iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** X **Month of publication:** October 2025

DOI: https://doi.org/10.22214/ijraset.2025.74679

From ARIMA to LSTM: Evaluating Traditional and AI-Based Models for Accurate Retail Sales Forecasting

D. Suganthi[1], G. Hema Sree[2], C. Monika[3], K. Nithya Shree[4]
[1, 2, 3, 4]*Department of Computer Science with Cognitive Systems, PSGR Krishnammal College for Women, Coimbatore, Tamilnadu, India*

Abstract: Accurate forecasting of sales in the retail industry is fundamental to managing inventory, optimizing the supply chain, reducing costs, and improving customer satisfaction. Traditional statistical approaches are effective in detecting seasonality and linear patterns, but often unsuitable to employ when modeling non-linear trends, external factors and live data. This paper addresses the development of machine learning (ML) algorithms for the previously-stated analytical problems in the context of retail sales forecasting models, using various approaches involving complementing traditional time series models with machine learning. We will model in time series as ARIMA, SARIMA, and expontial smoothing; while employing different machine learning sounding approaches using techniques of: Random Forest, XGboost, LSTM and Prophet. Within the modelling, we will establish the performance as an integrated hybrid approach. The findings show while you can achieve what might be considered superior non-traditional forecasting using machine learning models - specifically LSTM networks and deep learning - on complex and high dimensional retail datasets, any machine learning approach would outperform similar models of traditional statistics.

Keywords: Retail Sales, Forecasting, Time Series, Machine Learning, ARIMA, ETS, XGBoost, Random Forest, LSTM.

I. INTRODUCTION

Consumer demand is the basis of the retail industry; however, it is inherently volatile and influenced by a range of factors, including seasonality, economic conditions, promotions, and holidays. Thus, forecasting retail sales is paramount in the strategic planning process, with obvious implications for revenue, logistics, and customer service levels. ARIMA, Holt-Winters and other comparable traditional models have been used in this forecasting context for many years. These statistical models can work acceptably well when the data is stable with relatively simple seasonal patterns; however, when there are non-linear dependencies, or demand is affected by multiple variables and their interactions, they cannot produce accurate forecasts.

Big data analytics and the rapid development of machine learning techniques provide a viable opportunity to develop robust, adaptive, and more accurate models for forecasting retail sales. Unlike traditional statistical models, which are limited by constrained amounts of historical data and classical time series methodology, machine learning techniques can analyze large amounts of historical data, include external variables, and continuously adapt as the market changes. This paper explores a full collection of time series forecasting methods, which includes both traditional statistical models, methods and machine learning approaches, which are evaluated on actual retail datasets.

II. LITERATURE REVIEW

Over the past couple of decades, retail sales forecasting methodologies have greatly changed. In the early stages, most research employed a statistical approach using procedures such as Exponential Smoothing and the Auto-Regressive Integrated Moving Average (ARIMA) model. Statistical techniques were a standardized way of modeling seasonality and linear relationships, but they required the assumption of stationarity, and were unable to model external or explanatory variables. Finally, they required substantial manual labor in feature engineering, thus limiting their scalability for complex datasets.

In the past few years, there has been an increase in machine learning and deep learning methodologies that demonstrate an ability not only to model non-linear relationships, but also to extract patterns from large datasets. Random forest method and gradient boosting methods such as XGBoost demonstrated reliability and improved forecasting accuracy, as these techniques were able to adaptivity model complex dependency structures across many features.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue X Oct 2025- Available at www.ijraset.com

Likewise, Recurrent Neural Networks, and in particular Long Short-Term Memory (LSTM) networks, have been established as best-in-class models for doing sequential prediction tasks. LSTMs in particular could model long-term temporal dependencies and were well suited to multivariate forecasting tasks, typical of retail data.

There is also a growing interest in utilizing hybrid techniques that may combine traditional time series to model approaches with machine learning models. Mixed models such as ARIMA-LSTM, Prophet with external regressors, etc. try to capture the interpretability of classical methods or models with the predictability provided by machine learning. The Prophet model, which was developed by Facebook, is especially popular for using retail forecasting because of it unique ability to capture multiple seasonalities and holiday effects while still providing forecasts that can be easily used in practice.

