Enterprise Data Management

TOKOTRONIC Shipping

About the Company

TOKOTRONIC

is an e-commerce electronic platform in Indonesia that has its own courier service " TOKOTRONIC SHIPPING"

Arbi Indrawan

Syafira Rizka Kurniawati

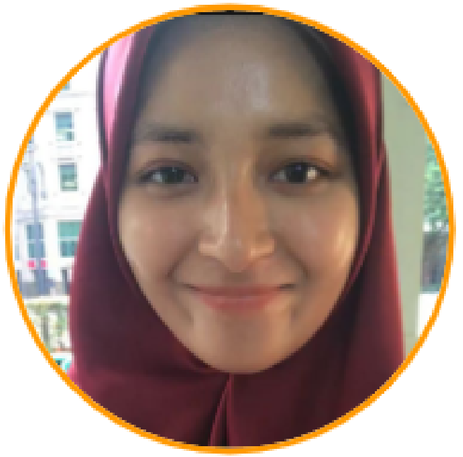
Irena Chandra

Denilson Pasaribu

Elprida Agustina

Arif Ibrahim Ginanjar

Nur Purwanto



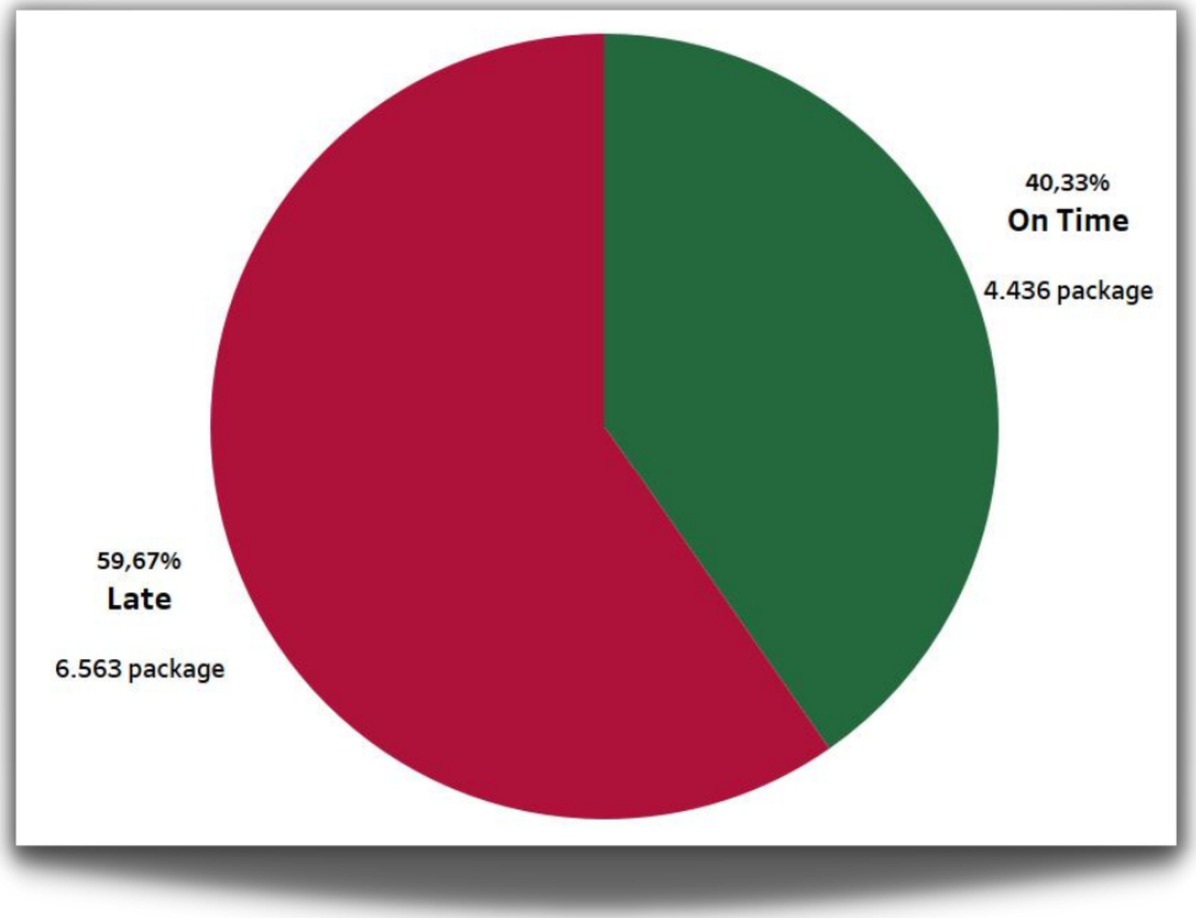
Intan Furqona T

We are Entreprise Data Management, an internal Data Scientist at TOKOTRONIC

Agenda

Problem Statement

Tokotronic Shipping has late delivery rate of 59.67% (6.563 packages)



