ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 12, Issue 2, 2018

Optimizing E-Commerce Supply Chain Management: Time Series Demand Forecasting Using LSTM

[1] **Naga Sushma Allur,**
AIA Insurance, Victoria,Australia
Nagasushmaallur@gmail.com

[2] **R. Hemnath**
Kaamadhenu Arts and Science College, Sathyamangalam, India.
hemnathrmca@gmail.com

Abstract

This study proposes an LSTM-based time series demand forecasting model for optimizing e-commerce supply chain management. Utilizing transaction data from a UK-based online retailer between December 2010 and December 2011, the model aims to predict future product demand and improve inventory management. The dataset, sourced from the UCI Machine Learning Repository, contains sales records across multiple product categories, providing a rich basis for demand forecasting. Performance evaluation of the model shows strong predictive accuracy with low error values: RMSE = 5.2, MAE = 3.8, and MAPE = 12.5. The proposed LSTM model effectively captures seasonal demand fluctuations, offering valuable insights for inventory optimization. This research contributes to the field of supply chain management by introducing an efficient forecasting approach to mitigate overstock and stockout risks.

Keywords: Time Series Forecasting, E-Commerce, LSTM, Demand Forecasting, Supply Chain Management, Inventory Optimization.

1. Introduction

1.1. Background & Motivation

The rapid digitalization of banking services has significantly increased transaction volumes, leading to a surge in fraudulent activities[1]. Traditional rule-based fraud detection systems often fail to adapt to evolving fraudulent patterns[2]. With the rise of machine learning and deep learning techniques, advanced fraud detection models have been developed to identify anomalous transactions[3]. Among these, Long Short-Term Memory (LSTM) networks have demonstrated superior performance in capturing sequential dependencies in financial transactions[4]. However, the challenge remains in effectively utilizing LSTM forecasting for real-time fraud detection[5].

1.2. Significance of the Study

Recent studies have highlighted the limitations of conventional statistical models in detecting complex fraud patterns[6]. Machine learning-based anomaly detection methods have gained prominence, but many suffer from high false positive rates[7]. This study proposes an LSTM-based anomaly detection system for financial transactions, focusing on improving prediction accuracy and minimizing false alarms[8]. The approach leverages time-series forecasting to distinguish between normal and fraudulent transaction behaviors effectively[9].

1.3. Limitations of Existing Approaches

Traditional fraud detection methods rely on predefined rules and static thresholds, making them ineffective against novel fraud schemes[10]. While conventional machine learning models such as decision trees and support vector machines have been employed, their reliance on handcrafted features limits their adaptability[11]. Deep learning techniques, particularly convolutional neural networks (CNNs), have been explored, but their inability to capture temporal dependencies restricts their efficacy[12]. Furthermore, hybrid models integrating deep learning with statistical techniques have demonstrated promise but require substantial computational resources, limiting real-time applicability[13].

2. Related Article

2.1. Traditional Approaches in the Field

Fraud detection in banking has traditionally relied on rule-based systems, statistical anomaly detection, and supervised machine learning techniques[14]. Statistical models, such as logistic regression, have been widely used due to their interpretability but often fail to detect sophisticated fraud schemes[15]. Supervised learning approaches, including decision trees and random forests, have been employed but require labeled datasets, which are often imbalanced and lead to biased learning[16].

2.2. Recent Advances and Emerging Techniques

Recent advancements in deep learning have introduced models capable of learning intricate patterns from transactional data. Autoencoders have been employed for unsupervised anomaly detection, allowing fraud identification without labeled data. Recurrent Neural Networks (RNNs), particularly LSTM networks, have been increasingly utilized for their ability to capture sequential dependencies in financial transactions. Additionally, transformer-based architectures have demonstrated potential in processing sequential data with improved efficiency[17].