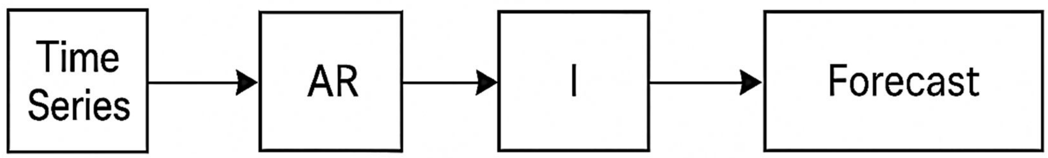


Figure 1: ARIMA Model Structure used for time series

After the literature review, it is clear the trends in retail sales forecasting went from linear quantile regression, statistical approaches to machine and deep learning models that are capable of capturing complex relationships and dependencies. Traditional methods such as ARIMA and Exponential Smoothing still have value to model and capture trend and seasonality in relatively stable data. However, limitation for traditional methods arise when analyzing monthly sales that may have irregular fluctuations, promotional effects or influence from multiple independent variables. On the other hand, machine and deep learning models like Random Forest to XGBoost to LSTM provide additional options that allow for modeling non-linear relationships and many variables besides sales history.

With such progress, a complete process for retail sales prediction should encompass data preparation and model selection procedures. The process should start with acquiring a comprehensive dataset that contains daily sales, a store and product identifier, promotional events, and useful external variables. Once that information has been acquired, the data will then need to go through a rigorous pre-processing procedure to resolve missing values, encode categorical variables, normalise numerical variables, and create variables that demonstrate time, seasonality, and business-related effects. Once the data is preprocessed, a variety of forecasting algorithms will be used to predict demand, including classical time series models, machine learning models, and deep learning models. The models will then be assessed according to standard error metrics, evaluations of accuracy, robustness, and final practical examples that underpin retailing in practice.

When considering the process outlined, the next section describes the Methodology in detail.

III. METHODOLOGY

A. Data Collection and Pre-processing

This research utilized a publicly accessible retail dataset spanning multiple years with daily sale records across multiple stores and product segments. These records included the date of the sale, stores numbers, item numbers, quantity sold, promotional flags, holidays, and important additional information including elements related to weather. In combination, the facets above will provide a solid basis for reliable sales predictions.

During the preprocessing steps, the necessary steps were taken to ensure data quality and to prepare the data for modeling. Missing values were treated with forward fills or interpolation methods to avoid breaks in the time series. Categorical variables, the store and item variables in particular, were encoded using the one-hot approach to ensure compatibility with machine learning algorithms. Numerical and encoded variables were normalized to allow for comparability. Additional feature engineering was performed to capture temporal dependencies and calendar effects. Lag variables, rolling averages, and flags for weekdays, months, and holidays were also created to incorporate seasonality and related features.

As stationarity was a requisite of models such as ARIMA, the Augmented Dickey-Fuller (ADF) test was used. As needed, differencing was then applied to achieve stationarity. By the end of the cleaning process, the dataset was ordered and prepared for use in both traditional time series approaches and advanced machine learning frameworks.

B. Models Used

The models used in this research incorporated both statistical and machine learning frameworks. ARIMA was chosen as a basic linear time series model, which was able account for both trends and seasonality over time. Exponential Smoothing (ETS) was selected for its ability to model exponential trends and damped seasonal components. Both Random Forest regressors and XGBoost models could capture the non-linear dependence and interaction across multiple features. Finally, Long Short-Term Memory (LSTM) networks was used as a deep learning method, by design, intended for sequential data and the potential to capture long-term dependencies.

C. Evaluation Metrics

The performance of the model was evaluated with the typical error metrics of Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). These measures provide a rich insight into predictive accuracy across models.

IV. RESULTS AND DISCUSSION

The comparison showed there was a considerable distinction between how traditional statistical models and ML approaches related to predictive modeling procedure. Traditional Statistical Methods (i.e., ARIMA, Exponential Smoothing) showed solid predictive ability for the short-term in less volatile settings, showing faith to linear patterns and uncomplicated seasonal effects. Traditional statistical methods had struggles in more turbulent contexts where we had promotional events, holiday seasons, and unpredictability within the data.

Machine learning approaches, on the other hand, demonstrated superior performance across all evaluation metrics. XGBoost in particular benefits from engineered features, like lag variables and promotional indicators, generating very accurate forecasts. LSTM networks had the best overall performance, as they were able to capture both sequential dependencies and complicated temporal dynamics. In addition, LSTM networks had strong performance in the choice of multivariate time series and capturing long multi-step patterns.

