

# **Artificial Intelligence & Machine Learning**

by Thio Perdana

## **Outline**

*Pendahuluan*

*Artificial Intelligence (AI)*

*Machine Learning (ML)*

*Pengenalan Machine Learning*

*Supervised Learning*

*Unsupervised Learning*

*Dataset*

*Deep Learning (DL)*

*Pengenalan Deep Learning*

*Perbedaan dengan Machine Learning Konvensional*

*Metode Deep Learning*

***Convolutional Neural Networks (CNN)***

***Recurrent Neural Networks (RNN)***

***Etika dalam AI***

---

---

## ***Pendahuluan***

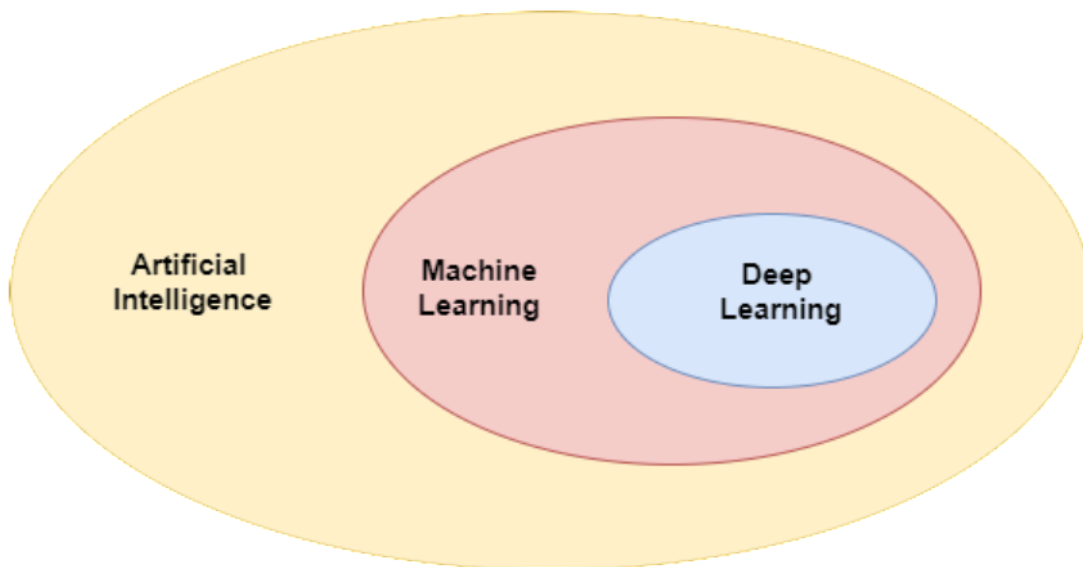
Artificial Intelligence (AI), Machine Learning (ML), dan Deep Learning (DL) merupakan konsep-konsep terkait dalam pengembangan teknologi kecerdasan buatan. AI merujuk pada upaya menyematkan kecerdasan manusia ke dalam mesin atau program komputer. Ini mencakup pengembangan algoritma untuk meniru kemampuan manusia seperti pemecahan masalah dan pengambilan keputusan. Di dalam AI, ML adalah subset yang fokus pada pengembangan sistem yang dapat belajar dari data tanpa pemrograman eksplisit. Algoritma ML menggunakan pendekatan berbasis data, memungkinkan sistem untuk meningkatkan kinerjanya seiring waktu. Selanjutnya, DL adalah sub-bidang khusus dalam ML yang menggunakan jaringan saraf tiruan dengan banyak lapisan (deep neural networks). DL meniru struktur otak manusia untuk mengenali pola kompleks secara hierarkis. Secara singkat, AI mencakup semua upaya kecerdasan buatan, ML adalah bagian dari AI yang berfokus pada pembelajaran dari data, dan DL merupakan teknik khusus dalam ML yang menggunakan jaringan saraf dalam skala besar.

Secara singkat:

**AI** adalah konsep umum yang mencakup semua upaya untuk membuat komputer melakukan tugas yang memerlukan kecerdasan manusia.

**ML** adalah subset dari AI yang fokus pada pembuatan sistem yang dapat belajar dari data. Semua ML adalah bagian dari AI, tetapi tidak semua AI adalah ML.

**DL** adalah subset dari ML yang menggunakan jaringan saraf tiruan dengan banyak lapisan (deep neural networks). Semua DL adalah ML, tetapi tidak semua ML adalah DL.



---

---

# *Artificial Intelligence (AI)*

Kecerdasan Buatan (Artificial Intelligence atau AI) merujuk pada kecerdasan yang dimiliki oleh mesin atau perangkat lunak, berbeda dengan kecerdasan yang dimiliki oleh manusia atau hewan. Ini merupakan bidang studi dalam ilmu komputer yang berkembang dan mengkaji mesin-mesin cerdas. Istilah ini juga dapat merujuk pada mesin-mesin cerdas tersebut.

Teknologi AI digunakan secara luas di berbagai industri, pemerintahan, dan ilmu pengetahuan. Beberapa aplikasi yang terkenal mencakup mesin pencari web canggih (seperti Google Search), sistem rekomendasi (digunakan oleh YouTube, Amazon, dan Netflix), pemahaman terhadap ucapan manusia (seperti Google Assistant, Siri, dan Alexa), mobil otonom (seperti Waymo), alat generatif dan kreatif (seperti ChatGPT dan seni AI), serta permainan strategi dengan analisis di atas kemampuan manusia (seperti catur dan Go).

Kecerdasan Buatan didirikan sebagai disiplin akademis pada tahun 1956. Bidang ini mengalami siklus optimisme yang diikuti oleh kekecewaan dan kekurangan pendanaan, namun setelah tahun 2012, ketika deep learning melampaui semua teknik AI sebelumnya, terjadi peningkatan dana dan minat yang besar.

Berbagai sub-bidang riset AI berfokus pada tujuan-tujuan tertentu dan penggunaan alat-alat khusus. Tujuan tradisional riset AI mencakup pemikiran, representasi pengetahuan, perencanaan, pembelajaran, pemrosesan bahasa alami, persepsi, dan dukungan untuk robotika. Intelijensi umum (kemampuan untuk menyelesaikan tugas apa pun yang dapat dilakukan oleh manusia) adalah salah satu tujuan jangka panjang bidang ini. Untuk menyelesaikan masalah-masalah ini, peneliti AI telah menyesuaikan dan mengintegrasikan berbagai teknik pemecahan masalah, termasuk pencarian dan optimisasi matematis,

logika formal, jaringan saraf buatan, dan metode berdasarkan statistik, riset operasi, dan ekonomi. AI juga bersandar pada psikologi, linguistik, filsafat, neurosains, dan banyak bidang lainnya.

