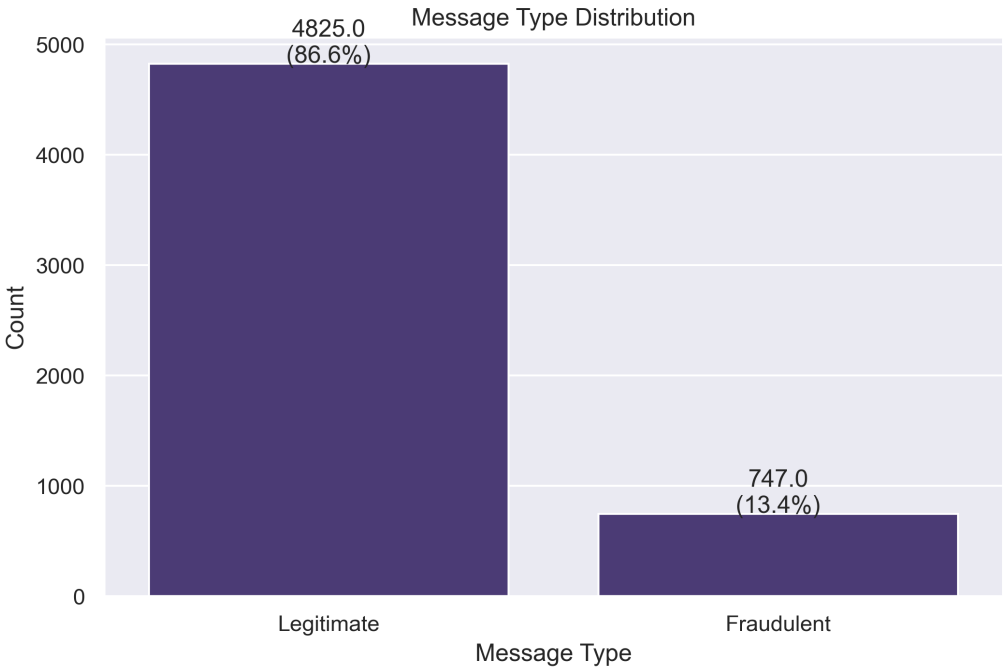
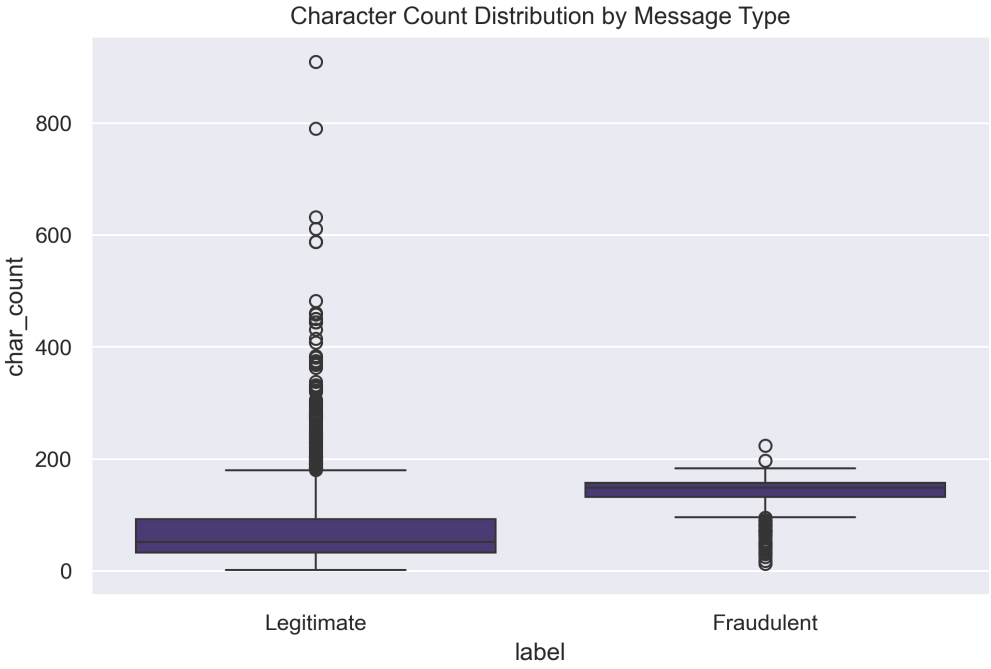
SMS Fraud Detection System - Executive Analysis

