Identifying Fraudulent Credit Card Transactions Using Ensemble Learning

vallamdasu shivani
(M.Tech Artificial Intelligence)
Aurora's Scientific and Technological Institute,Telangana,India
Email: shivanivallamdas21@gmail.com
Dr.M. Sridhar
Head Of The Department Computer Science and Engineering
Aurora's Scientific and Technological Institute,Telangana,India
Email: msridhar.msr@gmail.com

ABSTRACT

The Mastercard extortion is generally come in monetary administrations. The charge card extortion is created tremendous number of issues in each year. Absence of exploration on this Mastercard issue and presents this present reality charge card misrepresentation examines, that is issues. In this paper is presented best information mining calculation called "AI calculation", which is used to perceive the Visa extortion, so at first utilize this calculation and it is one of the standard model. Then, also apply the half and half techniques in particular, "AdaBoost and greater part vote strategy". Utilize this model adequacy, which is assessed, and afterward utilize the Mastercard informational index it is openly accessible one. The monetary establishment included genuine world informational index, so it is taking and dissected. In this vigor calculation moreover assess the clamor added information tests. This idea is utilized in examination and afterward produce the outcome decidedly show the crossover technique, that is larger part casting a ballot, it gives great precision rates in Visa extortion identification

Keywords: Credit Card Fraud Detection, Ensemble Learning, Machine Learning, Fraudulent Transactions, Predictive Analytics, Random Forest, Boosting Algorithms, Anomaly Detection, Financial Security, Cybersecurity.

I. INTRODUCTION

There are different fake exercises recognition strategies has carried out in Mastercard exchanges have been kept in scientist psyches to techniques to foster models in view of man-made consciousness, information mining, fluffy rationale and AI. Mastercard misrepresentation identification is altogether troublesome, yet in addition well known issue to address. In our proposed framework we assembled the Visa misrepresentation location utilizing AI. With the progression of AI methods. AI has been recognized as a fruitful measure for extortion identification. A lot of information is moved during on the web exchange processes, bringing about a double outcome: real or deceitful. Inside the example deceitful datasets, highlights are built. These are information focuses to be specific the age and worth of the client account, as well as the beginning of the charge card. There are many highlights and each contributes, to changing degrees, towards the misrepresentation likelihood. Note, the level in which each component adds to the misrepresentation score is produced by the man-made reasoning of the machine which is driven by the preparation set, not set in stone by an extortion expert. In this way, concerning the card extortion, if the utilization of cards to commit misrepresentation is demonstrated to be high, the extortion weighting of an exchange that utilizes a Mastercard will be similarly so. Notwithstanding, if this somehow managed to shrivel, the commitment level would resemble. Just make, these models self-learn without unequivocal programming, for example, with manual survey. Mastercard misrepresentation location utilizing AI is finished by sending the order and relapse calculations. We utilize regulated learning calculation, for example, Arbitrary backwoods calculation to arrange the misrepresentation card exchange in on the web or by disconnected. Arbitrary backwoods is progressed form of Choice tree. Arbitrary backwoods has improved effectiveness and exactness than the other AI calculations. Irregular timberland expects to lessen the recently referenced connection issue by picking just a subsample of the element space at each split. Basically, it means to make the trees de-related and prune the trees by fixing a stoping rules for hub parts, which I will be cover in more detail later.

II. LITERATURE SURVEY

[1] The Use of Predictive Analytics Technology to Detect Credit Card Fraud in Canada.

AUTHORES: “Kosemani Temitayo Hafiz, Dr. Shaun Aghili, Dr. Pavol Zavarsky.”

ABSTRACT

This research paper focuses on the creation of a scorecard from relevant evaluation criteria, features, and capabilities of predictive analytics vendor solutions currently being used to detect credit card fraud. The scorecard provides a side-byside comparison of five credit card predictive analytics vendor solutions adopted in Canada. From the ensuing research findings, a list of credit card fraud PAT vendor solution challenges, risks, and limitations was outlined.

[2] BLAST-SSAHA Hybridization for Credit Card Fraud Detection.