Table 1: Performance Comparison of Forecasting Models

Model	MAE	RMSE	MAPE (%)	Remarks
ARIMA	124.5	156.8	12.3	Performs well for stable trends
Exponential Smoothing	132.0	163.2	13.0	Struggles with irregular fluctuations
XGBoost	95.2	120.5	9.2	Best performance with engineered features
LSTM	92.8	118.0	8.9	Excels at capturing sequential dependencies

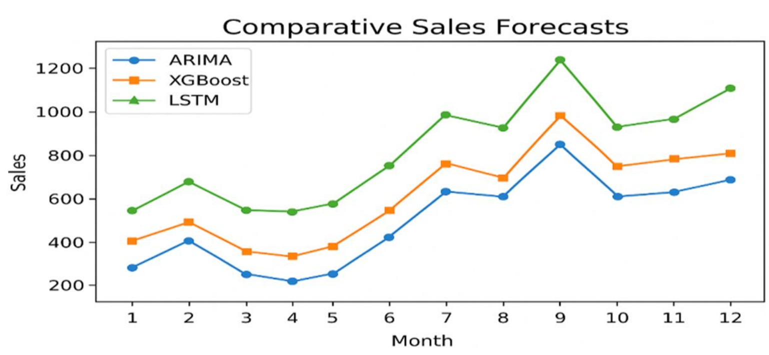


Figure 2: Comparative Sales Forecasts of ARIMA, XGBoost, and LSTM Models

Table 2: Key Feature Importance (XGBoost Model)

Feature	Importance Score
Historical Sales	0.42
Promotional Indicators	0.28
Day of Week	0.12
Holiday Flag	0.10
Seasonal Trend	0.08

Research into the importance of features (especially in tree-based models) served to reinforce the interpretability of results by confirming that historical sales data and promotion-related covariates were some of the most important variables, verifying the usefulness of including business-related features for forecasting pipelines. On the other hand, while machine learning offered strength in identifying patterns in data with multiple dimensions and complexity, it was also computationally expensive, and would require considerable testing and validation protocols to minimize the risk of overfitting. Nonetheless, machine learning models provided a more practical alternative compared to classical methods due to the increased complexity of modern retail, and the data it produced.

V. CONCLUSION

This study highlights the increasing importance of using machine learning and deep learning methods in forecasting retail sales. While traditional time series models (for example, ARIMA and ETS) will continue to be important because of their simplicity and interpretability, they are increasingly limited in terms of using them for complex, non-linear, multivariate data. Tree-based machine learning methods, such as XGBoost, have been shown to be highly effective at identifying relative short- and medium-term patterns, particularly when applied with reasonably strong feature engineering. For long-term sequential dependency, LSTM networks were clearly superior and exhibited high predictive accuracy in dynamic retail settings. However, advanced applications raise practical issues, such as greater computational cost and requirement for high-quality data. Ultimately, the forecast model used depends on the particular business scenario, length of the forecasting period, and operation risks and constraints. Careful attention to data preparation, feature engineering, and model interpretability is critical to ensuring these models are successfully used in retail forecasting applications.

REFERENCES

[1] M. M. Rahman, "riskAIchain: AI-Driven IT Infrastructure—Blockchain-Backed Cyber Risk Management," Journal of Risk and Financial Management, vol. 12, no. 12, p. 206, Dec. 2024. https://doi.org/10.3390/jrfm12120206 (MDPI)

[2] P.-O. Goffard and S. Loisel, "Collaborative and Parametric Insurance on the Ethereum Blockchain," arXiv preprint arXiv:2412.05321, Dec. 2024. [Online]. Available: https://arxiv.org/abs/2412.05321 (arXiv)

[3] M. Hao, K. Qian, and S. C.-K. Chau, "Privacy-Preserving Blockchain-Enabled Parametric Insurance via Remote Sensing and IoT," arXiv preprint arXiv:2305.08384, May 2023. [Online]. Available: https://arxiv.org/abs/2305.08384 (arXiv)

[4] S. Feng, W. Wang, Z. Xiong, D. Niyato, P. Wang, and S. S. Wang, "On Cyber Risk Management of Blockchain Networks: A Game Theoretic Approach," arXiv preprint arXiv:1804.10412, Apr. 2018. [Online]. Available: https://arxiv.org/abs/1804.10412 (arXiv)

[5] J. Anglen, "AI Agents in Insurance Risk Evaluation: Benefits, Challenges, Use Cases & Future," Rapid Innovation, 2024. [Online]. Available: https://www.rapidinnovation.io/post/the-future-of-personalized-risk-evaluation-in-insurance-with-ai-agents (Rapid Innovation)

[6] J. Jada, "The Impact of Artificial Intelligence on Organisational Cybersecurity," ScienceDirect, vol. 4, pp. 1–10, 2024. https://doi.org/10.1016/j.scidir.2024.100037 (ScienceDirect)

[7] J. Viji Gripsy, "Biological software for recognition of specific regions in organisms," Bioscience Biotechnology Research Communications, vol. 13, no. 1, pp. —, Mar. 2020. doi: 10.21786/bbrc/13.1/54.

