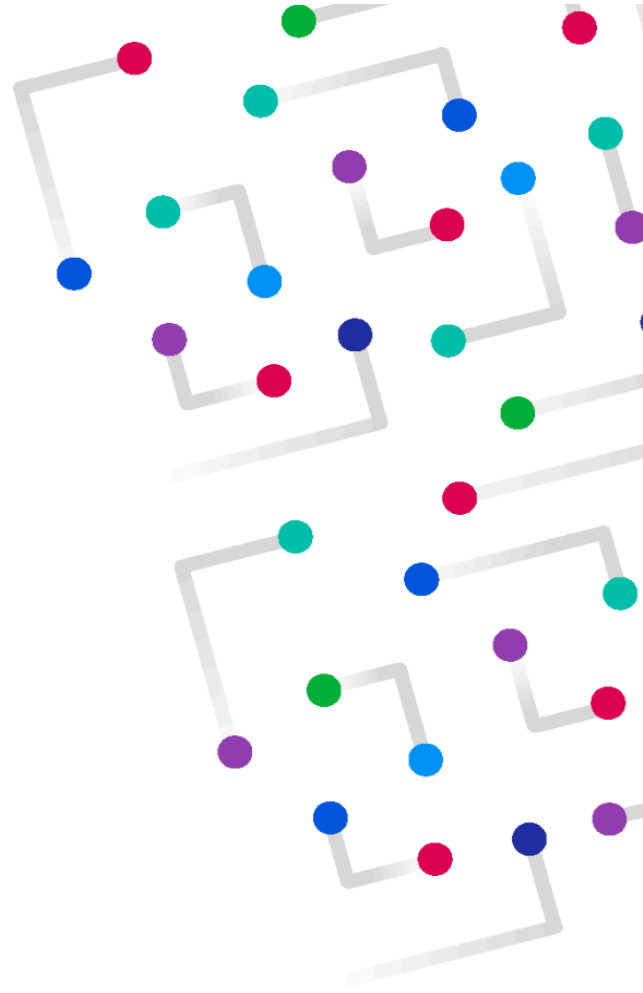




Raspberry Pi
Foundation

Experience AI

Glosari istilah AI



Experience AI – Glosari Istilah

Glosari ini menerangkan istilah utama kecerdasan buatan (AI) dan pembelajaran mesin (ML) yang digunakan dalam [Pelajaran Experience AI](#) dan selanjutnya

[Masuk ke dalam glosari](#)

Kami telah mereka penjelasan ini terutamanya untuk guru dan pendidik, dengan mempertimbangkan penonton muda. Dengan glosari ini, kami berhasrat untuk menyokong anda untuk mengukuhkan pemahaman anda tentang istilah utama ini, serta pengetahuan teknikal anda.

Perbendaharaan kata merupakan bahagian penting dalam pengajaran dan pembelajaran. Menggunakan perbendaharaan kata dengan betul boleh menyokong pelajar untuk mengembangkan pemahaman mereka, manakala penggunaannya yang tidak konsisten boleh membawa kepada konsep alternatif (miskonsepsi) yang boleh mengganggu pembelajaran pelajar. Anda boleh membaca lebih lanjut mengenai ini dalam [Bahan Bacaan Ringkas Pedagogi](#) kami tentang konsep alternatif. Sebagai seorang guru, menggunakan perbendaharaan kata yang tepat dan teknikal secara berkala dan konsisten dapat menyokong pemahaman konsep pelajar.

Kami telah menggunakan teori 'gelombang semantik' untuk membantu kami menulis penjelasan. Setiap penjelasan mengikut struktur tiga bahagian yang sama: bahagian pertama ialah penjelasan yang lebih abstrak tentang istilah, bahagian kedua menjelaskan maksud istilah menggunakan contoh biasa, dan bahagian ketiga mengumpul balik semula apa yang dijelaskan dalam contoh dalam bentuk yang lebih abstrak untuk menyambung semula dengan perbendaharaan kata tersebut. Anda boleh mengetahui lebih lanjut dalam [bahan bacaan ringkas Pedagogi kami tentang gelombang semantik](#).

Ini adalah Versi 1 glosari. Glosari ini akan ditambahkan, disemak dan dikemas kini bersama-sama evolusi Pelajaran Experience AI.

