, N., & Jayawickrama, U. (2022). Artificial intelligence and knowledge sharing: Contributing factors to organizational performance. *Journal of Business Research*, 145, 605-615.
- Pallathadka, H., Ramirez-Asis, E. H., Loli-Poma, T. P., Kaliyaperumal, K., Ventayen, R. J. M., & Naved, M. (2023). Applications of artificial intelligence in business management, e-commerce and finance. *Materials Today: Proceedings*, 80, 2610-2613.
- Paramesha, M., Rane, N., & Rane, J. (2024). Big data analytics, artificial intelligence, machine learning, internet of things, and blockchain for enhanced business intelligence. *Artificial Intelligence, Machine Learning, Internet of Things, and Blockchain for Enhanced Business Intelligence* (June 6, 2024).
- Perifanis, N. A., & Kitsios, F. (2023). Investigating the influence of artificial intelligence on business value in the digital era of strategy: A literature review. *Information*, 14(2), 85.
- Rana, G., Khang, A., Sharma, R., Goel, A. K., & Dubey, A. K. (Eds.). (2021). *Reinventing manufacturing and business processes through artificial intelligence*. CRC Press.
- Rana, N. P., Chatterjee, S., Dwivedi, Y. K., & Akter, S. (2022). Understanding dark side of artificial intelligence (AI) integrated business analytics: assessing firm's operational inefficiency and competitiveness. *European Journal of Information Systems*, 31(3), 364-387.
- Rane, N. (2023). Role and challenges of ChatGPT and similar generative artificial intelligence in business management. Available at SSRN 4603227.
- Ruiz-Real, J. L., Uribe-Toril, J., Arriaza Torres, J. A., & de Pablo Valenciano, J. (2021). Artificial intelligence in business and economics research: Trends and future. *Business Economics and Management (JBEM)*, 22(1), 98-117.
- Sestino, A., & De Mauro, A. (2022). Leveraging artificial intelligence in business: Implications, applications and methods. *Technology Analysis & Strategic Management*, 34(1), 16-29.
- Sipola, J., Saunila, M., & Ukko, J. (2023). Adopting artificial intelligence in sustainable business. *Journal of Cleaner Production*, 426, 139197.
- Sjödin, D., Parida, V., & Kohtamäki, M. (2023). Artificial intelligence enabling circular business model innovation in digital servitization: Conceptualizing dynamic capabilities, AI

capacities, business models and effects. *Technological Forecasting and Social Change*, 197, 122903.

Sun, W., Ren, S., & Tang, G. (2025). In the era of responsible artificial intelligence and digitalization: business group digitalization, operations and subsidiary performance. *Annals of Operations Research*, 1-23.

Taherdoost, H., & Madanchian, M. (2023). Artificial intelligence and sentiment analysis: A review in competitive research. *Computers*, 12(2), 37.

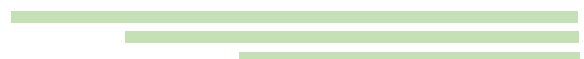
Thakur, R. (2024). Introduction to artificial intelligence and its importance in modern business management. In *Leveraging AI and emotional intelligence in contemporary business organizations* (pp. 133-165). IGI Global Scientific Publishing.

Wang, X., Lin, X., & Shao, B. (2022). How does artificial intelligence create business agility? Evidence from chatbots. *International journal of information management*, 66, 102535.

Wang, Z., Li, M., Lu, J., & Cheng, X. (2022). Business Innovation based on artificial intelligence and Blockchain technology. *Information Processing & Management*, 59(1), 102759.

Xu, J. J., & Babaian, T. (2021). Artificial intelligence in business curriculum: The pedagogy and learning outcomes. *The International Journal of Management Education*, 19(3), 100550.

Yathiraju, N., Raman, P., Madala, R., Patil, P. S., Kumar, A., & Ashwin, S. (2023, April). Research and Innovation to Market Development: Artificial Intelligence in Business. In *2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)* (pp. 1-6). IEEE.



ARTIFICIAL INTELLIGENCE (AI) DALAM BISNIS

Dr. Agus Wibowo, M.Kom, M.Si, MM.

BIO DATA PENULIS



Penulis memiliki berbagai disiplin ilmu yang diperoleh dari Universitas Diponegoro (UNDIP) Semarang. dan dari Universitas Kristen Satya Wacana (UKSW) Salatiga. Disiplin ilmu itu antara lain teknik elektro, komputer, manajemen dan ilmu sosiologi. Penulis memiliki pengalaman kerja pada industri elektronik dan sertifikasi keahlian dalam bidang Jaringan Internet, Telekomunikasi, Artificial Intelligence, Internet Of Things (IoT), Augmented Reality (AR), Technopreneurship, Internet Marketing dan bidang pengolahan dan analisa data (komputer statistik).

Penulis adalah pendiri dari Universitas Sains dan Teknologi Komputer (Universitas STEKOM) dan juga seorang dosen yang memiliki Jabatan Fungsional Akademik Lektor Kepala (Associate Professor) yang telah menghasilkan puluhan Buku Ajar ber ISBN, HAKI dari beberapa karya cipta dan Hak Paten pada produk IPTEK. Sejak tahun 2023 penulis tercatat sebagai Dosen luar biasa di Fakultas Ekonomi & Bisnis (FEB) Universitas Diponegoro Semarang. Penulis juga terlibat dalam berbagai organisasi profesi dan industri yang terkait dengan dunia usaha dan industri, khususnya dalam pengembangan sumber daya manusia yang unggul untuk memenuhi kebutuhan dunia kerja secara nyata.



YAYASAN PRIMA AGUS TEKNIK

PENERBIT :

YAYASAN PRIMA AGUS TEKNIK
Jl. Majapahit No. 605 Semarang
Telp. (024) 6723456. Fax. 024-6710144
Email : penerbit_ypat@stekom.ac.id

ISBN 978-634-7227-31-7 (PDF)