**Tujuan:** Masalah umum mensimulasikan (atau menciptakan) kecerdasan telah dibagi menjadi sub-masalah. Ini mencakup ciri atau kemampuan tertentu yang diharapkan oleh peneliti dari suatu sistem cerdas. Ciri-ciri yang dijelaskan di bawah ini mendapatkan perhatian paling banyak dan mencakup cakupan riset AI.

**Pemikiran, Pemecahan Masalah:** Peneliti AI awal mengembangkan algoritma yang meniru pemikiran langkah-demi-langkah yang digunakan manusia saat menyelesaikan teka-teki atau membuat deduksi logis. Pada akhir 1980-an dan 1990-an, metode dikembangkan untuk mengatasi informasi yang tidak pasti atau tidak lengkap, dengan menggunakan konsep-konsep dari probabilitas dan ekonomi.

Banyak dari algoritma-algoritma ini tidak mencukupi untuk menyelesaikan masalah-masalah pemikiran besar karena mengalami "ledakan kombinatorial": mereka menjadi eksponensial lebih lambat seiring dengan pertumbuhan masalah. Bahkan manusia jarang menggunakan deduksi langkah-demi-langkah yang dapat dimodelkan oleh penelitian AI awal. Mereka menyelesaikan sebagian besar masalah mereka menggunakan penilaian intuitif yang cepat. Pemikiran yang akurat dan efisien masih menjadi masalah yang belum terselesaikan.

**Representasi Pengetahuan:** Representasi pengetahuan dan rekayasa pengetahuan memungkinkan program AI untuk menjawab pertanyaan secara cerdas dan membuat deduksi tentang fakta-fakta dunia nyata. Representasi pengetahuan formal digunakan dalam indeks dan pengambilan berbasis konten, interpretasi adegan, dukungan keputusan klinis, penemuan pengetahuan (penambangan inferensi "menarik" dan dapat dilakukan dari basis data besar), dan bidang-bidang lainnya.

Basis pengetahuan adalah kumpulan pengetahuan yang direpresentasikan dalam bentuk yang dapat digunakan oleh program. Ontologi adalah kumpulan objek, relasi, konsep, dan properti yang digunakan oleh suatu domain pengetahuan tertentu. Basis pengetahuan perlu merepresentasikan hal-hal seperti:

objek, properti, kategori, dan hubungan antar objek; situasi, peristiwa, keadaan dan waktu; penyebab dan akibat; pengetahuan tentang pengetahuan (apa yang kita ketahui tentang apa yang orang lain ketahui); penalaran default (hal-hal yang manusia anggap benar sampai diberi tahu sebaliknya dan akan tetap benar bahkan ketika fakta lain berubah); dan banyak aspek dan domain pengetahuan lainnya.

**Perencanaan dan Pengambilan Keputusan:** Sebuah "agen" adalah apapun yang merasakan dan mengambil tindakan di dunia. Agen rasional memiliki tujuan atau preferensi dan mengambil tindakan untuk mewujudkannya. Dalam perencanaan otomatis, agen memiliki tujuan tertentu. Dalam pengambilan keputusan otomatis, agen memiliki preferensi - ada beberapa situasi yang ingin dia berada di dalamnya, dan ada beberapa situasi yang ingin dia hindari. Agen pengambil keputusan memberikan nomor ke setiap situasi (disebut "utilitas") yang mengukur seberapa banyak agen menyukainya. Untuk setiap tindakan yang mungkin, ia dapat menghitung "utilitas yang diharapkan": utilitas dari semua hasil yang mungkin dari tindakan tersebut, diboboti oleh probabilitas bahwa hasil tersebut akan terjadi. Kemudian ia dapat memilih tindakan dengan utilitas yang diharapkan maksimum.

**Pembelajaran:** Machine learning adalah studi tentang program-program yang dapat meningkatkan kinerjanya dalam suatu tugas tertentu secara otomatis. Ini telah menjadi bagian dari AI sejak awal.

Ada beberapa jenis machine learning. Unsupervised learning menganalisis aliran data dan menemukan pola serta membuat prediksi tanpa bimbingan lain. Supervised learning memerlukan manusia untuk memberi label data masukan terlebih dahulu dan hadir dalam dua varietas utama: klasifikasi (di mana program harus belajar memprediksi kategori mana masukan masuk) dan regresi (di mana program harus menyusun fungsi numerik berdasarkan masukan numerik). Dalam pembelajaran penguatan, agen diberi hadiah untuk respon yang baik dan dihukum untuk yang buruk. Agen belajar memilih respon yang diklasifikasikan sebagai "baik". Transfer learning adalah ketika pengetahuan yang didapat dari satu masalah diterapkan pada masalah baru. Deep learning adalah jenis machine learning yang menjalankan input melalui jaringan saraf buatan yang terinspirasi oleh biologi untuk semua jenis pembelajaran ini.

Teori pembelajaran komputasi dapat menilai pembelajar dengan kompleksitas komputasi, kompleksitas sampel (berapa banyak data yang diperlukan), atau dengan gagasan.

## Sejarah Perkembangan AI

**1642:** Mesin Penjumlah Mekanis oleh Blaise Pascal.

**1801:** Penggunaan Kartu Logam dalam Alat Tenun oleh Joseph Marie Jacquard.

**1847:** Pengembangan Logika Boolean oleh George Boole.

**1890:** Mesin Penghitung Statistik oleh Herman Hollerith.

**1937-1943:** Pengembangan Mesin Punch Card-Reading oleh IBM.

**1950-an:** Perkembangan Awal AI, Neural Network, dan Algoritma Nearest Neighbor.

**1952-Awal 90-an:** NSA membangun pusat data pertama.

**1960:** Teori Kontrol Banyak, Artificial Neural Network, dan Reinforcement Learning.

**1963:** Reinforcement Learning - Donald Michie menciptakan mesin yang dapat mengimplementasikan reinforcement learning.

**1967:** Algoritma Nearest Neighbor ditulis.

**1980-an:** Kenaikan dan Keruntuhan Minat Bisnis AI.

**1987:** AI Winter - Pasar perangkat keras AI spesialis runtuh setelah komputer desktop menjadi lebih powerful.

**1995:** Pengenalan Support Vector Machine (SVM).

**2000-an:** Kecenderungan Penghindaran Istilah "AI."

**2005-2009:** Istilah Big Data diperkenalkan oleh O'Reilly Media.

**2006:** Geoffrey Hinton memperkenalkan deep belief nets dan era Deep Learning.

**2010-an:**

**2011:** Sukses dalam DARPA Grand Challenge dan kemenangan Watson di Jeopardy!

**2020-an:** Munculnya Large Language Model (LLM) seperti GPT-3.

Dengan begitu, kita melihat perkembangan AI dan Machine Learning dari mesin mekanis pada abad ke-17 hingga kemunculan teknologi canggih seperti Large Language Models pada abad ke-21, dengan tantangan dan masa sulit di antaranya.