If late deliveries is not taken care of:
Dissatisfied experience and discourage customers to purchase products through our platform

Eventually, in the long run it will:
Decrease C*ustomer Retention Rate (CRR)*,
Decrease *Customer Lifetime Value (CLV)*.
Increase *Customer Acquisition Cost (CAC)* :

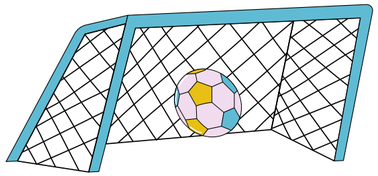
Impact:
Distrust in the community
High potential loss of revenue
Negative reviews
Increase in marketing cost

Statistics

69% of consumers are much less or less likely to shop with a retailer in the future if an item they purchased is not delivered within two days of the date promised [Voxware]

17% of respondents will stop shopping with a retailer after receiving a late delivery one time [Voxware]

55% of respondents will stop shopping with a retailer after receiving a late delivery two to three times [Voxware]

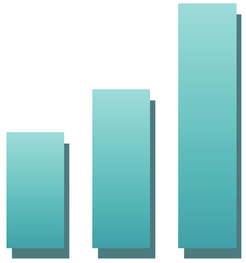


Goal

Increase on-time delivery rate to increase customer retention rate

Objectives

- Identify the potential causes of late deliveries
- Develop a Machine Learning model to predict whether an item will be delivered on time or not
- Evaluate and recommend sensible business solutions to increase on time delivery and improve overall delivery process



Business Metrics

On time delivery rate

Exploratory Data Analysis

Data Features

Our dataset contains
12 features and
10999 row of data

Missing & Duplicated Data

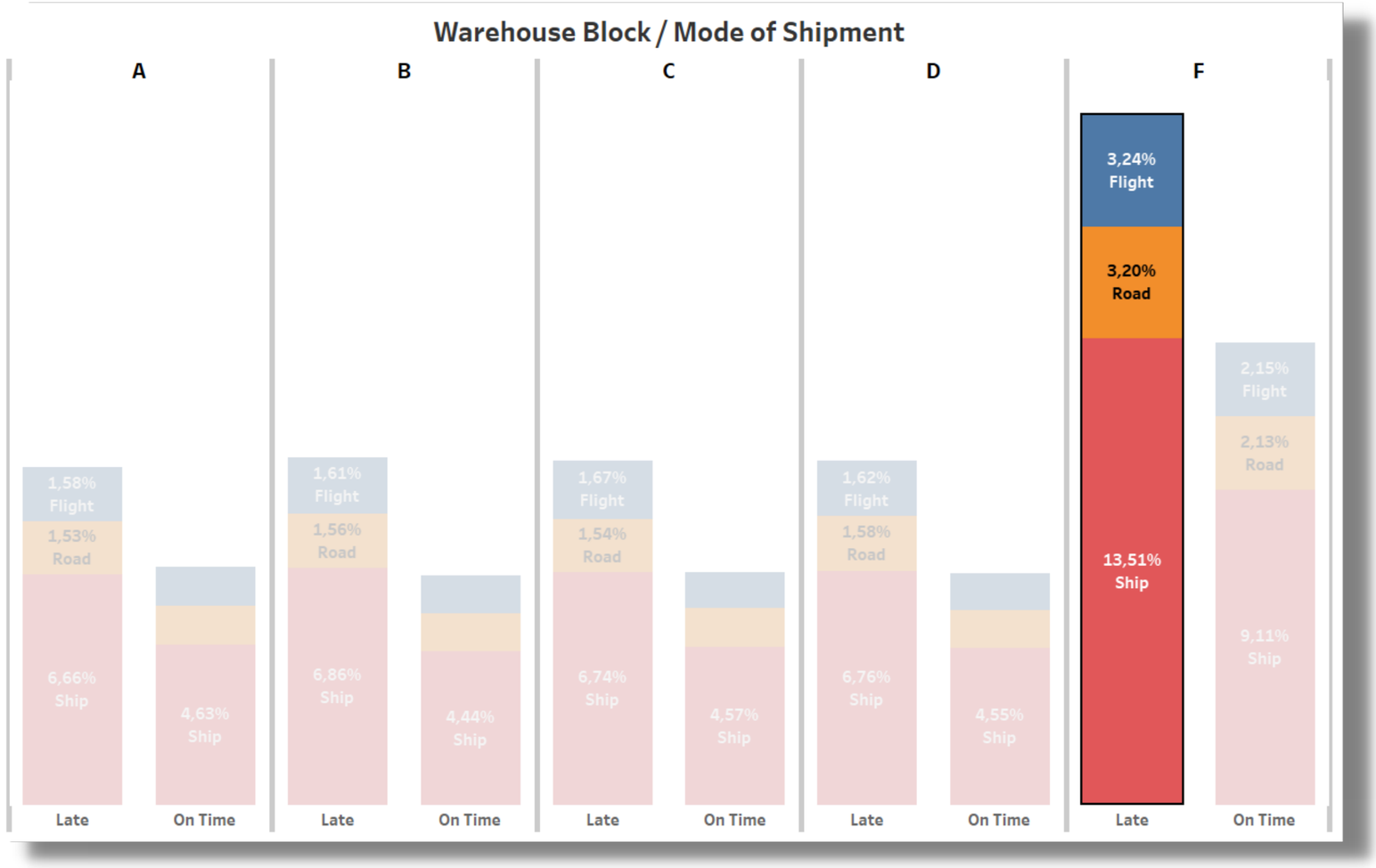
There is no Missing and
Duplicated Data

Data Type

Turning
Customer_rating data type
and Reached.on.Time_Y.N
from Integer to Object

Exploratory Data Analysis

- Warehouse F **contributes 2x the percentage of late deliveries**.
- This is presumably because the capacity of the warehouse is indeed large or because the warehouse is overloaded so the delivery management is not going well.

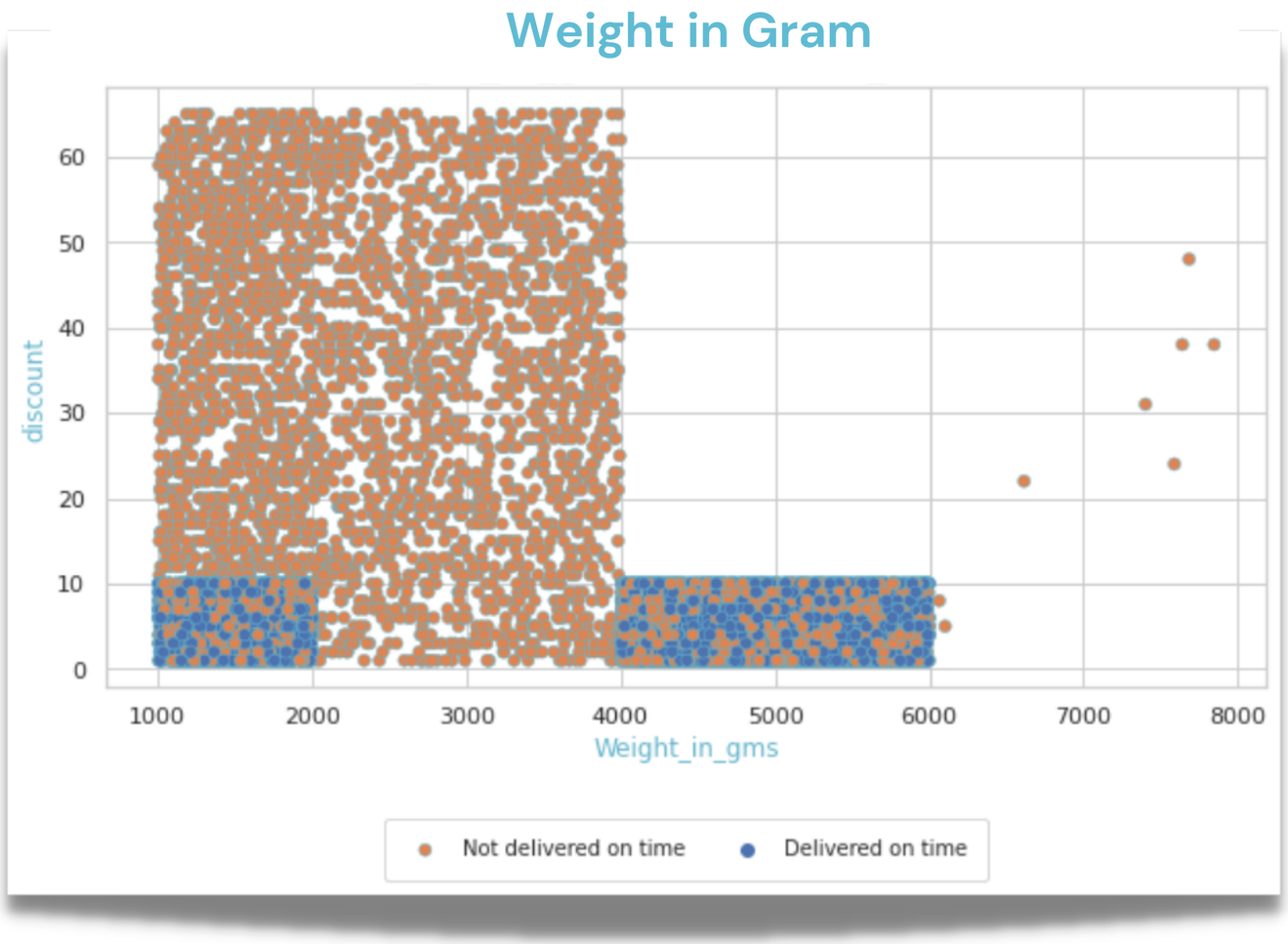


Overall the delay in delivery is by ship. In warehouse F, the shipping portion by ship is the highest 13.51%.

Exploratory Data Analysis

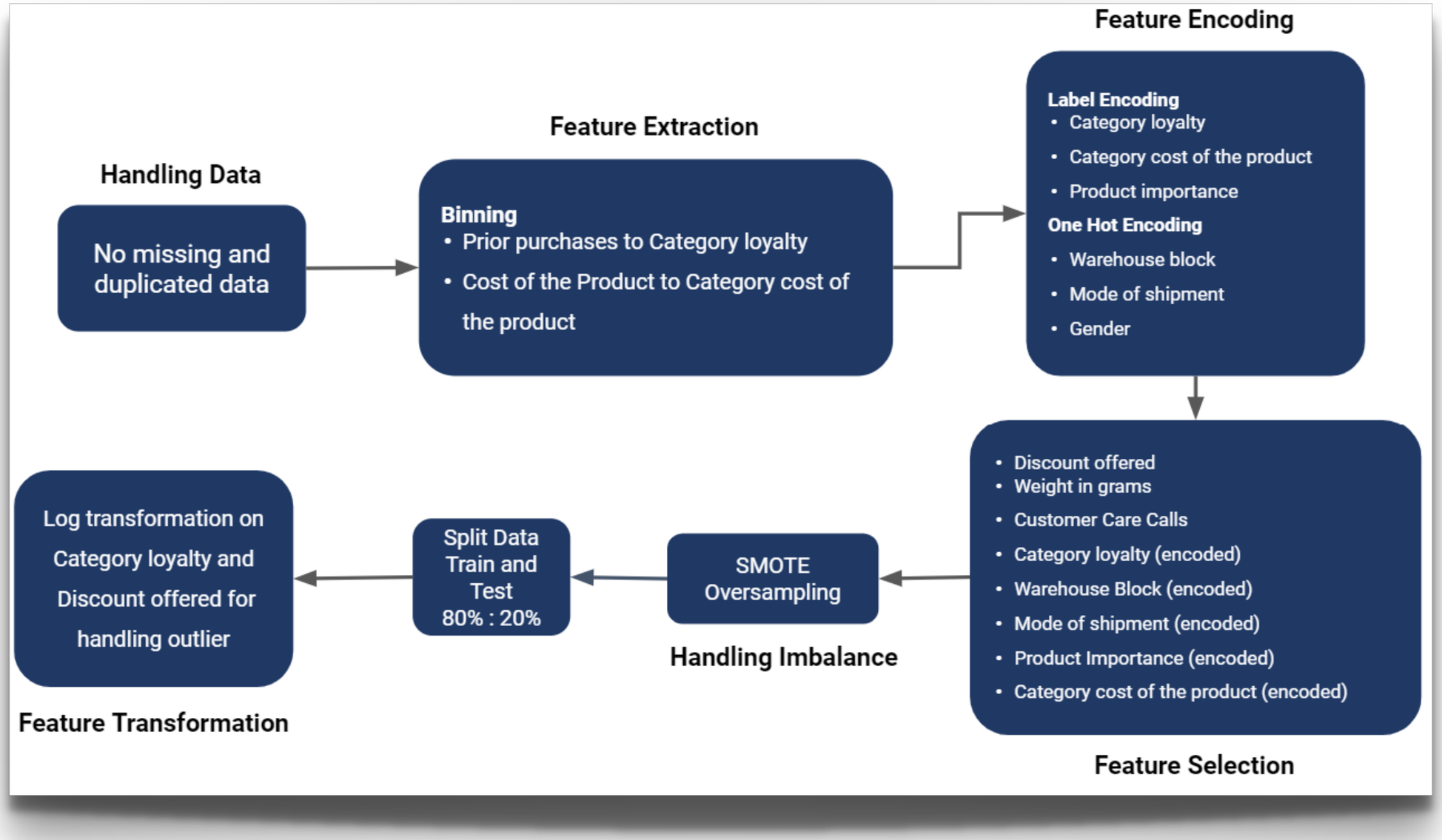
All discounted goods above 10% are more likely to experience delays

Exploratory Data Analysis



Products with weight 2000 – 4000 gram are more likely to experience delays

Data Preprocessing



Best Fit Model Machine Learning Evaluation

	Accuracy	Precision	Recall	ROC AUC Train	ROC AUC Test
Logistic Regression	0.69	0.70	0.66	0.80	0.79
KNN	0.72	0.77	0.62	0.89	0.78
Adaboost	0.92	0.89	0.93	**0.99**	**0.97**
Decision Tree	0.70	0.69	0.71	1	0.70
XGBoost	0.95	0.93	0.97	1	0.99

Best fit model : Adaboost

ROC AUC Score 0.97

High evaluation ROC AUC score value
=> High level of robust machine learning
=> High accurate prediction results
=> Making the right business decision

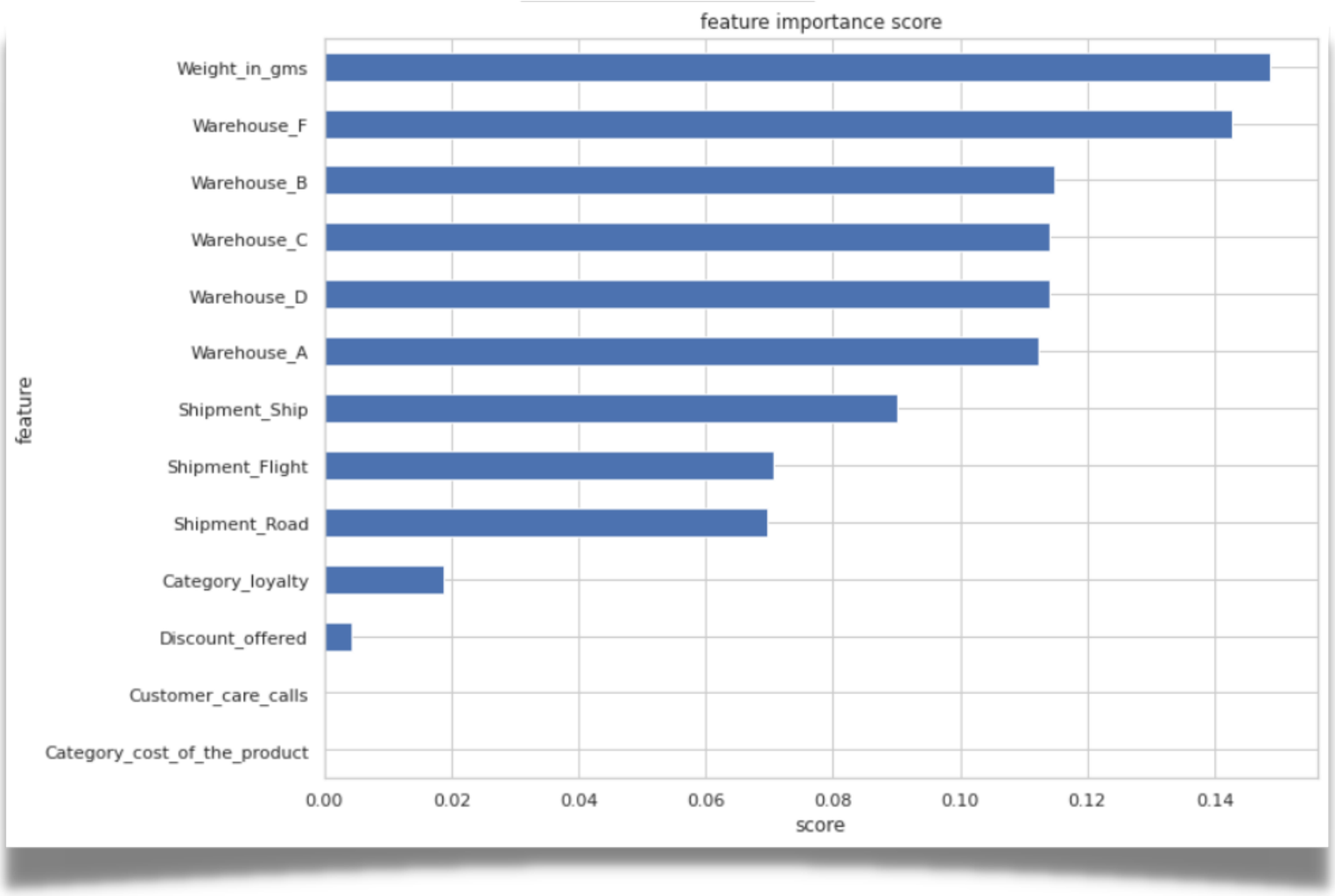
Hyperparameter Tunning

	Accuracy	Precision	Recall	ROC AUC Train	ROC AUC Test
Logistic Regression	0.69	0.77	0.53	0.72	0.72
KNN	0.73	0.78	0.63	0.94	0.79
Adaboost	0.93	0.91	0.93	**0.98**	**0.97**
Decision Tree	0.70	0.74	0.62	0.87	0.80
XGBoost	0.94	0.93	0.97	1	0.99

Feature Importance

Most important feature :

- **Weight in gms**
- **Warehouse**
- **Mode of shipment**

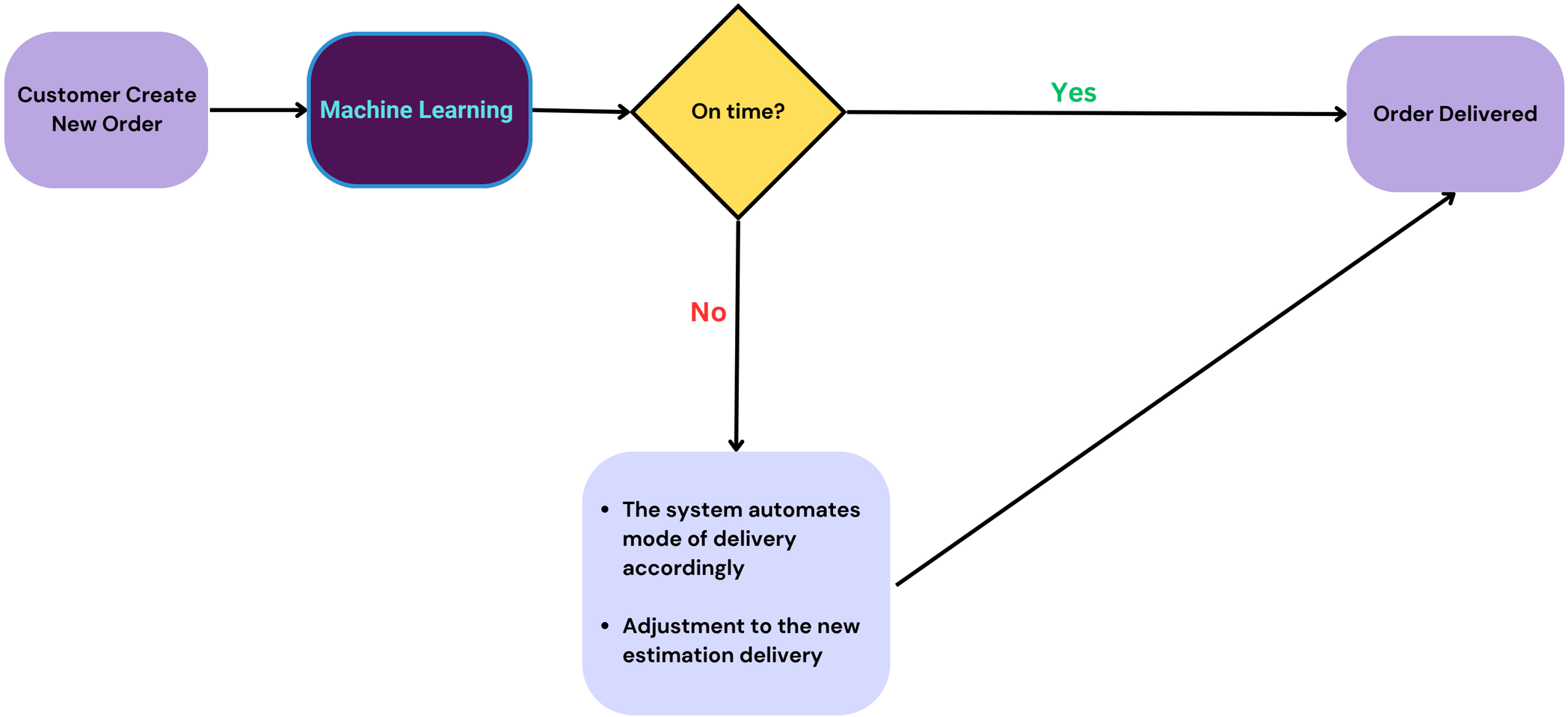


Implementation Machine Learning

Checkout Flow Before Machine Learning Model

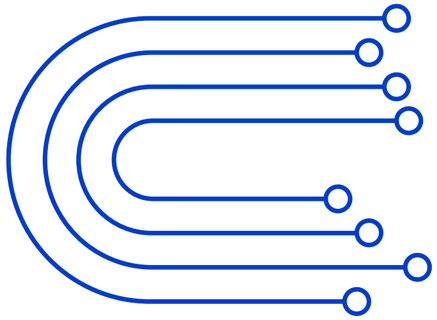
Implementation Machine Learning

Checkout Flow with Machine Learning Model



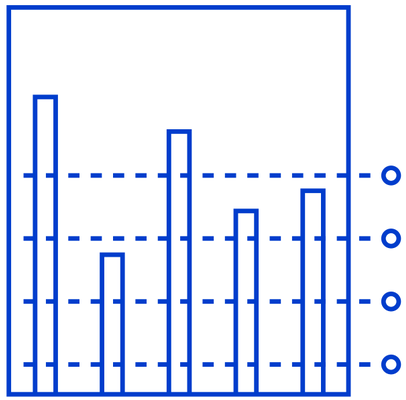
Business Recommendation

- Utilizing fulfillment centers to help ecommerce businesses in managing order fulfillment operations
- Employ freelancers to help operational process during discount season
- Drone deliveries for fast-moving goods and goods that are potentially to experience late deliveries with distance constrain



Expected Outcomes

Increase On Time Delivery rate --> increase Customer Retention Rate (CRR) --> increase Customer Lifetime Value (CLV) --> decrease Customer Acquisition Cost (CAC)



Impacts

High OTD rate can be a strong competitive advantage
Company gains a reputation for reliability and trustworthiness
Positive reviews / referrals / recommendations from customer
Save costs from decreasing of CAC
Attract new customers

Conclusion

Machine Learning Model → Adaboost with 97% accuracy prediction classifies on time and late items

Future Implementation →

- Improve on time delivery
- Increase CRR
- Increase revenue

Appendix

- *https://www.websitemagazine.com/blog/the-impact-of-late-and-inaccurate-deliveries-on-customer-loyalty*
- *https://www.oracle.com/corporate/pressrelease/returns-in-store-experiences-101519.html*
- *https://www.supplychainbrain.com/articles/14912-impact-of-late-or-inaccurate-deliveries-can-be-disastrous-study-shows*
- *https://neilpatel.com/blog/customer-acquisition-cost/*
- *https://www.salesforce.com/products/marketing-cloud/customer-interaction/*
- *https://blog.hubspot.com/service/how-to-calculate-customer-lifetime-value*

"Trust us, your data will be useful"

Enterprise Data Management

- Arbi Indrawan
- syafirark
- Irena Chandra
- denilson-joshua
- Intan Furqona Trisnaningrum
- Elprida Agustina
- nurpurwanto
- arif-ibrahim-g