

NAMA : Bernandiko Priyambodo
NIM : A11.2022.14673

Topik : Prediksi Penjualan Produk dengan Time Series Analysis

Deskripsi :

Prediksi penjualan produk adalah langkah strategis yang dapat membantu perusahaan memahami permintaan di masa mendatang, mengoptimalkan persediaan, dan merencanakan strategi pemasaran. Dengan memperkirakan penjualan di masa depan, perusahaan bisa mengurangi risiko *stock-out* atau *overstock*, yang sering kali berdampak pada penurunan pendapatan dan peningkatan biaya operasional. Target dari proyek ini adalah untuk menghasilkan prediksi penjualan yang akurat dalam periode mingguan atau bulanan agar perusahaan dapat mengambil keputusan yang lebih proaktif.

Masalah dan Tujuan yang ingin dicapai :

1. Masalah

Perusahaan sering mengalami kesulitan dalam mengelola persediaan secara efisien karena fluktuasi permintaan produk yang tidak terduga. Ketidaktepatan dalam memperkirakan permintaan dapat menyebabkan beberapa masalah, seperti:

- **Overstock:** Jika terlalu banyak stok yang disimpan, biaya penyimpanan meningkat, dan potensi kerugian akibat produk yang kedaluwarsa atau rusak menjadi lebih besar.
- **Stock-Out:** Ketika permintaan lebih tinggi dari perkiraan, perusahaan berisiko kehabisan stok. Hal ini dapat menurunkan kepuasan pelanggan dan berdampak negatif pada pendapatan.
- **Strategi Pemasaran yang Kurang Efektif:** Ketidaktahuan mengenai waktu puncak permintaan bisa membuat kampanye pemasaran kurang optimal.

2. Tujuan

Project ini bertujuan untuk membangun model prediksi penjualan yang dapat memberikan perkiraan permintaan produk secara lebih akurat untuk periode mendatang. Dengan menggunakan *time series analysis*, tujuan spesifiknya adalah:

- **Memprediksi Permintaan Produk:** Menghasilkan perkiraan penjualan untuk periode mendatang, seperti harian, mingguan, atau bulanan.
- **Optimalisasi Manajemen Persediaan:** Membantu perusahaan memutuskan berapa banyak stok yang perlu disiapkan untuk memenuhi permintaan tanpa terjadi *overstock* atau *stock-out*.
- **Meningkatkan Efektivitas Strategi Pemasaran dan Promosi:** Dengan memahami pola musiman atau tren permintaan, perusahaan dapat menyesuaikan waktu promosi agar sesuai dengan puncak permintaan pelanggan.
- **Penghematan Biaya Operasional:** Meminimalkan biaya terkait penyimpanan dan kerugian produk melalui manajemen persediaan yang lebih baik.

Alur/Tahapan/Eksperimen :

1. Business Understanding

Tahap pertama adalah mendefinisikan tujuan utama dari project, yaitu mengembangkan model prediktif yang dapat membantu perusahaan memproyeksikan penjualan mendatang berdasarkan data historis. Business understanding membantu kita dalam memahami bagaimana hasil model akan digunakan, misalnya untuk optimalisasi persediaan atau penyusunan strategi pemasaran.

2. Data Understanding

Di tahap ini, kita akan mengeksplorasi data untuk memahami karakteristiknya. Beberapa langkah utama yang perlu dilakukan:

- Memeriksa variabel penting, seperti **tanggal pembelian, kuantitas penjualan, jenis produk, dan cabang**.
- Melihat distribusi data, tren penjualan, dan pola musiman, jika ada.
- Mengetahui format data dan melakukan pembersihan data dasar, seperti memeriksa data yang hilang (*missing values*) dan memastikan data berada dalam format yang tepat.

3. Data Preparation

Setelah memahami data, kita melakukan persiapan data yang mencakup:

- **Data Cleaning:** Menghilangkan atau memperbaiki data yang rusak atau tidak valid.
- **Feature Engineering:** Membuat fitur tambahan yang relevan, seperti waktu dalam hari, bulan, atau hari dalam seminggu, yang dapat mempengaruhi pola penjualan.
- **Data Transformation:** Mengubah data ke format yang sesuai untuk analisis, seperti mengagregasi penjualan berdasarkan waktu (misalnya, harian, mingguan, atau bulanan) dan *scaling* data jika diperlukan.