Isi kandungan

[Experience AI — Glosari Istilah](#)

[AI Generatif](#)

[Ambang keyakinan ML](#)

[Berasaskan peraturan](#)

[Bias data](#)

[Bias masyarakat](#)

[Bias](#)

[Ciri ML](#)

[Data latihan ML](#)

[Data ujian ML](#)

[Data](#)

[Dipacu-data](#)

[Kad model ML](#)

[Kebolehjelasan ML](#)

[Kecerdasan buatan](#)

[Kelas ML](#)

[Ketepatan ML](#)

[Keyakinan ML](#)

[Kitaran hayat projek AI](#)

[Latihan ML](#)

[Literasi AI](#)

[ML label](#)

[Model ML](#)

[Nod pepohon keputusan ML](#)

[Pembelajaran diselia](#)

[Pembelajaran mesin](#)

[Pembelajaran pengukuhan](#)

[Pembelajaran tanpa seliaan](#)

[Pembersihan data](#)

[Pengelasan ML](#)

[Penglihatan komputer](#)

[Pepohon keputusan ML](#)

[Ramalan ML](#)

AI Generatif (*Generative AI*)

AI Generatif ialah sejenis **kecerdasan buatan (AI)** yang direka untuk menjana kandungan, seperti teks, imej atau bunyi. Terdapat banyak aplikasi yang menggunakan AI generatif, termasuk penghasilan seni atau muzik, atau menjana teks untuk chatbot. Contohnya, aplikasi seni AI generatif boleh menjana imej berdasarkan gesaan, seperti "hasilkan gambar naga sedang membaca buku". Seni AI Generatif dicipta menggunakan **model pembelajaran mesin yang dilatih** pada jutaan imej seni sedia ada. Imej yang terhasil mungkin meniru gaya artis, tanpa diketahui atau diluluskan oleh artis asal. Aplikasi AI Generatif menjadi semakin biasa dan selalunya anda tidak dapat mengetahui bahawa AI generatif sedang digunakan.

Ambang keyakinan ML (*ML confidence threshold*)

Ambang keyakinan merujuk kepada nilai yang ditetapkan sebagai tahap penerimaan untuk **ramalan model pembelajaran mesin (ML)**. Ambang keyakinan dipilih oleh pembangun ML semasa mereka bentuk model ML. Sebagai contoh, ramalan dijana oleh model ML dengan keyakinan 50% bahawa akan ada ribut salji esok. Walau bagaimanapun, jika ambang keyakinan ditetapkan kepada 60%, ramalan itu akan dianggap tidak tepat. Dalam erti kata lain, melainkan keyakinan ramalan adalah 60% atau lebih tinggi, ia tidak akan diterima sebagai tepat. Nilai ditetapkan mengikut sifat masalah yang sedang diselesaikan, dengan ramalan diagnosis perubatan memerlukan ambang keyakinan yang lebih tinggi daripada cadangan lagu. Pilihan nilai ambang menentukan tahap keyakinan ramalan yang boleh diterima.

Berasaskan peraturan (*Rule-based*)

Berasaskan peraturan ialah cara mereka bentuk sistem menggunakan set peraturan yang telah ditetapkan. Sebagai contoh, program permainan X O (tic-tac-toe) direka menggunakan peraturan tentang apa yang perlu dilakukan untuk cuba memenangi permainan. Peraturan ditakrifkan oleh manusia yang biasanya pakar dalam bidang masalah yang diselesaikan. Sistem **kecerdasan buatan (AI)** yang dibina menggunakan pendekatan berasaskan peraturan juga dikenali sebagai 'AI kuno yang baik'. Sistem berasaskan peraturan berbeza dengan sistem **dipacu data** di mana **data** digunakan sebagai contoh untuk menyelesaikan masalah. Sistem berasaskan peraturan berguna untuk menyelesaikan masalah di mana peraturan yang merangkumi kebanyakan situasi boleh dihasilkan dan diikuti.

