IJIRAE:: International Journal of Innovative Research in Advanced Engineering ISSN:2349-2163
Volume 13, Issue 04, April https://www.ijirae.com/archives
https://doi.org/10.26562/ijirae.2026.v1304.08

An Enhanced RNN-LSTM Model for Accurate and Real Time Fraud Detection in Online Advertising

Illendula Sai Krishna
Assistant Professor, Department of CSE
Guru Nanak Institute of Technology, Hyderabad, India
saikrishnai.csegnit@gniindia
https://orcid.org/0009-0007-4764-1082
Godala Ashwitha Bushodu,Laxman,Galla Poojitha
UG Student, Department of CSE,
Guru Nanak Institute of Technology, Hyderabad, India
ashwithagoudgodala@gmail.com, laxmanbushodu672@gmail.com, gallanani25@gmail.com

Publication History
Manuscript Reference No: IJIRAE/RS/Vol.13/Issue04/AEAP26.APAE10087
Research Article | Open Access | Double-Blind Peer-Reviewed| Article ID: IJIRAE/RS/Vol.13/Issue04/AEAP26.APAE10187
Received:02,March 2026, Revised: 29, March 2026, Accepted: 10,April 2026,Published Online: 22, April 2026.
https://www.ijirae.com/volumes/Vol13/iss-04/08.AEAP26.APAE10087.pdf
Article Citation: IllendulaSai,Godala,Laxman,Galla(2026),An Enhanced RNN-LSTM Model for Accurate and Real Time Fraud Detection in Online Advertising, IJIRAE: International Journal of Innovative Research in Advanced Engineering, Volume 13, Issue 04 of 2026 pages 779-784 **Doi:>** https://doi.org/10.26562/ijirae.2026.v1304.08
BibTeX Key: Illendula@2026Enhanced
IJIRAE papers should be cited as IJIRAE (International Journal of Innovative Research in Advanced Engineering, AM Publications, India 2025, ISSN 2349-2163, https://doi.org/10.26562/ijirae.2026.v1304.08 The journal's official abbreviation is IJIRAE. Orcid: https://orcid.org/0009-0004-9398-7488

Abstract: Click fraud remains a major threat in online advertising, increasing costs and reducing campaign effectiveness by directing budgets toward illegitimate activity. Although machine learning and deep learning approaches have shown promise, many struggle to detect subtle behavioural patterns in fraudulent clicks. This work proposes a robust LSTM-based Recurrent Neural Network (RNN) framework that improves fraud detection by modelling sequential patterns and time-dependent features in user interaction data. A comprehensive preprocessing pipeline comprising timestamp decomposition, feature scaling, and label encoding was implemented to ensure optimal data representation. The model was trained and evaluated on a carefully engineered dataset enriched with behavioural and contextual click features and compared with other architectures such as Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN).The RNN-LSTM model outperformed the others, achieving 99% accuracy along with high precision and recall, demonstrating the effectiveness of temporal modelling in detecting fraudulent click patterns and its suitability for real-time deployment, while also contributing to advancements in intelligent ad verification and fraud prevention.
Keywords: Click Fraud Detection, Online Advertising, RNN-LSTM, Deep Learning, Sequential Analysis, User Behavior, Feature Engineering, Real-Time Fraud Detection

I. INTRODUCTION

Electronic In the era of digital transformation, online advertising has become a key platform for businesses to reach global audiences, generate leads, and drive conversions. However, it faces a major challenge in the form of click fraud, where fake or automated clicks are generated on pay-per-click (PPC) ads using bots, scripts, or malicious users. This leads to financial losses, distorted analytics, and reduced trust in advertising platforms, while existing rule-based and traditional machine learning methods often fail to detect evolving and complex fraud patterns. To overcome these limitations, this project proposes a deep learning-based approach using Recurrent Neural Networks (RNN) with Long Short-Term Memory (LSTM) architecture. The model captures sequential and time-dependent user behavior using features like click frequency, session timing, and user interactions. Compared with other models such as Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN), the proposed RNN-LSTM model achieves superior performance with around 99% accuracy, effectively identifying fraudulent clicks with high precision, recall, and F1-score.

