Comparing Hard Neural Network based Undersampling and SMOTE Techniques for Handling Class Imbalance in Credit Card Fraud Detection

Lokiron Anda

Computer Science Department, Earlham College

Abstract

- Credit card fraud is a major issue that can result in significant financial losses for individuals and financial institutions.
- Machine learning techniques have been used to detect credit card fraud, but the imbalance of fraud transactions makes it challenging to develop accurate models.
- We are comparing two popular techniques: *Hard Neural Network based Undersampling* [1] and *SMOTE (Synthetic Minority Oversampling Technique)* [3] on an imbalanced credit card fraud dataset.
- We will apply both techniques to the training set and train a *Random Forest* [2] algorithm classifier.
- We will evaluate the performance of the classifiers compared each other and to the baseline performance of the system without any techniques for addressing class imbalance.

Introduction

- Credit card fraud is a significant problem for financial institutions and consumers worldwide.
- Machine learning algorithms have proven to be effective in detecting fraudulent transactions, but class imbalance is a common issue in credit card fraud detection, where the number of legitimate transactions heavily outnumber the number of fraudulent transactions.
- Class imbalance can result in poor performance for machine learning models because machine learning models would not be able to have an accurate representation of the minority class causing the model to be inaccurate because of the lack of examples from the minority class compared to the majority class leading the model to have bias and more likely to incorrectly identify the fraud.
- The goal to fix class imbalance is to not have an over-representation of one of the classes. This will help the model to have a fair assessment between both classes and improve the performance of the model and it's accuracy. Therefore, addressing class imbalance is critical in improving the accuracy of fraud detection systems.

Procedure

The dataset that is being used is *Credit Card Fraud Detection* by *The Machine Learning Group* [4].

- Spilt the data into training and testing sets. The data will be processed in a *Random Forest* algorithm. The Random Forest algorithm will be trained on the dataset to establish a baseline performance.
- Train the *Hard Neural Network based Undersampling* technique to the majority class in the training set to create a balanced dataset. The new dataset will train the Random Forest model on this balanced dataset and evaluate the performance.
- The same procedure with *SMOTE*. Apply the SMOTE technique to generate synthetic instances of the minority class in the training set to create a balanced dataset. The new dataset will train the Random Forest model with the balanced dataset and evaluate the performance.
- Evaluate the performance of Random Forest model with the techniques applied and the model without the techniques using the performance metrics.
- A statistical analysis determining if the techniques improved the performance of the model.

