Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Comparative Analysis of Machine Learning Models for Accurate Retail Sales Demand Forecasting

Rasel Mahmud Jewel[1], Ahmed Ali Linkon[2], Mujiba Shaima[3], Badruddowza[4], Md Shohail Uddin Sarker[5], Rumana Shahid[6] ✉ Norun Nabi[7], Md Nasir Uddin Rana[8] and Md Ahnaf Shahriyar[9], Mehedi Hasan[10], Md Jubayar Hossain[11]

[1]Department of Business Administration, Westcliff University, Irvine, California, USA.
[2]Department of Computer Science, Westcliff University, Irvine, California
[38]Department of Computer Science, Monroe College, New Rochelle, New York, USA
[45]Department of Computer & Info Science, Gannon University, Erie, Pennsylvania, USA
[6]Department of Management Science and Quantitative Methods, Gannon University, USA
[7]Department of Information Technology, Washington University and Science and Technology, Alexandria, Virginia
[9]Department of Information Technology, Westcliff University, Irvine, California, USA
[10]Master of Science, Management- Business Analytics, St. Francis College, USA
[11]Master of Science, Management- Business Analytics, St. Francis College, USA

Corresponding Author: Rumana Shahid, **E-mail**: shahid001@gannon.edu

| ABSTRACT

This article compares sales forecasting models, LSTM and LGBM, using retail sales data from an American multinational company. The study employs a meticulous methodology, optimizing memory, performing feature engineering, and adjusting model parameters for both LSTM and LGBM. Evaluation metrics, including RMSE, MAE, WMAPE, and WRMSEE, demonstrate that LGBM consistently outperforms LSTM in capturing and predicting sales patterns. The analysis favors LGBM as the preferred model for retail sales demand forecasting, emphasizing the importance of model selection. This study contributes to practical machine learning applications in retail sales forecasting, highlighting LGBM as an effective choice.

| KEYWORDS

Sales forecasting, LSTM, LGBM, RMSE, MAE, WMAPE, Retail sales data

| ARTICLE INFORMATION

ACCEPTED: 02 February 2024 **PUBLISHED:** 26 February 2024 **DOI:** 10.32996/jcsts.2024.6.1.23

1. Introduction

In the dynamic landscape of retail, accurate sales forecasting is imperative for effective inventory management, resource allocation, and strategic decision-making. As technology continues to evolve, machine learning models have become instrumental in enhancing the precision of sales predictions. This article presents a comprehensive comparative study of two powerful machine learning models, Long Short-Term Memory (LSTM) and Light Gradient Boosting Machine (LGBM), applied to retail sales data from a multinational company based in America.

The framework of our study, depicted in Figure 1, encompasses crucial stages of data collection, preparation, and analysis, with a keen focus on evaluating the performance of LSTM and LGBM models. The retail sales dataset utilized in this research is divided into grouped time series, offering a detailed perspective on calendar details, product prices, and historical sales data. Three distinct datasets provide a rich source of information, allowing us to explore the intricate patterns within the sales data.

The methodology unfolds in three main sections: Dataset Collection and Preprocessing (Section 3.1), where we detail the characteristics of the retail sales data and the steps taken to optimize it for model training; Long-Short Term Memory (LSTM) (Section 3.2); highlighting the specific techniques employed for memory optimization and feature engineering; and Light Gradient Boosting Machine (LGBM) (Section 3.3), outlining the principles of the gradient boosting framework and its parameters, along with the features crafted for prediction.

Following these methodological steps, we delve into the Evaluation Metrics (Section 3.4), emphasizing the criteria used for model comparison, including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Weighted Mean Absolute Percentage Error (WMAPE), and Weighted Root Mean Squared Scaled Error (WRMSEE). The results, systematically presented in Table 1 and illustrated in Chart 1, provide a detailed overview of the models' comparative performance.

The heart of our findings lies in the Comparative Evaluation section (Section 4), where we meticulously analyze the results of our experiments. Notably, Light GBM emerges as the frontrunner, consistently outperforming LSTM across multiple evaluation metrics. The lower WRMSSE, along with superior values in RMSE, MAE, and WMAPE, positions Light GBM as the more effective model for demand forecasting in this specific retail context.

In conclusion, this article not only contributes valuable insights into the practical application of machine learning algorithms for demand forecasting but also underscores the importance of choosing the right model for accurate predictions in the retail sector. Our study offers a robust methodology and comparative analysis, emphasizing Light GBM as a promising choice for retail sales forecasting. The findings open avenues for future research, encouraging the exploration of additional features, parameter tuning, and the impact of external factors on forecasting accuracy in similar domains.

