Delivery Outcome Classification: MLP & LSTM Analysis

Project By: Phat Vuong & Hieu Hua

Date: 12th Mar 2025

Introduction

***Objective**: Develop a deep learning model to predict delivery delays.*

***Problem Statement**:*

•Late deliveries impact operations & customer satisfaction.

•Multi-label classification challenge for accurate delay predictions.

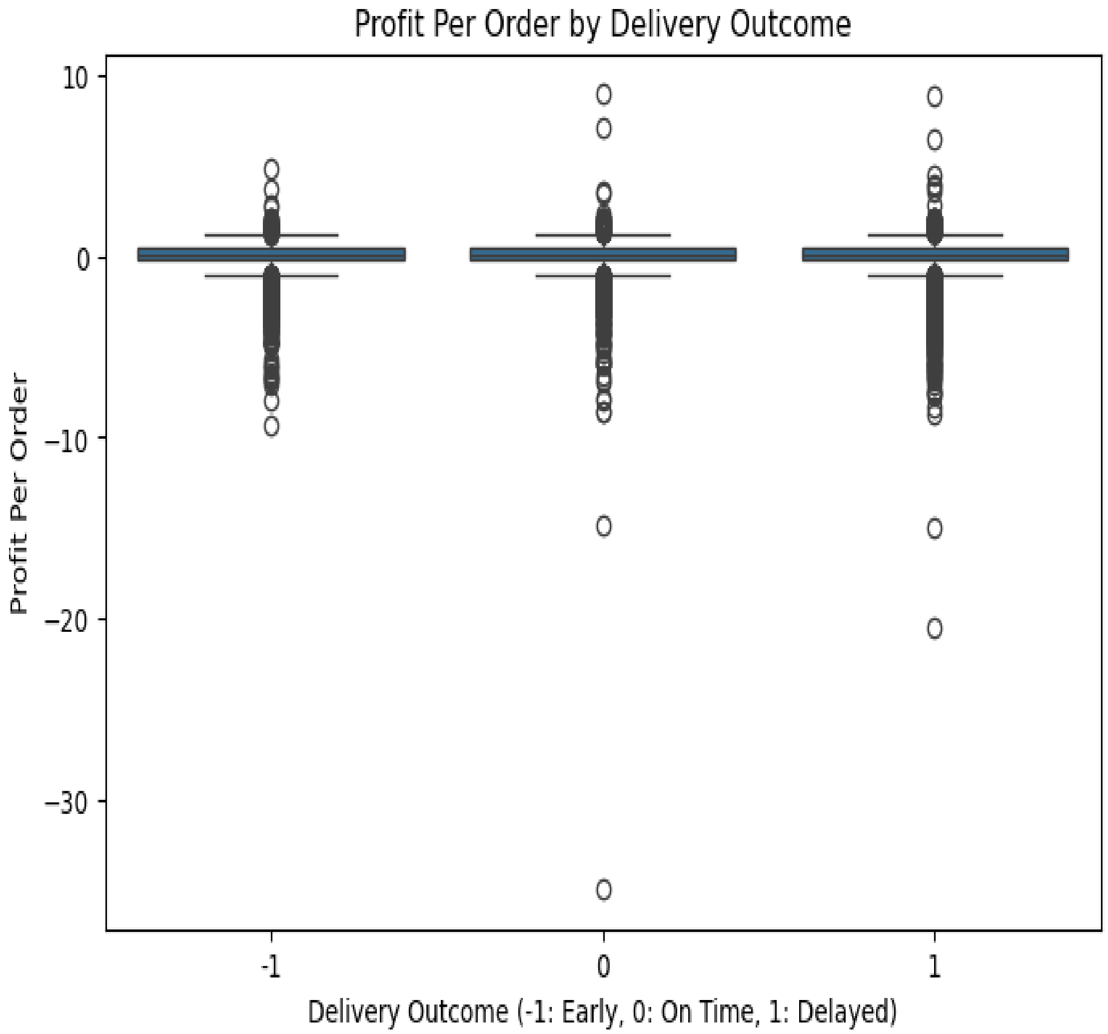
***Dataset**:*

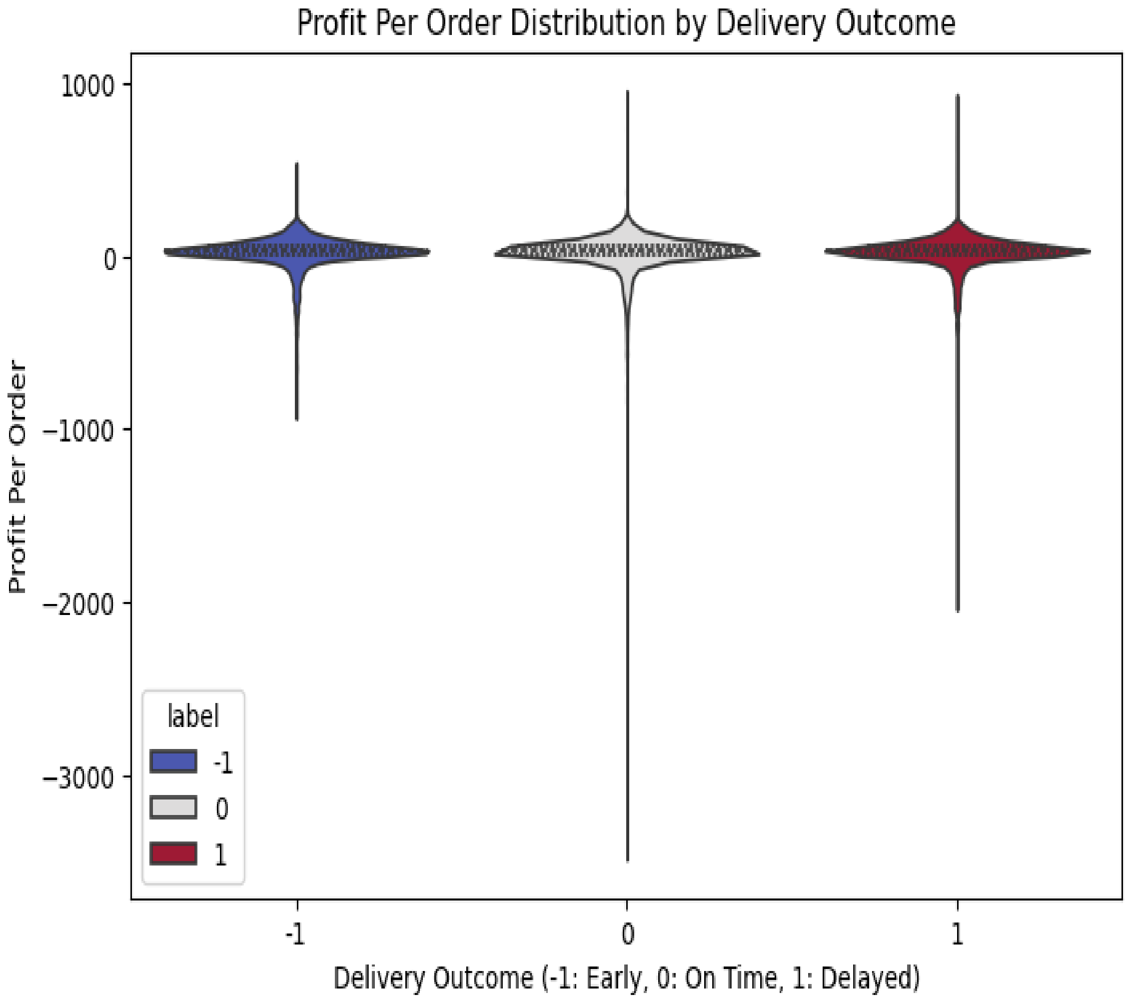
•Tabular logistics data (shipment time, distance, etc.).