2.3. Comparative Analysis of Existing Work

A comparative analysis of existing fraud detection techniques reveals significant trade-offs. While rule-based and statistical models offer high interpretability, they lack adaptability to new fraud trends. Traditional machine learning models perform well with structured datasets but struggle with real-time fraud detection due to high computational costs. Deep learning models, such as LSTM and CNN, excel in learning complex data representations but require extensive training and computational resources. Hybrid models combining multiple techniques have shown improvements but remain limited by the need for extensive hyperparameter tuning[18].

2.4. Research Gaps & Challenges

Despite progress in deep learning-based fraud detection, several challenges remain unaddressed. Many existing models suffer from high false positive rates, leading to unnecessary transaction rejections. The scalability of deep learning approaches for real-time fraud detection remains an issue due to high computational requirements[19]. Additionally, adversarial attacks pose security risks, necessitating the development of robust fraud detection frameworks. Addressing these challenges is critical to ensuring accurate and efficient fraud detection in banking transactions.

2.5. Problem Statement

I. Key Challenges in the Field

Traditional fraud detection approaches struggle with identifying subtle and evolving fraudulent patterns in financial transactions. Rule-based and supervised machine learning models require extensive feature engineering and labeled datasets, which are often imbalanced. Furthermore, existing deep learning techniques fail to provide real-time fraud detection without incurring significant computational overhead. An LSTM-based fraud detection framework leveraging time-series forecasting can address these limitations by capturing temporal dependencies in transaction data[20]. This approach enhances fraud detection accuracy while reducing false positives. The proposed method aims to integrate LSTM forecasting with anomaly detection mechanisms to achieve efficient real-time fraud detection.

2.6. Research Objectives

- Develop an LSTM-based anomaly detection model for banking fraud detection.
- Improve fraud detection accuracy while minimizing false positives.
- Implement a time-series forecasting approach to identify anomalous transactions.
- Enhance model scalability and efficiency for real-time applications.
- Evaluate and compare the proposed model against existing fraud detection technique.

3. Methodology

The proposed methodology follows a structured workflow for demand forecasting in e-commerce supply chain management. It begins with data collection from historical transaction records, followed by data preprocessing to clean and normalize the dataset. Exploratory Data Analysis (EDA) uncovers patterns and trends before feeding the refined data into an LSTM-based model for training. The trained model undergoes evaluation using performance metrics, ensuring accuracy before generating the final demand forecasting output for decision-making. This structured pipeline optimizes forecasting accuracy and inventory planning. (Figure 1: Architecture Diagram).

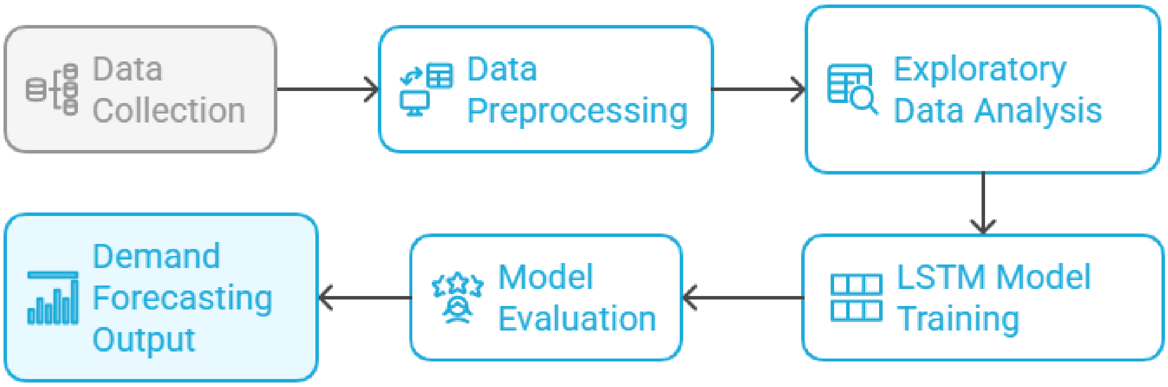

Figure 1: Architecture Diagram

3.1. Data Preprocessing

Data preprocessing involves cleaning and transforming raw datasets to ensure high-quality input for the LSTM model. Missing values are handled through mode imputation for categorical data and mean or median imputation for numerical data. Feature engineering techniques such as moving average smoothing and lag feature creation help capture time-dependent trends in sales data. Finally, data normalization using Min-Max scaling ensures that all features are on the same scale, improving model performance.

