Root Cause Analysis

Root Cause Analysis for Credit Card Fraud Detection Model Performance

1. Introduction

- **Brief overview of the problem:**
  In the context of credit card fraud detection, imbalanced datasets are a significant issue. In such datasets, fraudulent transactions represent a very small proportion compared to legitimate transactions. This leads to models that perform well overall (in terms of accuracy) but struggle to identify fraudulent transactions, which is the primary concern. The imbalance results in poor model performance specifically for detecting the rare class (fraudulent transactions), leading to ineffective fraud detection systems.
- **Importance of addressing the issue:**
  Addressing the imbalance is crucial for accurate fraud detection. If the model cannot reliably detect fraudulent transactions (which is the minority class), it defeats the purpose of the fraud detection system. Moreover, misidentifying fraud can result in significant financial losses, damage to the company's reputation, and negative customer experience. Thus, tackling this imbalance improves the model's ability to detect fraud and ultimately strengthens the reliability of the system.
-

2. Problem Statement

- **The dataset is highly imbalanced:**
  In fraud detection datasets, fraudulent transactions are a rare event, accounting for only a small percentage (often less than 1%) of the total dataset. This creates an imbalance, with the majority class being non-fraudulent transactions.
- **Models trained on imbalanced data:**
  Models trained on such imbalanced datasets tend to perform well in predicting the majority class (non-fraudulent transactions), but fail in detecting the minority class (fraudulent transactions). This is because the model is biased towards the majority class, learning to predict it more frequently, leading to inflated accuracy. However, the sensitivity (recall) for detecting fraudulent transactions tends to be low, as the model's focus is on correctly identifying legitimate transactions. This results in poor fraud detection performance despite high overall accuracy.

3. Root Cause Analysis

- Why is the dataset imbalanced?
  The imbalance exists because fraudulent transactions are rare events. Most customers make legitimate purchases, and fraudulent activities are comparatively few. This makes the fraudulent transactions significantly underrepresented in the data.
- Why does this affect model performance?
  When models are trained on imbalanced data, they are more likely to predict the majority class (non-fraudulent transactions). This results in a high accuracy but low sensitivity, meaning the model is not good at detecting the fraudulent transactions, which is the main goal of fraud detection systems. The model essentially learns to ignore the minority class, and this affects its ability to detect fraud effectively.
- Why are sensitivity and F1-score important?
  - Sensitivity (Recall): In fraud detection, sensitivity measures the model's ability to correctly identify fraudulent transactions. A high sensitivity means the model detects most of the fraud cases, which is crucial in minimizing losses.
  - F1-score: The F1-score is a metric that balances precision (how many predicted frauds were actually fraud) and recall (how many actual frauds were detected). Since both precision and recall are important, especially in fraud detection (where false positives and false negatives have serious consequences), the F1-score helps ensure a balance between not missing fraudulent transactions and not blocking legitimate ones.
- Why do some models perform better than others?
  Some models, like XGBoost and Logistic Regression, can handle imbalanced data better than others. This is often because these models can be adapted with techniques such as SMOTE (Synthetic Minority Oversampling Technique) or ADASYN, which generate synthetic samples for the minority class (fraudulent transactions). These techniques help the model learn better representations of the minority class, improving sensitivity and F1-score. On the other hand, models that do not implement such balancing techniques may be biased towards the majority class and underperform in fraud detection.

4. Findings

- SMOTE (Synthetic Minority Oversampling Technique) significantly improves performance:
  By generating synthetic instances of fraudulent transactions, SMOTE addresses the class imbalance and helps the model better detect fraud. This technique boosts sensitivity, enabling the model to detect more fraud cases, and also improves the F1-score, balancing precision and recall.
- XGBoost with SMOTE:
  Logistic Regression, when paired with SMOTE, achieves a good sensitivity and an F1-score.

- These scores indicate that the model is very effective at detecting fraudulent transactions (high recall) while maintaining a good balance between precision and recall (high F1-score), making it one of the best-performing models for fraud detection in this scenario.

5. Recommendations

Use SMOTE or ADASYN to balance the dataset:
These techniques should be used to oversample the minority class (fraudulent transactions) to ensure the model learns to recognize these rare events. This will improve sensitivity and the F1-score.

- Focus on models like XGBoost:
  These models, especially when combined with techniques like SMOTE, handle imbalanced data effectively and have demonstrated better fraud detection performance in terms of sensitivity, F1-score, and ROC Curve.
- Regularly evaluate model performance:
  Model performance should not be evaluated solely based on accuracy but should focus on metrics like sensitivity, F1-score, and ROC curve to ensure the model is effectively detecting fraud without sacrificing precision.

6. Conclusion

- Addressing dataset imbalance is crucial for improving fraud detection:
  Imbalanced datasets can lead to inaccurate fraud detection. SMOTE and similar techniques offer effective solutions for overcoming this issue and improving the model's ability to detect fraudulent transactions.
- SMOTE and proper model selection:
  When paired with appropriate algorithms like Logistic Regression and XGBoost, balancing the dataset can significantly enhance the model's performance. This leads to better fraud detection, minimizing the risk of fraud going undetected and improving the overall reliability of fraud detection systems.