Bias data (*Data bias*)

Bias data merujuk kepada **bias** yang ditunjukkan dalam **data** yang digunakan untuk **melatih model pembelajaran mesin (ML)**. Bias data boleh menyebabkan model ML dilatih untuk menjana **ramalan bias**. Sebagai contoh, sesetengah model pengesanan muka adalah bias terhadap wajah dengan warna kulit tertentu, kerana model ML telah dilatih menggunakan kebanyakan imej wajah dengan satu warna kulit. Terdapat beberapa potensi sumber bias data. Ini termasuk data yang tidak lengkap dan data yang mencerminkan **bias masyarakat**. Mengesan bias data adalah penting untuk mengelakkan model ML menjana ramalan bias.

Bias masyarakat (*Societal bias*)

Bias masyarakat merujuk kepada bias yang dipegang oleh sekumpulan besar orang, atau oleh masyarakat pada umumnya. Terdapat pelbagai jenis bias masyarakat, seperti berat sebelah kaum, berat sebelah jantina atau berat sebelah etnik. Contoh bias jantina ialah idea bahawa wanita kurang sesuai untuk kerjaya kejuruteraan berbanding lelaki. **Data** yang dikumpul daripada kumpulan besar orang boleh mencerminkan bias masyarakat, mengakibatkan **bias data**. Jika data yang mencerminkan bias masyarakat digunakan untuk **melatih model pembelajaran mesin (ML)**, ini boleh menyebabkan model menjana **ramalan bias**. Dalam ML, adalah penting untuk mengurangkan bias masyarakat yang ditunjukkan dalam **data latihan**, untuk mengelakkan hasil yang diskriminasi atau tidak adil.

Bias (*Bias*)

Bias merujuk kepada keutamaan kepada atau menentang sesuatu. Sebagai contoh, pelajar mungkin lebih suka pelajaran Bahasa Inggeris daripada pelajaran matematik dan menghabiskan lebih banyak

masa untuk kerja rumah Bahasa Inggeris mereka. Dengan kata lain, mereka mempunyai bias terhadap subjek Bahasa Inggeris. Terdapat banyak jenis bias, termasuk **bias masyarakat** dan **bias data**. Pembangun pembelajaran mesin (ML) perlu memikirkan dengan teliti sama ada **data** yang mereka gunakan untuk melatih model ML adalah bias atau tidak. Sikap berat sebelah boleh mengakibatkan memberi kelebihan kepada seseorang, kumpulan, atau set idea atau kepercayaan berbanding dengan yang lain.

Ciri ML (*ML feature*)

Dalam **pembelajaran mesin (ML)**, ciri mewakili ciri yang dikaitkan dengan **data**. Contohnya, set data muzik mungkin mempunyai ciri seperti tempo, pic, tenaga atau genre. Sesetengah **model ML dilatih** menggunakan ciri untuk mencari persamaan dalam data. Yang lain **meramalkan** ciri baharu dalam data yang tidak dapat diperhatikan dengan mudah oleh orang ramai. Memilih ciri yang hendak digunakan semasa melatih model ML boleh menghasilkan perbezaan kepada keberkesanan model tersebut.

Data latihan ML (*ML training data*)

Dalam **pembelajaran mesin (ML)**, data latihan merujuk kepada contoh dalam bentuk **data** yang digunakan untuk **melatih** model ML. Pembangun ML membina model untuk menyusun corak dalam data latihan, yang boleh digunakan untuk menjana **ramalan** tentang data baharu. Sebagai contoh, pembangun ML membina aplikasi pengecaman pertuturan. Data latihan mungkin termasuk banyak contoh orang bercakap, dalam loghat atau nada suara yang berbeza. Lebih banyak data latihan mewakili realiti, berkemungkinan lebih baik model itu berfungsi.

Data ujian ML (*ML test data*)

Dalam **pembelajaran mesin (ML)**, data ujian merujuk kepada data yang digunakan untuk menguji dan menilai **model ML terlatih**. Sebagai contoh, model ML dilatih untuk meramalkan diagnosis keadaan perubatan. Sebelum digunakan dalam situasi sebenar, model diuji dan dinilai menggunakan data ujian. Data ujian adalah berasingan daripada data latihan model ML dilatih. Data ujian digunakan untuk mengukur prestasi model ML dengan contoh di luar data latihan.

Data (*Data*)

Data merujuk kepada nilai, ukuran, fakta atau pemerhatian dalam bentuk yang sesuai untuk diproses oleh program komputer. Terdapat banyak jenis data, seperti teks, imej atau bunyi. Contoh data teks ialah mesej orang menghantar kepada rakan mereka dengan peranti digital. Dalam pembelajaran mesin (ML), data mewakili contoh model ML dilatih. Mengumpul, membersihkan dan menstrukturkan sejumlah besar data adalah bahagian penting dalam mereka bentuk model ML.