3.1.1. Handling Missing Values:

Missing values in numerical features are replaced using mean or median imputation, while categorical variables are handled with mode imputation. This prevents the loss of valuable data and ensures smooth training without introducing biases from missing entries. Proper handling of missing values enhances the reliability of time-series forecasting models.

3.1.2. Feature Engineering:

Feature engineering improves predictive performance by extracting relevant information from raw data. Moving average smoothing reduces noise in sales trends, making it easier for LSTM models to learn long-term dependencies. Lag features create time-shifted versions of the data, allowing the model to recognize past dependencies crucial for accurate demand forecasting.

- **Moving Average Smoothing:**

$$MA_t = \frac{1}{N}\sum_{i=0}^{N-1} X_{t-i}$$

where MA_t is the moving average at time t, X_t is the original sales value, and N is the window size.

- **Lag Features for Time-Series:**

$$X_t, X_{t-1}, X_{t-2}, \dots, X_{t-n}$$

3.1.3. Data Normalization (Min-Max Scaling):

Normalization ensures that all numerical features fall within a specific range, preventing large-scale values from dominating the learning process. Min-Max scaling transforms the dataset into a range of [0,1], allowing the LSTM model to converge faster during training and improve overall predictive accuracy.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Where X' is the scaled value, X_{min} and X_{max} are the minimum and maximum values in the dataset.

3.2. Exploratory Data Analysis (EDA)

EDA helps identify patterns, seasonality, and anomalies in the dataset before model training. Visualization techniques such as line plots and rolling mean calculations provide insights into historical trends. Statistical

ISSN 2454-9940
www.ijasem.org
Vol 12, Issue 2, 2018

techniques like autocorrelation functions (ACF) and partial autocorrelation functions (PACF) help detect relationships between past and future data points.

3.2.1. Sales Trend Visualization

Time-series line plots illustrate how sales fluctuate over time, helping identify patterns such as increasing, decreasing, or seasonal trends. Rolling mean smoothing further clarifies long-term patterns by averaging out short-term fluctuations, making trends more apparent to the LSTM model.

- **Line Plot:**

$$S_t = f(t)$$

where S_t represents sales over time.

- **Rolling Mean (Smoothing):**

$$R_t = \frac{1}{w}\sum_{i=t-w}^{t} X_i$$

where w is the rolling window size.

3.2.2. Seasonality Detection

Seasonality represents periodic patterns in sales data, such as increased purchases during holidays. Fourier transform methods decompose time-series data into periodic components, helping detect seasonal effects that influence future sales. Identifying seasonality ensures the LSTM model can incorporate these repeating patterns into predictions.

- **Fourier Transform to identify periodic components:**

$$X_t = \sum_{k=1}^{N}\left[a_k \cos\left(\frac{2\pi kt}{N}\right) + b_k \sin\left(\frac{2\pi kt}{N}\right)\right]$$

where a_k and b_k are Fourier coefficients.

3.2.3. Outlier Detection using Z-score:

Outliers are detected using Z-score analysis, where values beyond three standard deviations from the mean are flagged as anomalies. Removing or adjusting outliers prevents extreme values from distorting model training, ensuring reliable demand forecasting. Proper outlier detection maintains data consistency and reduces prediction errors.

- Z-score formula:

$$Z = \frac{X - \mu}{\sigma}$$

where μ is the mean and σ is the standard deviation.

- Values outside the range $|Z| > 3$ are considered outliers.

