

HW_FUAD HASYIM_Business Understanding with Data Metrics

Kasus 1: Evaluasi Performa Model dengan MAE

Kamu sedang mendesain model **regresi** untuk memprediksi harga saham UNVR dengan **Mean Absolute Error (MAE) sebesar Rp900,-.**

- Apakah MAE **Rp900** merupakan nilai yang cukup baik untuk performa model? Jelaskan alasanmu.
- Jika model ini mampu mempertahankan MAE **Rp900** selama satu bulan berturut-turut, apakah model ini dapat digunakan untuk mendapatkan keuntungan? Jelaskan bagaimana strategi yang dapat digunakan.

Goals: Memahami bagaimana nilai MAE dapat diinterpretasikan dalam konteks bisnis dan apakah akurasi model cukup untuk pengambilan keputusan.

Jawaban :

Mean Absolute Error (MAE) digunakan untuk mengukur seberapa besar rata-rata kesalahan prediksi model dibandingkan nilai sebenarnya. Misalnya, kalau MAE-nya Rp. 900, artinya secara rata-rata model meleset sekitar 900 rupiah dari harga saham aktual.

Apakah MAEnya bagus atau tidak?

Kalau kita lihat konteksnya, harga saham UNVR per 31 Oktober 2025 berada dikisaran Rp.2.580. Dengan begitu, kesalahan Rp.900 setara dengan sekitar 35% dari harga aktualnya, yang berarti errornya cukup besar.

Jadi, secara performa, model dengan MAE Rp. 900 **belum bisa dikatakan baik**, karena margin error kesalahannya masih terlalu tinggi untuk digunakan sebagai dasar keputusan investasi atau trading. Dalam konteks harga saham, model yang baik seharusnya punya error yang jauh lebih kecil, misalnya dibawah 5%-10% dari harga aktual sekitar MAE (Rp.100).

Selain itu, kita juga perlu membandingkan dengan baseline model sederhana, seperti:

“Harga besok = harga hari ini.”

Jika model sederhana seperti ini menghasilkan error yang hampir sama, berarti model machine learning kita belum memberikan nilai tambah yang signifikan.

Apakah model ini bisa digunakan untuk mendapatkan keuntungan?

Jika model dengan MAE RP.900 digunakan secara konsisten selama sebulan penuh, tetapi belum bisa dipastikan akan menghasilkan keuntungan.

Dalam konteks trading, akurasi prediksi harga bukan satu-satunya faktor yang menentukan profitabilitas.

Ada beberapa hal yang harus dipertimbangkan:

1. Arah pergerakan harga : meskipun selisihnya kecil, tapi kalau arah prediksinya salah (model prediksi naik padahal turun), tetap akan rugi
2. Biaya transaksi dan risiko volatilitas : pada saham seperti UNVR, fluktuasi harian bisa mencapai ratusan rupiah. Jadi error Rp.900 bisa membuat keputusan beli/jual jadi tidak akurat
3. Tujuan jenis model : kalau tujuannya untuk trend analisis error sebesar itu mungkin bisa di toleransi, tapi bukan untuk strategi jual-beli langsung.

Strategi yang bisa dilakukan:

- Gunakan model sebagai alat pendukung keputusan, bukan satu-satunya dasar. Misalnya, model bisa membantu mengidentifikasi tren jangka pendek atau potensi titik pembalikan harga.
- Gabungkan dengan indikator teknikal lain, seperti moving average, RSI, atau sentiment analysis dari berita dan media sosial untuk meningkatkan akurasi arah harga.

Kasus 2: Root Cause Analysis dalam Bisnis

Kamu bekerja di perusahaan properti, dan **manajer kamu menyampaikan bahwa volume transaksi mengalami penurunan.**

- Apa yang akan kamu lakukan untuk mencari **root cause** dari penurunan ini? Sebutkan minimal 3 faktor yang berpotensi menyebabkan penurunan tersebut dengan menggunakan metode **slice and dice**.

Goals: Menggunakan teknik analisis data untuk mengidentifikasi penyebab utama dari perubahan dalam metrik bisnis dan menyusun rekomendasi yang sesuai.

Jawaban:

Ketika volume transaksi menurun, langkah pertama yang perlu dilakukan adalah mencari penyebabnya secara sistematis. Pendekatan yang bisa digunakan adalah metode **slice and dice**, yaitu memecah data ke dalam beberapa dimensi untuk mencari pola penyebab.

Kita bisa menganalisis data berdasarkan beberapa dimensi seperti:

- **Waktu:** perbandingan bulan ke bulan atau tahun ke tahun

- **Lokasi:** wilayah kota tertentu (misalnya pusat kota vs pinggiran)
- **Tipe properti:** rumah, apartemen, tanah kavling, dsb
- **Channel penjualan:** online ads, agen, pameran properti, dll

Dengan cara ini, kita bisa “memotong” data dari berbagai sisi untuk menemukan area mana yang paling terpengaruh.

Beberapa faktor potensial penyebab penurunan transaksi:

1. Faktor Ekonomi dan Daya Beli Masyarakat

Penurunan volume transaksi sering kali disebabkan oleh faktor eksternal seperti naiknya suku bunga KPR, inflasi, atau ketidakpastian ekonomi.

Jika bunga KPR naik, cicilan rumah otomatis lebih mahal hal ini membuat masyarakat menunda pembelian properti.

Analisis yang bisa dilakukan:

- a. Lihat data leads dari setiap channel (Facebook Ads, Google Ads, agen, pameran).
- b. Bandingkan conversion rate bulan ini dengan bulan sebelumnya.
- c. Perhatikan juga engagement rate dari konten promosi.

2. Kinerja Pemasaran dan Channel Penjualan

Bisa jadi penyebabnya berasal dari sisi internal. Misalnya, anggaran iklan digital berkurang, algoritma media sosial berubah, atau kampanye promosi kurang menarik.

Akibatnya, jumlah prospek (leads) menurun dan berdampak langsung pada penjualan.

Analisis yang bisa dilakukan:

- a. Lihat data leads dari setiap channel (Facebook Ads, Google Ads, agen, pameran).
- b. Bandingkan conversion rate bulan ini dengan bulan sebelumnya.
- c. Perhatikan juga engagement rate dari konten promosi.

3. Masalah pada Produk (Properti) Itu Sendiri

Faktor lain bisa datang dari sisi produk, misalnya harga yang terlalu tinggi dibanding pesaing, lokasi yang kurang strategis, atau stok unit baru yang terbatas.

Akibatnya, jumlah prospek (leads) menurun dan berdampak langsung pada penjualan.

Analisis yang bisa dilakukan:

- a. Lihat tren harga properti kompetitor di wilayah yang sama.
- b. Evaluasi tingkat ketersediaan unit (supply) dan kecepatan penjualan (days on market).
- c. Gunakan feedback pelanggan untuk menilai apakah ada faktor kualitas atau desain yang memengaruhi minat beli.

Langkah Lanjutan Setelah Analisis

Setelah melakukan analisis “slice and dice” dan mengidentifikasi faktor penyebab, langkah berikutnya adalah **membuat rekomendasi strategis**.

Contohnya:

- Jika penurunan terjadi karena suku bunga tinggi : fokus pada strategi rent-to-own atau cashback program.
- Jika masalahnya ada di marketing : optimalkan kembali digital campaign dan buat konten promosi berbasis insight audiens.
- Jika dari sisi produk : pertimbangkan bundling harga, promo DP ringan, atau diversifikasi tipe unit yang lebih terjangkau.

Kasus 3: Peramalan Stok Barang di TokoShopea

Kamu bertugas untuk memberikan saran kepada **seller di TokoShopea** mengenai jumlah stok barang yang harus mereka siapkan di minggu depan. Tujuan dari peramalan ini adalah untuk **mengurangi kasus Out of Stock serta menghindari stok berlebih** yang dapat membebani logistik penyimpanan.

- Desainlah sebuah proyek machine learning yang dapat membantu dalam peramalan stok barang, termasuk pemilihan algoritma yang sesuai, success metrics yang digunakan, serta bagaimana model ini dapat membantu seller dalam meningkatkan efisiensi bisnis mereka.