2. Related Work

Gupta et al. [2014] developed a versatile framework using robust machine learning algorithms to enhance customer decision-making on e-commerce platforms, with a focus on inventory-centric companies. Their model is adaptable for online marketplaces without physical inventories, employing statistical and machine learning models to forecast purchase decisions. Emphasizing customer segments over individual buyers improves accuracy, and the framework integrates various data sources, such as visit attributes, visitor characteristics, purchase history, web data, and context understanding. The research progresses to personalize adaptive pricing and purchase forecasting by incorporating web mining, big data technologies, and machine learning algorithms, providing a cohesive solution to optimize pricing decisions in the e-commerce domain.

Li et al. [2021] presented a game theoretical model addressing pricing strategies for cause-related products. The study incorporates customer prosociality levels and reference behavior, using the overall utility of a regular product as a benchmark. It investigates pricing in scenarios with and without competitive responses, emphasizing the significant impact of consumer purchasing behaviors, including prosociality levels and recognition of reference points. The research underscores the need for adaptive pricing strategies for cause-related products, considering diverse consumer behaviors.

Ensafi et al. [2022] analyzed the sales history of a retail store from a public dataset to predict furniture sales. Various forecasting models are employed, starting with classical time-series techniques like SARIMA and Triple Exponential Smoothing. Advanced methods, such as Prophet, LSTM, and CNN, are subsequently applied. Model performances are compared using metrics like RMSE and MAPE, revealing the Stacked LSTM's superiority. Prophet and CNN also exhibit noteworthy performances in predicting furniture sales.

Lindfors et al. [2021] contrasted time series analysis, relying solely on historical sales data, with machine learning models that can incorporate additional data in forecasting retail sales. The comparison involves a comprehensive literature review and an empirical investigation using Walmart sales data. The forecasting methods applied include ARIMA, Holt-Winter’s exponential smoothing, linear regression, decision trees, and artificial neural networks. The empirical results indicate that, for this specific dataset, ARIMA and Holt-Winter’s exponential smoothing emerge as the most effective models.

Singha et al. [2022] explored the utilization of advanced Machine Learning algorithms, including the Multi-layered Perceptron model (MLP), Convolution Neural Network (CNN), and Long-Short Term Memory (LSTM) Networks, among others, for Time Series Forecasting. Subsequently, a comparative analysis is conducted to determine the most effective algorithm. The assessment is carried out using the 'Store Item Demand Forecasting' dataset from Kaggle.

3. Methodology

The model's framework is outlined in Figure 1. Following the stages of data collection, data preparation, and data analysis, we implement these processes on the two models and assess their performance. A detailed discussion of each step is provided in separate sections below.

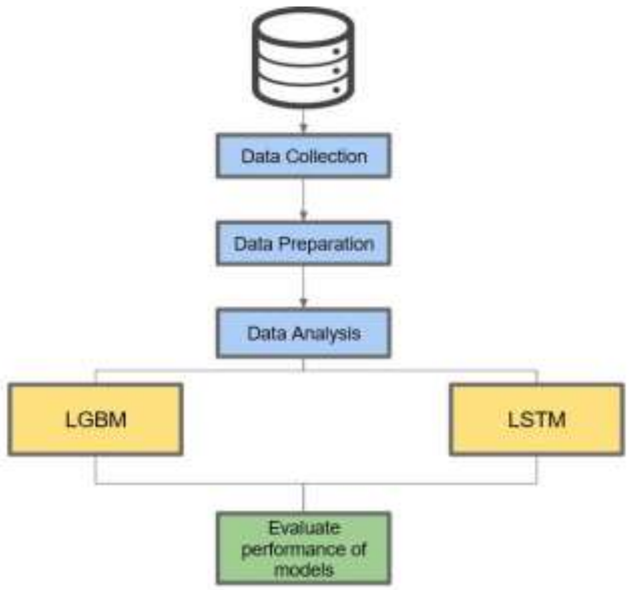

Fig 1: The entire workflow of our model

3.1 Dataset Collection and Preprocessing

In this paper, we have used retail sales data of a multinational company based in America. The data is divided into grouped time series. We use three datasets; one of the datasets consists of calendar details from 29th Jan of the year 2011 to 19th June of the year 2016. It contains details about the dates the products are sold. Our second dataset provides data on the prices of the products offered in each store along with their date. Our third dataset has details of historical data for sales of each product and store. The dataset involves the sales of 3075 products, classified into three product categories (Hobbies, Foods, and Household), and seven product departments distribute their items through ten locations situated across three states. Sales of products in individual stores, constituting the lower tier of the hierarchy, can be interconnected across various product categories.

