Literature Review: Optimized Neural Network Classification Framework with Advanced Feature Engineering and Dimensionality Reduction for Scalable Chain Anomaly Detection in Multi-Resource Billing Systems

Introduction

The increasing complexity of multi-resource billing systems has amplified the need for effective anomaly detection frameworks capable of identifying fraudulent activities and billing errors. The integration of advanced neural network architectures, sophisticated feature engineering techniques, and dimensionality reduction methods has emerged as a promising approach to enhance anomaly detection capabilities in these systems. This literature review synthesizes key research findings relevant to the development of an optimized neural network classification framework designed for this purpose.

Advanced Feature Engineering Techniques

Feature engineering plays a pivotal role in the performance of anomaly detection systems. Sakurada and Yairi (2014) emphasize the utility of autoencoders for anomaly detection, particularly through nonlinear dimensionality reduction. This approach is advantageous for multi-resource billing systems, as it allows for the detection of subtle anomalies that traditional linear methods, such as PCA, often overlook. The robustness of denoising autoencoders further strengthens this framework, making it applicable in real-world scenarios with noisy or incomplete data.

Deng and Hooi (2021) introduce a method that integrates structure learning with graph neural networks, effectively capturing complex interrelationships among billing metrics. This capability is crucial for enhancing feature engineering within the proposed classification framework, enabling improved accuracy in anomaly detection. The inclusion of attention weights for explainability addresses stakeholder concerns regarding transparency in detected anomalies, a critical requirement in billing systems.

Further, the iLearn toolkit, as discussed by Xiao et al. (2019), offers comprehensive feature extraction and dimensionality reduction capabilities. By automating optimal feature selection and construction processes, iLearn can significantly streamline the development of a scalable anomaly detection system tailored for multi-resource billing contexts.

Dimensionality Reduction Techniques

Dimensionality reduction is essential for managing high-dimensional data while retaining crucial information. Tuli et al. (2022) present the TranAD model, which employs attention-based sequence encoders to enhance feature extraction in multivariate time-series data. This model-agnostic approach is particularly beneficial for billing systems, where timely anomaly detection is critical amidst high data volatility. Similarly, Yang et al. (2021) highlight the importance of dimensionality reduction methods like autoencoders in improving classification performance, demonstrating high accuracy in their empirical studies.

The Temporal Hierarchical One-Class (THOC) network, as proposed by Wyatt et al. (2022), utilizes dilated recurrent neural networks to capture temporal dynamics in time-series data. This capability is indispensable in billing systems, where understanding temporal patterns can significantly enhance anomaly detection performance. Additionally, the use of denoising diffusion probabilistic models (DDPMs) for unsupervised anomaly detection, as discussed by Abdulhammed et al. (2019), offers a novel generative approach to learning normal reference data, further enriching the dimensionality reduction strategies applicable to this research topic.

Integration of Classification Frameworks

The integration of various classification frameworks is essential for achieving high-performance anomaly detection. Perera and Patel (2018) introduce a deep one-class transfer learning approach that focuses on feature learning in one-class classification scenarios. This method is particularly relevant for billing systems where normal transactions vastly outnumber anomalies. By employing descriptive features and maintaining low intra-class variance, this approach can substantially enhance the classification framework's efficacy.

Moreover, the few-shot learning model proposed by Ullah et al. (2020) demonstrates the potential to address the challenges of limited labeled data in anomaly detection. Utilizing a Siamese CNN for feature representation, this model can improve the framework's ability to identify anomalies in billing systems, especially in data-scarce environments.

Knowledge Gaps and Future Research Directions

Despite the advancements highlighted, several knowledge gaps persist regarding the integration of these methodologies into a cohesive framework for billing system anomaly detection. For instance, while various techniques for feature engineering and dimensionality reduction have been explored, comprehensive studies that integrate these approaches within a unified neural network classification framework remain limited. Future research should focus on the development of hybrid models that combine the strengths of autoencoders, graph neural networks, and attention mechanisms to achieve superior anomaly detection performance.

Additionally, the exploration of adversarial learning strategies, as discussed by Miller et al. (2020), could inform the development of more resilient classification frameworks capable of withstanding potential attacks, thereby enhancing security in billing systems. Investigating the scalability of these frameworks in real-time scenarios and their effectiveness in heterogeneous billing environments presents another promising avenue for future research.

Conclusion

The synthesis of advanced feature engineering techniques, dimensionality reduction methods, and robust classification frameworks holds significant promise for optimizing neural network classification systems aimed at anomaly detection in multi-resource billing systems. As the landscape of billing systems continues to evolve, ongoing research will be essential in addressing existing knowledge gaps and developing comprehensive solutions that ensure accuracy, efficiency, and transparency in anomaly detection efforts.

## References:

1. Feng, Cheng., Li, Tingting., & Chana, D.. (2017). Multi-level Anomaly Detection in Industrial Control Systems via Package Signatures and LSTM Networks. <i>2017 47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)</i> , 261-272 . http://doi.org/10.1109/DSN.2017.34

2. Perera, Pramuditha., & Patel, Vishal M.. (2018). Learning Deep Features for One-Class Classification. <i>IEEE Transactions on Image Processing</i> , 28 , 5450-5463 . http://doi.org/10.1109/TIP.2019.2917862

3. Monaco, G.., Chen, Hao., Poidinger, M.., Chen, Jinmiao., Magalhães, J. P.., & Larbi, A.. (2016). flowAI: automatic and interactive anomaly discerning tools for flow cytometry data. <i>Bioinformatics</i> , 32 16 ,

2473-80

. http://doi.org/10.1093/bioinformatics/btw191

4. Deng, Hanqiu., & Li, Xingyu. (2022). Anomaly Detection via Reverse Distillation from One-Class Embedding. <i>2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)</i> , 9727-9736 . http://doi.org/10.1109/CVPR52688.2022.00951

5. Choi, Shinhyun., Shin, J.., Lee, Jihang., Sheridan, P.., & Lu, Wei D.. (2017). Experimental Demonstration of Feature Extraction and Dimensionality Reduction Using Memristor Networks.. <i>Nano letters</i> , 17 5 ,

3113-3118

. http://doi.org/10.1021/acs.nanolett.7b00552

6. Xiao, Yihan., Xing, C.., Zhang, Taining., & Zhao, Zhongkai. (2019). An Intrusion Detection Model Based on Feature Reduction and Convolutional Neural Networks. <i>IEEE Access</i> , 7 , 42210-42219 . http://doi.org/10.1109/ACCESS.2019.2904620

7. Yang, Yi., Zhuang, Yueting., & Pan, Yunhe. (2021). Multiple knowledge representation for big data artificial intelligence: framework, applications, and case studies. <i>Frontiers of Information Technology & Electronic Engineering</i> , 22 , 1551 - 1558 . http://doi.org/10.1631/FITEE.2100463

8. Chen, Zhen., Zhao, Pei., Li, Fuyi., Marquez-Lago, T.., Leier, A.., Revote, Jerico., Zhu, Yan., Powell, D.., Akutsu, T.., Webb, Geoffrey I.., Chou, K.., Smith, Alexander Ian., Daly, R.., Li, Jian., & Song, Jiangning. (2019). iLearn : an integrated platform and meta-learner for feature engineering, machine-learning analysis and modeling of DNA, RNA and protein sequence data. <i>Briefings in bioinformatics</i> . http://doi.org/10.1093/bib/bbz041

9. Mukherjee, Sudipta. (2020). Anomaly Detection. , 60 . http://doi.org/10.1007/978-3-319-17885-1\_100064

10. Zhou, Xiaokang., Liang, Wei., Shimizu, Shohei., Ma, Jianhua., & Jin, Qun. (2021). Siamese Neural Network Based Few-Shot Learning for Anomaly Detection in Industrial Cyber-Physical Systems. <i>IEEE Transactions on Industrial Informatics</i> , 17 , 5790-5798 . http://doi.org/10.1109/TII.2020.3047675

11. Miller, David J.., Xiang, Zhen., & Kesidis, G.. (2020). Adversarial Learning Targeting Deep Neural Network Classification: A Comprehensive Review of Defenses Against Attacks. <i>Proceedings of the IEEE</i> , 108 , 402-433 . http://doi.org/10.1109/JPROC.2020.2970615

12. Sakurada, M.., & Yairi, T.. (2014). Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction. , 4 . http://doi.org/10.1145/2689746.2689747

13. Ullah, Waseem., Ullah, Amin., Haq, I.., Muhammad, Khan., Sajjad, M.., & Baik, S.. (2020). CNN features with bi-directional LSTM for real-time anomaly detection in surveillance networks. <i>Multimedia Tools and Applications</i> , 80 , 16979 - 16995 . http://doi.org/10.1007/s11042-020-09406-3

14. Tuli, Shreshth., Casale, G.., & Jennings, N.. (2022). TranAD: Deep Transformer Networks for Anomaly Detection in Multivariate Time Series Data. <i>Proc. VLDB Endow.</i> , 15 , 1201-1214 . http://doi.org/10.14778/3514061.3514067

15. Mansour, R.. (2017). Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy. <i>Biomedical Engineering Letters</i> , 8 , 41 - 57 . http://doi.org/10.1007/s13534-017-0047-y

16. https://www.semanticscholar.org/paper/fc36d50be4352afada233f55a59f0dca0a7a826b

17. Deng, Ailin., & Hooi, Bryan. (2021). Graph Neural Network-Based Anomaly Detection in Multivariate Time Series. , 4027-4035 . http://doi.org/10.1609/aaai.v35i5.16523

18. Wyatt, Julian., Leach, Adam., Schmon, Sebastian M.., & Willcocks, Chris G.. (2022). AnoDDPM: Anomaly Detection with Denoising Diffusion Probabilistic Models using Simplex Noise. <i>2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)</i> , 649-655 . http://doi.org/10.1109/CVPRW56347.2022.00080

19. Laptev, N.., Amizadeh, S.., & Flint, I.. (2015). Generic and Scalable Framework for Automated Time-series Anomaly Detection. <i>Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining</i> . http://doi.org/10.1145/2783258.2788611

20. Abdulhammed, Razan., Musafer, Hassan., Alessa, Ali., Faezipour, M.., & Abuzneid, Abdel-shakour. (2019). Features Dimensionality Reduction Approaches for Machine Learning Based Network Intrusion Detection. <i>Electronics</i> . http://doi.org/10.3390/ELECTRONICS8030322