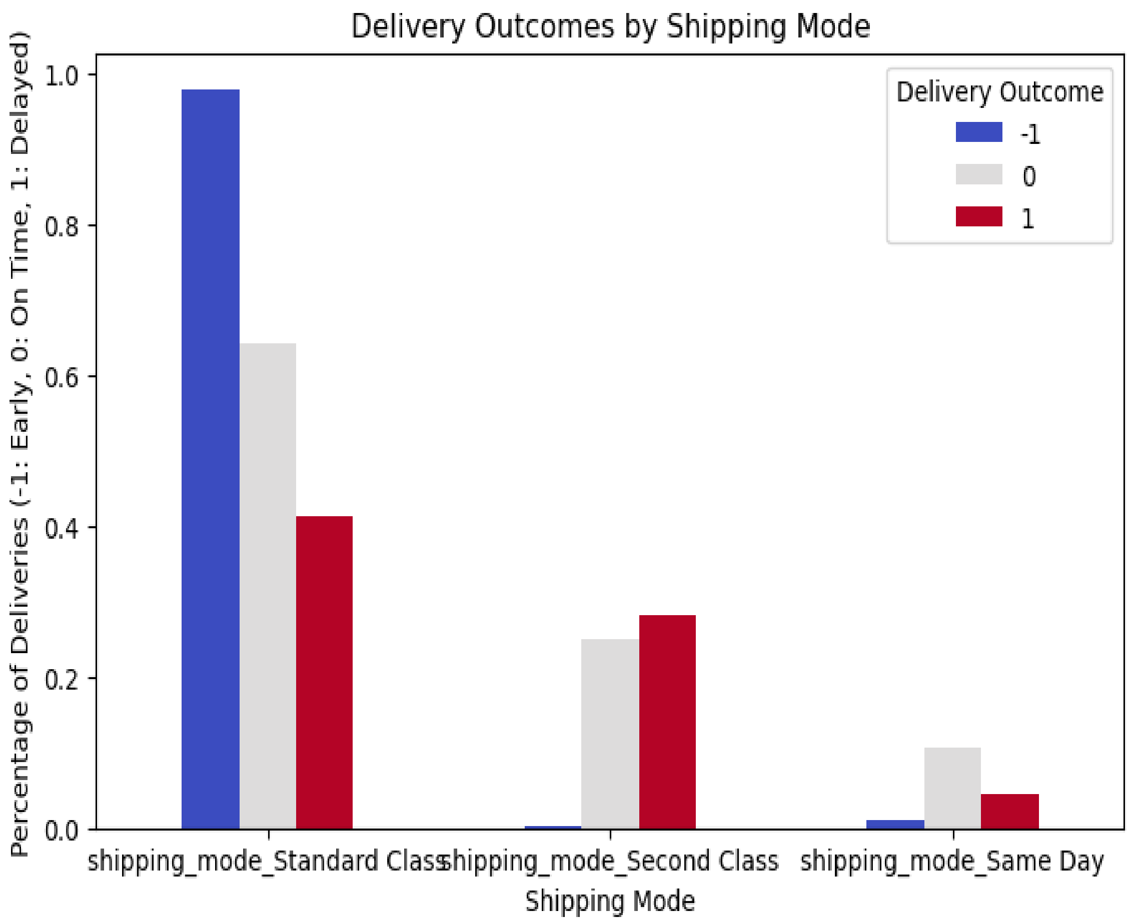
***Challenges & Importance**:*

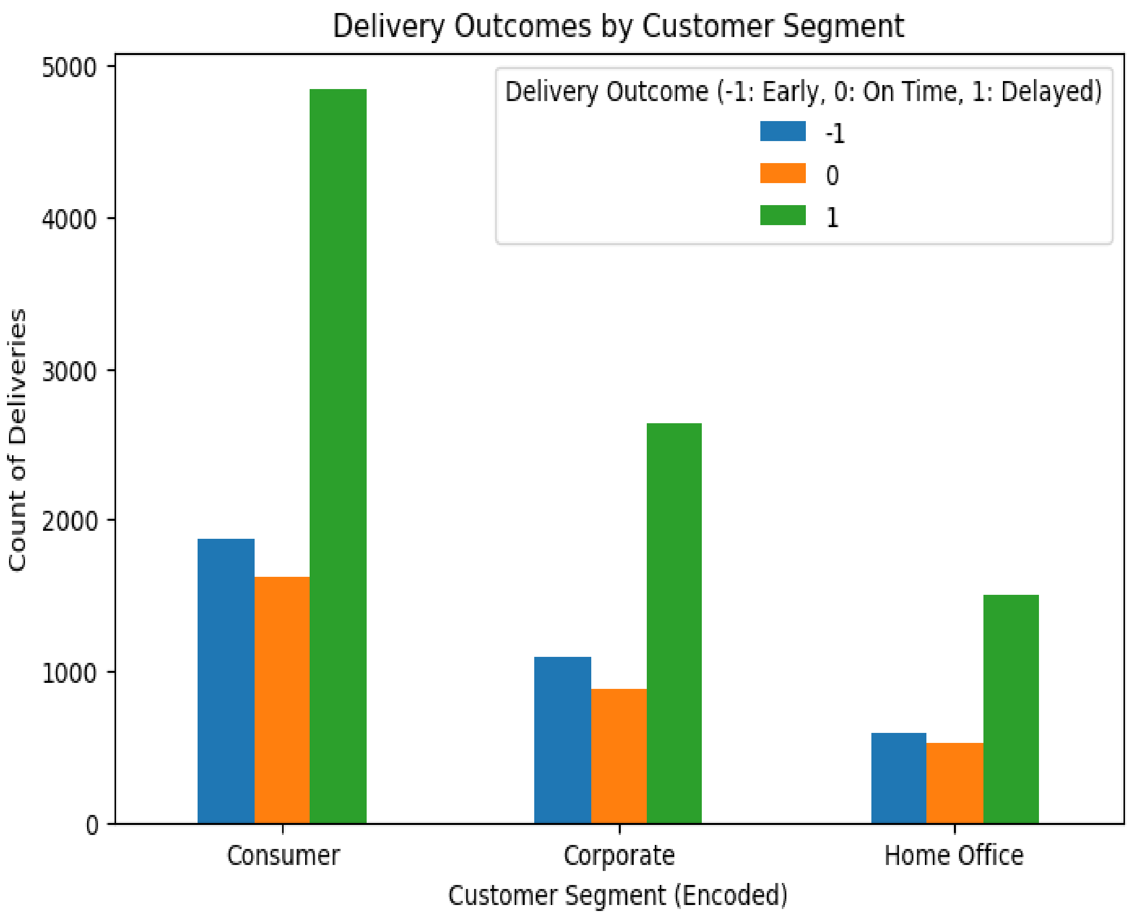
•Supply chain inefficiencies cause financial losses.