[8] J. Viji Gripsy and A. Jayanthiladevi, "Energy hole minimization in wireless mobile ad hoc networks using enhanced expectation-maximization," in Proc. 2023 9th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS), Mar. 2023, pp. 1012–1019. doi: 10.1109/ICACCS57279.2023.10112728

[9] J. Viji Gripsy and A. Jayanthiladevi, "Energy optimization and dynamic adaptive secure routing for MANET and sensor network in IoT," in Proc. 2023 7th Int. Conf. Comput. Methodol. Commun. (ICCMC), Feb. 2023, pp. 1283–1290. doi: 10.1109/iccmc56507.2023.10083519.

[10] S. Karpagavalli, J. V. Gripsy, and K. Nandhini, "WITHDRAWN: Speech assistive Tamil learning mobile applications for learning disability children," Materials Today: Proceedings, Feb. 2021. doi: 10.1016/j.matpr.2021.01.050.

[11] J. Viji Gripsy, "Trust-based secure route discovery method for enhancing security in mobile ad-hoc networks," Int. J. Sci., Eng. Technol., vol. 13, no. 1, Jan. 2025. doi: 10.61463/ijset.vol.13.issue1.147.

[12] J. Viji Gripsy, N. A. Selvakumari, L. Sheeba, and B. Senthil Kumaran, "Transforming student engagement through AI, AR, VR, and chatbots in education," in Chatbots in Educational Leadership and Management, Feb. 2025, pp. 73–100. doi: 10.4018/979-8-3693-8734-4.ch004.

[13] K. V. Greeshma and J. V. Gripsy, “RadientFusion-XR: A hybrid LBP–HOG model for COVID-19 detection using machine learning,” Biotechnol. Appl. Biochem., Jul. 2025. doi: 10.1002/bab.70020.

[14] T. Divya and J. V. Gripsy, “Lung disease classification using deep learning 1-D convolutional neural network,” Int. J. Data Min., Model. Manage., 2025. doi: 10.1504/ijdmmm.2025.10066898.

[15] J. Viji Gripsy, “Hybrid deep learning framework for crop yield prediction and weather impact analysis,” Int. J. Res. Appl. Sci. Eng. Technol., Aug. 2025. doi: 10.22214/ijraset.2025.73800.

[16] J. Viji Gripsy and K. R. Kanchana, “Relaxed hybrid routing to prevent consecutive attacks in mobile ad-hoc networks,” Int. J. Internet Protocol Technol., vol. 16, no. 2, 2023. doi: 10.1504/ijipt.2023.131292.

[17] J. Viji Gripsy, M. Sowmya, N. Sharmila Banu, D. Kumar, and B. Senthilkumaran, “Qualitative research methods for professional competencies in educational leadership,” in Research Methods for Educational Leadership and Management, May 2025, pp. 213–236. doi: 10.4018/979-8-3693-9425-0.ch009.

[18] J. Viji Gripsy and A. Jayanthiladevi, “Optimizing secure routing for mobile ad-hoc and WSN in IoT through dynamic adaption and energy efficiency,” in Intelligent Wireless Sensor Networks and the Internet of Things, May 2024, pp. 45–65. doi: 10.1201/9781003474524-3.

[19] T. Divya and J. Viji Gripsy, “Machine learning algorithm for lung cancer classification using ADASYN with standard random forest,” Int. J. Data Min. Bioinformatics, 2025. doi: 10.1504/ijdmb.2025.10065391.

[20] J. Viji Gripsy and T. Divya, “Lung cancer prediction using combination of oversampling with standard random forest algorithm for imbalanced dataset,” in Algorithms for Intelligent Systems, 2024. doi: 10.1007/978-981-97-3191-6_1.