---

---

## ***Machine Learning (ML)***



## ***Pengenalan Machine Learning***

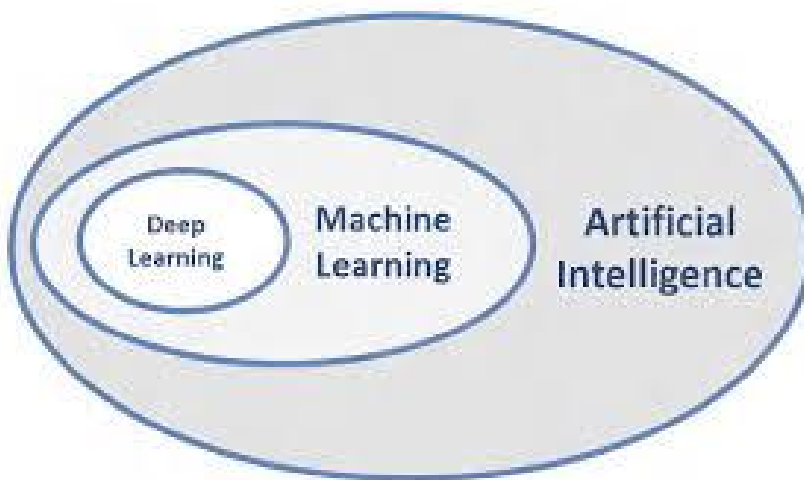
Machine Learning adalah ilmu yang mempelajari tentang algoritma komputer yang bisa mengenali pola-pola di dalam data, dengan tujuan untuk mengubah beragam macam data menjadi suatu tindakan nyata dengan sesedikit mungkin campur tangan manusia. Sederhananya, dengan campur tangan manusia seminimal mungkin, machine learning dapat belajar bagaimana suatu data diolah dari data itu sendiri.

Mengapa kita harus belajar *machine learning*? *machine learning* akan mempermudah kita dalam mengolah banyak data atau suatu proses yang membutuhkan algoritma yang kompleks. Menariknya adalah, ketika data baru yang memiliki pola baru muncul, maka suatu model dalam machine learning akan belajar bagaimana untuk mengenali pola baru tersebut.

Machine Learning sendiri adalah cabang dari *artificial intelligence*. Apa itu artificial ? menurut John McCarthy (2014)

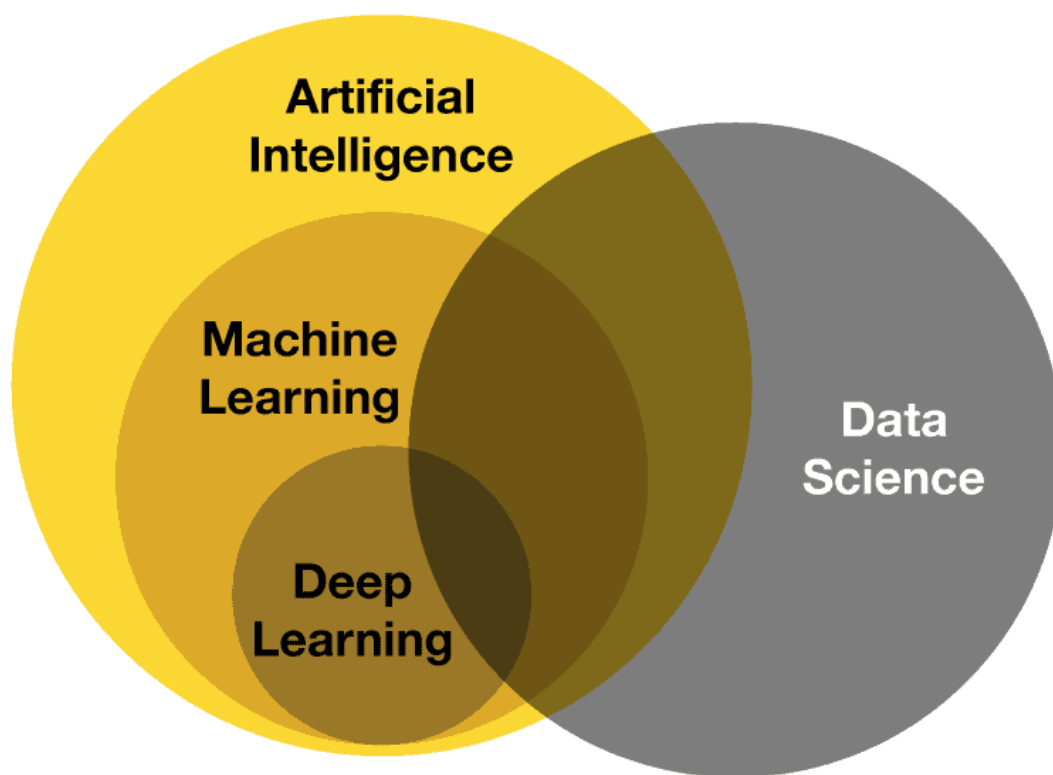
*It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.*

Jadi AI sendiri adalah usaha untuk membuat sebuah mesin dapat menyerupai bagaimana manusia berpikir, sedangkan machine learning adalah upaya agar mesin tersebut dapat belajar secara mandiri seperti manusia. Secara garis besar, diagram hubungan antara machine learning dan AI dapat ditunjukkan pada diagram di bawah.



sumber:human-centered.ai

bagaimana hubungan antara *machine learning* dan *data science*?