•AI-driven predictions enhance efficiency & decision-making.

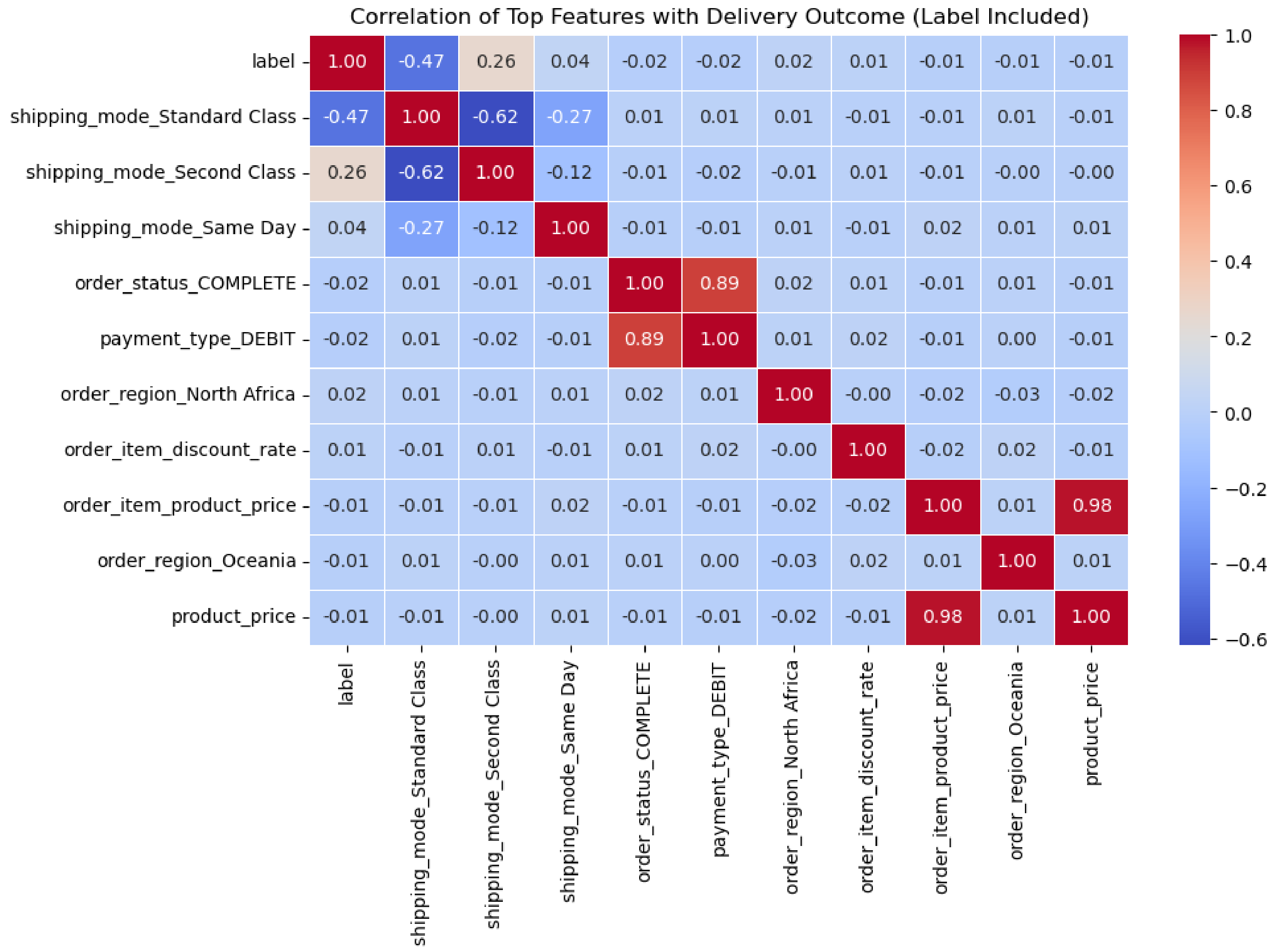

Profit Per Order by Delivery Outcome
10
0
−10
−20
−30
Profit Per Order
-1
0
1
Delivery Outcome (-1: Early, 0: On Time, 1: Delayed)


Profit Per Order Distribution by Delivery Outcome
1000
0
−1000
−2000
−3000
Profit Per Order
label
-1
0
1
-1
0
1
Delivery Outcome (-1: Early, 0: On Time, 1: Delayed)

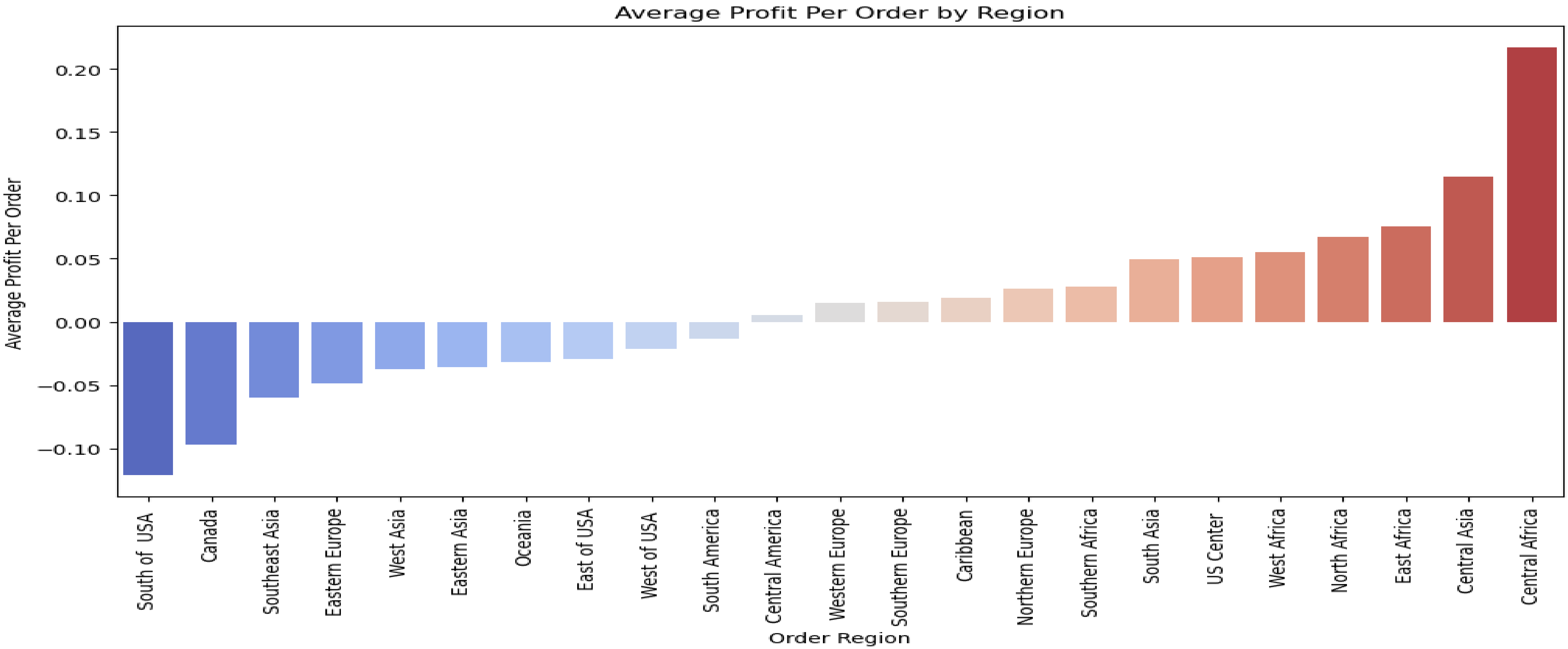

Delivery Outcomes by Shipping Mode
Percentage of Deliveries (-1: Early, 0: On Time, 1: Delayed)
1.0
0.8
0.6
0.4
0.2
0.0
Delivery Outcome
-1
0
1
shipping_mode_Standard Class
shipping_mode_Second Class
shipping_mode_Same Day
Shipping Mode


