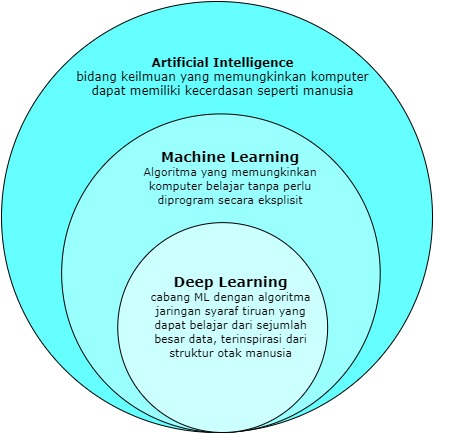
**Review Basic Machine Learning**

1. **Intro To Machine Learning**
   1. **Definition**



* + 1. **Artificial Intellegence (AI) atau Kecerdasan buatan adalah simulasi dari kecerdasan yang dimiliki oleh manusia yang dimodelkan di dalam mesin dan diprogram agar bisa berpikir seperti halnya manusia.**
    2. **Machine learning adalah cabang dari artificial intelligence. Kecerdasan buatan memiliki pengertian yang sangat luas tapi secara umum dapat dipahami sebagai komputer dengan kecerdasan layaknya manusia. Sedangkan ML memiliki arti lebih spesifik yaitu menggunakan metode statistika untuk membuat komputer dapat mempelajari pola pada data tanpa perlu diprogram secara eksplisit.**
    3. **Deep learning adalah cabang machine learning dengan algoritma jaringan syaraf tiruan yang dapat belajar dan beradaptasi terhadap sejumlah besar data. Algoritma jaringan syaraf tiruan pada deep learning terinspirasi dari struktur otak manusia.**
  1. **Mengapa Machine Learning**
     1. **Masalah yang Solusinya Membutuhkan Banyak Penyesuaian dan Aturan**
     2. **Masalah Rumit yang Tidak Bisa Diselesaikan dengan Pemrograman Tradisional**
  2. **Jenis Machine Learning**
     1. **Supervised Learning**
     2. **Unsupervised Learning**
     3. **Semi-supervised Learning**
     4. **Reinforcement Learning**
  3. **Library Populer**
     1. **Numpy**
     2. **Panda**
     3. **Matplotlib**
     4. **Scikit Learn**
     5. **PyTorch**
     6. **TensorFlow**
     7. **Keras**
  4. **Data Collecting**

**Cara mengumpulkan data, Ada tiga cara yang bisa kita lakukan untuk mengumpulkan data, yaitu.**

* **Mengekstrasi data (misal dari internet, riset, survei, dll).**
* **Mengumpulkan dan membuat dataset Anda sendiri dari nol.**
* **Menggunakan dataset yang telah ada.**
  1. **Data Cleaning**

**Data cleaning penting sebab proses ini meningkatkan kualitas data yang juga berpengaruh terhadap produktivitas kerja secara keseluruhan. Data yang tidak akurat bisa berpengaruh buruk terhadap akurasi dan performa model. Saat kita melakukan proses data cleaning, kita membuang data dan informasi yang tidak dibutuhkan sehingga kita akan mendapatkan data yang berkualitas. Data yang akurat dan berkualitas akan berpengaruh positif terhadap proses pengambilan keputusan. Pernahkah mendengar ungkapan “Garbage In - Garbage Out?” Dalam konteks machine learning, jika input yang Anda masukkan itu buruk, sudah barang tentu hasil olah data Anda pun akan buruk.**

**Hal yang harus diperhatikan dalam proses Data Cleaning :**

* **Konsistensi Format**
* **Skala data**
* **Duplikasi data**
* **Missing value (data tak lengkap)**
* **Skewness Distribution (data tak seimbang)**
  1. **Data Processing**

**Pada tahap ini, setelah data diambil dari sumber tertentu, ia dimasukkan pada suatu environment. Lantas data diproses agar bisa diolah oleh model machine learning.**

* 1. **Data Preparation**

**Setelah dataset dibersihkan, masih ada beberapa tahap yang perlu dilakukan agar dataset benar-benar siap untuk diproses oleh model machine learning. Biasanya, dataset Anda akan terdiri dari dua jenis data: kategorik dan numerik.**

* 1. **Data Storage**

**Data warehouse tujuan awalnya adalah untuk membantu proses aliran data dari sistem operasional ke dalam sistem pendukung keputusan atau decision-support system (DSS). Seiring berjalannya waktu, data warehouse berkembang menjadi lebih efisien. Ia berevolusi dari penyimpanan informasi pendukung platform business intelligence menjadi infrastruktur analitis luas yang mendukung berbagai macam aplikasi.**

