

## Rangkuman Materi Perkuliahan Machine learning Kelompok 21

Muhammad Shaadam haidar Yuwono<sup>1</sup>, Muhammad labib<sup>2</sup>

### 1.Deskripsi singkat

Pertemuan ini menekankan bahwa keberhasilan sebuah proyek *Machine Learning* tidak hanya terletak pada akurasi model, tetapi pada kualitas **Pipeline Engineering** (alur kerja teknik) dan pemahaman mendalam terhadap konteks bisnis. Dosen menggeser fokus mahasiswa dari sekadar "belajar algoritma" menjadi "membangun solusi siap pakai."

### Poin-poin Utama:

#### 1.Diferensiasi Peran (Science vs. Engineering):

- Science:** Berfokus pada eksperimen untuk menemukan jawaban fundamental dari data ("Mengapa ini terjadi?").
- Engineering:** Berfokus pada membangun produk akhir yang stabil, cepat, dan bisa diintegrasikan ke aplikasi atau server nyata.

#### 2. Fase Pengembangan Solusi ML: Dosen merinci alur kerja profesional yang terdiri dari:

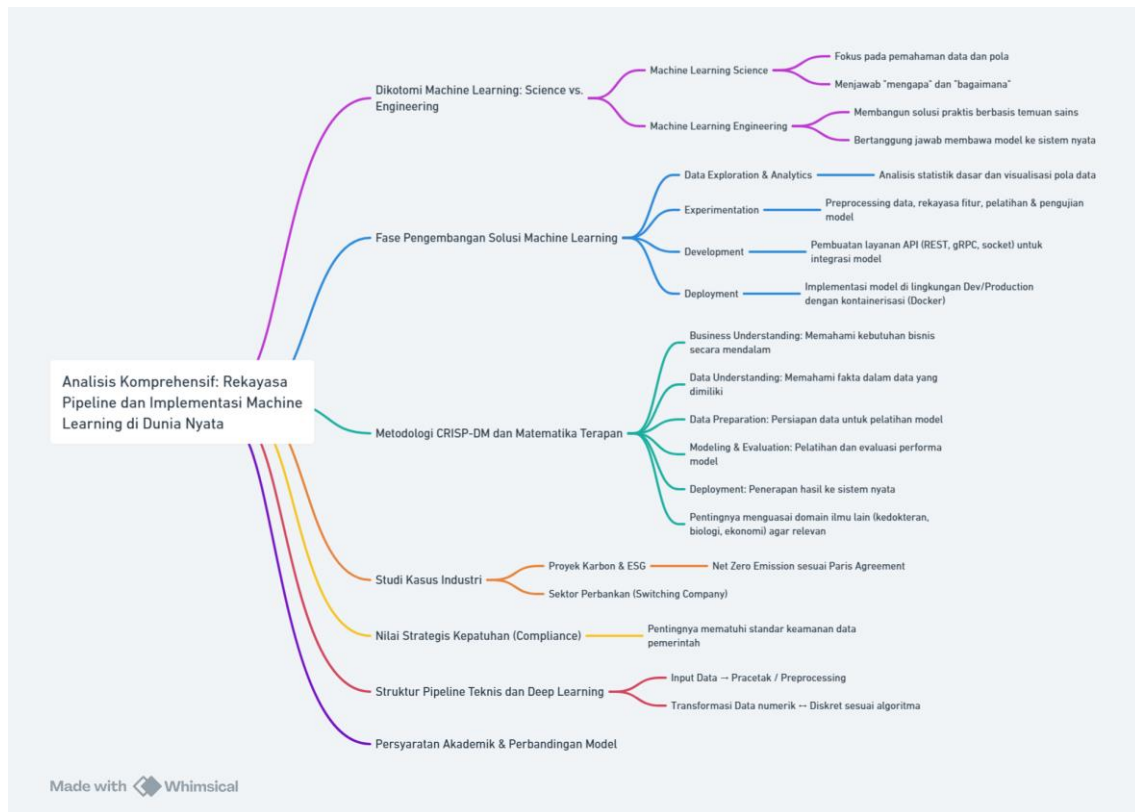
- Eksplorasi & Analitik:** Memahami karakteristik dan fakta lapangan dari data.
- Eksperimen:** Tahap prapemrosesan, rekayasa fitur (*feature engineering*), dan pelatihan model.
- Development:** Mengonversi model menjadi layanan (*service*) yang fungsional (API).
- Deployment:** Menerapkan model ke lingkungan produksi (disarankan menggunakan **Docker** agar lebih profesional).

#### 3. ML sebagai Matematika Terapan: Ditekankan bahwa informatika adalah alat. Praktisi ML harus memahami domain ilmu lain (seperti bisnis karbon/ESG atau perbankan) agar model yang dibuat memiliki relevansi dan memberikan solusi nyata bagi masalah di industri tersebut.