Delivery Outcomes by Customer Segment
Count of Deliveries
5000
4000
3000
2000
1000
0
Delivery Outcome (-1: Early, 0: On Time, 1: Delayed)
-1
0
1
Consumer
Corporate
Home Office
Customer Segment (Encoded)

[redacted]



Correlation of Top Features with Delivery Outcome (Label Included)
label
shipping_mode_Standard Class
shipping_mode_Second Class
shipping_mode_Same Day
order_status_COMPLETE
payment_type_DEBIT
order_region_North Africa
order_item_discount_rate
order_item_product_price
order_region_Oceania
product_price
1.00 -0.47 0.26 0.04 -0.02 -0.02 0.02 0.01 -0.01 -0.01 -0.01
-0.47 1.00 -0.62 -0.27 0.01 0.01 0.01 -0.01 -0.01 0.01 -0.01
0.26 -0.62 1.00 -0.12 -0.01 -0.02 -0.01 0.01 -0.01 -0.00 -0.00
0.04 -0.27 -0.12 1.00 -0.01 -0.01 0.01 -0.01 0.02 0.01 0.01
-0.02 0.01 -0.01 -0.01 1.00 0.89 0.02 0.01 -0.01 0.01 -0.01
-0.02 0.01 -0.02 -0.01 0.89 1.00 0.01 0.02 -0.01 0.00 -0.01
0.02 0.01 -0.01 0.01 0.02 0.01 1.00 -0.00 -0.02 -0.03 -0.02
0.01 -0.01 0.01 -0.01 0.01 0.02 -0.00 1.00 -0.02 0.02 -0.01
-0.01 -0.01 -0.01 0.02 -0.01 -0.01 -0.02 -0.02 1.00 0.01 0.98
-0.01 0.01 -0.00 0.01 0.01 0.00 -0.03 0.02 0.01 1.00 0.01
-0.01 -0.01 -0.00 0.01 -0.01 -0.01 -0.02 -0.01 0.98 0.01 1.00
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6


Average Profit Per Order by Region
Average Profit Per Order
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
South of USA
Canada
Southeast Asia
Eastern Europe
West Asia
Eastern Asia
Oceania
East of USA
West of USA
South America
Central America
Western Europe
Southern Europe
Caribbean
Northern Europe
Southern Africa
South Asia
US Center
West Africa
North Africa
East Africa
Central Asia
Central Africa
Order Region

Apply Model 1: MLP (Multilayer Perception)

***- Model Evaluated**: Multilayer Perceptron (MLP)*
*- **Accuracy**: 60.26%*
*- **Training Time**: 52.89 seconds*

- Strengths:

*•**Good at predicting delayed deliveries (Class 1) → 77% Precision & 73% F1-score***

*•**Fast training time**, making it a good baseline model*

- Weaknesses & Challenges:

*•**Struggles with on-time deliveries (Class 0) → Low recall (7%)***

*•**Earlier deliveries (Class 2) have high recall (83%) but low precision (42%)***

*•**Confuses earlier and delayed deliveries →** Model favors predicting delays*

- ***Observations from Confusion Matrix:***

•***Many on-time deliveries misclassified as delayed***

•***Earlier deliveries often confused with delays***

•***MLP lacks time-awareness & struggles with sequence-based trends***

How to Improve?
+ ***Combine with LSTM*** → *Capture sequential patterns for better predictions*
+ ***Feature Engineering*** → *Include real-time data (traffic, seasonal trends)*
+ ***Class Balancing*** → *Improve on-time delivery predictions*
+ ***Hyperparameter Tuning*** → *Optimize learning rate & hidden layers*

Apply Model 2: LSTM (Long Short-Term Memory)

- Model Evaluated: Long Short-Term Memory (LSTM)
- Accuracy: 58.84%
Training Time: 88.49 seconds

- Strengths:

- ***Better recall for delayed deliveries (Class 1) → 74% Recall & 72% F1-score***
- ***Performs slightly better than MLP in identifying earlier deliveries (Class 2)***