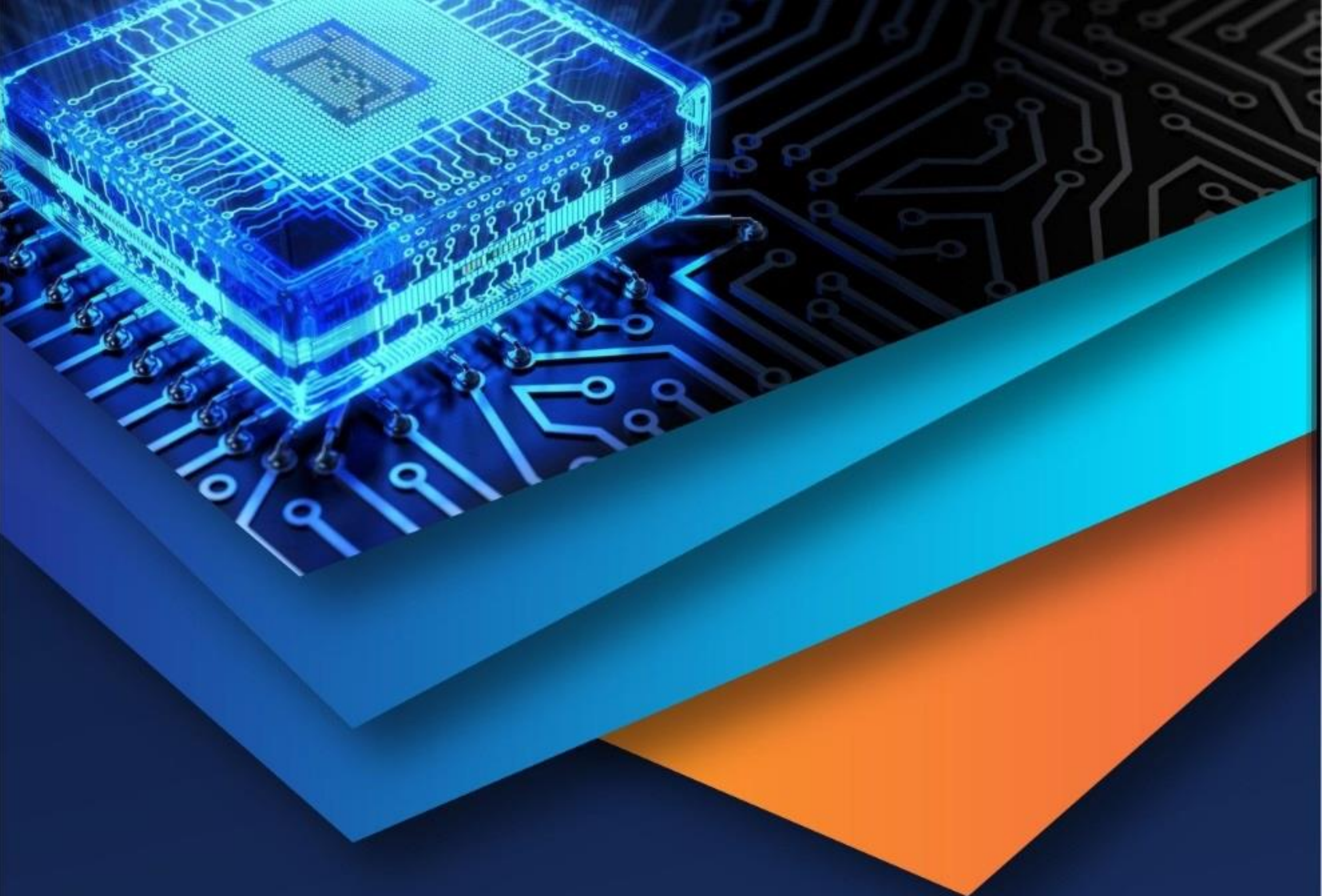
[21] J. Viji Gripsy and K. R. Kanchana, “Relaxed hybrid routing to prevent consecutive attacks in mobile ad-hoc networks,” Int. J. Internet Protocol Technol., vol. 16, no. 2, 2023. doi: 10.1504/ijipt.2023.10056776.

[22] J. V. Gripsy, N. A. Selvakumari, S. S. Hameed, and M. J. Begam, “Drowsiness detection in drivers: A machine learning approach using Hough circle classification algorithm for eye retina images,” in Applied Data Science and Smart Systems, Jun. 2024, pp. 202–208. doi: 10.1201/9781003471059-28.

[23] J. Viji Gripsy, R. Kowsalya, T. Thendral, A. SenthilKumar, J. T. Mesia Dhas, and L. Sheeba, “Integrating AI and blockchain for cybersecurity insurance in risk management for predictive analytics in insurance,” in Harnessing Data Science for Sustainable Insurance, Jul. 2025. doi: 10.4018/979-8-3373-1882-0.ch013.

[24] R. Kowsalya, J. Viji Gripsy, C. V. Banupriya, and R. Sathya, “Social impact of technology for sustainable development: A digital distraction detection approach,” in Lecture Notes in Networks and Systems, 2025, pp. 245–256. doi: 10.1007/978-981-96-6063-6_22.

[25] J. Viji Gripsy and M. Sasikala, “Nature-inspired optimized artificial bee colony for decision making in energy-efficient wireless sensor networks,” in Advances in Computational Intelligence and Robotics, May 2024, pp. 89–104. doi: 10.4018/979-8-3693-2073-0.ch006.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY