Enhancing Credit Card Fraud Detection through Deep Learning Techniques and Imbalanced Data Strategies

[1]Ahmed Olanrewaju [2] Salawou Musodiq Adebayo

[1]University of Ibadan [2]Hill City University

Introduction

- Alarming rise in credit card fraud in the digital age.
- Deep learning offers promise in improving fraud detection.
- Focus on RNNs, CNNs, and Ensemble methods.
- Addressing imbalanced datasets crucial for robustness.

Problem Statement

- Traditional methods struggle with evolving fraud tactics.
- Imbalanced data skews model outcomes.
- Need for sophisticated techniques to detect fraud while minimizing false positives.
- Deep learning methods can offer solutions.

Dataset Description

- Comprehensive credit card transaction dataset.
- Simulated data covering 1 Jan 2019 - 31 Dec 2020.
- 1000 customers, 800 merchants.
- Contains legitimate and fraud transactions.

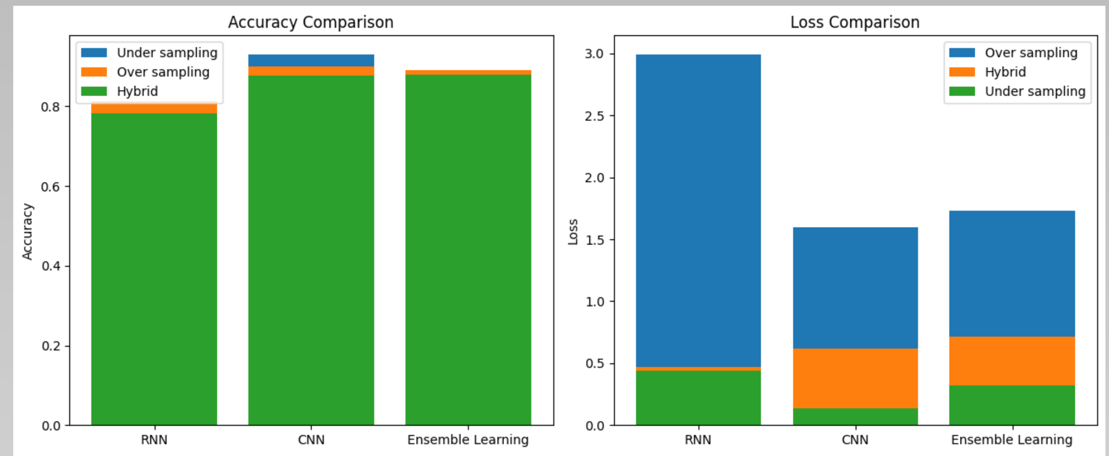
Methodology

- RNNs for sequential data (time-series transactions).
- CNNs for spatial patterns and sequences.
- Ensemble of RNNs and CNNs for robustness.
- Under-sampling, over-sampling, and hybrid methods for class imbalance.

Experimental Setup

- The experimental setup entails evaluating the effectiveness of various deep learning models for fake review detection. The models include Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), and an Ensemble of RNNs and CNNs.
- The study employs different sampling methods, including under-sampling, over-sampling, and hybrid techniques, to address class imbalance in the dataset.
- Performance metrics such as loss, accuracy, and Mean Squared Error (MSE) are measured across multiple epochs. The results highlight the hybrid approach's superiority in terms of accuracy and MSE.
- The research showcases the potential of deep learning techniques in combating fake reviews and underscores the importance of handling imbalanced data.

Results



Model Name	Sampling Method	Loss	Accuracy	MSE	Time	Epochs
RNN	Hybrid	0.46755334734916687	0.7832167744636536	0.151466265320779	62.61452841758728	10
RNN	Under sampling	0.43592366576194763	0.8121878121878122	0.139168873429298	22.37613010406494	10
RNN	Over sampling	2.994174331327035	0.8121878121878122	0.03527341789080050 4		10
CNN	Hybrid	0.6178991794586182	0.8766279816627502	0.1187646836042404 2	4006.089469671249 4	10
CNN	Under sampling	0.13772976398468018	0.9311960339546204	0.04427330568432808	3777.27960610389 7	10
CNN	Over sampling	1.5948631255077401	0.8999608275091726	0.01000783603264952 3		10
Ensemble Learning	Hybrid	0.7130040526390076	0.8787878751754761	0.10402944684028625	30.707619428634644	10
Ensemble Learning	Under sampling	0.32114464044570923	0.8767898678779602	0.08952364325523376	31.287129163742065	10
Ensemble Learning	Over sampling	1.7306752843695496	0.8914418914418915	0.01178486293371408 4		10

- Experimental findings highlight diverse strengths of RNNs, CNNs, and their ensembles in credit card fraud detection. RNNs excel in temporal pattern recognition, evident in hybrid and under-sampling methods with accuracies of 0.7832 and 0.8122.
- CNNs, skilled in spatial pattern detection, achieve 0.8766 accuracy in the hybrid approach and 0.9312 with under-sampling.
- Ensemble learning fuses RNNs and CNNs, yielding 0.8788 accuracy. However, caution is warranted with oversampling due to potential overfitting.
- These insights underscore tailored model selection and balancing class distribution to enhance fraud detection.
- A significant step towards innovative advancements in artificial intelligence and machine learning.

Conclusion

The research presents a comprehensive exploration of fake review detection techniques, focusing on the application of deep learning methodologies. The study emphasizes the critical challenge of distinguishing authentic and deceptive reviews in the digital landscape. By employing advanced techniques such as Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), and Ensemble Learning, the research contributes to the enhanced identification of fake reviews. The investigation delves into addressing the class imbalance issue, an essential consideration for effective model generalization. The findings underscore the significance of adeptly handling imbalanced data to uncover concealed patterns of deception. The research not only advances the field of fake review detection but also highlights the pivotal interplay between deep learning strategies and the nuanced treatment of imbalanced datasets, with potential implications for broader applications in artificial intelligence and machine learning innovation.

References

Abri, F., Gutierrez, L. F., Namin, A. S., Jones, K. S., & Sears, D. R. (2020). Fake reviews detection through analysis of linguistic features. arXiv preprint arXiv:2010.04260.

Bansal, M., Zhou, Z.-H., & Wang, R. (2021). A Survey on Fake Review Detection: Methods, Challenges, and Opportunities. arXiv preprint arXiv:2103.09377.

Zhang, Y., & Liu, Z. (2019). A deep learning framework for fake review detection. IEEE Transactions on Knowledge and Data Engineering, 31(1), 133-146.

Zhou, Z., Zhang, F., Li, Z., & Yu, H. (2019). Unsupervised detection of fake reviews using text features. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2158-2166).

Acknowledgement

The authors would like to express their gratitude to the faculty and staff of the University of Ibadan and Hill City University for their support and guidance throughout the research process.

Contact Information:

- Ahmed Olanrewaju University of Ibadan
  Email: abono2000@gmail.com
- Salawou Musodiq Adebayo Hill City University