

Nama : Felicia Amanda Cahyadewi

NIM : A11.2022.14090

Matkul : Data Mining A11.4519

## **Prediksi Permintaan Produk pada E-Commerce untuk Optimalisasi Stok**

### **Deskripsi**

**Business Understanding:** Fokus dari penelitian ini adalah untuk memahami pola permintaan produk pada platform e-commerce dengan tujuan meningkatkan efisiensi pengelolaan stok. Dengan memahami pola ini, perusahaan dapat mengurangi kejadian \*stock-out\* yang dapat menyebabkan hilangnya peluang penjualan, serta mengurangi kelebihan stok yang bisa membebani biaya penyimpanan.

**Data Understanding:** Dataset yang digunakan mencakup data penjualan harian, faktor musiman, promosi diskon, dan informasi dari pemasok. Sumber data ini berasal dari sistem manajemen inventaris perusahaan e-commerce atau data historis penjualan yang dimiliki. Atribut-atribut ini diharapkan dapat memberikan gambaran menyeluruh terkait faktor-faktor yang mempengaruhi permintaan produk.

### **Masalah dan Tujuan yang Ingin Dicapai**

Masalah yang ingin diselesaikan dalam penelitian ini adalah ketidakmampuan perusahaan e-commerce dalam memprediksi permintaan produk secara akurat, yang menyebabkan inefisiensi dalam pengelolaan stok. Tujuan utama dari penelitian ini adalah untuk mengembangkan model prediksi permintaan yang dapat memproyeksikan kebutuhan stok secara akurat sehingga dapat membantu perusahaan dalam mengoptimalkan ketersediaan produk di gudang, meningkatkan efisiensi logistik, dan pada akhirnya meningkatkan kepuasan pelanggan.

### **Alur / Tahapan / Kerangka Eksperimen**

1. Data Collection: Mengumpulkan data penjualan historis, data promosi, dan informasi pemasok dari sistem manajemen inventaris perusahaan.
2. Data Preprocessing: Membersihkan data untuk menghapus nilai \*outlier\*, menangani data yang hilang, dan melakukan normalisasi data jika diperlukan.
3. Feature Engineering: Mengidentifikasi fitur penting yang relevan, seperti tren musiman, efek promosi, dan interaksi antara fitur.

4. Model Selection: Menggunakan beberapa model prediksi, seperti *\*time series models\** (ARIMA, SARIMA) dan model berbasis *\*machine learning\** (Random Forest, XGBoost), untuk membandingkan performa dalam memprediksi permintaan.
5. Model Training & Validation: Melatih model menggunakan data historis yang telah dibersihkan dan divalidasi untuk mendapatkan hasil prediksi terbaik.
6. Evaluation: Mengevaluasi performa model menggunakan metrik akurasi seperti *\*Mean Absolute Error\** (MAE) dan *\*Root Mean Square Error\** (RMSE) untuk memilih model terbaik.
7. Implementation Plan: Menyusun rencana implementasi model prediksi terbaik ke dalam sistem manajemen stok perusahaan.

### **Penjelasan Datasets (Sumber data dan penjelasan atribut)**

- Sumber Data: Data akan diambil dari sistem manajemen inventaris e-commerce yang mencakup data penjualan historis dan data pemasok.
- Atribut Data:
  - Tanggal Penjualan: Tanggal penjualan produk yang digunakan untuk analisis *\*time series\**.
  - Jumlah Produk Terjual: Jumlah unit produk yang terjual setiap hari.
  - Musim/Tanggal Spesial: Informasi terkait musim atau hari-hari spesial (misalnya, Hari Raya, diskon liburan) yang dapat mempengaruhi lonjakan permintaan.
  - Diskon: Informasi mengenai diskon yang diterapkan pada produk tersebut.
  - Data Pemasok: Informasi terkait pemasok, seperti waktu pengiriman dan kecepatan restock.