Dipacu-data (*Data-driven*)

Dipacu data ialah cara mereka bentuk sistem menggunakan **data** dan bukannya arahan langkah demi langkah. Sebagai contoh, ianya sukar untuk mengetahui punca penyakit tertentu, tetapi terdapat banyak contoh data. Oleh itu, pereka bentuk menggunakan data perubatan orang menghidap penyakit tersebut untuk mendiagnosisnya. Sistem dipacu data berbeza dengan sistem **berasaskan peraturan**. Sistem dipacu data sesuai untuk menyelesaikan masalah di mana peraturan yang merangkumi setiap situasi sukar dihasilkan. Sebaliknya, contoh yang mencukupi boleh dikumpulkan untuk memaklumkan penyelesaian.

Kad model ML (*ML model card*)

Kad model **pembelajaran mesin (ML)** ialah cara mendokumentasikan maklumat penting tentang **model** ML secara berstruktur. Kad model ML ditulis oleh pembangun ML untuk kedua-dua pengguna pakar dan bukan pakar. Sebagai contoh, aplikasi ML dibangunkan untuk menterjemah bahasa yang berbeza, seperti dari Arab ke Perancis dan sebaliknya. Kad model termasuk maklumat tentang ketepatan terjemahan model, serta prestasi model di sekitar jargon, slanga dan dialek. Maklumat kad model lain mungkin termasuk jenis model ML, penunjuk prestasi yang berbeza dan juga **bias** yang diketahui. Kad model dicipta semasa peringkat penjelasan kitaran hayat projek AI untuk mendedahkan maklumat tentang keupayaan dan had model, dengan cara yang mudah difahami.

Kebolehjelasan ML (*ML explainability*)

Kebolehjelasan merujuk kepada sejauh mana sesuatu boleh difahami. Dalam **pembelajaran mesin (ML)**, kebolehjelasan membantu orang memahami cara **ramalan** dihasilkan. Sebagai contoh, **model pepohon keputusan** ML boleh dijelaskan kerana **nod** boleh dianalisis dengan cara yang boleh difahami oleh orang ramai. Kebanyakan model ML tidak dapat dijelaskan sepenuhnya, dan ada yang lebih boleh dijelaskan daripada yang lain. Meningkatkan kebolehjelasan model boleh membantu menyelesaikan masalah dan melawan bias.

Kecerdasan buatan (*Artificial intelligence*)

Kecerdasan buatan (AI) ialah reka bentuk dan kajian sistem yang kelihatan meniru tingkah laku pintar. Sesetengah aplikasi AI adalah berdasarkan peraturan. Lebih kerap sekarang, aplikasi AI dibina menggunakan **pembelajaran mesin** yang dikatakan 'belajar' daripada contoh dalam bentuk **data**. Sebagai contoh, beberapa aplikasi AI dibina untuk menjawab soalan atau membantu mendiagnosis penyakit. Aplikasi AI lain boleh dibina untuk tujuan berbahaya, seperti menyebarkan berita palsu. Aplikasi AI tidak berfikir. Aplikasi AI dibina untuk menjalankan tugas dengan cara yang kelihatan pintar.

Kelas ML (*ML class*)

Untuk **melatih model pengelasan**, pembangun pembelajaran mesin (ML) menyusun **data** ke dalam kumpulan yang dipratentukan yang dipanggil kelas. Kelas ditakrifkan lebih awal berdasarkan perkara yang orang dapati berguna untuk mengumpulkan perkara. Bayangkan aplikasi ML yang direka untuk mengenal pasti buah-buahan di pasar raya. Data mungkin disusun dalam kelas epal, pisang, oren, beri biru, dsb. Kelas ialah sekumpulan perkara yang digunakan model pengelasan untuk mengenal pasti persamaan dalam data.