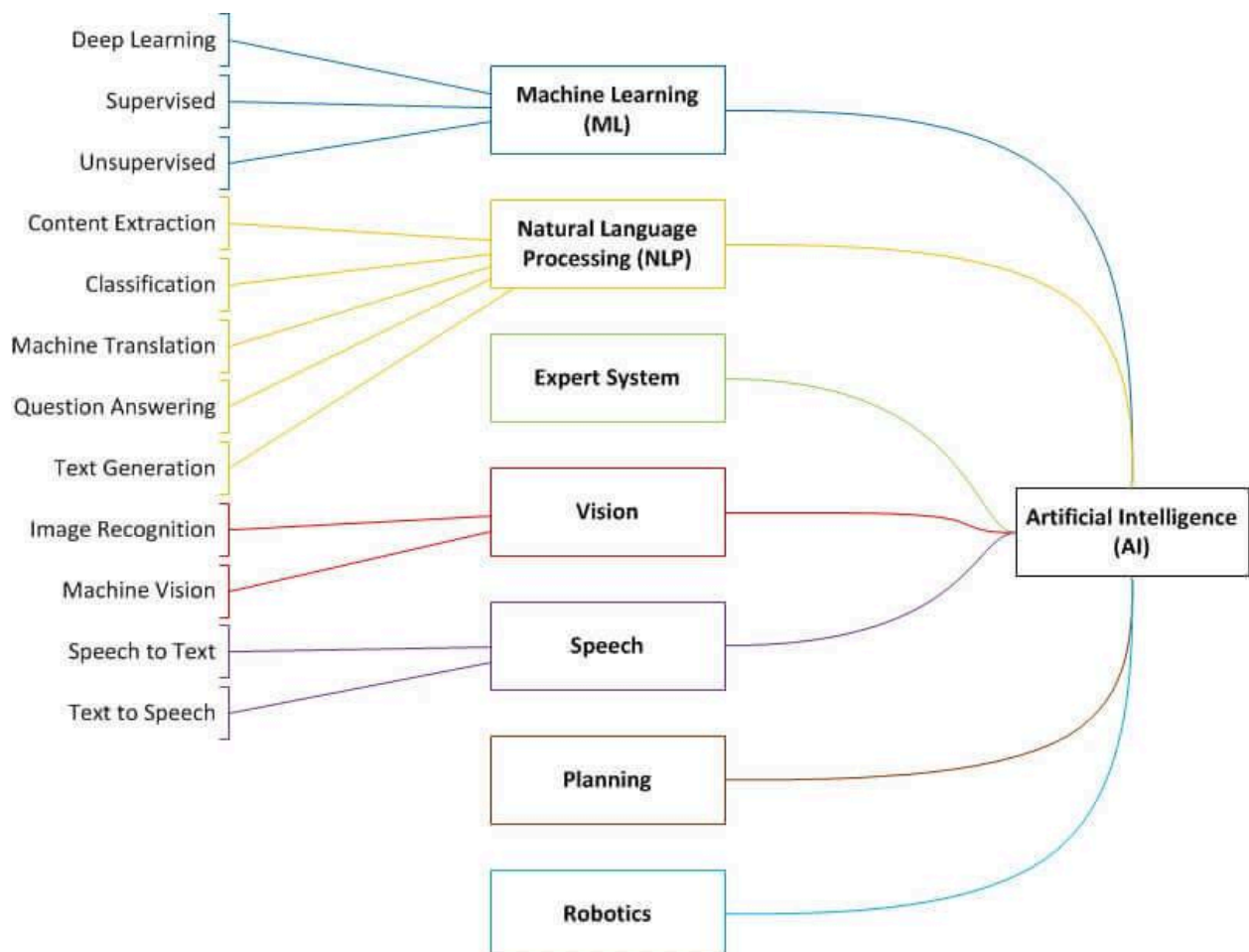


<https://deviq.io/>

Jadi apa yang dikerjakan oleh seorang data scientist itu sendiri belum tentu harus berhubungan dengan machine learning. Akan tetapi, dengan menggunakan machine learning maka beberapa pekerjaan data

scientist akan dipermudah. Ini adalah konsep yang harus diingat, *Data Scientist* tidaklah harus selalu bermain dengan machine learning.

Misalkan, untuk menarik kesimpulan tentang pencapaian akademik siswa SMA di suatu kelas maka menggunakan statistika sederhana sudah cukup. Beberapa cabang dari AI selain machine learning bisa dilihat pada bagai di bawah



sumber : dicoding.com

## Hulu-Hilir Alur Kerja Machine Learning

Saat bekerja dengan proyek *machine learning*, maka terdapat alur kerja yang dapat kita ikuti. Dengan berbasiskan pada *Data Science Methodology*, maka terdapat 6 langkah utama saat kita mengerjakan

project machine learning. Perbedaannya atau detail tambahan ada pada langkah *data processing*, yaitu modeling dan evaluation.

### **Modeling**

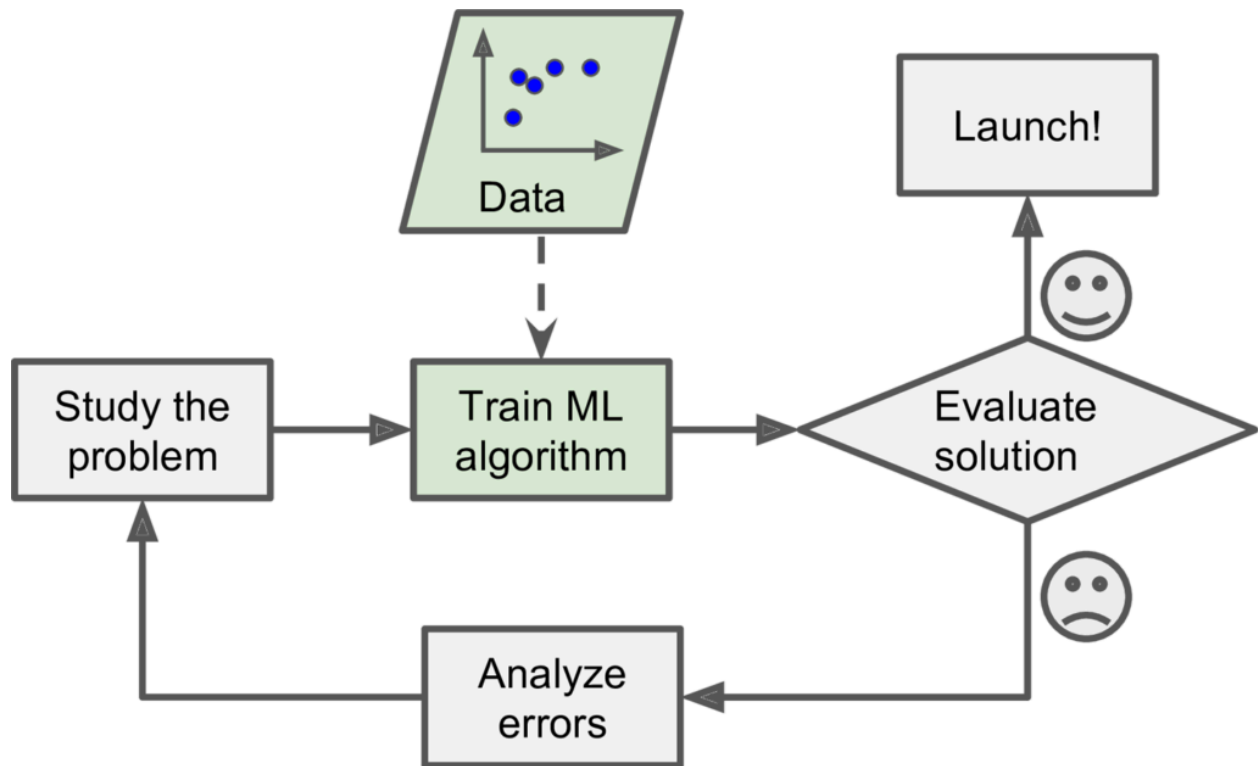
Seperti yang disebutkan di atas bahwa pada *machine learning*, mesin akan belajar sendiri tanpa perlu banyak campur tangan manusia. Hasil dari usaha mesin untuk belajar ini yang kita sebut dengan model. Misalnya, saat sebuah mesin kita ajarkan cara untuk membedakan mana siswa yang lulus dan tidak melalui data, maka hasilnya adalah model kelulusan siswa. Apakah bisa dipakai untuk data lain, misal pemilahan bunga? Tidak, karena untuk hal tersebut, sama seperti manusia, mesin kita harus diajarkan kembali dengan data pemilahan bunga.