The calendar dataset encompasses various details for each date, including the date itself, the corresponding week ID, the day of the week, the weekday number starting from Saturday, the month, and the year. Additionally, it includes information about events, specifying both the event name and type, along with a binary variable indicating whether store purchases took place on the examined date. The second dataset, focusing on product prices across stores and dates, includes features like item ID and the product's price relative to the week. Notably, if this information is unavailable, it signifies that the product remained unsold during that particular week. The sales dataset incorporates features such as product ID, department ID, store ID at the location of the product sale, the state in which the store is located, and the total number of products sold on a specific day.

3.2 Long-Short Term Memory

To optimize dataset memory usage, our primary step involves downcasting memory and leveraging memory compression techniques to save both space and time during the training process, given the substantial size of our dataset. Furthermore, we exclude entries from our initial data where sales are zero, or items are out of stock, contributing to memory reduction. Notably, SNAP days exhibit a significant impact on food sales, attributed to the Supplemental Nutrition Assistance Program (SNAP) benefit provided by the government. Consequently, these days play a crucial role in sales prediction.

The inclusion of national, religious, and sports events in our analysis reveals heightened sales on the day preceding such events. Therefore, we incorporate this feature into our primary dataset, extracting information from our calendar dataset regarding SNAP days and the day before events labeled as "eventname_1." These details are encoded and seamlessly merged with our dataset.

As part of our methodology, we eliminate specific store identifying features, such as "id." Subsequently, we transpose the dataset to ensure the variation of product sales over time, and the data is normalized for further analysis and model training.

3.3 Light Gradient Boosting Machine

The Light Gradient Boosting Machine (LGBM) is a gradient boosting framework that relies on the Decision Tree Algorithm, primarily employed for classification and ranking purposes. Operating on the principle of ensemble learning, LGBM combines multiple weak learners to form a robust single learner. This algorithm is particularly favored for handling extensive datasets, given its rapid processing capabilities. Notably, LGBM outperforms XGBoost in terms of both time efficiency and accuracy.

To minimize memory consumption in our dataset, we opt for downcasting, employing memory compression techniques to efficiently utilize space and reduce training time, given the dataset's substantial size. Additionally, Figure 2 analysis reveals that initial data points with zero sales or out-of-stock items are excluded from our analysis. For validation prediction, the target data is set 28 days from the training dataset, and the evaluation prediction target is established after 56 days.

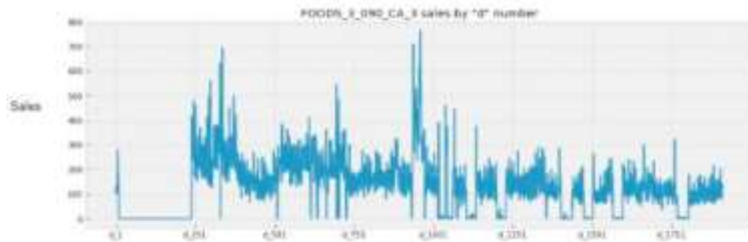

Figure 2: Sales vs Days

In a Light Gradient Boosting Machine (LGBM), Decision Trees (DT) play a crucial role by employing leaf splitting methods to divide nodes, determining the nodes with the highest information gain within each time segment. Due to LGBM's inclination to over-fit, controlling the tree depth becomes essential. The algorithm utilizes a Decision Tree technique wherein features are categorized into bins and stored in a histogram, facilitating subsequent splits based on these bins. This approach effectively reduces both storage and computation costs. Optimization of the LightGBM algorithm is paramount, and this involves adjusting various parameters. Key parameters include the learning rate, minimum amount of data in a leaf, feature fractions, number of leaves per tree, and bagging fractions, among others. While LightGBM encompasses more than fifty parameters in total, the ones are considered the most crucial and commonly used.

In crafting features for prediction, a straightforward lag feature is implemented, considering a time frame of 28 days prior to the target. Specifically, lag values are computed for 7, 14, and 28 days. Furthermore, rolling mean lags with window sizes of (7, 28) are incorporated, along with individual lags of 7, 14, and 28 days. To enhance the predictive model, data from both the calendar and price datasets are shifted by 28 days and seamlessly merged with the sales evaluation training data.