AUTHORES: “Amlan Kundu, Suvasini Panigrahi, Shamik Sural, Senior Member, IEEE, and Arun K. Majumdar”

ABSTRACT:

This paper propose to use two-stage sequence alignment in which a profile

Analyser (PA) first determines the similarity of an incoming sequence of transactions on a given credit card with the genuine cardholder's past spending sequences. The unusual transactions traced by the profile analyser are next passed on to a deviation analyser (DA) for possible alignment with past fraudulent behaviour. The final decision about the nature of a transaction is taken on the basis of the observations by these two analysers. In order to achieve online response time for both PA and DA, we suggest a new approach for combining two sequence alignment algorithms BLAST and SSAHA.

[3] Research on Credit Card Fraud Detection Model Based on Distance Sum.

AUTHORES: "Wen-Fang YU, Na Wang".

ABSTRACT:

Along with increasing credit cards and growing trade volume in China, credit card fraud rises sharply. How to enhance the detection and prevention of credit card fraud becomes the focus of risk control of banks. It proposes a credit card fraud detection model using outlier detection based on distance sum according to the infrequency and unconventionality of fraud in credit card transaction data, applying outlier mining into credit card fraud detection. Experiments show that this model is feasible and ac

[4] Fraudulent Detection in Credit Card System Using SVM & Decision Tree.

AUTHORES: "Vijayshree B. Nipane, Poonam S. Kalinge, Dipali Vidhate, Kunal War, Bhagyashree P. Deshpande".

ABSTRACT:

With growing advancement in the electronic commerce field, fraud is spreading all over the world, causing major financial losses. In current scenario, Major cause of financial losses is credit card fraud; it not only affects trades person but also individual clients. Decision tree, Genetic algorithm, Meta learning strategy, neural network, HMM are the presented methods used to detect credit card frauds. In contemplate system for fraudulent detection, artificial intelligence concept of Support Vector Machine (SVM) & decision tree is being used to solve the problem. Thus by implementation of this hybrid approach, financial losses can be reduced to greater extend.

[5] Supervised Machine (SVM) Learning for Credit Card Fraud Detection.

AUTHORES: "Sitaram patel, Sunita Gond".

ABSTRACT:

This thesis propose the SVM (Support Vector Machine) based method with multiple kernel involvement which also includes several fields of user profile instead of only spending profile. The simulation result shows improvement in TP (true positive), TN (true negative) rate, & also decreases the FP (false positive) & FN (false negative) rate.

[6] Detecting Credit Card Fraud by Decision Trees and Support Vector Machines.

AUTHORES: "Y. Sahin and E. Duman"

ABSTRACT:

In this study, classification models based on decision trees and support vector machines (SVM) are developed and applied on credit card fraud detection problem. This study is one of the firsts to compare the performance of SVM and decision tree methods in credit card fraud detection with a real data set.

III. EXISTING SYSTEM

In existing Framework, an exploration about a contextual analysis including charge card extortion identification, where information standardization is applied before Group Examination and with results got from the utilization of Bunch Investigation and Counterfeit Brain Organizations on misrepresentation recognition has shown that by bunching credits neuronal data sources can be limited. Also, encouraging outcomes can be gotten by utilizing standardized information and information ought to be MLP prepared. This examination depended on unaided learning. Meaning of this paper was to track down new strategies for extortion identification and to expand the exactness of results. The informational index for this paper depends on genuine value-based information by an enormous European organization and individual subtleties in information is kept private. Precision of a calculation is around half. Meaning of this paper was to track down a calculation and to decrease the expense measure. The outcome got was by 23% and the calculation they find was Bayes least gamble.

IV. PROPOSED SYSTEM

In proposed Framework, we are applying arbitrary woodland calculation for order of the charge card dataset. Arbitrary Woodland is a calculation for order and relapse. Immediately, it is an assortment of choice tree classifiers. Arbitrary woodland enjoys upper hand over choice tree as it amends the propensity for over fitting to their preparation set. A subset of the preparation set is examined haphazardly so that to prepare every individual tree and afterward a choice tree is constructed, every hub then parts on a component chose from an irregular subset of the completely unlocked set. In any event, for enormous informational indexes with many highlights and information examples preparing is very quick in arbitrary woodland and on the grounds that each tree is prepared freely of the others. The Irregular Backwoods calculation has been found to give a decent gauge of the speculation blunder and to be impervious to over fitting.