II. LITERATURE SURVEY

A. Alzahrani, Malak Aljabri, Rami Mustafa A. Mohammad (2025). This paper focuses on detecting click fraud using Machine Learning (ML) and Deep Learning (DL) models, highlighting the impact of fraudulent clicks on online advertising. It applies feature engineering and Recursive Feature Elimination (RFE) to enhance model performance. Results show ML models support real-time detection, while DL models better capture complex patterns, improving accuracy and reducing false positives.

Zainab A. Abbas, Zahraa M. Hilal, Hanan G. Jabbar (2024) This paper compares machine learning models like Decision Tree, Random Forest, Gradient Boosting, XGBoost, and RNN for click fraud detection using click stream data. Tree-based models achieve accuracy above 96%, while RNN captures temporal behaviour with slightly lower accuracy. It highlights the need to combine ML and DL techniques to improve accuracy and adapt to evolving fraud patterns.

Tangavelou Karthikeyan, Veeramani Vijayakumar (2024) This paper proposes a hybrid model called Chimp-Optimized LSTMfor fraud detection. It combines a metaheuristic optimization algorithm with LSTM to enhance feature selection and classification performance. The model is tested on financial datasets such as credit card and insurance data. The approach improves precision, recall, and reduces error rates, making it a reliable solution for detecting complex fraud patterns.

Ranjeet Vishwakarma, Rajesh Dhakad (2024) – This paper provides a survey on online advertising and click fraud, highlighting different types of fraud such as click farms, bots, and fake interactions. It discusses traditional detection methods like rule-based systems and modern approaches like machine learning. The study emphasizes the need for intelligent and real-time detection systems to improve accuracy and reliability in online advertising platforms.

B. Kirkwood, M. Vanamala, N. Seliya (2024) – This paper focuses on machine learning techniques for click fraud detection using models like Decision Tree, SVM, and Random Forest. It highlights the importance of feature engineering and real-time data analysis. The study shows that Random Forest performs best in handling noisy and imbalanced data, improving overall detection accuracy

III. METHODOLOGY

The proposed click fraud detection system follows a structured workflow using a deep learning RNN-LSTM model to identify fraudulent clicks effectively.

A.Data Collection: The system collects click stream data such as timestamps, IP addresses, device information, and user interaction features like click frequency and session duration. Both legitimate and fraudulent click data are gathered to ensure balanced training.

B.Data Preprocessing: The collected data is cleaned and formatted by handling missing values, removing inconsistencies, and applying techniques like timestamp decomposition, label encoding, and feature scaling to prepare structured input.

C.Feature Engineering: Important features such as time-based patterns, user behavior metrics, and session-level attributes are extracted to enhance model performance and capture fraud-related patterns.

D.Model Development (RNN-LSTM): The processed data is reshaped into sequential format and fed into the LSTM model, which learns temporal dependencies and behavioral patterns in user clicks to detect fraud effectively.

E.Fraud Detection (Prediction): The trained model analyzes new input data and classifies clicks as fraudulent or legitimate based on learned patterns, supporting real-time detection.

F.Performance Evaluation: The model is evaluated using metrics such as accuracy, precision, recall, and F1-score, along with confusion matrix analysis to validate its effectiveness.

G.Deployment and Result Display: The final model is deployed in a web-based system where results are displayed clearly to the user, enabling real-time monitoring and detection of fraudulent activities.