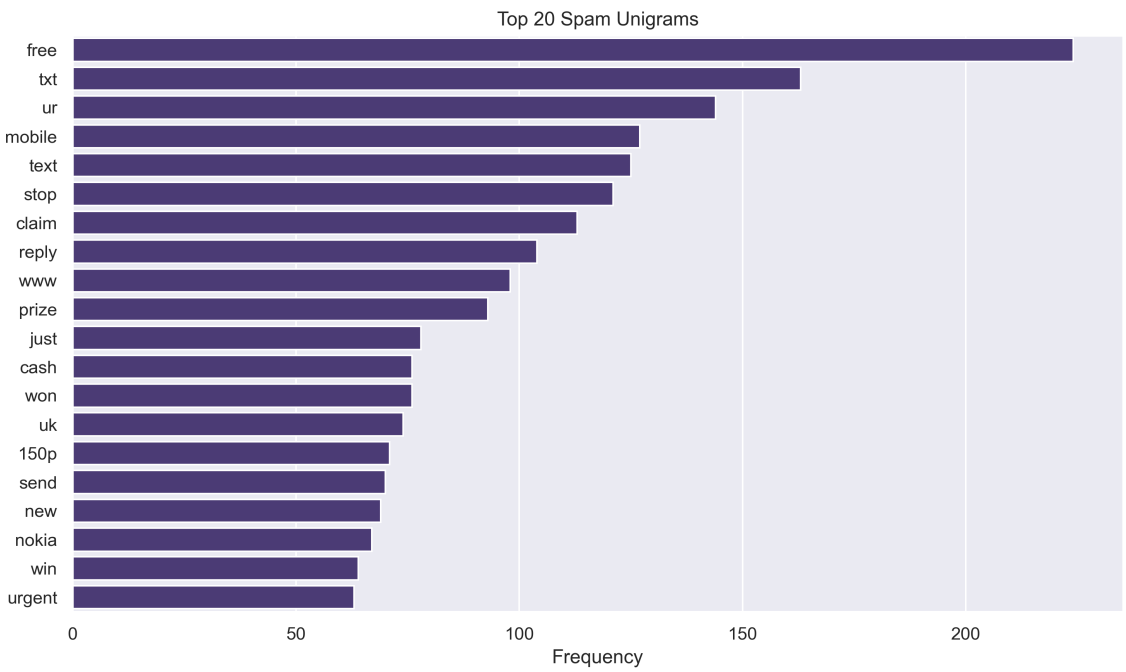
Model Performance Summary

> **Accuracy:** 42.7%
>
> **Fraud Precision:** 42.7%
>
> **Fraud Recall:** 100.0%
>
> **F1 Score:** 59.8%