* 1. **Datasets**

**Dataset yang telah dibersihkan dan diproses kemudian siap kita latih dengan machine learning. Satu-satunya cara untuk mengetahui apakah model machine learning kita bagus atau tidak adalah dengan mengujinya pada kasus atau data baru yang belum dikenali oleh model. Kita bisa saja membuat model dan langsung mengujinya pada tahap produksi lalu memonitor kualitasnya. Tapi jika ternyata model yang kita kembangkan bekerja dengan buruk, pelanggan dan klien kita akan komplain. Selain itu, cara ini tentu akan memakan sumber daya dan biaya yang lebih besar. Pilihan yang lebih baik adalah dengan membagi dataset menjadi 2 bagian yaitu data training dan data testing.**

* 1. **Data Evaluation**

**Hal ini menyebabkan model tidak bekerja dengan baik ketika menemui data baru. Solusi paling umum dari masalah ini adalah dengan menambahkan validation set pada model machine learning. (Train, Test, Validation Set) dan Cross Validation**

1. **Supervised and Unsupervised Learning**
   1. **Pada supervised learning kita melatih sekumpulan data yang memiliki label. Label adalah pengenal dari sebuah data.**

**Berikut adalah beberapa algoritma supervised learning yang penting Anda ketahui, dan akan dibahas di kelas Belajar Machine Learning untuk Pemula:**

* **Linear Regression**
* **Logistic Regression**
* **Classification**
* **Decision Trees**
* **Support Vector Machines**
* **Neural Networks.**
  1. **Model unsupervised learning akan menentukan segmen market dan mengelompokkan pengunjung ke dalam segmen market yang berbeda. Dengan output dari model ini, pengelola ecommerce dapat menentukan strategi untuk meningkatkan penjualan atau strategi lain yang dirasa perlu diambil untuk kelanjutan bisnis.**

1. **Support Vector Machine**

* **Support Vector Machine adalah model ML multifungsi yang dapat digunakan untuk menyelesaikan permasalahan klasifikasi, regresi, dan pendeteksian outlier. Termasuk ke dalam kategori supervised learning, SVM adalah salah satu metode yang paling populer dalam machine learning. Siapa pun yang tertarik untuk masuk ke dalam dunia ML, perlu mengetahui SVM.**
* **Tujuan dari algoritma SVM adalah untuk menemukan hyperplane terbaik dalam ruang berdimensi-N (ruang dengan N-jumlah fitur) yang berfungsi sebagai pemisah yang jelas bagi titik-titik data input. Untuk lebih jelasnya, perhatikan gambar berikut.**

1. **Basic Machine Learning**
   1. **Machine Learning Workflow**
   2. **Machine Learning Use Case**
   3. **Overfitting dan Underfitting**
   4. **Model Selection**
2. **Neural Network**

**Pembahasan mengenai neural network, sebuah model machine learning yang terinspirasi dari cara kerja otak manusia. Di modul ini akan dibahas mengenai artificial neural network, multi layer perceptron, dan convolutional neural network. Setelah mempelajari modul ini Anda akan memahami bagaimana neural network dapat dipakai dalam bidang yang rumit untuk dikerjakan oleh model machine learning biasa, seperti computer vision.**

1. **TensorFlow**

TensorFlow(TF) adalah end-to-end open source platform yang dikembangkan oleh Google Brain dan sangat populer untuk pengembangan machine learning berskala besar. Ia memiliki ekosistem tools, library, dan sumber daya komunitas yang komprehensif dan fleksibel, yang memungkinkan para peneliti dan pengembang membangun dan menerapkan (deploy) aplikasi machine learning dengan mudah.

Keunggulan TensorFlow

* Bisa dijalankan di hampir semua platform: GPU, CPU, dan TPU (TensorFlow Processing Units) yang secara khusus dimanfaatkan untuk mengerjakan matematika tensor.
* Memberikan performa terbaik dengan kemampuan melakukan iterasi dan melatih model secara cepat sehingga mampu menjalankan lebih banyak eksperimen.
* Skalabilitas komputasi yang tinggi pada kumpulan data yang sangat besar.
* Pembuatan model yang mudah dengan beberapa level abstraksi sesuai kebutuhan.
* Menyediakan jalur langsung ke produksi, baik itu pada server, perangkat mobile atau web sehingga memudahkan kita melakukan pipeline machine learning hingga ke level produksi.