Ketepatan ML (*ML accuracy*)

Ketepatan merujuk kepada sejauh mana sesuatu itu betul. Dalam **pembelajaran mesin (ML)**, ketepatan ialah cara untuk mengukur kekerapan **model** ML membuat **ramalan** yang betul. Sebagai contoh, model **pengelasan** direka untuk mengelaskan epal. Daripada 100 imej epal, 90 dikelaskan dengan betul. Ketepatan pengelasan model ialah 90%. Ketepatan ialah satu cara untuk menilai model ML. Selalunya, ketepatan digunakan bersama-sama dengan langkah-langkah lain untuk menilai kualiti model.

Keyakinan ML (*ML confidence*)

Keyakinan merujuk kepada kepastian sesuatu. Dalam **pembelajaran mesin (ML)**, keyakinan adalah cara untuk mengukur kepastian sesuatu **ramalan**. Sebagai contoh, **model pengelasan** direka untuk meramalkan sama ada hujan akan turun esok. Model itu meramalkan dengan keyakinan 90% bahawa akan ada hujan esok. Dengan kata lain, terdapat 90% kepastian bahawa esok akan hujan. Menggunakan keyakinan untuk mengukur kepastian ramalan membantu menilai kualiti model ML.

Kitaran hayat projek AI (*AI project lifecycle*)

Kitaran hayat projek **kecerdasan buatan (AI)** merujuk kepada langkah berbeza yang mungkin diambil untuk mereka bentuk dan **membina model pembelajaran mesin (ML)**. Langkah-langkah tersebut termasuk mentakrifkan masalah, menyediakan **data**, **melatih** model, menguji model, menilai model, dan menerangkan model. Contohnya, model ML direka untuk menjana senarai lagu baharu. Pembangun ML mungkin terlebih dahulu mempertimbangkan jenis senarai lagu yang mereka ingin hasilkan. Seterusnya, mereka mungkin mengumpul dan menyediakan data lagu. Model ML dilatih dan diuji dengan data lagu. Model ML dinilai untuk melihat sama ada ia berfungsi seperti yang diharapkan. Akhir sekali, model ML dijelaskan supaya orang lain boleh menggunakannya. Biasanya, langkah kitaran hayat projek AI digunakan secara berulang dan bukannya satu demi satu. Kitaran hayat projek AI ialah satu siri langkah berulang yang digunakan untuk membina dan menambah baik model ML.

Label ML (*ML label*)

Dalam **pembelajaran diselia**, **model pembelajaran mesin (ML)** dilatih menggunakan data **berlabel**. Setiap bahagian data dicatatkan dengan satu atau lebih label yang memberikan maklumat tentang data tersebut. Sebagai contoh, model ML direka untuk mengenal pasti bunyi burung. Setiap bunyi dilabelkan dengan nama burung yang mengeluarkan bunyi tersebut. Model ML dilatih dengan bunyi berlabel dan boleh **meramalkan** label (nama burung) bunyi baharu. Data paling kerap dilabelkan oleh orang untuk memberikan contoh yang tepat untuk melatih model ML.

Latihan ML (*ML training*)

Model pembelajaran mesin (ML) dilatih menggunakan contoh dalam bentuk **data** untuk mencari corak dan membuat **ramalan**. Semasa latihan, corak diperhalusi untuk menambah baik ramalan. Sebagai contoh, pembangun ML mungkin membina model untuk mengesyorkan lagu. Model ML akan dilatih tentang pilihan lagu ramai orang untuk mencari persamaan antara perkara yang orang berbeza suka mendengar. Lebih banyak pilihan lagu yang pelbagai model dilatih, lebih baik ramalan lagu yang disyorkan mungkin. Terdapat banyak cara yang berbeza untuk melatih model ML, menggunakan jenis data yang berbeza. Seorang pembangun akan memilih antara jenis latihan yang tersedia bergantung pada masalah yang mereka cuba selesaikan dan data yang tersedia untuk menyelesaikannya. Kualiti latihan sebahagian besarnya bergantung kepada kualiti data yang digunakan.

Literasi AI (*AI literacy*)

Literasi AI ialah satu set kecekapan dan cara berfikir yang membolehkan orang ramai terlibat secara bermakna dengan aplikasi **kecerdasan buatan (AI)**, serta dalam situasi di mana aplikasi AI digunakan di sekeliling mereka. Kecekapan tersebut termasuk memahami AI, terlibat dalam membangunkan sistem AI, dan mempunyai pendapat yang dimaklumkan tentang cara sistem AI digunakan di dunia.