V. SYSTEM ARCHITECTURE

The above diagram represents the working process of a fraud detection system using machine learning and ensemble learning techniques. Initially, data is collected from the real network environment, which includes credit card transaction details such as transaction amount, time, user information, and location. This collected data may contain missing values, noise, or unnecessary information. Therefore, the data is passed to the preprocessing stage, where cleaning, normalization, feature selection, and transformation are performed to improve the quality of the dataset.

After preprocessing, the data is divided into two parts: training data and monitoring data. The training data is used in the training phase to build the machine learning model. During this phase, ensemble learning algorithms such as Random Forest, Gradient Boosting, or AdaBoost learn patterns from both legitimate and fraudulent transactions. The trained system model stores the learned behavior and relationships among transaction features, which helps in identifying suspicious activities effectively.

The monitoring data is then sent to the detection module, where real-time transaction analysis is carried out. The detection system compares incoming transactions with the trained system model to determine whether a transaction is genuine or fraudulent. Ensemble learning improves detection accuracy by combining the predictions of multiple classifiers, thereby reducing false alarms and increasing reliability.

Finally, the system generates decision

results based on the detection outcome. If the transaction is identified as normal, it is approved; otherwise, it is marked as fraudulent and can be blocked or flagged for further investigation. This framework helps financial institutions enhance transaction security, reduce financial losses, and provide safer online payment services.

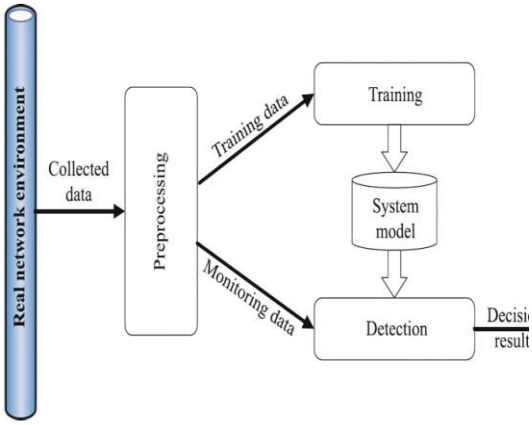

Fig 5.1: System Architecture Of Proposed System

VI. IMPLEMENTATION

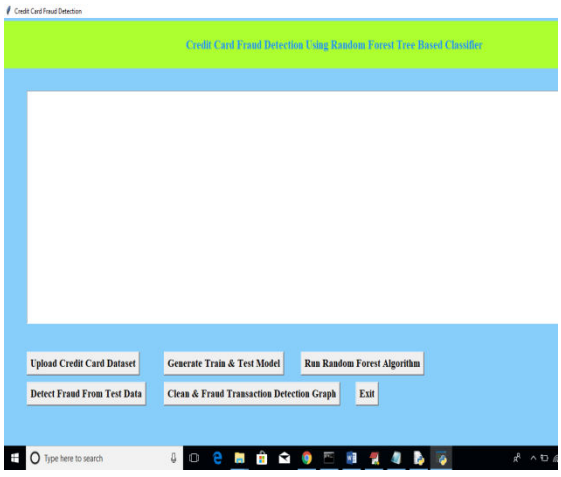

Fig 6.1: Home Page

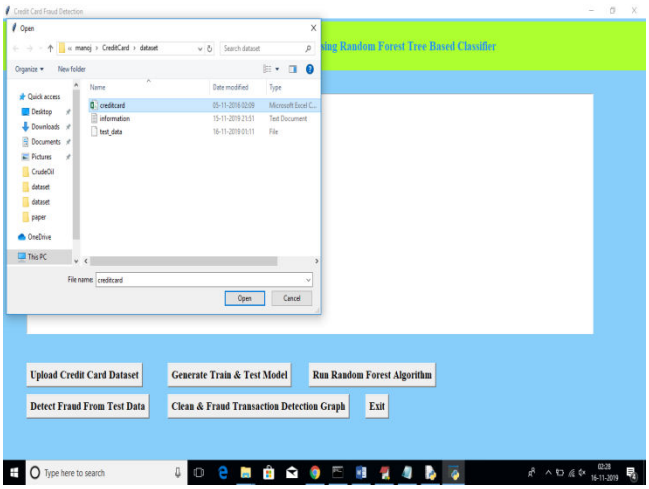

Fig 6.2: Load Dataset

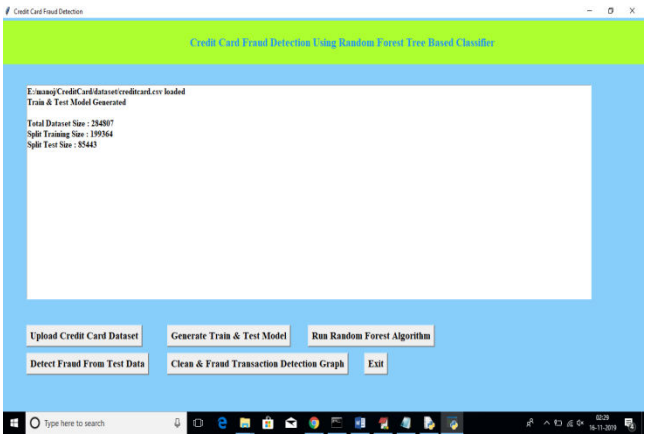

Fig 6.3: Model Training

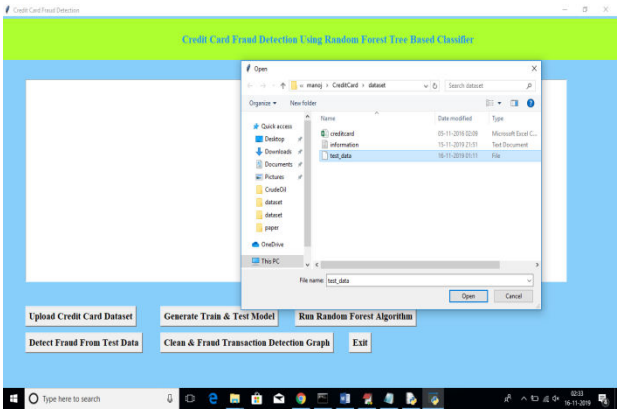

Fig 6.4: Prediction Page

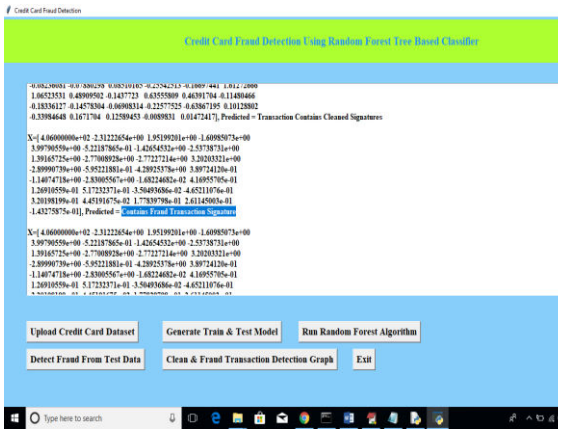

Fig 6.5: Result Page

VII. CONCLUSION

The Random forest algorithm will perform better with a larger number of training data, but speed during testing and application will suffer. Application of more pre-processing techniques would also help. The SVM algorithm still suffers from the imbalanced dataset problem and requires more preprocessing to give better results at the results shown by SVM is great but it could have been better if more preprocessing have been done on the data.

VIII. FUTURE SCOPE

The future scope of Identifying Fraudulent Credit Card Transactions Using Ensemble Learning is highly promising due to the rapid growth of digital payments and online banking systems. Future systems can be enhanced by integrating advanced deep learning techniques such as Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN), and Transformer models to improve the detection of complex fraud patterns in real-time transaction data. These models can learn user behavior more accurately and identify hidden fraudulent activities with higher precision.

The proposed system can also be extended by incorporating real-time big data analytics and cloud computing technologies. This will enable financial institutions to process millions of transactions instantly with reduced computational delay. Integration with distributed computing frameworks such as Hadoop and Spark can further improve scalability and performance for large-scale financial datasets.

Another important future enhancement is the use of adaptive and self-learning fraud detection systems. Fraud techniques continuously evolve, and future models can utilize reinforcement learning and online learning methods to automatically update themselves based on new fraud patterns. This will help in reducing false positives and maintaining high detection accuracy over time.

In addition, biometric authentication, blockchain technology, and Internet of Things (IoT)-based transaction monitoring can be integrated with ensemble learning models to provide stronger security mechanisms. Mobile banking applications and e-commerce platforms can benefit from multi-layer security frameworks that combine behavioral analysis, geolocation tracking, and transaction history analysis.

Future research can also focus on explainable artificial intelligence (XAI) techniques to make fraud detection systems more transparent and understandable. This will help banks and financial organizations understand why a transaction was classified as fraudulent, thereby improving trust, reliability, and decision-making processes in financial security systems.

IX. REFERENCES

1. [1] S. Jha, M. Guillen, and J. C. Westland, “Employing transaction aggregation strategy to detect credit card fraud,” Expert Systems with Applications, vol. 39, no. 16, pp. 12650–12657, 2012. DOI: 10.1016/j.eswa.2012.05.078
2. [2] A. Dal Pozzolo, O. Caelen, Y. A. Le Borgne, S. Waterschoot, and G.

Bontempi, “Learned lessons in credit card fraud detection from a practitioner perspective,” Expert Systems with Applications, vol. 41, no. 10, pp. 4915–4928, 2014.
DOI: 10.1016/j.eswa.2014.02.026

3. [3] V. Phua, C. Lee, K. Smith, and R. Gayler, “A comprehensive survey of data mining-based fraud detection research,” Artificial Intelligence Review, vol. 34, no. 1, pp. 1–14, 2010.
DOI: 10.1007/s10462-009-9129-y
4. [4] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, “Data mining for credit card fraud: A comparative study,” Decision Support Systems, vol. 50, no. 3, pp. 602–613, 2011.
DOI: 10.1016/j.dss.2010.08.008
5. [5] T. Sahin and E. Duman, “Detecting credit card fraud by decision trees and support vector machines,” International MultiConference of Engineers and Computer Scientists, pp. 442–447, 2011.
DOI: 10.48550/arXiv.1109.6897
6. [6] C. Whitrow, D. J. Hand, P. Juszczak, D. Weston, and N. M. Adams, “Transaction aggregation as a strategy for credit card fraud detection,” Data Mining and Knowledge Discovery, vol. 18, no. 1, pp. 30–55, 2009.
DOI: 10.1007/s10618-008-0116-z
7. [7] R. Jurgovsky, M. Granitzer, S. Ziegler, et al., “Sequence classification for credit-card fraud detection,” Expert Systems with Applications, vol. 100, pp. 234–245, 2018.
DOI: 10.1016/j.eswa.2018.01.037
8. [8] A. Carcillo, Y. A. Le Borgne, O. Caelen, et al., “Combining unsupervised and supervised learning in credit card fraud detection,” Information Sciences, vol. 557, pp. 317–331, 2021.
DOI: 10.1016/j.ins.2019.05.042
9. [9] D. Dua and C. Graff, “UCI Machine Learning Repository,” University of California, Irvine, 2019.
DOI: 10.24432/C5NC77
10. [10] L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
DOI: 10.1023/A:1010933404324
11. [11] J. H. Friedman, “Greedy function approximation: A gradient boosting machine,” Annals of Statistics, vol. 29, no. 5, pp. 1189–1232, 2001.
DOI: 10.1214/aos/1013203451
12. [12] Y. Freund and R. E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting,” Journal of Computer and System Sciences, vol. 55, no. 1, pp. 119–139, 1997.
DOI: 10.1006/jcss.1997.1504
13. [13] N. Japkowicz and S. Stephen, “The class imbalance problem: A systematic study,” Intelligent Data Analysis, vol. 6, no. 5, pp. 429–449, 2002.
DOI: 10.3233/IDA-2002-6504
14. [14] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
DOI: 10.5555/3086952
15. [15] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, 2016, pp. 785–794.
DOI: 10.1145/2939672.2939785
16. 10. Todupunuri, A. (2024). Explore How AI Can Be Used To Create Dynamic And Adaptive Fraud & Rules That Improve The Detection And Prevention Of Fraudulent & Activities In Digital Banking. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5014699
17. 11. Babburi, S. Privacy-Preserving Collaborative Framework with Auditable Federated Learning.
18. 12. Gaddam, S. (2024). Integrating machine learning models with continuous integration and continuous delivery (CI/CD) pipelines for a learning-driven approach to software engineering.
19. 13. Immadi, S. K. (2025).

Optimizing ERP for Human Capital Management. Applied Research for Growth, Innovation and Sustainable Impact, 377–384. https://doi.org/10.1201/9781003684657-63

20. 14. Reddy, S. K. R. Developing a Modular AI Framework to Enhance Scalability and Personalization in Next-Generation Reward Platforms.
21. 15. Poojari, R. INTELLIGENT SYSTEMS+B108 AND APPLICATIONS IN ENGINEERING.
22. 16. Vasagam, M. (2024, August 30). Ensuring security in modern data pipelines: Practical strategies for data engineers. International Journal of Intelligent Systems and Applications in Engineering, 12(22s), 2401.
23. 17. Santthosh Saai Reddy Purmani. (2026). Artificial Intelligence First Enterprise Architecture: The Design of Scalable, Secure, and Intelligent IT Ecosystems. American Journal of AI Cyber Computing Management, 6(1(2)), 1–8. https://doi.org/10.64751/ajaccm.2026.v6.n1(2).pp1-8
24. 18. Purmani, S. S. R. (2024). Aligning IT investment decisions with overall business strategy from an enterprise program management perspective, focusing on the integration of IT leadership in strategic decision-making processes. International Journal of Communication Networks and Information Security, 16(5), 1213–1219
25. 19. Kumara, S. (2026, February). A Lightweight Deep Learning Based Classification Models for Non-Human Identity Threat Detection. In 2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC) (pp. 1-6). IEEE.
26. 20. Kotte, G. (2025). Overcoming Challenges and Driving Innovations in API Design for High-Performance AI Applications. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5283649
27. 21. Kotte, G. (2025). Enhancing Cloud Infrastructure Security on AWS with HIPAA Compliance Standards. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5283660
28. 22. Ranjbareslamloo, S., Dzukeya, G. A., Muhit, M. M. I., & Qattawi, A. (2025). Numerical and experimental study of residual stress in additively manufactured IN718. Manufacturing Letters, 44, 915–927. https://doi.org/10.1016/j.mfglet.2025.915927
29. 23. Viswanathan, V. (2023). AI-Augmented Decision Intelligence for Enterprise Systems: Integrating Cognitive Analytics for Resource and Talent Optimization.
30. 24. Viswanathan, V. Generative AI for Smarter Workforce Planning and Enterprise Resource Decisions.
31. 25. Mudusu, S. (2025). Health Insurance Fraud Detection: The Role Of Advanced It Systems In Preventing And Identifying Fraud. International Journal, 16(1), 3769-3777
32. 26. Mudusu, S. K. (2026, April 15). The secure intelligence framework: Architecting AI systems for a data-driven world. CIO (Foundry Expert Contributor Network).
33. 27. Agrawal, A. M., Gajula, S., Shinde, R. P., Shah, H., & Ghosh, H. (2025, July). Machine Translation for Long Sequences with Enhanced Attention Mechanisms. In 2025 5th International Conference on Electrical, Computer and Energy Technologies (ICECET) (pp. 1-6). IEEE.
34. 28. Gajula, S. (2026, March). Two Pillars of Banking Intelligence: A Comparative Analysis of AI Techniques for Fraud Prevention and Churn Mitigation. In 2026 14th

International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-6). IEEE.

35. 29. Maturi, S. Y. (2021). Blockbond hardening: Securing pooled-hash protocols against traffic tampering, MITM hash-rate hijacking, and template coercion. International Journal of Communication Networks and Information Security, 13(3), 718–728.
36. 30. Maturi, S. Y. (2023). Crowdsourced frontier: Unveiling autonomous adversarial cybercapabilities via open AI competition. International Journal of Intelligent Systems and Applications in Engineering, 11(1s), 275–284.
37. 31. Sikder, M. Z., Shakil, M. A. I., Ahad, A., Karim, M. F., Intakhab, B., & Islam, D. A. (2025, June). Microwave-Based Detection of Early-Stage Renal Cell Carcinoma Using UHF Range Antenna. In 2025 International Conference on Computer Systems and Technologies (CompSysTech) (pp. 1-6). IEEE.
38. 32. Manoharan, D. (2024). Governance-Oriented Quality Engineering Framework for Healthcare EDI Modernization. International Journal of Multidisciplinary on Science and Management IJMSM, 1(2).
39. 33. Manoharan, D. (2026). Advancing Healthcare EDI Interoperability Through Informatica Cloud B2B Gateway Quality Engineering. Available at SSRN 6385719.
40. 34. Ravishankara, M. (2026, February). PlotChain: Deterministic Checkpointed Evaluation of Multimodal LLMs on Engineering Plot Reading. In SoutheastCon 2026 (pp. 1-8). IEEE.
41. 35. Doragacharla, V. R. (2026). Building Real-Time Pricing Systems for Modern Retail. Available at SSRN 6451760.
42. 36. Adabala, P. K. (2024). Utilizing predictive analytics to improve efficiency and decision-making in ERP-connected supply chains. International Journal of Intelligent Systems and Applications in Engineering, 12(22s), 2465
43. 37. Venkata Ramana, P. (2024). AI-driven predictive analytics in ERP systems for proactive supply chain optimization. International Journal of Research in Information Technology and Computing, 8(4).
44. 38. Kavuri, S. (2026). An Explainable Machine Learning Framework for Predicting Software Defects in Large-Scale Software Systems. 2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC), 1–6. https://doi.org/10.1109/icaic67076.2026.11395777
45. 39. Srikanth Kavuri. (2025). AI-DRIVEN TEST AUTOMATION FRAMEWORKS: ENHANCING EFFICIENCY AND ACCURACY IN SOFTWARE QUALITY ASSURANCE. International Journal of Applied Mathematics, 38(10s), 699–710. https://doi.org/10.12732/ijam.v38i10s.990
46. 40. Venkata Pavan Kumar Gummadi. (2023). MuleSoft Batch Processing: High-Volume Streaming Architecture. Computer Fraud and Security, 50–57. https://doi.org/10.52710/cfs.886
47. 41. Venkata Pavan Kumar Gummadi. (2026). Infrastructure Optimization Techniques for Enterprise Integration Platforms: A Comprehensive Analysis. Computer Fraud and Security, 37–44. https://doi.org/10.52710/cfs.875
48. 42. Shashank, A. (2025). Self-Healing Data Pipelines for Enhanced Reliability: A Paradigm Shift in Enterprise Data Management. Journal of Computer Science and Technology Studies, 7(8), 1097-1104.
49. 43. Harshitha, G. K., Nandigama, C., & Thiripalu, P. (2026). An exploration into identification of opportunities and

challenges of establishing and running an enterprise in the area of biofuels. Minnesota Journal of Business Law and Entrepreneurship, 2026(1), 1159–1168.

50. 44. Ghali Krishna Harshitha & P. Thiripalu. (2025). Assessing the influence of age and gender on soft skills among emerging Gen Z HR professionals. Advances in Consumer Research, 2(2), 991–999.