3.4 Evaluation metrics

Diverse functions are employed to gauge the disparity between predicted and actual values to assess the reliability of the algorithmic results. The magnitude of error serves as a benchmark for evaluating the effectiveness of prediction models, with lower errors indicating superior performance. Several criteria are utilized for model comparison. In the realm of model evaluation, computations are carried out for Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Weighted Mean Absolute Percentage Error (WMAPE), and Weighted Root Mean Squared Scaled Error (WRMSEE). The RMSSE, a modified version of the original Mean Absolute Scaled Error (MASE) statistic, is incorporated into the evaluation metrics.

4. Result

All experiments were meticulously conducted using Python 3.0 and the Scikit-learn (sklearn) library to harness machine learning algorithms. The experimentation process took place on an HP workstation equipped with an Intel Core i5-4210U processor featuring 2 cores and 4GB RAM. The primary objective of these experiments was to highlight the superior performance of sales forecasting models based on Long Short-Term Memory (LSTM) and Light Gradient Boosting Machine (LGBM). The study involved a comprehensive comparison conducted on a dataset obtained from an American Multinational Retail company.

Within the context of time-series forecasting, the experiments focused on predicting future sales for a specific store by considering various features discussed in the preceding sections. The chosen evaluation method for conducting a thorough comparison was the Weighted Root Mean Squared Scaled Error (WRMSSE), meticulously detailed in Section 3.4. The assessment of model performance was based on WRMSSE, where a lower value is indicative of a more effective and accurate model.

The results of the experiments are systematically presented and analyzed in Table 1, providing a detailed overview of how the models performed in terms of WRMSSE. Notably, the findings underscore that Light GBM demonstrates superior performance compared to LSTM in our dataset. This outcome highlights the efficacy of Light GBM in the context of sales forecasting for the specific parameters and dataset under consideration. The comprehensive experimental approach and detailed evaluation methodology contribute valuable insights into the comparative performance of LSTM and LGBM models, specifically tailored for sales forecasting in the context of the American Multinational Retail company's dataset.

Table 1: Comparative Evaluation of different machine learning models

Model	WRMSSE	RMSE	MAE	WMAPE
Light GBM	0.7255	4233	2325.28	0.1158
LSTM	0.8966	5124	2459.05	0.1445

In the realm of demand forecasting, a key consideration is the model's ability to provide accurate predictions, minimizing error metrics such as Weighted Root Mean Squared Scaled Error (WRMSSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Weighted Mean Absolute Percentage Error (WMAPE). In the presented comparison between Light GBM and LSTM models, it is evident that Light GBM outperforms LSTM across multiple metrics.

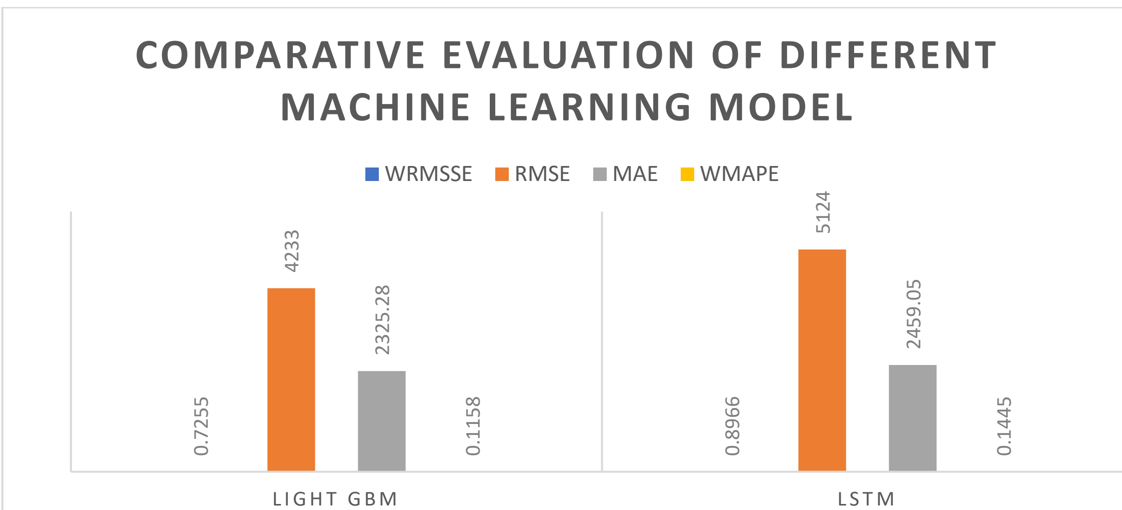

Chart 1: Evaluation metrics evaluation of different machine learning Algorithms

Light GBM achieves a significantly lower WRMSSE value of 0.7255 compared to LSTM's 0.8966, indicating a superior ability to capture and predict intricate patterns in the demand data. Additionally, when examining individual error metrics, Light GBM demonstrates better performance with lower RMSE (4233 vs. 5124), MAE (2325.28 vs. 2459.05), and WMAPE (0.1158 vs. 0.1445) values.