Contohnya, seseorang menggunakan literasi AI apabila mereka menilai ketepatan maklumat yang diberikan oleh aplikasi chatbot AI. Selain cara yang berbeza untuk melibatkan diri dengan aplikasi AI, literasi AI termasuk potensi orang ramai untuk mengambil bahagian secara aktif dalam menentukan cara sistem AI boleh digunakan di sekeliling mereka.

Model ML (*ML model*)

Model **pembelajaran mesin (ML)** digunakan oleh aplikasi ML untuk menyelesaikan tugas atau menyelesaikan masalah. Model ML ialah perwakilan masalah yang sedang diselesaikan. Pembangun ML menggunakan sejumlah besar data yang mewakili masalah tertentu untuk melatih model untuk mengesan corak. Hasil latihan ialah model, yang digunakan untuk membuat ramalan tentang data baharu dalam konteks yang sama. Sebagai contoh, kereta pandu sendiri dibina menggunakan model ML untuk meramalkan masa untuk berhenti. Model dilatih menggunakan berjuta-juta contoh situasi di mana kereta perlu berhenti. Terdapat pelbagai jenis model, menggunakan pelbagai jenis data, dan cara yang berbeza untuk melatih model. Semua model ML dilatih untuk mengesan corak dalam **data latihan** untuk membuat ramalan tentang data baharu.

Nod pepohon keputusan ML (*ML decision tree node*)

Pepohon keputusan pembelajaran mesin (ML) terdiri daripada nod. Nod disambungkan untuk membentuk struktur berdasarkan **ramalan** yang boleh dijana. Terdapat dua jenis nod: nod keputusan dan nod daun. Sebagai contoh, pertimbangkan pokok keputusan yang dibina untuk meramalkan jenis bintang dalam sistem suria kita. Nod keputusan mewakili **ciri data** seperti suhu, jejari, warna atau kecerahan bintang. Nod daun mewakili jenis bintang dalam bentuk **label** ramalan, seperti 'Red Dwarf', 'White Dwarf' atau 'Brown Dwarf'. Nod pokok keputusan membentuk struktur yang diperlukan untuk **model** ML untuk menjana ramalan.

Pembelajaran diselia (*Supervised learning*)

Pembelajaran diselia ialah satu pendekatan yang digunakan untuk **melatih model pembelajaran mesin (ML)**. Pendekatan pembelajaran yang diselia menggunakan sejumlah besar **data yang dilabelkan** oleh orang yang mempunyai maklumat yang berkaitan. Satu jenis pembelajaran diselia adalah pengelasan. Contoh masalah pengelasan ialah mengenal pasti harimau di alam liar. Data tersebut terdiri daripada banyak imej, dengan imej yang mengandungi harimau dilabelkan sedemikian. Model ML dilatih dengan imej berlabel dan **meramalkan** jika terdapat harimau dalam imej tersebut. Mempunyai imej yang dilabel dengan betul membolehkan pembangun mengetahui sejauh mana ramalan model adalah **tepat** dan menyesuaikan latihan model. Berikutan itu, model ML boleh digunakan untuk meramalkan jika terdapat harimau dalam imej yang baharu. Pendekatan pembelajaran yang diselia bergantung pada mempunyai cukup data yang dilabelkan dengan betul untuk menghasilkan ramalan yang tepat.

Pembelajaran mesin (*Machine learning*)

Pembelajaran mesin (ML) ialah pendekatan yang digunakan untuk mereka bentuk dan membina sistem **kecerdasan buatan (AI)**. ML dikatakan 'belajar' dengan menggunakan contoh dalam bentuk data, bukannya melaksanakan arahan langkah demi langkah. Dalam erti kata lain, aplikasi ML adalah **dipacu data**. Sebagai contoh, aplikasi ML digunakan untuk mengenali pertuturan. Ia berdasarkan banyak contoh orang bercakap dalam loghat dan nada suara yang berbeza. Aplikasi ML lain termasuk mengenal pasti objek dalam imej atau bermain permainan yang kompleks. Setiap aplikasi ML direka untuk menyelesaikan masalah tertentu.