IV. SYSTEM ARCHITECTURE

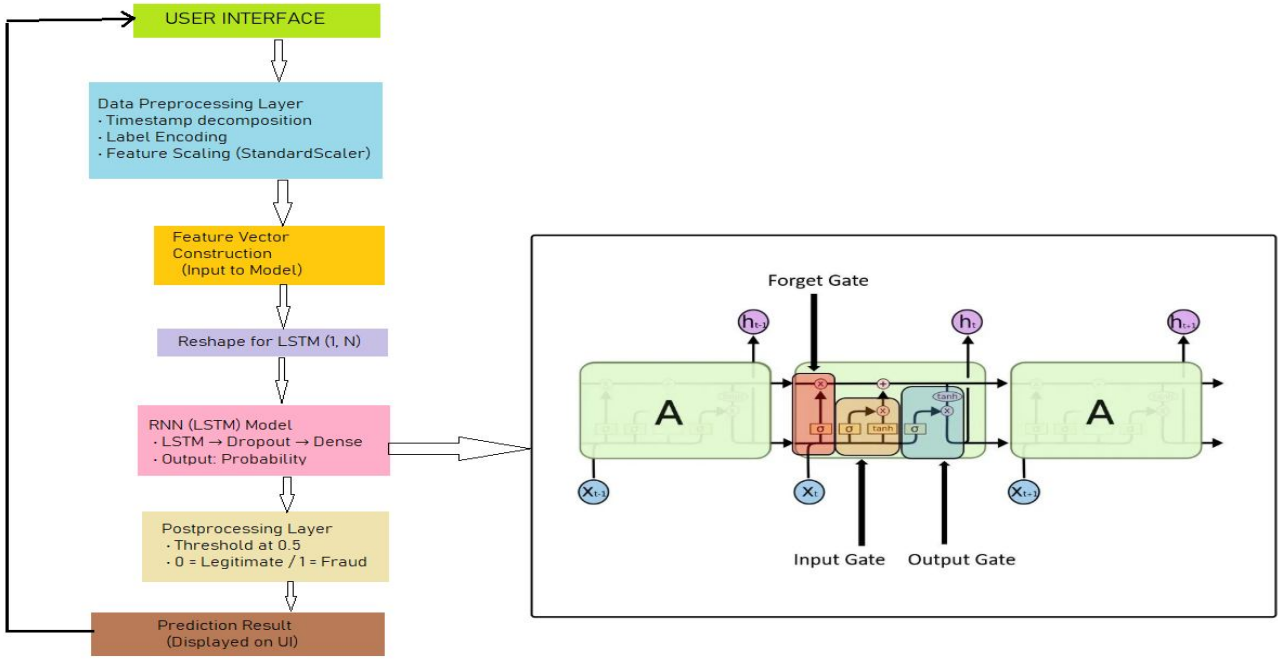

Fig 1: System Architecture

V. IMPLEMENTATION

The implementation process of the proposed system is carried out through a web-based click fraud detection framework using deep learning techniques, specifically an RNN-LSTM model. The system integrates data preprocessing, feature engineering, and real-time prediction to ensure accurate identification of fraudulent clicks in online advertising. It leverages sequential data analysis to capture user behavior patterns over time, improving the reliability and effectiveness of fraud detection.

The architecture is divided into multiple modules, including data collection, preprocessing, and model training, prediction, and result visualization. This modular design ensures scalability, flexibility, and efficient performance. The trained LSTM model is used to analyze clickstream data and classify it as fraudulent or legitimate. The system is implemented using a combination of front-end technologies for user interaction and back-end frameworks (such as Python-based libraries) along with a database for storing data and results. The architecture also supports handling large volumes of clickstream data and enables real-time fraud detection. It ensures smooth integration between the user interface, processing modules, and the database system. Overall, the system provides an efficient and intelligent platform for detecting click fraud, helping advertisers safeguard their campaigns and improve decision-making.

Existing System

Current click fraud detection systems mainly use traditional machine learning algorithms such as Decision Tree, Random Forest, SVM, and rule-based methods. These systems rely on manually engineered features like IP address, click frequency, and device information to identify fraudulent activity. While they provide moderate accuracy, they analyze clicks independently and fail to capture sequential and temporal user behavior. As a result, they struggle to detect complex and evolving fraud patterns such as bot attacks and coordinated clicks. Additionally, these systems face challenges like high false positives, poor handling of imbalanced data, and limited real-time adaptability, reducing their overall effectiveness in fraud detection.

Proposed System

The proposed click fraud detection system uses a deep learning approach based on an RNN-LSTM model to accurately identify fraudulent clicks. It processes clickstream data by applying preprocessing techniques such as timestamp decomposition, feature scaling, and label encoding to generate high-quality input. The system focuses on analyzing sequential and time-dependent user behaviour, enabling it to capture complex patterns that traditional models cannot detect.The LSTM model learns temporal dependencies in user interactions such as click frequency, session duration, and activity patterns to distinguish between legitimate and fraudulent clicks. The system supports real-time prediction through a web-based interface and provides accurate classification results. Overall, it offers improved accuracy, adaptability, and efficiency in detecting evolving click fraud activities compared to existing methods.

VI. RESULTS

Login Page

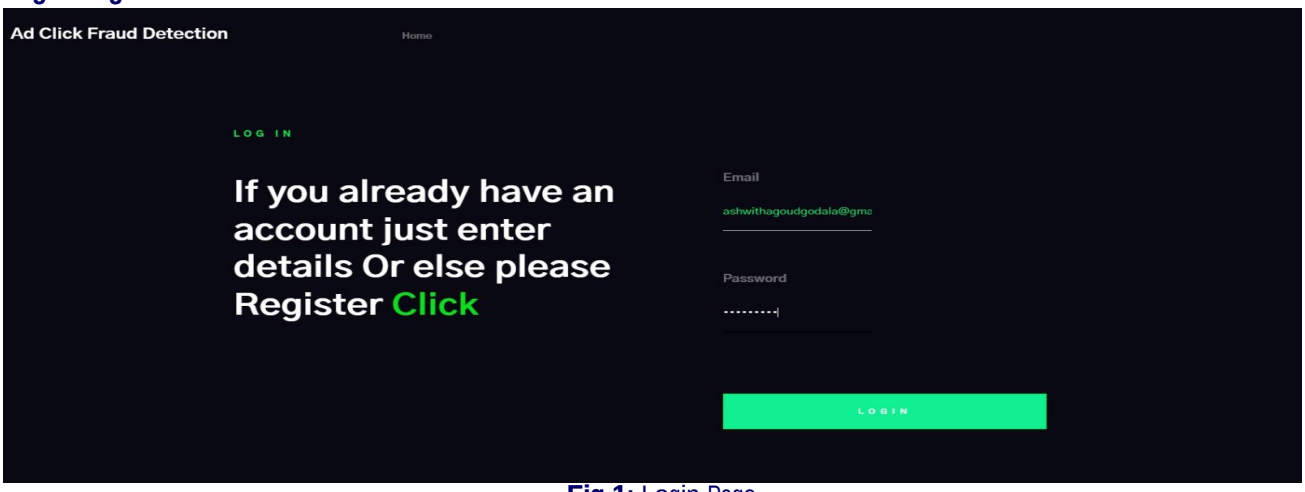

Fig 1: Login Page

The Login Page displays the interface used by registered users to access the click fraud detection system. It allows users to enter their username and password for secure authentication. The design is simple and clean, ensuring easy navigation and quick access to the platform. Overall, the page is user-friendly and provides safe and efficient system access.

REGISTRATION PAGE

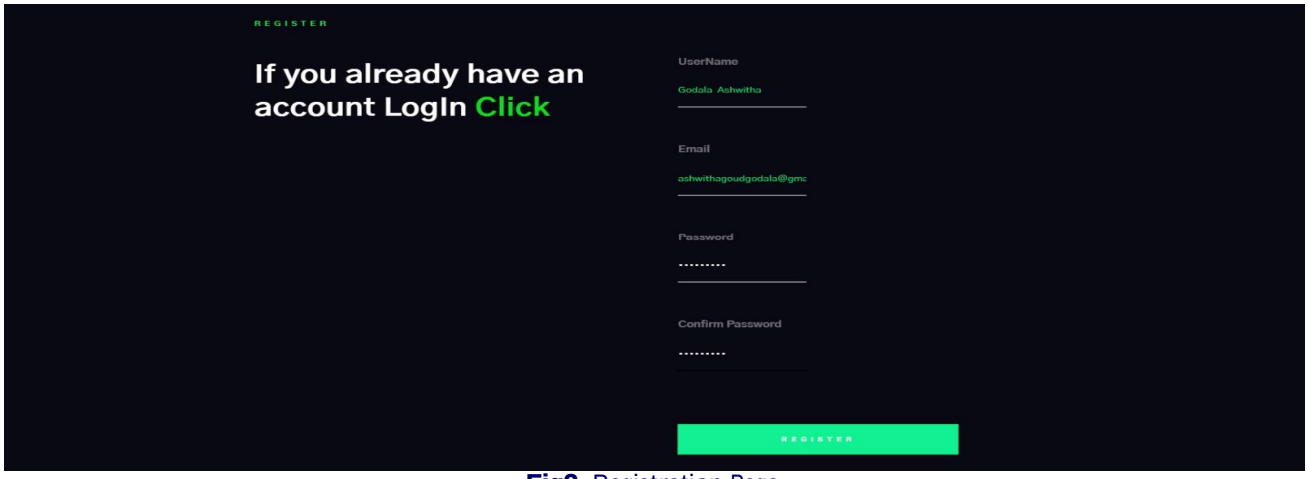

Fig2: Registration Page

The Registration Page displays the sign-up interface of the click fraud detection system. It provides a simple form for users to enter details such as username, password, and confirm password to create an account. The design is clean and focused, ensuring ease of use and smooth registration. Overall, the page is user-friendly, secure, and designed for quick access to the system.