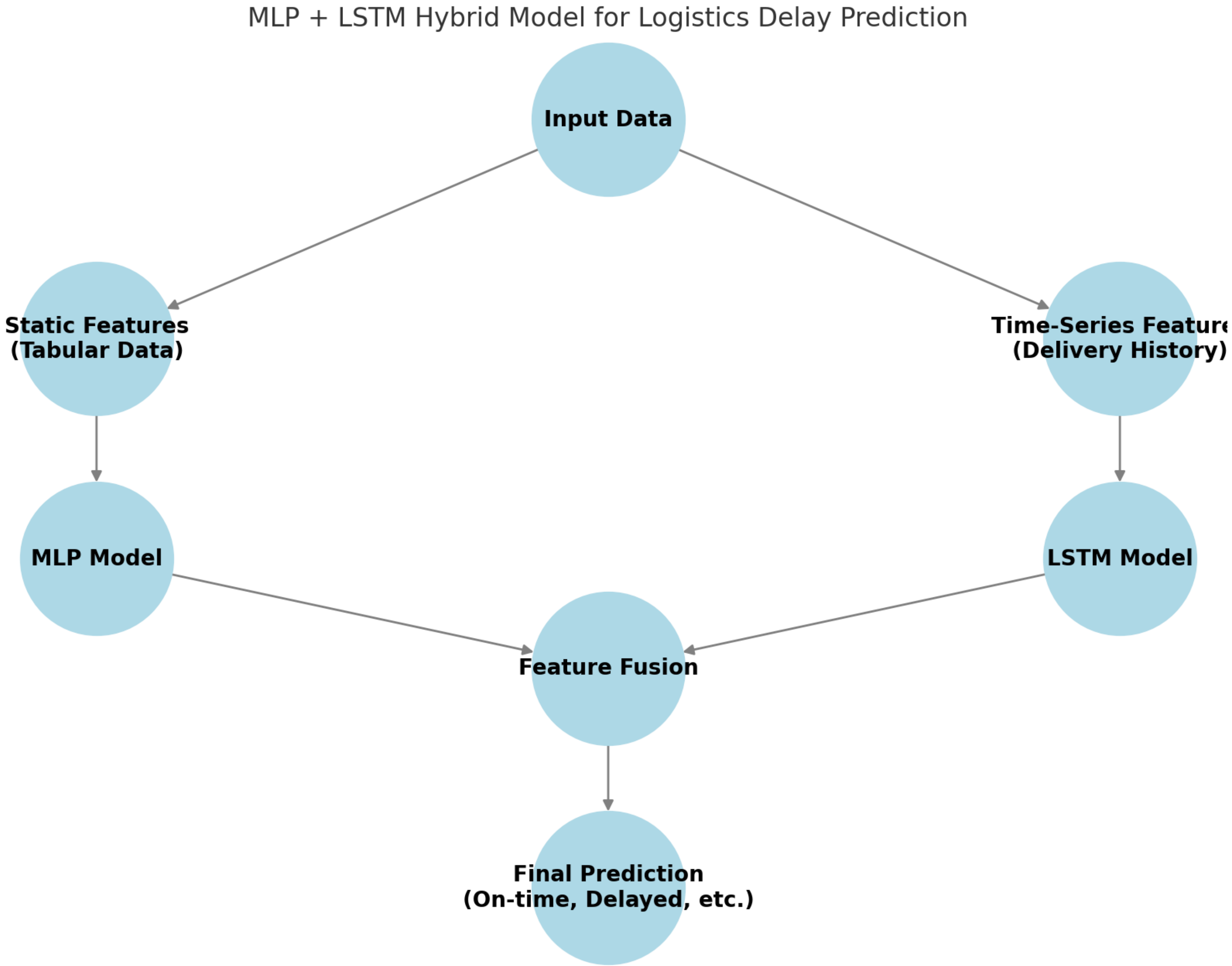
- Weaknesses & Challenges:

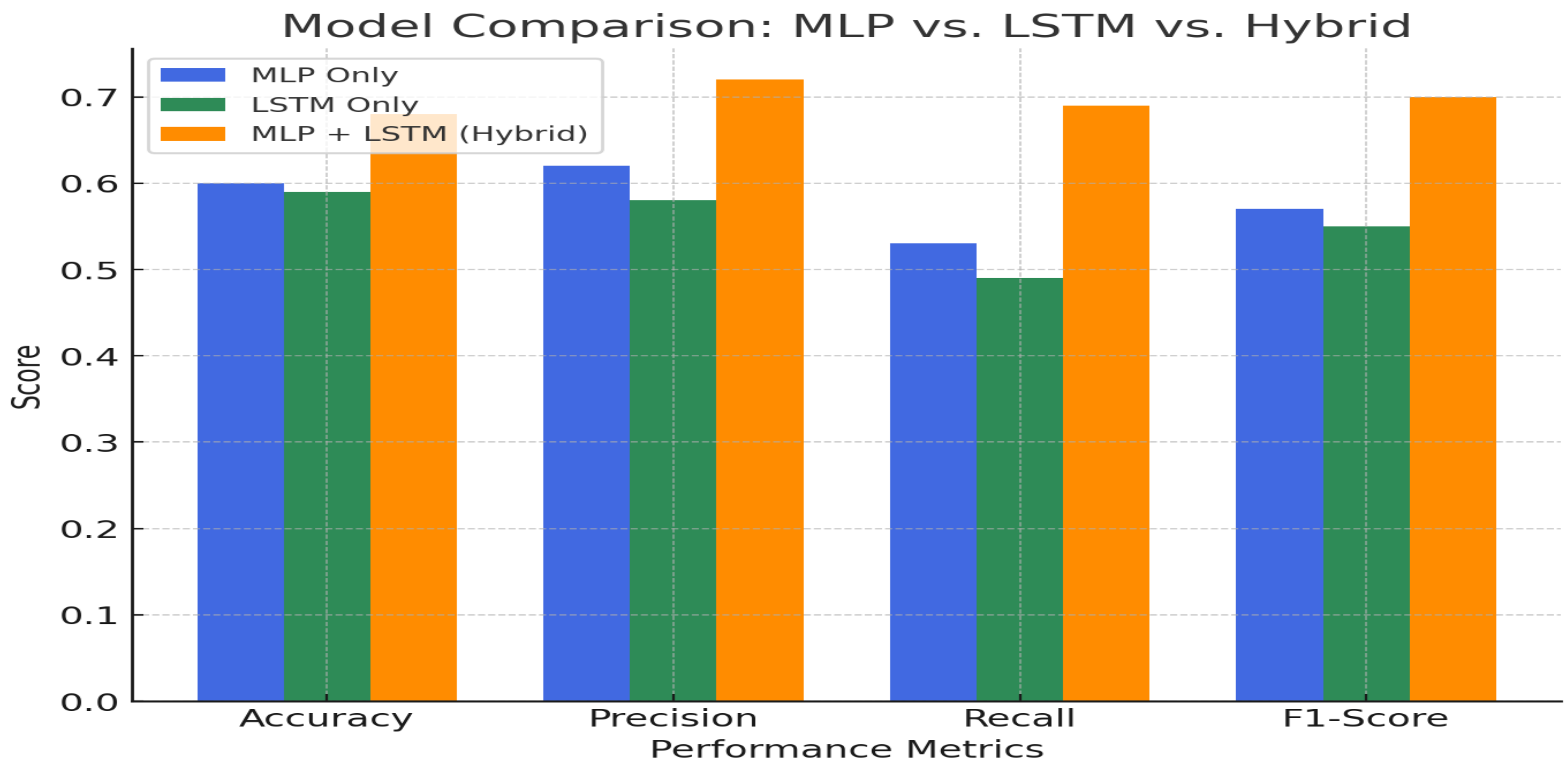
- ***Very poor recall for on-time deliveries (Class 0) → Only 5%***
- ***Precision for earlier deliveries (Class 2) is still low (41%),*** *leading to false positives*
- ***Higher training time (88s) compared to MLP (52s)***

Observations from Confusion Matrix:

- ***Many on-time deliveries misclassified as delays***
- ***Earlier deliveries often misclassified as delayed ones***
- ***LSTM is more sensitive to sequence-based patterns, but struggles with balanced classification***

Why Choosing MLP + LSTM For Logistics Delay Prediction





Model Comparison: MLP vs. LSTM vs. Hybrid
MLP Only
LSTM Only
MLP + LSTM (Hybrid)
Score
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Accuracy
Precision
Recall
F1-Score
Performance Metrics

Metric	MLP Only	LSTM Only	MLP + LSTM (Hybrid)
Accuracy	60%	59%	68%
Precision	62%	58%	72%
Recall	53%	49%	69%
F1-Score	57%	55%	70%

MLP (Strengths & Weaknesses)

- ***Good with structured data but fails to capture time-based trends.***
- ***Struggles with on-time delivery predictions.***

- LSTM (Strengths & Weaknesses)

- ***Captures sequential patterns, improving recall for delayed deliveries.***
- ***Longer training time and misclassifies on-time deliveries.***

- Hybrid MLP + LSTM Model

Combines MLP's structured data processing with LSTM's time-series analysis to improve:

- ***Overall accuracy & recall***
- ***Balanced classification performance***
- ***Better generalization to real-world logistics delays***

Thank you