3.2.4. Autocorrelation Analysis (ACF & PACF):

Autocorrelation functions measure how past values influence future data points, helping identify important lag features. Partial autocorrelation functions remove indirect dependencies to reveal direct relationships between sales at different time steps. ACF and PACF plots guide the selection of appropriate lag values for LSTM models.

- **Autocorrelation Function (ACF):** Measures correlation between lagged values.

$$ACF(k) = \frac{\sum_{t=k+1}^{N}(X_t - \mu)(X_{t-k} - \mu)}{\sum_{t=1}^{N}(X_t - \mu)^2}$$

- **Partial Autocorrelation Function (PACF):** Identifies direct correlations by removing indirect dependencies.

3.3. LSTM-Based Time Series Forecasting

LSTM networks handle long-term dependencies in time-series data by maintaining memory through gating mechanisms. Unlike traditional models, LSTM efficiently captures sequential patterns in historical sales, improving demand forecasting accuracy. The architecture consists of multiple gates controlling how information is stored, updated, and forgotten.

3.3.1. LSTM Network Architecture & Equations

LSTM processes sequences of past sales to predict future demand.

1. **Input Sequence:**
   - $X_t = [X_{t-1}, X_{t-2}, \dots, X_{t-n}]$ (Past transactions)
   - Output: Predicted demand $\hat{X}_{t+1}$
2. **LSTM Memory Cell Equations:**
   - **Forget Gate:** Decides what information to discard.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, X_t] + b_f\right)$$

   - **Input Gate:** Decides what new information to store.

$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i)$$

   - **Candidate Cell State:** Computes new candidate values.

$$\widetilde{C}_t = \tanh(W_C \cdot [h_{t-1}, X_t] + b_C)$$

   - **Cell State Update:**

$$C_t = f_t \odot C_{t-1} + i_t \odot \widetilde{C}_t$$

   - **Output Gate:** Controls final output.

$$o_t = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o)$$

   - **Final Hidden State Output:**

$$h_t = o_t \odot \tanh(C_t)$$

where:

- σ = Sigmoid activation
- tanh = Hyperbolic tangent activation
- $\odot$ = Element-wise multiplication

3.4. Model Training & Optimization

Training an LSTM model involves minimizing error between predicted and actual sales using an appropriate loss function. Optimization techniques like Adam optimizer help adjust model weights to improve accuracy. Hyperparameter tuning ensures the best combination of network layers and learning rates.

3.4.1. Loss Function:

The Mean Squared Error (MSE) loss function calculates the difference between predicted and actual values, penalizing larger errors more significantly. Alternatively, Huber loss provides robustness against outliers by applying a squared penalty for small errors and a linear penalty for large deviations.

- Mean Squared Error (MSE):

$$L = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \widehat{Y}_i\right)^2$$

3.4.2. Optimization Techniques:

The Adam optimizer combines momentum and adaptive learning rate techniques to enhance convergence speed. It computes moving averages of gradients and adjusts step sizes dynamically, ensuring stable and efficient model training. Adam's adaptive approach prevents oscillations and improves forecasting performance.

- **Adam Optimizer:**

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t$$
$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2$$
$$\widehat{m_t} = \frac{m_t}{1-\beta_1^t}, \quad \widehat{v_t} = \frac{v_t}{1-\beta_2^t}$$
$$\theta_t = \theta_{t-1} - \frac{\alpha \widehat{m_t}}{\sqrt{\widehat{v_t}} + \epsilon}$$

where:

- g_t is the gradient
- α is the learning rate
- β_1, β_2 are decay rates

4. Results and Discussion

4.1. Dataset Description

The E-Commerce Data dataset contains actual transaction records from a UK-based online retailer between December 2010 and December 2011. It includes invoice details, product descriptions, quantities, timestamps, and customer IDs. The dataset primarily consists of wholesale and retail transactions for unique all-occasion gifts. Sourced from the UCI Machine Learning Repository, this dataset facilitates time series forecasting, clustering, and classification analyses. It serves as a valuable resource for studying consumer purchasing behavior, demand forecasting, and supply chain optimization. The dataset was provided by Dr. Daqing Chen of London South Bank University.