### **Evaluation**

Kadang Model yang dibuat oleh mesin kita tidaklah sempurna, terutama karena data yang kita berikan belum tentu selalu dapat merepresentasikan data di masa depan dengan sempurna. Maka dari itu kita butuh evaluasi. Evaluasi adalah kegiatan dimana kita akan mengubah parameter pada model kita agar dapat merepresentasikan data umum secara lebih baik.

Contoh sederhananya seperti saat kita belajar membuat sebuah kopi di cafe, tentu kita akan melihat kopi yang sudah ada. Evaluasi adalah ketika kita mencoba mencicipi dan mencari kekurangan kopi kita dari contoh, lalu kita ubah parameter yang kurang sesuai, misalnya gula. agar bisa merepresentasikan contoh dengan baik.

Penggambaran langkah kerja saat membuat proyek machine learning adalah sebagai berikut.



sumber : Buku *Hands-On Machine Learning with Scikit Learn, Keras, & TensorFlow* by Aurelian Geron (2017)

---

## ***Supervised Learning***

Supervised Learning adalah algoritma yang digunakan saat data kita "berpasangan". Selain adanya input, algoritma tersebut memiliki juga output. Output ini sering juga disebut "Label". Label menjadi

*penuntun/supervisor* yang mengarahkan bagaimana algoritma ini harus memperlakukan dataset yang dimiliki.

Data set (D) yang digunakan oleh algoritma supervised learning didefinisikan sebagai berikut:

$$D=\{(x_0,y_0),(x_1,y_1),\dots,(x_n,y_n)\}$$

Model Supervised Learning yang menggunakan data terlabel (y) untuk belajar melakukan mapping fungsi. fungsi ini akan menghubungkan label/output (y) dengan input (x). Kaitan kedua digambarkan dengan fungsi

$$y = f(x)$$

Berbeda dengan unsupervised learning, pada supervised learning, model yang telah kita latih harus kita evaluasi dengan menggunakan dataset baru yang tidak digunakan untuk melakukan mapping fungsi di awal. Ini dilakukan agar kita mengetahui gambaran bagaimana model dapat bekerja dengan baik atau tidak. Dua hal yang sering dilakukan oleh algoritma ini adalah

- Classification
- Regression

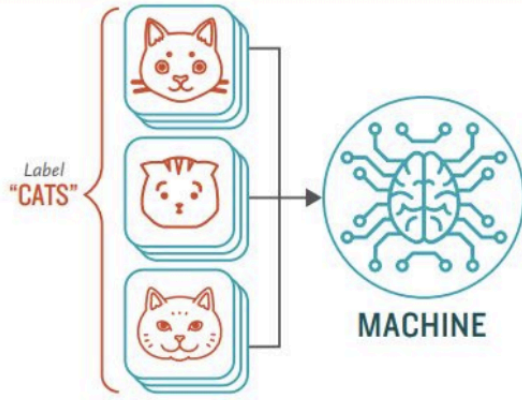
Beberapa algoritma yang termasuk supervised learning adalah:

- Support Vector Machine
- Linear Regression
- Logistic Regression
- Naïve Bayes
- Decision Tree
- K-Nearest Neighbor
- Dan Sebagainya

## How **Supervised** Machine Learning Works

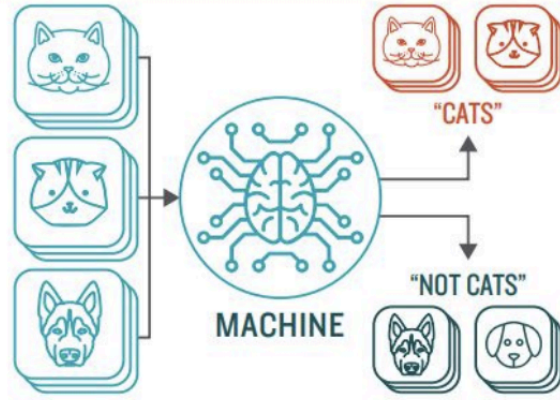
### STEP 1

Provide the machine learning algorithm categorized or "labeled" input and output data from to learn

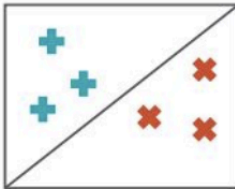


### STEP 2

Feed the machine new, unlabeled information to see if it tags new data appropriately. If not, continue refining the algorithm

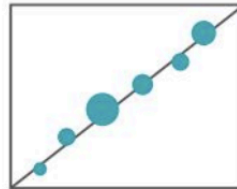


## TYPES OF PROBLEMS TO WHICH IT'S SUITED



### CLASSIFICATION

Sorting items into categories



### REGRESSION

Identifying real values (dollars, weight, etc.)

## *Unsupervised Learning*

Unsupervised Learning adalah algoritma yang --seperti namanya-- belajar dengan tidak adanya "tuntunan". Pada algoritma ini mesin kita dibiarkan untuk benar-benar belajar sendiri, karena mesin kita hanya akan menerima data input tanpa ada data output yang digunakan sebagai "penuntun".

dataset yang digunakan pada unsupervised learning didefinisikan sebagai berikut:

$$D = \{x_0, x_1, \dots, x_n\}$$

Unsupervised Learning akan membagi data tersebut tergantung pada pola yang terlihat dan manusialah yang nantinya harus menginterpretasikan pembagian tersebut. Karena inilah Unsupervised Learning biasanya berkuat dengan dua hal utama, yaitu:

- Association
- Clustering

Beberapa Algoritma yang termasuk Unsupervised Learning antara lain:

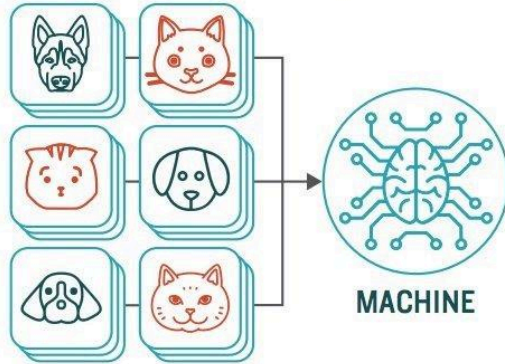
- K-Means
- Hierarchical Clustering
- DBSCAN
- Fuzzy C-Means
- Dan Sebagainya



## How **Unsupervised** Machine Learning Works

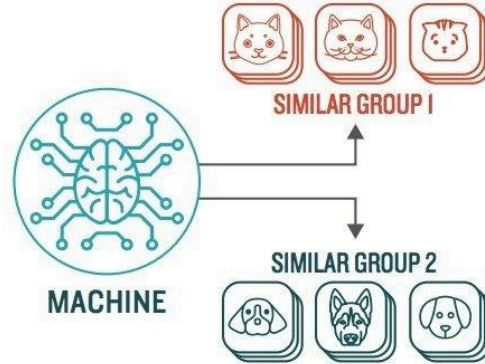
### STEP 1

Provide the machine learning algorithm uncategorized, unlabeled input data to see what patterns it finds

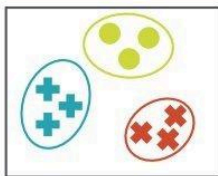


### STEP 2

Observe and learn from the patterns the machine identifies



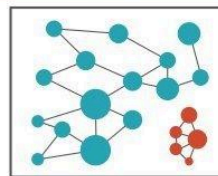
### TYPES OF PROBLEMS TO WHICH IT'S SUITED



#### CLUSTERING

Identifying similarities in groups

*For Example: Are there patterns in the data to indicate certain patients will respond better to this treatment than others?*



#### ANOMALY DETECTION

Identifying abnormalities in data

*For Example: Is a hacker intruding in our network?*

---

## Data Set

Algoritma machine learning sangat bergantung pada data yang diinputkan padanya karena sifatnya yang belajar "mandiri". Semakin baik data yang masuk maka model yang diciptakan akan baik juga. Mayoritas algoritma machine learning juga akan bekerja dengan baik jika dataset yang baik itu berada dalam jumlah yang banyak. Data set sering disimbolkan dengan huruf *D*.

Untuk lebih mudah menggambarkan bagaimana sebuah dataset, kita dapat melihat tabel di bawah ini.

age	height	gender	class
18	165	M	0
18	160	F	0
17	164	M	1
18	163	M	1
19	170	M	0
19	160	F	0
17	161	F	1
16	162	F	1

Setiap baris dan kolom menunjukkan relasi tertentu. beberapa istilah dataset yang dapat digambarkan tabel tersebut adalah:

- **Instance:** setiap baris disebut dengan instance. Biasanya instance merupakan satu buah data hasil pengamatan yang terpisah dari data pengamatan lainnya.
- **Feature:** setiap kolom disebut dengan feature. feature menunjukkan komponen dari pengamatan pada data tersebut. Feature biasanya terdiri keseluruhan input data, atau dibagi menjadi input dan output data (label)
- **Dataset:** keseluruhan tabel atau data
- **Data Type:** tipe data yang dimiliki oleh feature. Tipe data ini bisa string, integer, real, ordinal, dsb. Pada machine learning, tipe data ini dikelompokkan menjadi dua, yaitu bilangan real dan kategorikal.

## Pembagian Dataset

Setiap algoritma biasanya memiliki pembagian datasetnya sendiri agar bisa bekerja dengan optimal.

Unsupervised Learning biasanya menggunakan keseluruhan datanya untuk melakukan *training*.

Supervised Learning tidak demikian, Data set yang digunakan biasanya dibagi kedalam 3 bagian, yaitu:

- **Training Dataset** Data set yang digunakan untuk melatih model.
- **Test Dataset** Data set yang digunakan untuk menguji model, biasanya digunakan untuk evaluasi apakah model kita apakah sudah dapat mengolah data yang "asing" untuk model tersebut.
- **Validation Dataset** digunakan untuk melakukan validasi pada dataset yang kita miliki pada saat kita masih melakukan perubahan pada parameternya.

Pembagian Dataset biasanya dilakukan dengan dua langkah. Pertama dengan membagi kedalam dataset training dan dataset test. Lalu dataset training kita bagi menjadi dataset test dan dataset validation.

Untuk pembagian besaran porsi setiap dataset itu tergantung setiap case yang dikerjakan. Ada yang menggunakan perbandingan (50%:25%:25%). Ada juga yang menggunakan perbandingan (80%:10%:10%). Profesor Andrew Ng dari Stanford University menganjurkan pembagian sebagai berikut:

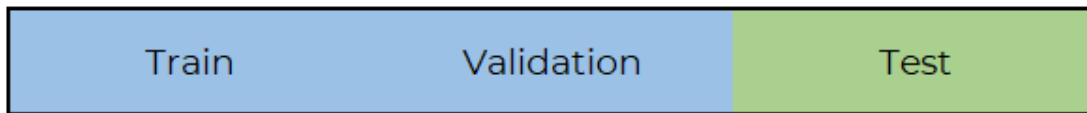
- Train Set :60%
- Validation Set: 20%
- Test Set : 20%

Atau bila kita masukkan Train dan Validation ke dalam fase training dan test set kita masukan dalam fase testing, maka pembagiannya biasa dilakukan sebesar 75%:25% atau 80%:20%



Fase Training

Fase Testing



### ***Unbalanced Dataset***

Hal lain yang perlu diperhatikan saat melakukan pembagian adalah, kesetimbangan dataset kita. Hal ini biasanya terjadi pada saat kita akan melakukan klasifikasi Data set yang tidak setimbang akan membuat model yang kita buat menjadi kurang dapat diandalkan, biasanya ini dikarekan model kita cenderung belajar pada suatu label. Penggambarannya adalah seperti seorang anak yang kita berikan lebih banyak gambar pesawat daripada mobil. Maka dia dapat lebih mengenali jenis2 pesawat daripada jenis2 mobil.

Penggambarannya sederhanaya diperlihatkan seperti dibawah, misalkan kita memiliki 10 data yang terdiri dari 2 label, yaitu A dan B yang memiliki jumlah yang sama (masing-masing 5)

A	A	B	B	A	A	B	B	B	A
---	---	---	---	---	---	---	---	---	---

Lalu kita melakukan pembagian data train dan data test dengan kombinasi 70:30 dan didapatkan dataset seperti di bawah ini

Fase Training

A	A	A	B	A	A	B
---	---	---	---	---	---	---

Fase Test

B	B	B
---	---	---

Dataset di atas kurang ideal, karena data pada fase training cenderung ke arah A. Nantinya model kita akan lebih "pintar" dalam memprediksi A dan kurang baik dalam memprediksi B. Kekurangan lainnya adalah Data pada Fase Testing ternyata berisi B semua padahal kita tahu model kita kurang baik memprediksi label B. Akhirnya model dan hasil evaluasi model kita cenderung kurang baik.

Dataset yang lebih baik ditunjukkan seperti di bawah ini

### Fase Training

A	A	A	B	B	A	B
---	---	---	---	---	---	---

### Fase Test

A	B	B
---	---	---

Data pada fase Training lebih seimbang sehingga model kita belajar tentang pola label A dan label B hampir sama baiknya, Data Test kita pun sedikit lebih baik daripada data test sebelumnya sehingga hasil evaluasi yang diberikan akan lebih baik dibanding hasil evaluasi dataset sebelumnya yang hanya memiliki label A.

---

---

## ***Deep Learning (DL)***

### ***Pengenalan Deep Learning***

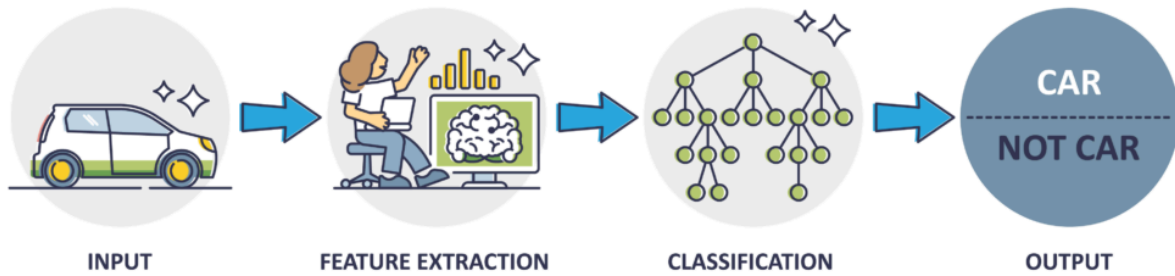
Deep learning adalah suatu kelas algoritma machine learning yang menggunakan multiple layer untuk secara progresif mengekstrak fitur-fitur tingkat tinggi dari input awal. Sebagai contoh, dalam pengolahan citra, lapisan-lapisan rendah mungkin mengidentifikasi tepi, sementara lapisan-lapisan tinggi dapat mengidentifikasi konsep-konsep yang relevan bagi manusia seperti angka, huruf, atau wajah.

Dari sudut pandang lain terhadap deep learning, deep learning merujuk pada "komputer-mensimulasikan" atau "mengotomatiskan" proses pembelajaran manusia dari suatu sumber (misalnya, gambar anjing) menjadi suatu objek yang dipelajari (anjing). Oleh karena itu, istilah yang diciptakan sebagai "pembelajaran yang lebih dalam" atau "pembelajaran yang paling dalam" memiliki makna. Pembelajaran yang paling dalam mengacu pada pembelajaran sepenuhnya otomatis dari suatu sumber menjadi objek yang dipelajari akhir. Pembelajaran yang lebih dalam, oleh karena itu, mengacu pada suatu proses pembelajaran campuran: proses pembelajaran manusia dari suatu sumber menjadi objek semi-dipelajari, yang diikuti oleh proses pembelajaran komputer dari objek semi-dipelajari manusia menjadi objek yang dipelajari akhir.

---

## ***Perbedaan dengan Machine Learning Konvensional***

## MACHINE LEARNING



## DEEP LEARNING



Perbedaan utama antara Deep Learning dan Machine Learning Konvensional terletak pada struktur dan cara mereka menangani data. Berikut adalah beberapa perbedaan kunci:

### Struktur Model:

**Machine Learning Konvensional (ML):** Model dalam machine learning konvensional umumnya terdiri dari algoritma yang dirancang untuk mengekstrak fitur dari data dan membuat prediksi.

**Deep Learning (DL):** Deep learning menggunakan neural network dengan multiple layers (deep neural network). Setiap layer dalam neural network memproses data untuk mengekstrak representasi yang semakin abstrak.



### **Ekstraksi Fitur Otomatis:**

**ML Konvensional:** Memerlukan ekstraksi fitur manual dari data yang dilakukan oleh manusia sebelumnya.

**DL:** Dapat melakukan ekstraksi fitur otomatis dari data tanpa memerlukan banyak kerja manual.

### **Representasi Tingkat Tinggi:**

**ML Konvensional:** Cenderung untuk mengekstrak fitur-fitur yang lebih sederhana dan langsung dari data.

**DL:** Mampu mengekstrak fitur-fitur tingkat tinggi dan kompleks dari data, membuatnya sangat efektif untuk tugas-tugas seperti pengenalan wajah atau bahasa alami.

### **Kuantitas Data:**

**ML Konvensional:** Dalam beberapa kasus, memerlukan jumlah data yang lebih banyak untuk mencapai kinerja yang baik.

**DL:** Meskipun memerlukan data yang besar, deep learning dapat memberikan hasil yang baik bahkan dengan jumlah data yang lebih sedikit, terutama karena kemampuannya untuk mengekstrak fitur secara otomatis.

### **Fleksibilitas dan Umum:**

**ML Konvensional:** Lebih umum digunakan untuk tugas-tugas khusus dan memerlukan penyesuaian manual.

**DL:** Lebih fleksibel dan dapat digunakan untuk berbagai macam tugas tanpa perlu penyesuaian manual yang signifikan.

#### **Kompleksitas Model:**

**ML Konvensional:** Model cenderung lebih sederhana dengan fokus pada algoritma yang lebih tradisional.

**DL:** Model cenderung lebih kompleks dengan banyak layer, memungkinkan mereka untuk memahami struktur yang lebih kompleks dalam data.

#### **Pelatihan Model:**

**ML Konvensional:** Proses pelatihan mungkin memerlukan waktu lebih lama dan hasilnya tergantung pada pemilihan fitur yang tepat.

**DL:** Meskipun pelatihan bisa memakan waktu, deep learning dapat secara otomatis menyesuaikan bobot dalam neural network untuk meningkatkan kinerja tanpa banyak campur tangan manusia.