4. Modeling

Di tahap ini, berbagai algoritma *time series* akan diterapkan untuk membuat model prediksi. Beberapa model yang dapat dipertimbangkan meliputi:

- **ARIMA/SARIMA:** Model statistik yang dapat menangani data dengan tren atau musiman.
- **Prophet:** Model dari Facebook yang efektif untuk data musiman dan sangat cocok untuk bisnis kecil hingga menengah.
- **LSTM (Long Short-Term Memory):** Algoritma berbasis jaringan saraf yang mampu menangani dependensi jangka panjang pada data *time series*.

Model-model tersebut dilatih dan diuji untuk menentukan mana yang paling akurat dalam memprediksi penjualan berdasarkan *metrics* seperti MSE (Mean Squared Error) atau MAE (Mean Absolute Error).

5. Evaluation

Mengevaluasi model menggunakan data uji (data yang belum digunakan saat pelatihan) untuk melihat performanya. Selain menggunakan metrik evaluasi, kita juga bisa melakukan *back-testing* untuk melihat prediksi historis dan mengidentifikasi kekuatan serta kelemahan model.

6. Deployment & Monitoring

Jika model prediksi memberikan hasil yang memadai, tahap selanjutnya adalah mengimplementasikan model tersebut dalam lingkungan produksi agar dapat digunakan secara rutin. Ini bisa dilakukan dengan menyimpan model di dalam aplikasi atau sistem yang mengotomatisasi proses prediksi setiap periode tertentu. Monitoring performa model secara berkala juga diperlukan untuk memastikan akurasi tetap konsisten dan melakukan retraining jika diperlukan.

7. Interpretasi & Kesimpulan

Di tahap akhir, hasil prediksi digunakan untuk mengambil keputusan bisnis. Hasilnya bisa disajikan dalam bentuk visualisasi yang mudah dimengerti, seperti grafik tren penjualan yang bisa membantu divisi manajemen dalam menyusun rencana bisnis atau persediaan.

Penjelasan Datasets :

Dataset ini adalah data penjualan historis dari sebuah supermarket besar yang memiliki tiga cabang di berbagai kota. Data ini mencakup periode tiga bulan, dengan informasi rinci tentang transaksi setiap cabang yang dapat digunakan untuk analisis prediktif penjualan.

Dataset ini memiliki beberapa atribut, antara lain:

- **ID Faktur:** Kode unik untuk setiap transaksi yang dihasilkan secara otomatis.
- **Cabang:** Identifikasi cabang supermarket, yang terdiri dari tiga lokasi, masing-masing diberi kode A, B, atau C.
- **Kota:** Lokasi kota dari setiap cabang supermarket.
- **Jenis Pelanggan:** Tipe pelanggan yang terdiri dari "Anggota" bagi pelanggan dengan kartu anggota dan "Normal" untuk pelanggan tanpa kartu anggota.
- **Jenis Kelamin:** Jenis kelamin pelanggan, yang membantu dalam analisis demografis.
- **Lini Produk:** Kategori barang, termasuk Aksesori Elektronik, Makanan & Minuman, Kesehatan & Kecantikan, dan lainnya.
- **Harga Satuan:** Harga per unit dari produk yang dibeli.

- **Kuantitas:** Jumlah produk yang dibeli dalam setiap transaksi.
- **Pajak:** Biaya pajak sebesar 5% yang dikenakan pada setiap pembelian.
- **Total:** Total nilai pembelian setelah pajak ditambahkan.
- **Tanggal & Waktu:** Tanggal dan jam pembelian.
- **Pembayaran:** Metode pembayaran yang digunakan, seperti Tunai, Kartu Kredit, atau E-Wallet.
- **Harga Pokok Penjualan (HPP):** Harga modal produk.
- **Persentase Margin Kotor:** Persentase margin keuntungan kotor untuk produk.
- **Pendapatan Kotor:** Pendapatan total dari penjualan.
- **Peringkat:** Peringkat yang diberikan oleh pelanggan untuk pengalaman belanja mereka, pada skala 1-10.

Dataset ini diambil dari platform *Kaggle*, sebuah sumber populer untuk dataset publik yang mendukung eksplorasi data, analisis tren, dan prediksi penjualan.