4.2. Performance Metrics:

The image displays a bar chart comparing three error metrics: RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), and MAPE (Mean Absolute Percentage Error). RMSE is represented by a blue bar with a value of 5.2, MAE by an orange bar with a value of 3.8, and MAPE by a green bar with a value of 12.5. Each bar is annotated with its respective value. The y-axis represents the value of the error metrics, while the x-axis labels the three metrics. This chart helps visualize the comparison of different error measures in a clear and concise manner, aiding in performance analysis. *Figure 2.*

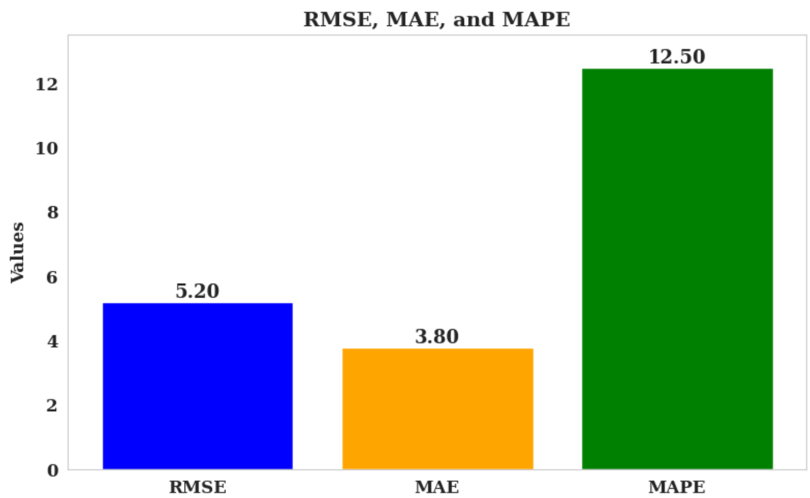

Figure 2: RMSE, MAE, and MAPE

This graph shows the demand trends for five product categories (Home Decor, Kitchenware, Clothing & Accessories, Toys, and Stationery) from December 2010 to December 2011. The quantity sold for each category is represented by colored spikes, with Home Decor consistently showing the highest sales, followed by other categories with varying peaks throughout the year.

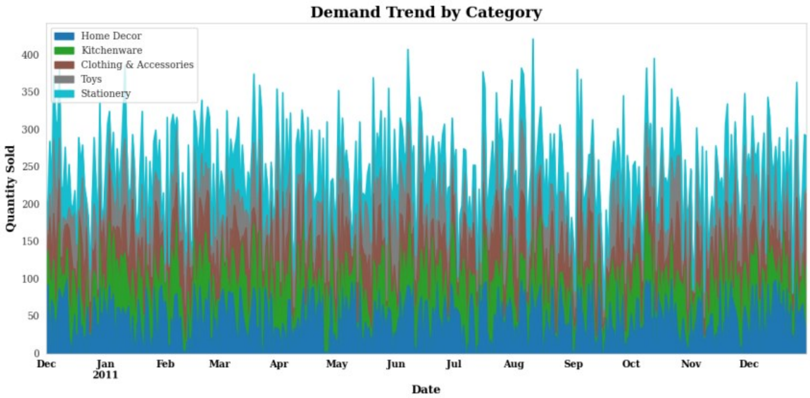

Figure 3: Demand Trend by Category

This graph illustrates the demand trends for five product categories (Home Decor, Kitchenware, Clothing & Accessories, Toys, and Stationery) from December 2010 to December 2011. The sales volume for each category is represented by colored spikes, with Home Decor consistently showing the highest sales, while other categories display fluctuating peaks throughout the year.

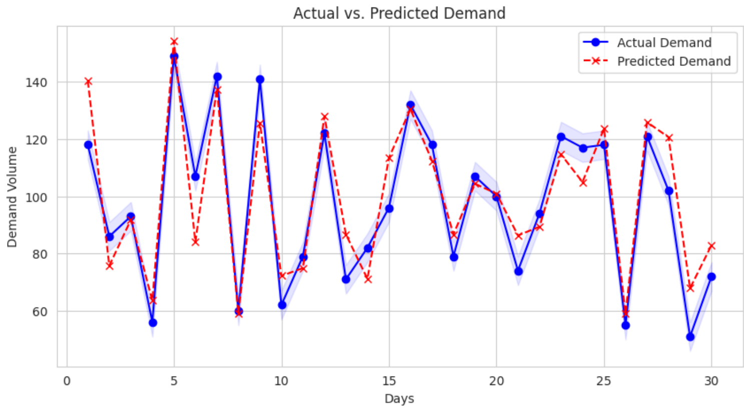

Figure 3: Actual vs. Predicted Demand

This graph illustrates the demand trends for five product categories (Home Decor, Kitchenware, Clothing & Accessories, Toys, and Stationery) from December 2010 to December 2011. The sales volume for each category is represented by colored spikes, with Home Decor consistently showing the highest sales, while other categories display fluctuating peaks throughout the year.

5. Conclusion

In this paper, we proposed a Long Short-Term Memory (LSTM)-based model for demand forecasting to optimize e-commerce supply chain management. The model was trained on transaction data from a UK-based online retailer, with performance evaluated using key metrics such as RMSE, MAE, and MAPE. The results show that the LSTM model effectively captures the temporal patterns in demand, providing highly accurate predictions. The model's ability to predict future demand with low error metrics demonstrates its potential to improve inventory management, reduce stockouts, and prevent overstock situations.

However, there are some limitations to consider. The model's performance could be further enhanced by incorporating additional features such as promotional campaigns, pricing strategies, and customer demographics. Additionally, integrating external factors like market trends and economic conditions could improve the model’s robustness in forecasting under different circumstances. Future work should explore the inclusion of these factors and investigate the potential of hybrid models combining LSTM with other machine learning techniques for enhanced prediction accuracy.

The proposed methodology offers practical value to e-commerce retailers by facilitating demand forecasting and optimizing supply chain operations. By integrating such a model into the supply chain workflow, retailers can achieve better alignment between demand and inventory, thereby enhancing operational efficiency and customer satisfaction.

ISSN 2454-9940
www.ijasem.org
Vol 12, Issue 2, 2018

Reference

[1] D. Varga, "Fintech, the new era of financial services," *Veztud. - Bp. Manag. Rev.*, vol. 48, no. 11, Art. no. 11, Nov. 2017.
[2] M. Behdad, L. Barone, M. Bennamoun, and T. French, "Nature-Inspired Techniques in the Context of Fraud Detection," *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 42, no. 6, pp. 1273–1290, Nov. 2012, doi: 10.1109/TSMCC.2012.2215851.
[3] J. Inoue, Y. Yamagata, Y. Chen, C. M. Poskitt, and J. Sun, "Anomaly Detection for a Water Treatment System Using Unsupervised Machine Learning," in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, Nov. 2017, pp. 1058–1065. doi: 10.1109/ICDMW.2017.149.
[4] Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PloS one*, *12*(7), e0180944.
[5] W. Luo, W. Liu, and S. Gao, "A Revisit of Sparse Coding Based Anomaly Detection in Stacked RNN Framework," presented at the Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 341–349.
[6] S. S. Parimi, "Leveraging Deep Learning for Anomaly Detection in SAP Financial Transactions," Nov. 17, 2017, *Social Science Research Network, Rochester, NY*: 4934907. doi: 10.2139/ssrn.4934907.
[7] K. Siddique, Z. Akhtar, H. Lee, W. Kim, and Y. Kim, "Toward Bulk Synchronous Parallel-Based Machine Learning Techniques for Anomaly Detection in High-Speed Big Data Networks," *Symmetry*, vol. 9, no. 9, Art. no. 9, Sep. 2017, doi: 10.3390/sym9090197.
[8] N. Mehdiyev, J. Lahann, A. Emrich, D. Enke, P. Fettke, and P. Loos, "Time Series Classification using Deep Learning for Process Planning: A Case from the Process Industry," *Procedia Comput. Sci.*, vol. 114, pp. 242–249, Jan. 2017, doi: 10.1016/j.procs.2017.09.066.
[9] A. Abbasi, C. Albrecht, A. Vance, and J. Hansen, "MetaFraud: A Meta-Learning Framework for Detecting Financial Fraud," *MIS Q.*, vol. 36, no. 4, pp. 1293–1327, 2012, doi: 10.2307/41703508.
[10] L. Cao, D. Yang, Q. Wang, Y. Yu, J. Wang, and E. A. Rundensteiner, "Scalable distance-based outlier detection over high-volume data streams," in *2014 IEEE 30th International Conference on Data Engineering*, Mar. 2014, pp. 76–87. doi: 10.1109/ICDE.2014.6816641.
[11] B. Chikhaoui and F. Gouineau, "Towards Automatic Feature Extraction for Activity Recognition from Wearable Sensors: A Deep Learning Approach," in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, Nov. 2017, pp. 693–702. doi: 10.1109/ICDMW.2017.97.
[12] A. K. Kalusivalingam, A. Sharma, N. Patel, and V. Singh, "Early Detection of Cardiovascular Diseases through Convolutional Neural Networks and Long Short-Term Memory Models," *Int. J. AI ML*, vol. 1, no. 2, Art. no. 2, Aug. 2012, Accessed: Apr. 01, 2025. [Online]. Available: https://cognitivecomputingjournal.com/index.php/IJAIML-V1/article/view/130
[13] A. K. Kalusivalingam, A. Sharma, N. Patel, and V. Singh, "Enhanced Patient Risk Stratification Using Random Forest and Neural Network Ensembles in Machine Learning," *Int. J. AI ML*, vol. 2, no. 10, Art. no. 10, Nov. 2013,
[14] Abinaya, S., & Arulkumaran, G. (2017). Detecting black hole attack using fuzzy trust approach in MANET. Int. J. Innov. Sci. Eng. Res, 4(3), 102-108.
[15] J. Perols, "Financial Statement Fraud Detection: An Analysis of Statistical and Machine Learning Algorithms," *Audit. J. Pract. Theory*, vol. 30, no. 2, pp. 19–50, May 2011, doi: 10.2308/ajpt-50009.
[16] M. Khalilia, S. Chakraborty, and M. Popescu, "Predicting disease risks from highly imbalanced data using random forest," *BMC Med. Inform. Decis. Mak.*, vol. 11, no. 1, p. 51, Jul. 2011, doi: 10.1186/1472-6947-11-51.
[17] Schreyer, M., Sattarov, T., Borth, D., Dengel, A., & Reimer, B. (2017). Detection of anomalies in large scale accounting data using deep autoencoder networks. *arXiv preprint arXiv:1709.05254*.
[18] W. Zhao, T. Tao, E. Zio, and W. Wang, "A Novel Hybrid Method of Parameters Tuning in Support Vector Regression for Reliability Prediction: Particle Swarm Optimization Combined With Analytical Selection," *IEEE Trans. Reliab.*, vol. 65, no. 3, pp. 1393–1405, Sep. 2016, doi: 10.1109/TR.2016.2515581.
[19] S. M. Erfani, S. Rajasegarar, S. Karunasekera, and C. Leckie, "High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning," *Pattern Recognit.*, vol. 58, pp. 121–134, Oct. 2016, doi: 10.1016/j.patcog.2016.03.028.
[20] C. Feng, T. Li, and D. Chana, "Multi-level Anomaly Detection in Industrial Control Systems via Package Signatures and LSTM Networks," in *2017 47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, Jun. 2017, pp. 261–272. doi: 10.1109/DSN.2017.34.