---

## ***Metode Deep Learning***

### ***Convolutional Neural Networks (CNN)***

Convolutional Neural Network (CNN) adalah jenis jaringan saraf maju yang diatur dengan baik dan mempelajari rekayasa fitur secara otomatis melalui optimisasi filter (atau kernel). Gradient yang

menghilang dan gradient yang meledak, yang terlihat selama backpropagation pada jaringan saraf sebelumnya, dapat dicegah dengan menggunakan bobot yang diatur dengan baik melalui koneksi yang lebih sedikit. Sebagai contoh, untuk setiap neuron dalam lapisan terhubung penuh, diperlukan 10.000 bobot untuk memproses gambar berukuran  $100 \times 100$  piksel. Namun, dengan menerapkan kernel konvolusi bertingkat, hanya diperlukan 25 neuron untuk memproses ubin berukuran  $5 \times 5$ . Fitur pada lapisan lebih tinggi diekstraksi dari jendela konteks yang lebih luas, dibandingkan dengan fitur lapisan yang lebih rendah.

### ***Recurrent Neural Networks (RNN)***

Recurrent Neural Network (RNN) adalah jenis jaringan saraf tiruan yang menggunakan data sekuensial atau data deret waktu. Algoritma deep learning ini umumnya digunakan untuk masalah ordinal atau temporal, seperti terjemahan bahasa, pemrosesan bahasa alami (NLP), pengenalan suara, dan penjelasan gambar; mereka diintegrasikan ke dalam aplikasi populer seperti Siri, pencarian suara, dan Google Translate. Seperti halnya jaringan saraf maju dan konvensional (CNN), jaringan saraf rekuren menggunakan data pelatihan untuk belajar. Mereka dibedakan oleh "memori" mereka karena mereka mengambil informasi dari masukan sebelumnya untuk memengaruhi masukan dan keluaran saat ini. Sementara jaringan saraf mendalam tradisional mengasumsikan bahwa masukan dan keluaran independen satu sama lain, keluaran dari jaringan saraf rekuren bergantung pada elemen-elemen sebelumnya dalam urutan. Meskipun kejadian masa depan juga akan membantu menentukan keluaran dari urutan yang diberikan, jaringan saraf rekuren unidireksional tidak dapat memperhitungkannya dalam prediksinya.

Mari ambil idiom, seperti "feeling under the weather", yang sering digunakan ketika seseorang sakit, untuk membantu kita menjelaskan RNN. Agar idiom itu masuk akal, perlu diungkapkan dalam urutan tersebut. Oleh karena itu, jaringan rekuren perlu memperhitungkan posisi setiap kata dalam idiom dan menggunakan informasi tersebut untuk memprediksi kata berikutnya dalam urutan.

## CNN vs RNN

**Aplikasi:** CNNs cocok untuk tugas-tugas pengenalan gambar dan pengolahan gambar, sedangkan RNNs lebih cocok untuk tugas-tugas berurutan seperti pemrosesan bahasa alami dan pengenalan suara.

**Representasi Data:** CNNs memahami representasi spasial dan fitur visual, sementara RNNs memahami representasi temporal dan konteks sekuensial.

**Struktur Jaringan:** CNNs memiliki struktur feedforward dengan lapisan konvolusi, sedangkan RNNs memiliki struktur yang mempertahankan informasi dari langkah sebelumnya.

**Pengolahan Data:** CNNs dapat memproses seluruh data pada satu waktu, sedangkan RNNs memproses data sekuensial satu langkah pada satu waktu.

Pilihan antara CNN dan RNN tergantung pada jenis data dan tugas tertentu yang ingin diatasi. Kombinasi keduanya, dikenal sebagai Convolutional Recurrent Neural Networks (CRNNs), juga digunakan untuk menangani data sekuensial dalam konteks spasial.

---

---

## *Etika dalam AI*

Kecerdasan Buatan (AI) telah menjadi bagian tak terpisahkan dari kehidupan kita, mengubah cara kita berinteraksi dengan teknologi. Namun, di balik kemajuan ini, penting untuk memahami dan menerapkan

prinsip-prinsip etika agar pengembangan dan penggunaan AI berlangsung secara adil, aman, dan sesuai dengan norma-norma sosial.

## **1. Bias dalam Algoritma**

**Apa itu Bias?** Bias adalah ketidakseimbangan atau ketidakadilan dalam pengambilan keputusan yang dapat terjadi dalam algoritma AI. Faktor penyebabnya dapat berasal dari data pelatihan yang tidak representatif atau desain algoritma yang inheren.

**Mengapa Ini Penting?** Keputusan yang dipengaruhi oleh bias dapat merugikan kelompok tertentu, seperti dalam rekrutmen, penilaian kredit, atau sistem peradilan.

**Bagaimana Mengatasi?** Pengembang memiliki tanggung jawab untuk mengidentifikasi dan mengurangi bias dengan menggunakan data yang beragam dan melakukan pengujian yang cermat.

## **2. Privasi dan Keamanan Data**

**Apa itu Privasi dan Keamanan Data?** Privasi dan Keamanan Data berkaitan dengan perlindungan informasi pribadi pengguna oleh sistem AI. Pengguna memiliki hak untuk privasi dan perlindungan data pribadi mereka.

**Mengapa Ini Penting?** Data pribadi dapat digunakan untuk pengambilan keputusan yang signifikan dalam kehidupan seseorang. Penting untuk memastikan bahwa data tidak disalahgunakan atau bocor.

**Bagaimana Mengatasi?** Pengembang perlu menerapkan enkripsi data, mengikuti regulasi privasi, dan memastikan bahwa data hanya digunakan untuk tujuan yang sah.

## **3. Tanggung Jawab Pengembang dan Pengguna**

**Apa itu Tanggung Jawab?** Tanggung Jawab Pengembang dan Pengguna menekankan pentingnya bertanggung jawab dalam menciptakan, mengimplementasikan, dan menggunakan teknologi AI.

Kesalahan dalam sistem AI dapat memiliki dampak besar, terutama dalam bidang kesehatan atau keamanan.

**Mengapa Ini Penting?** Kesalahan dalam pengembangan atau penggunaan AI dapat berdampak luas dan merugikan.

**Bagaimana Mengatasi?** Pengembang harus melakukan pengujian yang cermat, mengikuti prinsip etika, dan memastikan pengguna memahami batasan dan risiko AI. Penyadaran dan pelatihan etika bagi pengembang dan pengguna juga penting.

### **Kesimpulan:**

Menerapkan etika dalam AI melibatkan penanganan bias dalam algoritma, perlindungan privasi dan keamanan data, serta menetapkan tanggung jawab bagi pengembang dan pengguna. Dengan memahami dan mengimplementasikan prinsip-prinsip ini, kita dapat memastikan bahwa AI digunakan secara etis, bertanggung jawab, dan memberikan manfaat positif bagi masyarakat dan dunia secara keseluruhan. Etika menjadi dasar untuk membangun masa depan AI yang mendukung nilai-nilai kemanusiaan.