#### 4. Pentingnya Kepatuhan (Compliance): Salah satu nilai jual tertinggi seorang *ML Engineer* adalah pemahamannya terhadap aturan hukum dan keamanan data. Di dunia kerja, ketidakpatuhan terhadap regulasi bisa berakibat fatal secara hukum dan finansial.

#### 5.Struktur Pipeline Deep Learning: Penjelasan mengenai alur data di mana input (X) melalui tahap transformasi (numerik ke diskret), ekstraksi fitur, hingga tahap *Pre-training* (opsional namun krusial pada Deep Learning) sebelum masuk ke pemodelan.

## 2.Mindmap



### 3. penjelasan detail

#### Analisis Komprehensif: Rekayasa Pipeline dan Implementasi Machine Learning di Dunia Nyata

Pertemuan ini memberikan wawasan mendalam mengenai pergeseran paradigma dari sekadar memahami algoritma menjadi kemampuan membangun solusi *Machine Learning* (ML) yang utuh dan fungsional di industri. Inti dari bahasan ini adalah pemahaman bahwa seorang praktisi AI tidak hanya bekerja di laboratorium riset, tetapi harus mampu melakukan rekayasa (*engineering*) yang menjawab kebutuhan bisnis serta mematuhi regulasi yang berlaku.

##### 1. Dikotomi Machine Learning: Science vs. Engineering

Dosen membuka wawasan dengan membedakan dua pilar utama dalam bidang kecerdasan buatan:

- **Machine Learning Science:** Fokus utamanya adalah menjawab pertanyaan fundamental terkait data. Ilmuwan data bekerja untuk mencari tahu "mengapa" suatu pola terjadi, "bagaimana" hubungan antar variabel, dan melakukan estimasi waktu atau durasi tertentu.

- **Machine Learning Engineering:** Fokusnya adalah rekayasa untuk membangun solusi praktis. *Engineering* menggunakan temuan dari sains untuk menciptakan produk yang dapat digunakan, seperti sistem yang memprediksi performa akademik mahasiswa berdasarkan parameter kehadiran dan data historis lainnya.

Dalam konteks profesional, peran *engineer* jauh lebih luas karena mereka bertanggung jawab membawa model dari lingkungan eksperimen menuju sistem yang siap pakai di server atau aplikasi pengguna.

## 2. Fase Pengembangan Solusi Machine Learning

Membangun solusi ML yang tangguh memerlukan siklus hidup pengembangan yang sistematis. Dosen merinci fase-fase tersebut sebagai berikut:

- **Data Exploration & Analytics:** Tahap awal untuk melihat karakteristik data, seperti nilai rata-rata, distribusi statistik, dan pola-pola awal yang terlihat secara visual.
- **Experimentation:** Tahap ini merupakan jantung dari pengembangan model. Di dalamnya terdapat aktivitas prapemrosesan data (*preprocessing*), rekayasa fitur (*feature engineering*), serta proses *modeling* yang melibatkan pelatihan (*training*) dan pengujian (*testing*).
- **Development:** Setelah model tervalidasi, *engineer* membangun layanan (*service*) di atasnya. Ini melibatkan pembuatan antar muka seperti REST API, gRPC, atau koneksi *socket* agar model dapat berkomunikasi dengan sistem lain.
- **Deployment:** Tahap akhir di mana model diterapkan secara langsung di lingkungan pengembangan (*Dev*) atau produksi (*Production*). Dosen menyarankan penggunaan teknologi kontainerisasi seperti **Docker** untuk memastikan model berjalan konsisten di berbagai server.

## 3. Metodologi CRISP-DM dan Matematika Terapan

Dosen memperkenalkan kembali kerangka kerja **CRISP-DM** (*Cross-Industry Standard Process for Data Mining*) yang menjadi standar industri. Struktur utamanya mencakup:

1. **Business Understanding:** Memahami masalah dari sudut pandang bisnis.
2. **Data Understanding:** Memahami fakta-fakta yang ada di dalam data.
3. **Data Preparation:** Menyiapkan data agar siap dilatih.
4. **Modeling & Evaluation:** Melatih model dan mengevaluasi hasilnya.
5. **Deployment:** Menerapkan model ke sistem nyata.

Ditekankan bahwa informatika, khususnya ML, adalah **matematika terapan**. Praktisi AI harus keluar dari zona nyaman teknis mereka dan mempelajari domain ilmu lain—seperti kedokteran, biologi, atau ekonomi—agar solusi yang mereka bangun benar-benar relevan dengan fakta lapangan.

#### 4. Studi Kasus Industri: Karbon (ESG) dan Sektor Keuangan

Untuk mengilustrasikan pentingnya *Business Understanding*, dosen memberikan dua contoh nyata:

- **Proyek Karbon & ESG:** Berdasarkan *Paris Agreement*, negara-negara di dunia berkomitmen untuk mencapai *Net Zero Emission*. Dalam domain ini, seorang *engineer* harus belajar tentang kapasitas pohon dalam menyerap karbon dan bagaimana data polusi diukur agar bisa menciptakan solusi IT yang membantu perencanaan pengurangan emisi polusi.
- **Sektor Perbankan (Switching Company):** Dalam perusahaan yang menangani transaksi ATM atau EDC, seorang praktisi AI tidak hanya mengurus algoritma, tetapi juga harus memahami alur data dari perangkat fisik ke sistem pusat. Tantangan utamanya bukan hanya pada akurasi model, tetapi pada keamanan dan konsistensi data.