The lower WRMSSE, along with consistently lower values across other error metrics, positions Light GBM as the more effective model for demand forecasting in this specific context. These results suggest that Light GBM provides more accurate predictions with reduced error, making it a preferable choice for applications requiring precise demand forecasting.

5. Conclusion and Discussion

The methodology presented in this article outlines a comprehensive approach to sales forecasting, utilizing both Long Short-Term Memory (LSTM) and Light Gradient Boosting Machine (LGBM) models. The initial steps involve data collection, preparation, and analysis, followed by the application of these processes to both models for performance evaluation. The dataset used comprises retail sales data from a multinational company in America, categorized into grouped time series. Three key datasets provide information on calendar details, product prices, and historical sales data. The hierarchical structure of departments and the integration of various features, such as event details and SNAP days, enrich the dataset for accurate forecasting.

In the case of LGBM, its decision tree algorithm, coupled with efficient leaf splitting methods, contributes to its effectiveness in handling the dataset's complexity. Memory optimization techniques, such as downcasting and exclusion of zero sales or out-of-stock items, are applied to enhance the model's efficiency. Notably, the incorporation of features like national events and SNAP days further refines the model's predictive capabilities. Concurrently, the LSTM model leverages its ability to capture temporal dependencies in time-series data. Transposing the dataset, lag features, and rolling mean lags are implemented to enhance the LSTM model's predictive accuracy.

The results from the experiments, as presented in Table 1 and Chart 1, showcase that Light GBM consistently outperforms LSTM across multiple evaluation metrics. Notably, the WRMSSE values are substantially lower for Light GBM, indicating its superior ability to capture and predict sales patterns. This suggests that, in the context of demand forecasting for retail sales, Light GBM is a more effective and accurate model compared to LSTM.

In conclusion, this study demonstrates the efficacy of both LGBM and LSTM models in the domain of sales forecasting. The detailed methodology, encompassing data collection, preprocessing, and model optimization, provides a robust framework for future studies in similar domains. The comparative analysis reveals that Light GBM stands out as the preferred model for demand forecasting in this specific scenario. Its superior performance, indicated by lower WRMSSE, RMSE, MAE, and WMAPE values, showcases its capability to handle the complexities of retail sales data.

This research contributes valuable insights into the practical application of machine learning algorithms for demand forecasting in the retail sector. Future research could explore additional features, model parameter tuning, and the impact of external factors on forecasting accuracy. Overall, the findings emphasize the significance of choosing the right model for accurate demand forecasting, with Light GBM emerging as a promising choice in this context.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Abdur R R., Tanvir I., Duc M., Maliha T., Bishnu P G., Eftekhar H A., Nur N., Taslima A., Mamunur R & Mohammad S. B. (2024). Comparing Machine Learning Techniques for Detecting Chronic Kidney Disease in Early Stage. *Journal of Computer Science and Technology Studies, 6*(1), 20–32. https://doi.org/10.32996/jcsts.2024.6.1.3

[2] Bhuiyan, M. S., Chowdhury, I. K., Haider, M., Jisan, A. H., Jewel, R. M., Shahid, R., & Ferdus, M. Z. (2024). Advancements in Early Detection of Lung Cancer in Public Health: A Comprehensive Study Utilizing Machine Learning Algorithms and Predictive Models. *Journal of Computer Science and Technology Studies*, *6*(1), 113-121.

[3] Ensafi, Y., Amin, S. H., Zhang, G., & Shah, B. (2022). Time-series forecasting of seasonal items sales using machine learning–A comparative analysis. *International Journal of Information Management Data Insights*, *2*(1), 100058.

[4] Gupta, R., & Pathak, C. (2014). A machine learning framework for predicting purchase by online customers based on dynamic pricing. *Procedia Computer Science*, *36*, 599-605.

[5] Islam, M. T., Ayon, E. H., Ghosh, B. P., Shahid, R., Rahman, S., & Nguyen, T. N. (2024). Revolutionizing Retail: A Hybrid Machine Learning Approach for Precision Demand Forecasting and Strategic Decision-making in Global Commerce. *Journal of Computer Science and Technology Studies*, *6*(1), 33-39.

[6] Li, X., Xu, M., & Yang, D. (2021). Cause-related product pricing with regular product reference. *Asia-Pacific Journal of Operational Research*, *38*(04), 2150002.

[7] Lindfors, A. (2021). Demand Forecasting in Retail: A Comparison of Time Series Analysis and Machine Learning Models.

[8] Rana, M. S., Hossain, M. M., Jewel, R. M., & Islam, M. R. (2017). Evaluating Customers Satisfaction of Electronic Banking: An Empirical Study in Bangladesh. *The SIJ Transactions on Industrial, Financial & Business Management*, *5*(03), 07-12.

[9] Ray, R. K., Linkon, A. A., Bhuiyan, M. S., Jewel, R. M., Anjum, N., Ghosh, B. P., ... & Shaima, M. (2024). Transforming Breast Cancer Identification: An In-Depth Examination of Advanced Machine Learning Models Applied to Histopathological Images. *Journal of Computer Science and Technology Studies*, *6*(1), 155-161

[10] Singha D and Panse, C. (2022). Application of different Machine Learning models for Supply Chain Demand Forecasting: Comparative Analysis, *2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM)*, Gautam Buddha Nagar, India. 312-318, doi: 10.1109/ICIPTM54933.2022.9753864.

[11] Uddin, B., Al Mamun, A., Haque, A., & Jewel, R. M. (2017). FACTORS INFLUENCING SELECTION OF HIGHER LEARNING INSTITUTES: AN EMPIRICAL INVESTIGATION IN BANGLADESH. *Aktual'ni Problemy Ekonomiky= Actual Problems in Economics*, (196), 27-101.

[12] Bhuiyan, M. S., Chowdhury, I. K., Haider, M., Jisan, A. H., Jewel, R. M., Shahid, R., & Ferdus, M. Z. (2024). Advancements in Early Detection of Lung Cancer in Public Health: A Comprehensive Study Utilizing Machine Learning Algorithms and Predictive Models. *Journal of Computer Science and Technology Studies*, *6*(1), 113-121.

[13] Ahmed, A. H., Ahmad, S., Sayed, M. A., Ayon, E. H., Mia, M. T., Koli, T., & Shahid, R. (2023). Predicting the Possibility of Student Admission into Graduate Admission by Regression Model: A Statistical Analysis. *Journal of Mathematics and Statistics Studies*, *4*(4), 97-105.

[14] Ghosh, B. P., Bhuiyan, M. S., Das, D., Nguyen, T. N., Jewel, R. M., Mia, M. T., & Cao, D. M. (2024). Deep Learning in Stock Market Forecasting: Comparative Analysis of Neural Network Architectures Across NSE and NYSE. *Journal of Computer Science and Technology Studies*, *6*(1), 68-75.

[15] Bhuiyan, M. S., Chowdhury, I. K., Haider, M., Jisan, A. H., Jewel, R. M., Shahid, R., ... & Siddiqua, C. U. (2024). Advancements in Early Detection of Lung Cancer in Public Health: A Comprehensive Study Utilizing Machine Learning Algorithms and Predictive Models. *Journal of Computer Science and Technology Studies*, *6*(1), 113-121.

[16] Islam, M. T., Ayon, E. H., Ghosh, B. P., MD, S. C., Shahid, R., Rahman, S., ... & Nguyen, T. N. (2024). Revolutionizing Retail: A Hybrid Machine Learning Approach for Precision Demand Forecasting and Strategic Decision-Making in Global Commerce. *Journal of Computer Science and Technology Studies*, *6*(1), 33-39.

[17] ZHANG, Mengzhong; SHAHID, Rumana. Enlightening Bangladesh: Navigating power sector challenges through PPP excellence. **Journal of Infrastructure, Policy and Development**, [S.l.], v. 8, n. 3, feb. 2024. ISSN 2572-7931. Available at: <https://systems.enpress-publisher.com/index.php/jipd/article/view/2529>. Date accessed: 25 feb. 2024. doi:http://dx.doi.org/10.24294/jipd.v8i3.2529

[18] Jewel, R. M., Chowdhury, M. S., Al-Imran, M., Shahid, R., Ray, R. K., & Ghosh, S. K. (2024). Revolutionizing Organizational Decision-Making for Stock Market: A Machine Learning Approach with CNNs in Business Intelligence and Management. *Journal of Business and Management Studies*, *6*(1), 230-237.