

FINAL PROJECT - STAGE 3

Kelompok Datactives:

- Muhammad Syahril
- Glody Syah Rabbynawa
- Yudha Adi Putra
- Dendi Pratama Putra Pamungkas
- Sailina Karimah Syarifudin
- Susi Amelia
- Ummul Khaeria Rasyid
- Sindi Mustika Putri



Summary

Objective: Mengurangi jumlah false negative untuk menghindari customer yang pengirimannya terlambat tetapi diprediksi sampai tepat waktu

Metrics model: Recall

		Actual	
		1 (terlambat) Positive	0 (tepat waktu) Negative
Prediksi	1 (terlambat) Positive	TP	FP
	0 (tepat waktu) Negative	FN	TN



Target mengurangi
False Negative

Summary Modeling Score

Before (Default Parameters):

Classification	Train Score	Test Score	Accuracy	Precision	Recall	F1 Score	AUC ROC
Random Forest	1.0	0.64	0.64	0.74	0.62	0.67	0.64
KNN	0.71	0.63	0.63	0.77	0.56	0.65	0.70
Decision Tree	1.0	0.64	0.62	0.67	0.68	0.67	0.65
XGBoost	0.90	0.64	0.64	0.71	0.64	0.68	0.64
LightGBM	0.80	0.65	0.65	0.75	0.60	0.67	0.73
Logistic Regression	0.63	0.62	0.62	0.66	0.72	0.69	0.59

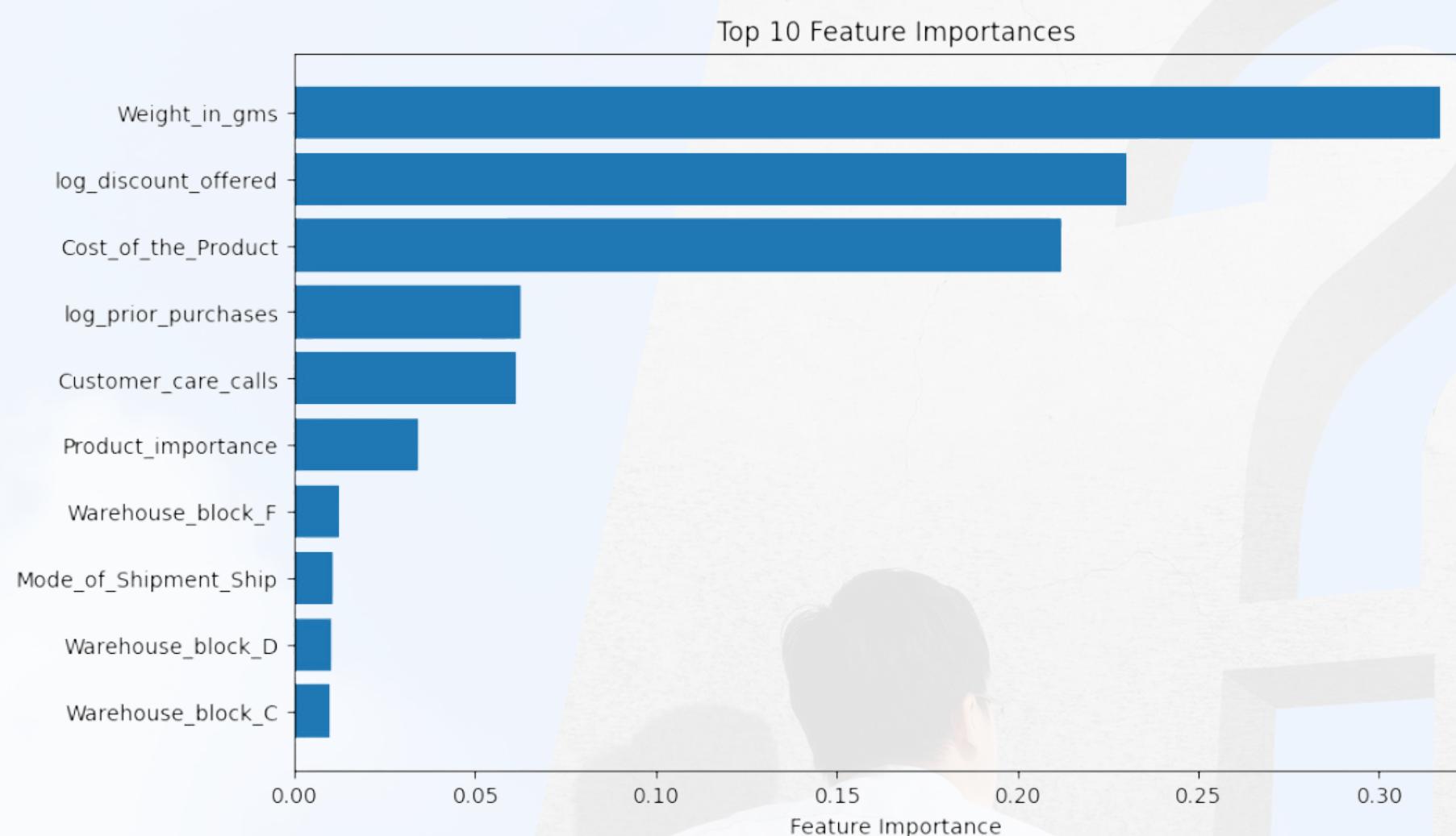
Summary Modeling Score

After (Tuned Parameters) :

Classification	Train Score	Test Score	Accuracy	Precision	Recall	F1 Score	AUC ROC
Random Forest	0.64	0.67	0.63	0.73	0.62	0.67	0.64
KNN	0.71	0.63	0.63	0.77	0.56	0.65	0.70
Decision Tree	0.79	0.73	0.66	0.72	0.69	0.69	0.65
XGBoost	0.86	0.72	0.66	0.78	0.58	0.67	0.68
LightGBM	0.80	0.65	0.68	0.89	0.51	0.64	0.74
Logistic Regression	0.71	0.69	0.60	0.61	0.99	0.75	0.51

Berdasarkan skor metrics dari beberapa modeling yang telah kami lakukan, kami memutuskan untuk memilih model **Decision Tree**

Feature Importance



Dengan menggunakan asumsi bahwa fitur yang memiliki tingkat kepentingan yang tinggi adalah fitur dengan nilai **lebih dari 0.05**, pada data berikut memiliki 5 fitur penting:

1. Weight in grams
2. Discount offered
3. Cost of the product
4. Prior purchase
5. Customer care calls

Berdasarkan korelasi dari kolom tersebut dengan kolom reach on time, dapat disimpulkan bahwa:

1. Semakin **berat barang** yang dikirim, semakin **rendah** kemungkinan pengiriman akan terlambat
2. Semakin **besar diskon** yang diberikan pada transaksi tersebut, semakin **tinggi** kemungkinan pengiriman akan terlambat
3. Semakin **tinggi harga** barang yang dikirimkan, semakin **rendah** kemungkinan pengiriman akan terlambat
4. Semakin **banyak histori pembelian** dari pembelian yang dilakukan, semakin **rendah** kemungkinan pengiriman akan terlambat
5. Semakin **banyak jumlah customer care calls** yang dilakukan, semakin **rendah** kemungkinan pengiriman akan terlambat

Business Insight

- 1. Weight in Grams (Berat Barang):** Semakin berat barang yang dikirim, semakin rendah kemungkinan pengiriman akan terlambat. Hal ini bisa menjadi informasi yang berguna bagi perusahaan dalam perencanaan pengiriman, terutama jika terdapat pengelompokan untuk barang-barang yang lebih berat memerlukan proses pengiriman khusus dengan mode transportasi khusus yang menyebabkan waktu pengiriman yang berbeda.
- 2. Discount Offered (Diskon yang Ditawarkan):** Semakin besar diskon yang diberikan pada transaksi, semakin tinggi kemungkinan pengiriman akan terlambat. Ini mungkin mengindikasikan bahwa diskon yang besar mungkin mengakibatkan peningkatan permintaan, yang kemudian dapat memengaruhi ketersediaan atau kelancaran pengiriman.
- 3. Cost of the Product (Harga Produk):** Semakin tinggi harga barang yang dikirimkan, semakin rendah kemungkinan pengiriman akan terlambat. Ini bisa berarti bahwa barang-barang dengan harga lebih tinggi mungkin mendapatkan prioritas dalam proses pengiriman atau bahwa pelanggan yang membayar lebih banyak mungkin mendapatkan layanan yang lebih baik.
- 4. Prior Purchase (Histori Pembelian Sebelumnya):** Semakin banyak histori pembelian dari pembelian yang dilakukan, semakin rendah kemungkinan pengiriman akan terlambat. Ini menunjukkan bahwa pelanggan yang sering berbelanja atau memiliki sejarah pembelian yang baik mungkin lebih cenderung menerima pengiriman tepat waktu dikarenakan alamat tujuan yang sudah tercatat sehingga meningkatkan efektivitas pengiriman.
- 5. Customer Care Calls (Panggilan Layanan Pelanggan):** Semakin banyak jumlah customer care calls yang dilakukan, semakin rendah kemungkinan pengiriman akan terlambat. Ini bisa berarti bahwa pelanggan yang aktif dalam berkomunikasi dengan layanan pelanggan cenderung mendapatkan perhatian lebih dan solusi yang lebih baik untuk masalah pengiriman.

Wawasan-wawasan ini dapat membantu perusahaan untuk mengoptimalkan operasi mereka, memahami preferensi pelanggan, dan meningkatkan layanan pengiriman dengan lebih baik. Namun, perlu diingat bahwa wawasan-wawasan ini didasarkan pada asumsi bahwa fitur-fitur yang teridentifikasi adalah yang paling penting, dan hasilnya mungkin dapat bervariasi dengan analisis lebih lanjut.

Rekomendasi



- 1. Optimisasi Berat Pengiriman:** Dapat mempertimbangkan untuk mengoptimalkan pengiriman barang dengan mempertimbangkan beratnya. Mungkin bisa mengkategorikan barang-barang berdasarkan beratnya dan mengembangkan strategi pengiriman yang berbeda untuk setiap kategori.
- 2. Manajemen Diskon:** Perhatikan dampak diskon pada kinerja pengiriman dan pastikan bahwa diskon yang ditawarkan sesuai dengan kemampuan untuk memenuhi janji pengiriman tepat waktu.
- 3. Manajemen biaya produk:** mengelompokkan prioritas pengiriman berdasarkan biaya yang perlu dikeluarkan oleh konsumen. Perusahaan dapat memberikan prioritas pengiriman yang tepat waktu kepada barang-barang dengan harga lebih tinggi. Perusahaan juga dapat membuat layanan asuransi penjaminan barang sampai tepat waktu dengan menambahkan sejumlah biaya.
- 4. Manajemen histori pembelian:** Perusahaan dapat memanfaatkan riwayat pembelian pelanggan untuk meningkatkan efektivitas pengiriman. Informasi tentang alamat tujuan yang sudah tercatat atau preferensi pengiriman pelanggan dapat membantu dalam perencanaan pengiriman yang lebih baik. Perusahaan bisa menerapkan strategi yang berfokus pada pelanggan yang sering berbelanja dengan memberikan layanan eksklusif atau penawaran pengiriman yang lebih baik, misalnya dengan membuat program untuk customer loyal.
- 5. Optimisasi layanan call center:** Peningkatan jumlah panggilan layanan pelanggan sebagai peluang untuk memberikan solusi yang lebih baik. Perusahaan dapat meningkatkan proses penanganan keluhan dan pertanyaan pelanggan yang aktif berkomunikasi dengan layanan pelanggan.

Rekomendasi ini dapat membantu perusahaan dalam mengoptimalkan operasional pengiriman, meningkatkan kepuasan pelanggan, dan menghindari masalah pengiriman terlambat. Dengan memperhatikan setiap insight ini, perusahaan dapat mengidentifikasi cara-cara yang lebih baik untuk mengelola pengiriman dan layanan pelanggan mereka.



Kelompok: Dataactive

Stage: 3

Mentor: Muhammad Hendrawan Hidayat

Pukul/ Tanggal: 21.00 - 21.00 WIB / 2 November 2023

Pembagian tugas di stage ini:

- Random Forest (Sailina)
- KKN (Dendi)
- Decision Tree (Yudha)
- XGBoost (Sindi, Aril)
- LightGBM (Ummul, Susi)
- Logistic Regression (Gloody)

Poin pembahasan:

- Modeling
- Feature Importance

Hasil Diskusi:

- Waspada nilai 1 dan 0 pada kolom reached on time
- Gunakan scaling di data test dan data train setelah melakukan split data
- Gunakan feature importance untuk mencari business insight

Tindak Lanjut:

1. Menggunakan recall untuk metrics score
2. Melakukan scaling di data test pada seluruh kolom
3. Menggunakan data baru yang sudah clean