Pembelajaran pengukuhan (*Reinforcement learning*)

Pembelajaran pengukuhan ialah satu pendekatan yang digunakan untuk **melatih model pembelajaran mesin (ML)**. Pendekatan ini digunakan untuk menyelesaikan masalah dengan matlamat yang jelas, di mana ganjaran dan penalti digunakan untuk mencapai matlamat tersebut. Pendekatan pembelajaran pengukuhan digunakan dalam reka bentuk kereta pandu sendiri atau untuk bermain permainan yang rumit. Sebagai contoh, model pembelajaran pengukuhan mungkin digunakan untuk mereka bentuk aplikasi untuk bermain catur. Model ini dilatih untuk **meramalkan** langkah yang memaksimumkan ganjaran dan meminimumkan denda ke arah kemenangan. Pendekatan pembelajaran pengukuhan menggunakan ganjaran dan denda untuk mengenal pasti strategi mencapai matlamat yang ditetapkan.

Pembelajaran tanpa seliaan (*Unsupervised learning*)

Pembelajaran tanpa seliaan ialah satu pendekatan yang digunakan untuk **melatih model pembelajaran mesin (ML)**. Pembangun ML melatih model pembelajaran tanpa seliaan untuk menyusun **data** berdasarkan persamaan. Proses ini menyebabkan mencari corak tersembunyi dalam data. Satu jenis pembelajaran tanpa pengawasan adalah pengelompokan. Contoh masalah pengelompokan ialah **meramalkan** cara data kesihatan boleh dikumpulkan untuk membantu mendiagnosis penyakit. Kumpulan ini dipanggil kelompok, yang tidak diketahui terlebih dahulu. Model ML boleh digunakan untuk meramalkan jika data kesihatan baharu termasuk dalam salah satu kelompok. Pendekatan pembelajaran tanpa seliaan berguna dalam menyelesaikan masalah di mana orang mungkin tidak tahu apa yang perlu dicari.

Pembersihan data (*Data cleaning*)

Pembersihan data ialah satu langkah dalam menyediakan **data** yang digunakan untuk **melatih model pembelajaran mesin (ML)**. Pembersihan data melibatkan mengenal pasti dan membetulkan ralat dalam data. Contohnya, membetulkan ralat menaip atau mengalih keluar pendua dalam data teks adalah dua tugas pembersihan data yang mudah. Selalunya, data tidak kemas memerlukan pembersihan yang lebih kompleks sebelum digunakan untuk melatih model ML. Terdapat banyak cara untuk membersihkan data bergantung pada masalah dan jenis data. Menggunakan data bersih adalah penting untuk membina model ML yang **tepat**.

Pengelasan ML (*ML classification*)

Pengelasan merujuk kepada tugas mengagihkan perkara ke dalam kumpulan yang telah ditetapkan, dipanggil **kelas**. Kelas ditakrifkan lebih awal berdasarkan perkara yang orang dapati berguna untuk mengumpulkan perkara. Contoh masalah pengelasan ialah analisis sentimen ulasan lagu. **Model pengelasan pembelajaran mesin (ML) dilatih** dengan ulasan yang **dilabelkan** oleh orang sebagai 'positif' atau 'negatif'. Selepas latihan, model ML boleh digunakan untuk **meramalkan** sama ada semakan baharu harus dikelaskan sebagai sama ada 'positif' atau 'negatif'. Model pengelasan meramalkan satu atau lebih label kelas. Pendekatan pengelasan berguna untuk menyelesaikan masalah yang jawapannya termasuk dalam kumpulan yang telah ditetapkan.

Penglihatan komputer (*Computer vision*)

Penglihatan komputer ialah kajian sistem yang direka untuk memproses maklumat daripada imej atau video digital. Contoh aplikasi penglihatan komputer termasuk pengesanan muka, pengimejan perubatan dan pengawasan video. Sebagai contoh, penglihatan komputer digunakan dalam reka bentuk kereta pandu sendiri untuk mengesan dan mengelakkan terlanggar objek. Selalunya, sistem penglihatan komputer menggunakan model pembelajaran mesin (ML) untuk mengenal pasti corak dalam data imej dan video. Sistem penglihatan komputer berguna di mana maklumat daripada imej atau video digital boleh digunakan untuk menyelesaikan masalah.