ABOUT PAGE

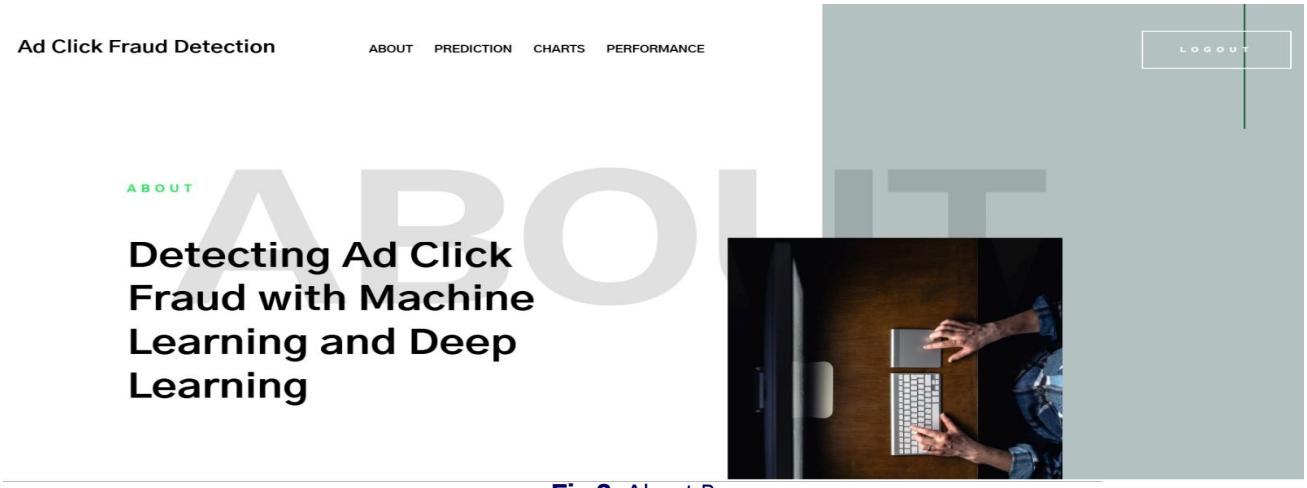

Fig 3: About Page

PREDICTION PAGE

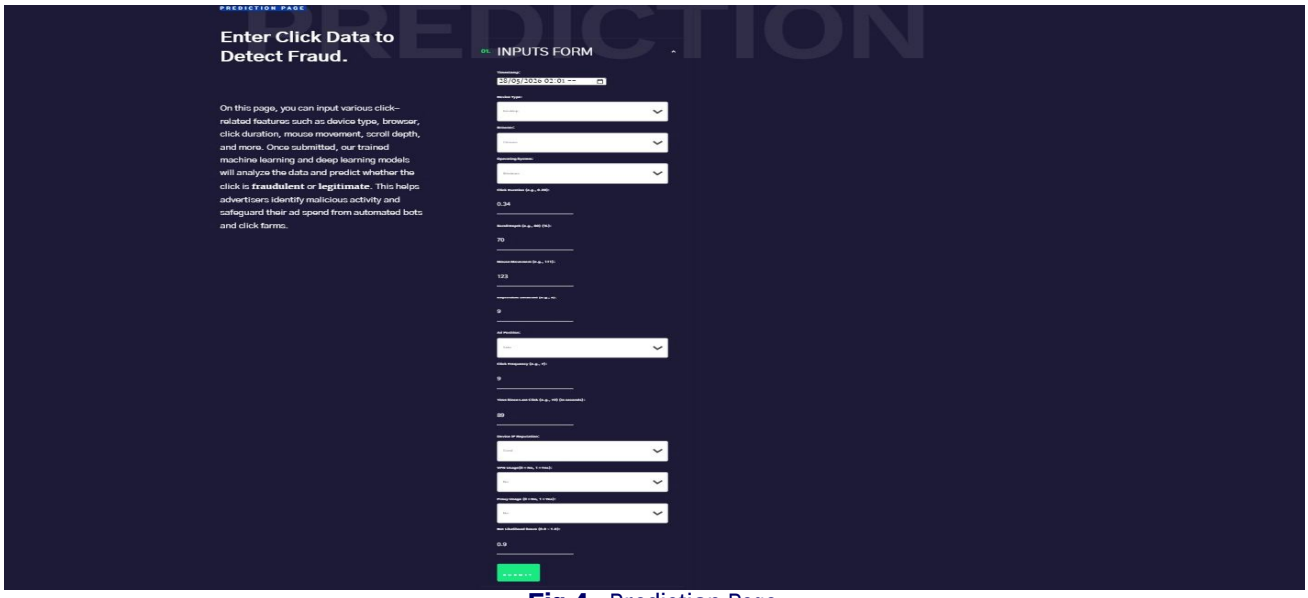

Fig 4 : Prediction Page

The module is responsible for handling the logging and managing of admins in the system, giving the logged-in admins secure access to the system. Admins can be able to monitor and keep track of voting processes, verifying voters and tracking results.

RESULT PAGE:

Fig 5: Result Page

The Detection displays the result after analyzing user input in the click fraud detection system. It shows the entered data along with the predicted result, indicating whether the activity is fraudulent or legitimate. The outcome is clearly highlighted for easy understanding. Overall, the page is simple and focused on delivering accurate prediction results effectively.

PERFORMANCE PAGE:

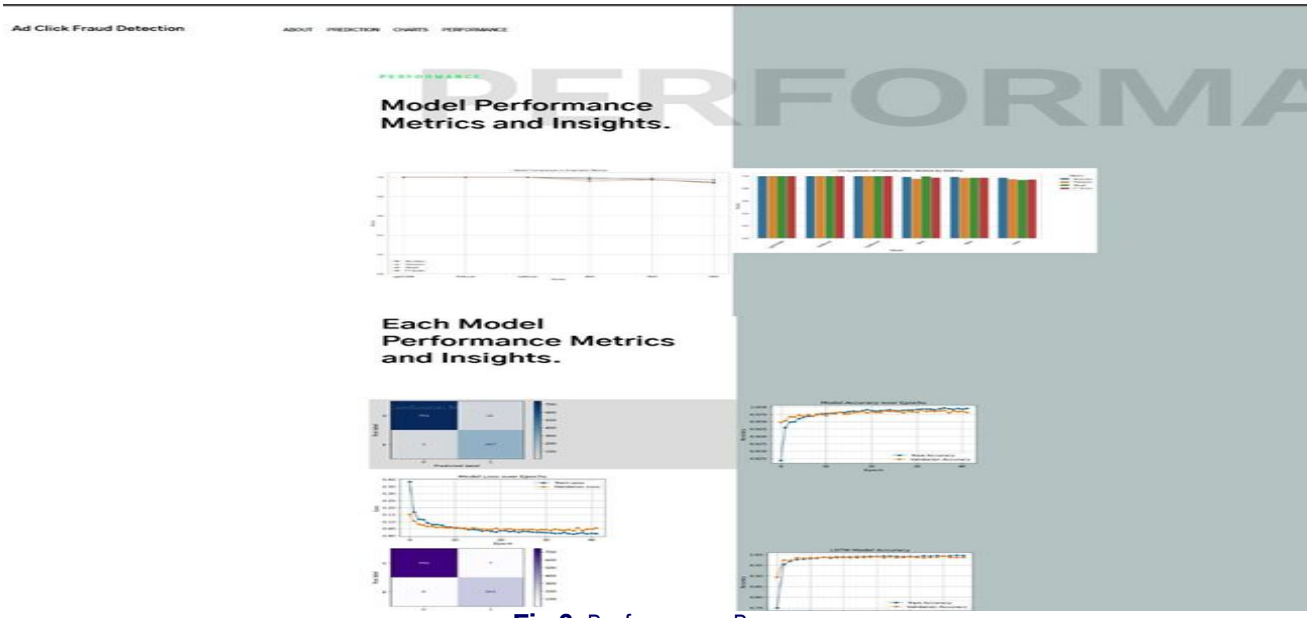

Fig 6: Performance Page

The Performance Page shows model metrics and insights for the click fraud detection system using graphs and charts. It includes accuracy, loss curves, and confusion matrices to evaluate performance and reliability. It also compares different models and highlights training progress over epochs. Overall, it provides clear insights for improving model efficiency.

CHART PAGE

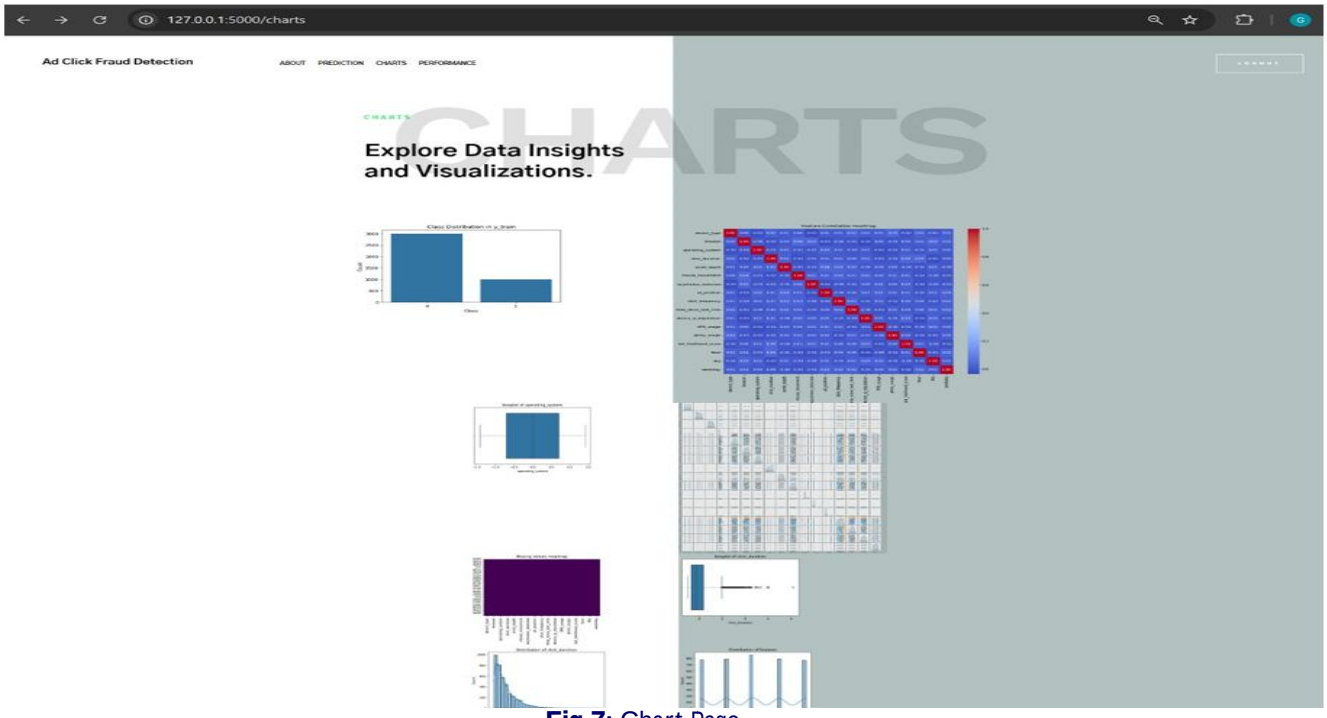

Fig 7: Chart Page

VII. CONCLUSION

The proposed click fraud detection system uses a deep learning approach based on an RNN-LSTM model to accurately identify fraudulent clicks in online advertising. It applies preprocessing techniques such as timestamp decomposition, feature scaling, and label encoding to prepare high-quality input data. The LSTM model effectively captures sequential and temporal patterns in user behavior, enabling it to detect both simple and complex fraud activities. This results in higher accuracy compared to traditional machine learning methods. The system also performs well on imbalanced and real-world datasets, ensuring reliability and robustness.

It can be integrated into web-based platforms for real-time fraud detection and monitoring, helping advertisers protect their budgets and improve campaign performance. Overall, the proposed system enhances fraud detection efficiency and can be further extended for large-scale deployment and evolving fraud patterns in online advertising environments.

VIII. FUTURE ENHANCEMENT

The next generation to improve the robustness and real-time effectiveness of click fraud detection, future enhancements can include integrating real-time streaming platforms like Apache Kafka or Spark Streaming for instant fraud detection. Incorporating unsupervised learning methods such as autoencoders and clustering can help identify new fraud patterns beyond labeled data. Advanced models like Transformers can better capture sequential and temporal user behavior, while user profiling and behavioral analytics can enhance contextual understanding. Additionally, deploying the system in a cloud-based, containerized environment ensures scalability, reliability, and efficient handling of large-scale traffic.

REFERENCES

1. Juniper Research,Hampshire,U.K.Quantifying the Cost of Ad Fraud: 2023–2028. Accessed: Jul. 12, 2024. https://www.fraudblocker.com/wpcontent/uploads/2023/09/Ad-Fraud-Whitepaper_Juniper-Research.pdf
2. X.Zhu, H.Tao, Z.Wu, J.Cao, K.Kalish, and J.Kayne, Fraud Prevention in Online Digital Advertising. Cham, Switzerland: Springer, 2017.
3. A.K.Wood and A.M.Ravel, ‘‘Fool me once: Regulating fake news and other online advertising,’’ S. Cal. L. Rev., vol. 91, p. 1223, Jan. 2017.
4. B.Stone-Gross, R. Stevens, A. Zarras, R. Kemmerer, C. Kruegel, and G. Vigna, ‘‘Understanding fraudulent activities in online ad exchanges,’’ in Proc. ACM SIGCOMM Conf. Internet Meas. Conf., Nov. 2011, pp. 279–294.(2024).
5. Wasted Ad Spend Report 2024. https://www.lp.lunio.ai/wpcontent/ uploads/2023/09 /Lunio_Wasted_ Ad_Spend Report_2024_V2.pdf 12762
6. D.Berrar,‘‘Random forests for the detection of click fraud in online mobile advertising,’’ in Proc. Int. Work. Fraud Detect. Mob. Advert. (FDMA), Singapore, 2012, pp. 1–10. http://berrar.com/resources/Berrar_FDMA2012.pdf
7. J.H.Yan and W.R.Jiang,‘‘Research on information technology with detecting the fraudulent clicks using classification method,’’Adv.Mater.Res.,vol.859, pp.586–590, Dec. 2013, https://www.doi.org/10.4028/www.scientific.net/amr.859.586
8. K.S.Perera,B.Neupane,M.A.Faisal,Z.Aung,andW.L.Woon, ‘‘A novel ensemble learning-based approach for click fraud detection in mobile advertising,’’ in Mining Intelligence and Knowledge Exploration (Lecture Notes in Computer Science: Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 8284. Berlin, Germany: Springer, 2013, pp. 370–382, https://doi.org/10.1007/978-3-319-03844-5_38.
9. C.Phua,E.-Y.Cheu, G.-E. Yap, K. Sim, and M.-N. Nguyen, ‘‘Feature engineering for click fraud detection,’’ in Proc. Work. Fraud Detect. Mob. Advert., 2012, pp. 1–10. http://palanteer.sis.smu.edu.sg/fdma2012/doc/FirstWinner www.gnithyd.ac.in
10. E.-A.Minastireanu and G.Mesnita, ‘‘Light GBM machine learning algorithm to online click fraud detection,’’ J. Inf. Assurance Cybersecur., vol. 2019, pp. 1–12, Apr. 2019, doi: 10.5171/2019.263928.