Model

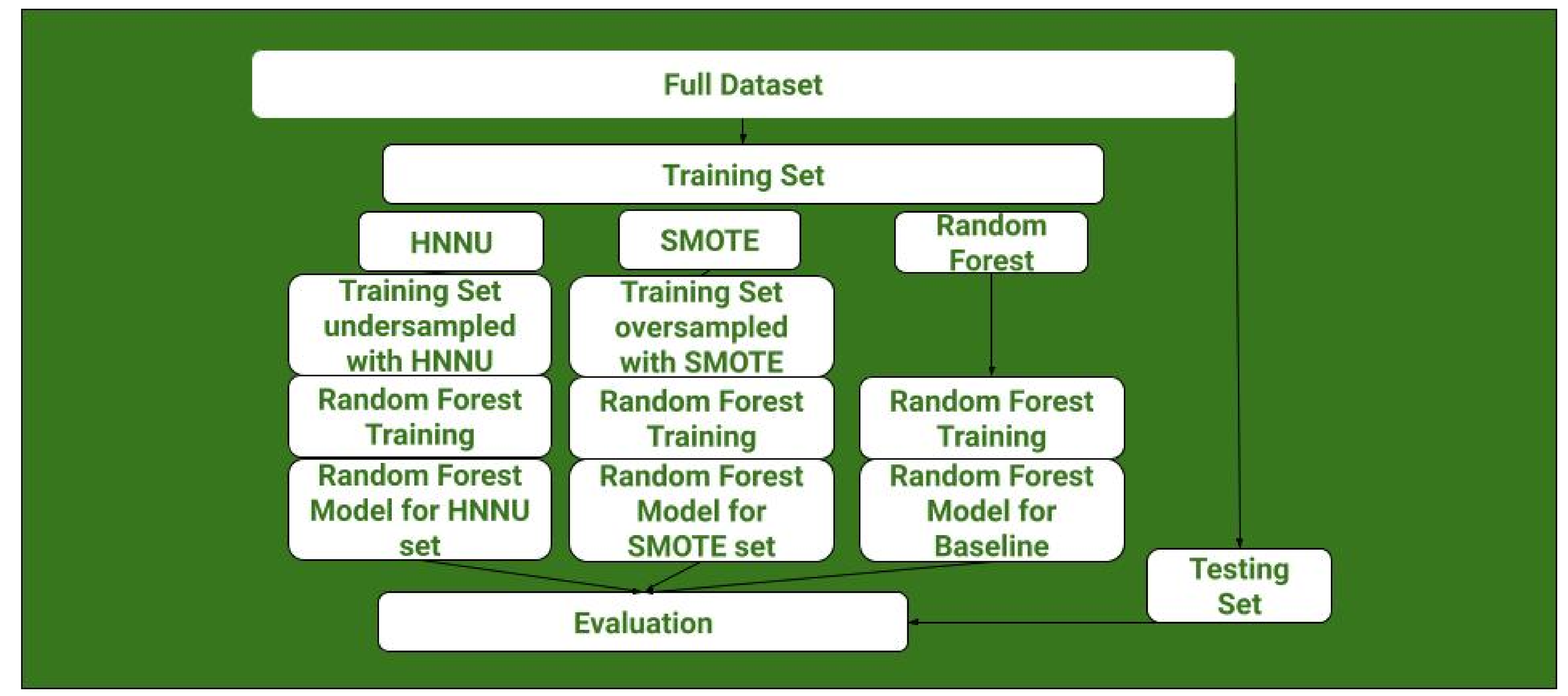

Figure 1. Framework Model

Evaluation Scores

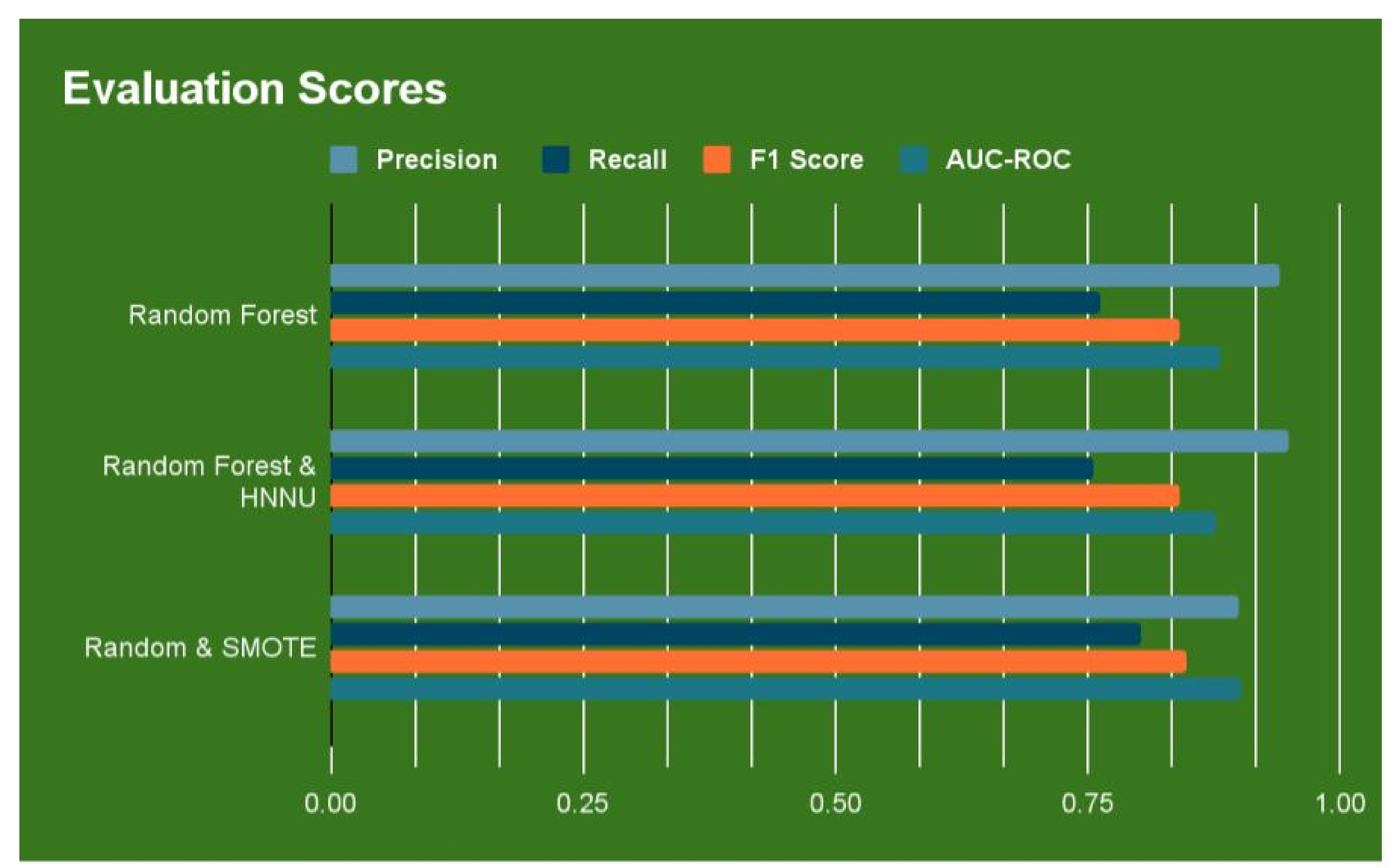

Figure 2. Evaluation Scores Table

Results

- Precision: The dataset without sampling adjustments and the HNNU technique achieved similar high precision scores, indicating a low false positive rate in identifying fraudulent transactions. SMOTE had slightly lower precision, suggesting it may generate more false positives.
- Recall: SMOTE showed a higher recall score, indicating a better ability to capture true fraudulent transactions. The dataset without sampling adjustments and the HNNU had lower scores than SMOTE, but very similar scores.
- F1 Score: All three methods achieved comparable F1 scores, striking a balance between precision and recall.
- AUC-ROC Score: The AUC-ROC scores for all three approaches were relatively close, indicating good overall model performance.

Conclusion

- In summary, Performance was fairly high with the Random Forest classifier even without any rebalancing.
- HNNU slightly improved precision score and SMOTE had a slight increased for the recall scores and AUC-ROC scores, but the techniques notably didn't make any significant differences to the overall performance of the Random Forest model.
- SMOTE and HNNU are two techniques that can be helpful for specific needs in imbalanced data, but are not necessary when put through the Random Forest classifier.

Acknowledgment

I would like to thank David Barbella for his support and feedback on my project idea and research proposal.

References

[1] Md. Adnan Arefeen, Sumaiya Tabassum Nimi, and M. Sohel Rahman.
Neural network-based undersampling techniques.
IEEE Transactions on Systems, Man, and Cybernetics: Systems, 52(2):1111–1120, 2020.

[2] Gérard Biau and Erwan Scornet.
A random forest guided tour.
Test, 25:197–227, 2016.

[3] Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer.
Smote: Synthetic minority oversampling technique.
Journal of artificial intelligence research, 16:321–357, 2002.

[4] Machine Learning Group ULB.
Credit card fraud detection, Mar 2018.