Pepohon keputusan ML (*ML decision tree*)

Pepohon keputusan **pembelajaran mesin (ML)** ialah satu jenis **model ML**. Pembangun ML menggunakan pepohon keputusan untuk menstruktur satu set keadaan berdasarkan **ramalan** yang boleh dibuat. Syarat diperoleh daripada **ciri** dalam **data**. Sebagai contoh, pepohon keputusan mungkin digunakan untuk membina sistem pengesyoran filem. Model pokok keputusan **dilatih** menggunakan pilihan filem ramai orang. Semasa latihan, keadaan dijana berdasarkan ciri seperti jenis filem, panjang masa atau pelakon utama. Model ML menjana ramalan filem mana yang mungkin seseorang ingin tonton seterusnya, berdasarkan cara pilihan mereka mengikut syarat dalam model. Struktur pepohon keputusan ML dijana berdasarkan sejumlah besar data dan mungkin berubah jika dilatih semula dengan data yang berbeza.

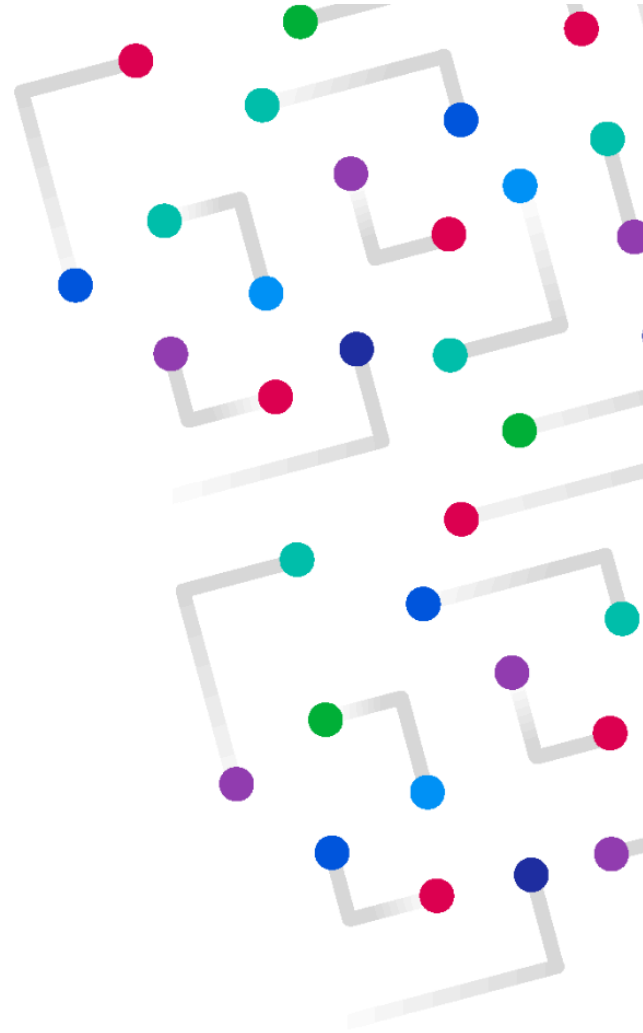
Ramalan ML (*ML prediction*)

Model pembelajaran mesin (ML) dilatih untuk membuat ramalan. Ramalan yang dihasilkan oleh model ML mencadangkan perkara yang diwakili oleh **data**, atau perkara yang mungkin berguna untuk tugas. Sebagai contoh, pembangun ML mungkin melatih model untuk meramalkan filem yang seseorang mungkin mahu tonton seterusnya, berdasarkan tabiat tontonan mereka. Model itu akan menjana ramalan selepas dilatih mengenai pilihan filem ramai orang. Tugas utama model ML ialah membuat ramalan. Semua model ML membuat ramalan, walaupun dalam beberapa kes ramalan ini tidak nyata bagi pengguna.



This resource is licensed by the [Raspberry Pi Foundation](https://www.raspberrypi.org/) under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License (CC BY-NC-ND 4.0). For more information on this licence, see creativecommons.org/licenses/by-nc-nd/4.0/.

Experience AI



Glosari ini telah dicipta oleh Raspberry Pi Foundation sebagai sebahagian daripada program pendidikan kami, Experience AI. Ketahui lebih lanjut di www.experience-ai.org

This resource is licensed by the Raspberry Pi Foundation under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License (CC BY-NC-ND 4.0). For more information on this licence, see creativecommons.org/licenses/by-nc-nd/4.0.