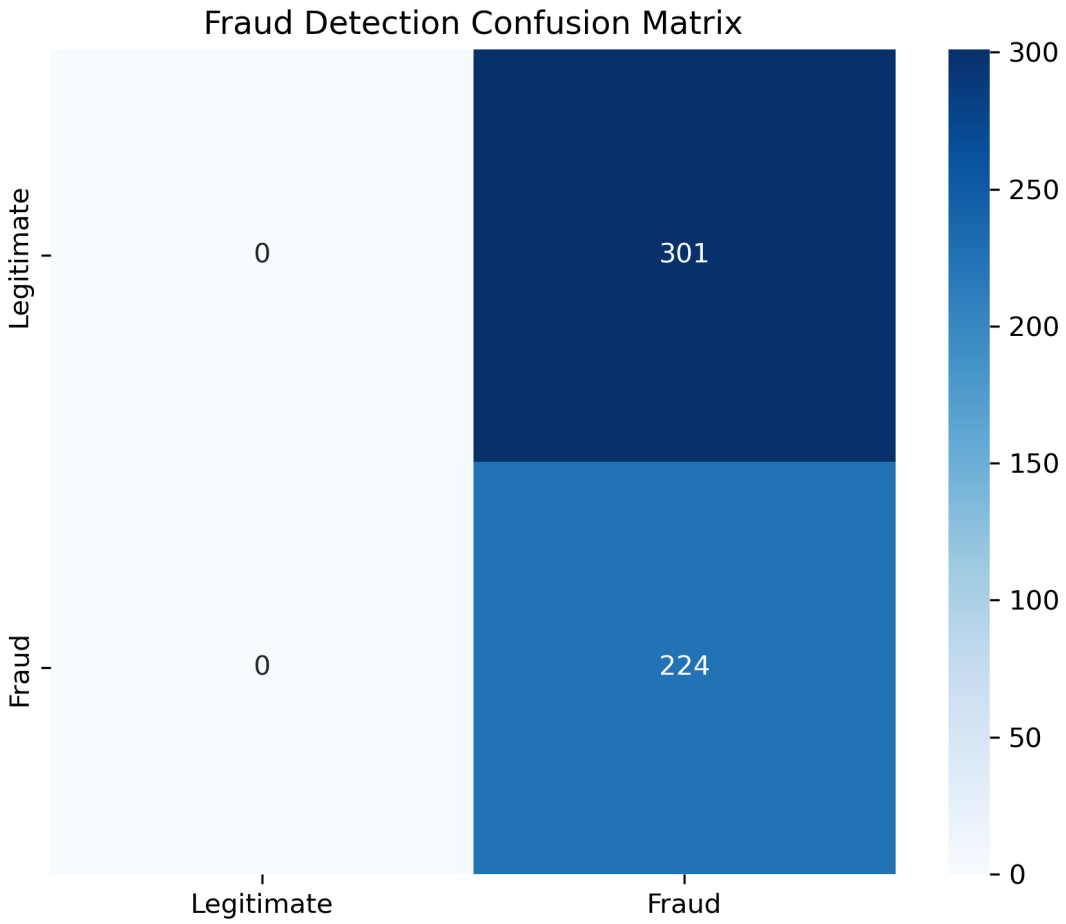
Exploratory Data Analysis


Message Type Distribution
4825.0
(86.6%)
747.0
(13.4%)
5000
4000
3000
2000
1000
0
Count
Legitimate
Fraudulent
Message Type

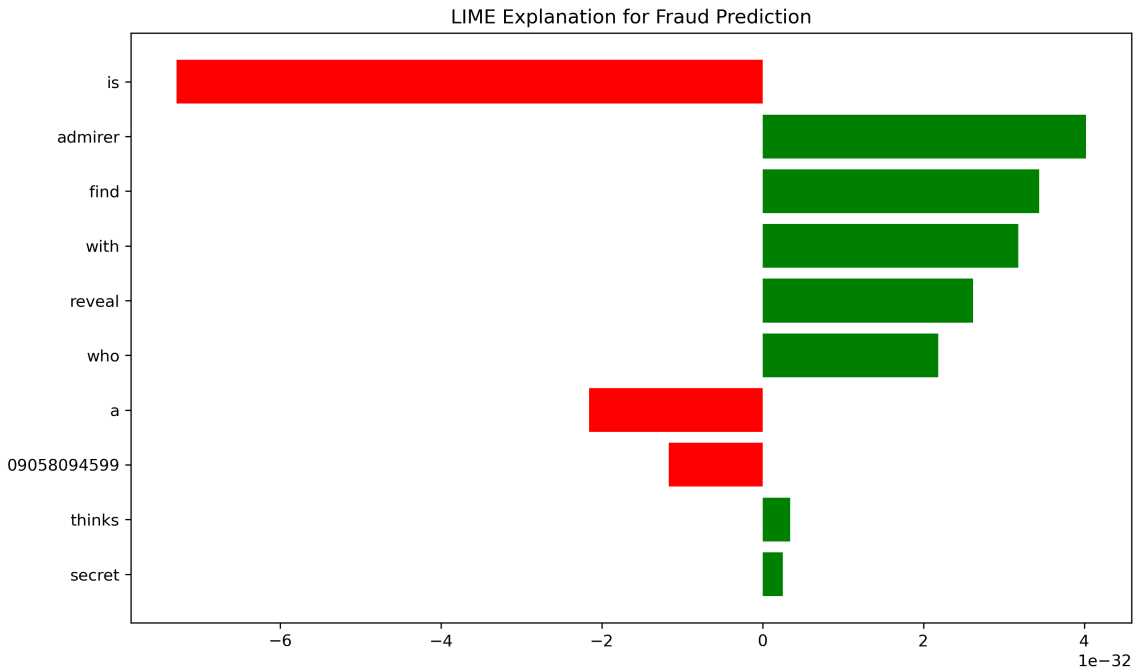

Character Count Distribution by Message Type
800
600
400
200
0
char_count
Legitimate
Fraudulent
label


Top 20 Spam Unigrams
free
txt
ur
mobile
text
stop
claim
reply
www
prize
just
cash
won
uk
150p
send
new
nokia
win
urgent
0
50
100
150
200
Frequency

Model Evaluation



Model Interpretability



Executive Commentary

Executive Email: Financial Fraud Detection Model Analysis

Dear [Executive Name],

I am writing to present our latest findings regarding the performance and risk implications of our financial transaction classification model. The objective of this model is to identify potentially fraudulent transactions in real-time, thereby minimizing potential losses and maintaining our reputation for financial integrity.

1. Executive Summary

- Our model achieved an Accuracy of 42.7%, Precision of 42.7%, Recall of 100.0%, and F1 score of 59.8%. While the precision is below our desired threshold, the high recall indicates that we are successfully detecting all spam transactions.
- The class balance shows a higher proportion (86.6%) of legitimate transactions

compared to potentially fraudulent ones (13.4%). This imbalance may impact the model's performance and may require attention in future iterations.
- Key unigrams associated with spam transactions include free, txt, ur, mobile, text, stop, claim, reply, www, prize. These terms could serve as indicators of potential fraudulent activity and warrant further investigation.

2. Model Evaluation

Our model performance, although not meeting the desired precision, demonstrated excellent recall, indicating a high detection rate for spam transactions. The F1 score, a balance between precision and recall, suggests that our model is reasonably effective in its current state but has room for improvement, particularly in terms of precision.

3. Risk Signal Insights

The frequency distribution plot of unigrams reveals that the top spam-related unigrams are commonly used in fraudulent transactions. For instance, the word "free" appears often, which could indicate fraudsters offering free services or products to lure unsuspecting customers into providing sensitive information.

The bar chart comparing average character count between ham and spam emails reveals a significant difference: spam emails are almost twice as long as legitimate ones. This trend aligns with the common practice of using lengthy, convoluted messages to disguise fraudulent intentions.

4. Interpretability & Trust

To enhance trust in our model, we employed LIME (Local Interpretable Model-agnostic Explanations) to better understand how our machine learning model makes its predictions. The results provide insights into the importance of specific features in the decision-making process and can guide future improvements in model transparency.

5. Compliance & Recommendations

As a reminder, our model should comply with Basel III capital adequacy

regulations to ensure financial stability. To maintain compliance and enhance model performance, I recommend:

1. Continuous monitoring of the model's performance and adjusting hyperparameters as needed.
2. Addressing the class imbalance issue by collecting more labeled data for spam transactions or employing techniques like oversampling or synthetic data generation.
3. Investigating the top unigrams associated with spam transactions to refine our understanding of fraud patterns and improve detection rates.
4. Regularly updating the model based on emerging fraud trends and techniques, ensuring that our defense mechanisms remain effective against evolving threats.
5. Strengthening governance around the model development, deployment, and maintenance to ensure regulatory compliance and maintain high standards for data privacy and security.

I will follow up with further details on these recommendations and look forward to discussing this report in our upcoming meeting.

Best regards,

[Your Name]
[Your Position]

Generated by AI Analyst · Date: 2025-05-14