



# PART 6

## Machine Learning

VERSI 2025

# Definisi

- Machine Learning (ML) adalah cabang dari *Artificial Intelligence (AI)* yang berfokus pada pengembangan algoritma yang dapat belajar dari data dan membuat prediksi atau keputusan tanpa harus diprogram secara eksplisit.



# Tujuan

- Mempelajari pola dari data historis.
- Membuat prediksi terhadap data baru.
- Mengotomatiskan pengambilan keputusan berbasis data.



# Komponen Utama

- **Data** : sumber pengetahuan bagi model.
- **Model** : representasi matematis dari pola dalam data.
- **Algoritma** : cara model belajar dari data (misalnya: *gradient descent*).
- **Evaluasi** : mengukur kinerja model menggunakan metrik seperti *accuracy*, *precision*, atau *RMSE*.

# Gradient Descent

- Gradient Descent adalah algoritma optimasi yang digunakan untuk *mencari nilai parameter terbaik (misalnya bobot dan bias)* agar error (kesalahan prediksi) menjadi sekecil mungkin.
- Anda sedang berada di atas bukit dan ingin turun ke lembah paling bawah. Lembah itu adalah titik “kesalahan terkecil”
- Misal marketing ingin membuat model sederhana untuk memprediksi harga rumah berdasarkan luas tanah. Awalnya model mungkin menghasilkan prediksi terlalu tinggi atau terlalu rendah.
- Dengan *gradient descent*, model akan belajar dari kesalahan itu dan menyesuaikan bobot sedikit demi sedikit agar prediksi makin akurat.

# RMSE

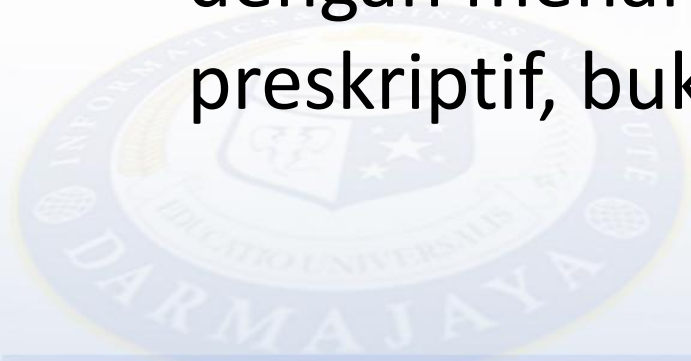
- RMSE adalah salah satu metode evaluasi model regresi yang mengukur rata-rata kesalahan prediksi antara nilai aktual dan nilai prediksi.
- Makna RMSE dalam Praktik:
  - Kecil = bagus, artinya prediksi mendekati nilai aktual.
  - Besar = buruk, artinya model masih salah besar dalam prediksi.
  - RMSE sering dipakai dalam: forecasting penjualan, prediksi harga, analisis risiko finansial, dan data bisnis.

# Jenis

Jenis	Penjelasan	Contoh
<b>Supervised Learning</b>	Model belajar dari data yang sudah diberi label (ada input dan output yang diketahui).	Prediksi penjualan, klasifikasi pelanggan, deteksi penipuan.
<b>Unsupervised Learning</b>	Model menemukan pola dalam data tanpa label.	Segmentasi pelanggan, analisis cluster, reduksi dimensi.
<b>Semi-Supervised Learning</b>	Kombinasi data berlabel dan tidak berlabel.	Pengenalan wajah, analisis teks.

# Hubungan Machine Learning dan Business Intelligence (BI) (1)

- Business Intelligence (BI) adalah proses mengubah data mentah menjadi informasi yang berguna untuk pengambilan keputusan bisnis.
- Machine Learning memperluas kemampuan BI dengan menambahkan analisis prediktif dan preskriptif, bukan hanya deskriptif.



# Hubungan Machine Learning dan Business Intelligence (BI) (2)

Aspek	BI Tradisional	BI dengan Machine Learning
Fokus	Analisis data historis	Prediksi tren masa depan
Pendekatan	Deskriptif: apa yang terjadi?	Prediktif & Preskriptif: apa yang akan terjadi? apa yang sebaiknya dilakukan?
Teknologi	SQL, OLAP, dashboard	Python, R, TensorFlow, Scikit-learn
Contoh Aplikasi	Laporan penjualan bulanan	Prediksi penjualan bulan depan berdasarkan pola historis

# Keterkaitan Gradient Descent & RMSE

- Saat melatih model, kita menggunakan gradient descent untuk meminimalkan RMSE (atau error function lainnya).
- Dengan setiap iterasi, model belajar dan menurunkan nilai RMSE sedikit demi sedikit.
- Bayangkan Anda menuruni bukit (gradient descent) untuk mencapai titik terendah (error terkecil).
  - RMSE adalah tinggi bukitnya, semakin rendah anda turun, semakin kecil RMSE-nya, dan semakin baik modelnya.

# Aplikasi Machine Learning dalam Business Intelligence (1)

## a. Customer Segmentation

- Menggunakan *clustering* (misalnya *K-Means*) untuk mengelompokkan pelanggan berdasarkan perilaku belanja.
- Hasilnya: strategi pemasaran lebih tepat sasaran.

## b. Sales Forecasting

- Model *regresi linier* atau *time series* (*ARIMA*, *LSTM*) digunakan untuk memprediksi penjualan masa depan.
- Membantu manajemen stok dan perencanaan anggaran.

# Aplikasi Machine Learning dalam Business Intelligence (2)

## c. Churn Prediction

- Memprediksi pelanggan yang berpotensi berhenti menggunakan layanan (churn) dengan *classification models* (misalnya *Logistic Regression* atau *Random Forest*).
- Dapat mengarahkan strategi retensi pelanggan.

## d. Fraud Detection

- Menggunakan *anomaly detection* untuk menemukan transaksi

# Aplikasi Machine Learning dalam Business Intelligence (3)

## e. Sentiment Analysis

- Analisis opini pelanggan di media sosial dengan *Natural Language Processing (NLP)*.
- Membantu perusahaan memahami persepsi publik terhadap produk.



# Tools dan Teknologi yang Digunakan

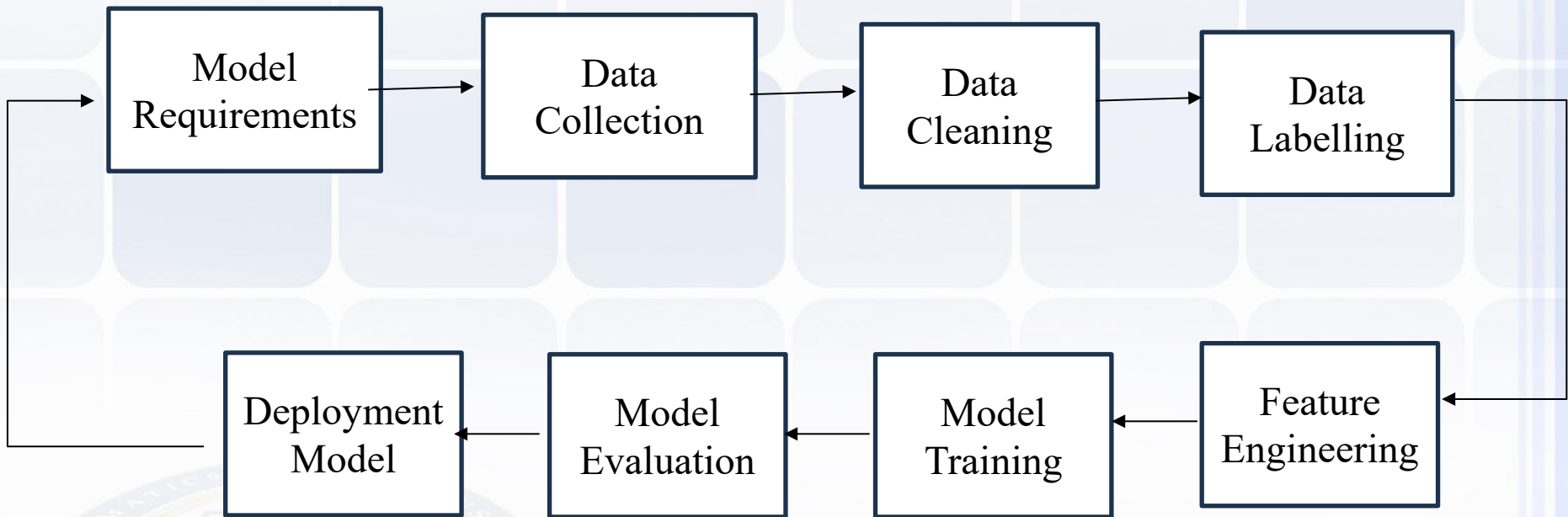
Tujuan	Tools/Platform
Data Integration & Visualization	Power BI, Tableau, Google Data Studio
Machine Learning	Python (Scikit-Learn, TensorFlow), R
Data Storage	SQL Server, BigQuery, Hadoop
Automation	Azure ML, AWS Sagemaker, Google AutoML

# Contoh Studi Kasus

- Kasus: Perusahaan e-commerce ingin meningkatkan penjualan produk elektronik.
- Langkah-langkah BI + ML:
  - Kumpulkan data: riwayat pembelian, usia pelanggan, lokasi, dan rating produk.
  - Gunakan BI: analisis produk paling laku dan waktu pembelian tertinggi.
  - Gunakan ML:
    - Model *classification* untuk memprediksi pelanggan yang berpotensi membeli.
    - Model *recommendation system* untuk menampilkan produk serupa.
  - Hasil: penjualan meningkat 25% setelah penerapan sistem rekomendasi berbasis ML.

- Machine Learning memperkaya Business Intelligence dengan kemampuan prediksi dan otomatisasi keputusan.
- BI menjawab *“apa yang terjadi”*, sedangkan ML menjawab *“apa yang akan terjadi dan apa yang harus dilakukan.”*
- Kombinasi BI dan ML menciptakan Data-Driven Decision Making yang meningkatkan efisiensi, akurasi, dan keuntungan bisnis.

# Machine Learning Life Cycle



# Model Requirements

**Tujuan:** Menentukan kebutuhan dan tujuan model sebelum mulai bekerja dengan data.

Tahap ini menjawab pertanyaan seperti:

- Masalah apa yang ingin diselesaikan?
- Jenis output apa yang diharapkan (prediksi nilai, klasifikasi, rekomendasi, dll)?
- Sumber data apa yang tersedia?

**Contoh (BI):**

- Perusahaan ingin mengetahui produk mana yang akan laris bulan depan.
- Tujuannya: membuat model *sales forecasting* berdasarkan data penjualan historis

# Data Collection

**Tujuan:** Mengumpulkan data yang relevan dari berbagai sumber.

Data dapat berasal dari:

- Sistem internal: ERP, CRM, database penjualan, laporan keuangan.
- Sumber eksternal: media sosial, sensor IoT, survei pelanggan, cuaca, ekonomi, dll.
- Data harus representatif terhadap masalah bisnis.

**Contoh (BI):**

- Mengambil data transaksi pelanggan dari SQL server dan data kampanye iklan dari Google Ads untuk dianalisis bersama.

# Data Cleaning

**Tujuan:** Membersihkan data agar kualitasnya baik dan siap digunakan untuk pelatihan model.

Proses ini mencakup:

- Menghapus data duplikat atau kosong.
- Memperbaiki kesalahan format (tanggal, angka, teks).
- Menangani *outlier* atau nilai ekstrem.
- Menstandarkan satuan atau skala data.

**Contoh (BI):**

- Menghapus entri penjualan dengan harga 0 atau pelanggan tanpa ID.

# Data Labelling

**Tujuan:** Memberi label pada data (khusus untuk supervised learning).

Label adalah nilai target yang ingin diprediksi model.

Misalnya, jika kita ingin memprediksi apakah pelanggan akan churn (berhenti berlangganan), label-nya adalah “Yes” atau “No.”

## **Contoh (BI):**

- Menandai data pelanggan sebagai “loyal” atau “churn” berdasarkan riwayat transaksi 6 bulan terakhir.

# Feature Engineering

**Tujuan:** Mengubah data mentah menjadi fitur (variabel) yang lebih bermakna bagi model.

Langkah ini sering disebut “otak” dari Machine Learning, karena menentukan kualitas input model.

- Kegiatan yang dilakukan:
- Membuat fitur baru dari data yang ada.
- Mengubah kategori ke numerik (*encoding*).
- Melakukan *normalisasi* atau *scaling* data.
- Menghapus fitur yang tidak relevan.

## **Contoh (BI):**

- Mengubah kolom “Tanggal Transaksi” menjadi fitur seperti “Hari dalam seminggu”, “Bulan”, atau “Periode libur”.

# Model Training

**Tujuan:** Melatih model menggunakan data berlabel agar dapat mengenali pola.

Model akan menyesuaikan parameter (bobot) untuk meminimalkan kesalahan prediksi, biasanya menggunakan algoritma seperti:

- *Linear Regression, Decision Tree, Random Forest, Neural Network*, dsb.
- Proses:  
Melibatkan konsep Gradient Descent untuk mengoptimalkan parameter dan menurunkan error.

**Contoh (BI):**

- Melatih model untuk memprediksi penjualan berdasarkan harga, promosi, dan musim.

# Model Evaluation

**Tujuan:** Mengukur performa model dengan data uji yang belum pernah dilihat model.

Gunakan metrik yang sesuai, misalnya:

- Untuk *regresi*: RMSE, MAE,  $R^2$ .
- Untuk *klasifikasi*: Accuracy, Precision, Recall, F1-score.

## **Contoh (BI):**

- Model churn prediction diuji dan menghasilkan akurasi 92% serta recall 0.85 → artinya model cukup baik mendeteksi pelanggan yang berpotensi berhenti.

# Deployment Model

**Tujuan:** Mengimplementasikan model ke sistem nyata agar bisa digunakan dalam pengambilan keputusan otomatis.

Model bisa diintegrasikan ke:

- Dashboard BI (Power BI, Tableau).
- Aplikasi prediksi internal.
- Sistem rekomendasi atau CRM otomatis.
- Setelah deployment, model perlu dimonitor dan diperbarui secara berkala karena data dan pola bisnis berubah seiring waktu (*model drift*).

**Contoh (BI):**

- Model prediksi penjualan dihubungkan langsung ke Power BI dashboard untuk memperkirakan stok mingguan secara real time.

# Keterkaitan dengan Business Intelligence

Aspek	Business Intelligence (BI)	Machine Learning
Fokus	Menganalisis data historis	Memprediksi data masa depan
Hasil	Insight deskriptif	Insight prediktif/preskriptif
Alat	Power BI, Tableau, SQL	Python, R, TensorFlow
Nilai tambah	Laporan & visualisasi	Otomatisasi keputusan

**Akhir bagian Part 6**