#### 5. Nilai Strategis Kepatuhan (Compliance)

Salah satu poin paling krusial dalam pertemuan ini adalah pembahasan mengenai **Compliance** (ketaatan terhadap peraturan). Di industri nyata, kepatuhan terhadap standar keamanan data (seperti peraturan pemerintah) memiliki nilai ekonomi yang sangat tinggi.

- Seorang *Machine Learning Engineer* yang memahami *compliance* akan dihargai jauh lebih tinggi dan dibayar lebih mahal karena mereka bisa menjamin bahwa sistem yang dibangun aman dan legal.
- Dosen memberikan peringatan bahwa di dunia kerja, kesalahan teknis yang melanggar kepatuhan bisa mengakibatkan kerugian finansial yang nyata bagi orang lain (misalnya uang nasabah hilang), berbeda dengan lingkungan kampus yang merupakan tempat belajar paling aman.

#### 6. Struktur Pipeline Teknis dan Deep Learning

Dosen menjelaskan secara visual bagaimana data mengalir dalam sebuah *pipeline*:

- **Input Data (X)** masuk ke tahap prapemrosesan.
- **Data Transformation:** Mengubah data dari numerik menjadi diskret atau sebaliknya agar sesuai dengan kebutuhan algoritma.
- **Feature Extracting:** Mengekstrak fitur-fitur penting yang memberikan informasi paling kaya bagi model.
- **Deep Learning Pipeline:** Berbeda dengan ML tradisional, dalam *Deep Learning* seringkali terdapat tahap **Pre-training** yang bersifat opsional namun sangat membantu dalam menginisialisasi bobot jaringan saraf sebelum pelatihan utama dilakukan.

## 7. Persyaratan Akademik dan Perbandingan Model

Dosen menetapkan standar profesional untuk tugas UTS dan UAS sebagai bentuk simulasi kerja di industri. Mahasiswa diwajibkan untuk:

- Menggunakan repositori **GitHub** dengan struktur folder yang rapi (folder UTS dan UAS terpisah).
- Menyertakan proposal, dataset, *Jupyter Notebook*, serta laporan resmi dalam format PDF.
- Melakukan analisis perbandingan antara *Supervised Learning* konvensional dengan *Deep Learning*. Mahasiswa harus mampu memberikan hipotesis yang kuat mengapa *Deep Learning* diperlukan untuk kasus tertentu dan apa perbedaan hasil yang didapatkan.

## 4.pseudocode

### // ALGORITMA REKAYASA SOLUSI MACHINE LEARNING (ML ENGINEERING)

#### BEGIN ML\_PROJECT\_WORKFLOW

##### // 1. TAHAP PEMAHAMAN BISNIS (Business Understanding)

INPUT business\_domain (Contoh: "ESG/Carbon Project" atau "Banking")

IDENTIFY business\_rules AND compliance\_standards

DEFINE problem\_statement ("Apa yang ingin diselesaikan?")

##### // 2. EKSPLORASI DATA (Data Exploration & Analytics)

LOAD raw\_dataset

FOR EACH feature IN dataset:

CALCULATE statistics (Mean, Median, Distribution)

CHECK data\_facts (Apakah data sesuai dengan kenyataan lapangan?)

END FOR

##### // 3. FASE EKSPERIMEN (Experiment & Engineering)

START experiment\_phase:

// Prapemrosesan & Transformasi

FOR EACH data\_input (X):

APPLY Data\_Transformation (Numeric <-> Discrete)

APPLY Feature\_Extracting (Ambil informasi paling penting)

END FOR

```
// Pemilihan Jalur Model
IF model_type == "Deep Learning" THEN
    APPLY Optional_Pre_training // Inisialisasi bobot awal
    EXECUTE Deep_Learning_Algorithm (Neural Networks)
ELSE
    EXECUTE Supervised_Learning_Algorithm (ML Tradisional)
END IF

// Evaluasi
COMPARE performance(ML_Tradisional, Deep_Learning)
SAVE best_model f(x)

END experiment_phase

// 4. PENGEMBANGAN LAYANAN (Development)
BUILD service_interface:
    IF use_api THEN
        CREATE REST_API_Endpoint
    ELSE IF use_high_performance THEN
        CREATE gRPC_Service
    END IF

// 5. PENERAPAN (Deployment & Compliance)
CREATE container_image (Using Docker):
    SET environment_variables
    PACK code_and_trained_model

START container ON Production_Server

// Verifikasi Kepatuhan (Compliance Check)
IF system_is_secure AND follows_government_regulations THEN
    SET status = "LIVE"
ELSE
    REVISE security_and_compliance
END IF

END ML_PROJECT_WORKFLOW
```