Goals: Menyusun solusi berbasis machine learning yang efektif untuk membantu seller dalam pengelolaan stok barang.

Jawaban:

Salah satu tantangan utama bagi para penjual (*seller*) di platform e-commerce seperti **TokoShopea** adalah menentukan jumlah stok barang yang ideal untuk minggu berikutnya. Jika stok terlalu sedikit, produk bisa cepat habis (*out of stock*) dan penjualan hilang. Tapi kalau stok terlalu banyak, biaya penyimpanan dan risiko barang tidak laku justru meningkat. Untuk membantu masalah ini, kita bisa merancang **solusi machine learning** yang mampu melakukan **peramalan stok (stock forecasting)** berdasarkan data historis penjualan.

Tujuan Proyek

Tujuan utama proyek ini adalah untuk memperkirakan jumlah stok optimal per produk (SKU) agar:

- Kasus *out of stock* bisa diminimalkan.
- Stok berlebih (*overstock*) dapat dikurangi untuk efisiensi biaya gudang.
- Seller dapat mengambil keputusan restock dengan lebih data-driven.

Dengan kata lain, model ini diharapkan bisa membantu seller menjaga keseimbangan antara permintaan pasar dan kapasitas penyimpanan.

Data yang Diperlukan

Agar model dapat bekerja dengan baik, beberapa jenis data yang dibutuhkan antara lain:

- **Data penjualan historis** per produk (harian atau mingguan).
- **Data stok gudang** saat ini.
- **Lead time** (waktu yang dibutuhkan supplier untuk restock).
- **Harga dan promo** (diskon, campaign, flash sale).
- **Event atau musim** (misalnya Ramadan, akhir tahun, atau promo 11.11).
- **Kategori produk dan rating** (karena kategori dan ulasan sering memengaruhi tingkat permintaan).

Dengan data tersebut, kita bisa menangkap pola permintaan, tren musiman, dan pengaruh faktor eksternal terhadap penjualan.

Rancangan Solusi Machine Learning

1. Preprocessing & Feature Engineering

- Membersihkan data penjualan dari anomali (misalnya lonjakan ekstrem karena promo besar).
- Membuat fitur tambahan seperti:
 - a. Rata-rata penjualan 7 hari terakhir.
 - b. Jumlah hari sejak stok terakhir diisi ulang.
 - c. Indikator event promo (1 jika ada promo, 0 jika tidak).
 - d. Trend musiman berdasarkan minggu atau bulan.

2. Pemilihan Algoritma

Random Forest Regressor atau **LightGBM** untuk menangani banyak variabel. Model ini akan memprediksi *demand* (permintaan) untuk minggu depan per produk.

3. Output dari model

Model akan memberikan prediksi jumlah permintaan per SKU untuk minggu berikutnya. Dari situ, kita bisa menghitung rekomendasi stok sebagai:

$$\text{Rekomendasi_Stok} = \text{Prediksi_Demand} + \text{Safety_Stock} - \text{Stok_Sekarang}$$

Di mana *Safety Stock* digunakan sebagai cadangan jika permintaan tiba-tiba naik.

4. Metrics dan evaluasi

Untuk mengukur performa model dan dampak bisnisnya, beberapa metrik yang bisa digunakan adalah:

- a. **MAE (Mean Absolute Error)**: mengukur rata-rata selisih antara permintaan aktual dan prediksi.
- b. **RMSE (Root Mean Squared Error)**: menilai tingkat error dengan penalti lebih besar untuk kesalahan besar.

5. Dampak dan Manfaat Bisnis

Dengan menerapkan sistem peramalan stok ini, seller di TokoShopea bisa mendapatkan berbagai manfaat nyata:

- **Efisiensi operasional meningkat**, karena keputusan restock tidak lagi didasarkan pada intuisi, melainkan pada prediksi berbasis data.
- **Penjualan meningkat**, karena produk tidak kehabisan stok saat permintaan tinggi.
- **Biaya gudang lebih terkendali**, sebab stok berlebih dapat diminimalkan.
- **Perencanaan logistik lebih baik**, terutama untuk seller yang memiliki lead time panjang dari supplier.