Journal of Next Generation Technology (ISSN: 2583-021X)
Vol. 5, Issue 6, Oct 2025

Distributed Big-Data Analytics with PySpark for Personalized Restaurant Recommendation Systems

Thedla Vasavi Krishna[1], N Rupa Rani[2], Bandla Siva Ranjani[3], Sai Srinivas Vellela[4], P.Muthukumar[5]
[1]Dept. of MBA, Chalapathi Institute of Technology, Guntur-522016, A.P, India.
[2]Dept. of MCA, Kallam Haranadhareddy Institute of Technology, Guntur-522019, A.P, India.
[3,4]Dept. of CSE-Data Science, Chalapathi Institute of Technology, Guntur-522016, A.P, India.
[5]Dept. of EEE, Saveetha school of Engineering, Saveetha University, Chennai, India-602105

Email id: vasavikrishna.kalari@gmail.com[1], rupanaidu89@gmail.com[2], bandlasivaranjani2016@gmail.com[3], sais1916@gmail.com[4], muthukumarvlsi@gmail.com[5]

Article Received: 3 Aug 2025 *Article Accepted: 25 Sep 2025* *Article Published: 02 Oct 2025*

Citation
Thedla Vasavi Krishna, N Rupa Rani, Bandla Siva Ranjani, Sai Srinivas Vellela, P.Muthukumar (2025), "Distributed Big-Data Analytics with PySpark for Personalized Restaurant Recommendation Systems", Journal of Next Generation Technology,5(6), pp. 23-34, October 2025.

Abstract

Technology has become an integral part of people's lives since social networks help people disclose ideas, activities, reviews, and interests. The goal of our work here is to design a recommendation system for restaurants based on machine learning that will predict the preference of a particular user. The system collects information from user reviews, ratings, and restaurant tables to be able to provide an appropriate prediction about the restaurants the user might be interested in. Another advantage of development is scaling due to the use of the PySpark system while incorporating large data, and quickly working on the recommendations. The added course uses the feedback to learn and adapt over time, thus making the dining experience for the users unique. Here, our purpose is to describe how the machine learning and PySpark allows to achieve efficient recommendation with high accuracy and relevance.

Keywords: Recommendation System, PySpark, Machine Learning, Support Vector Machine, Natural Language Processing.

I. INTRODUCTION

Technology is a very central aspect as it actually helps to shape our day today lives in the current world. Technology has affected almost every field with particular prominence in the food industries. Given the large number of restaurants nearly in every city, a consumer faces a problem of that which of them is worth to visit. A restaurant recommendation system resolves this problem by providing suggestions of restaurants which best meet the user's experience, options, and alternatives. Although this system makes decision making easier, it results in customers being prompted to make what can be perceived as more satisfying choices.

The recommendation system is aimed to present the users with the list of restaurants that will be of interest by considering the evaluation of ratings, user feedback, restaurant descriptions, and previous preferences. The system tends to use artificial intelligence to learn from past interaction and make future recommendations of restaurants from the data collected. This brings out an efficiency of the application since the users are offered restaurant choices, which they would prefer, making the options presented to them most relevant to them.

It operates by evaluating a vast amount of data gotten from reviews and ratings of restaurants, and other attributes like type of restaurant, and orientation of patrons. Using this information, the system can predict regularities in the users' actions, and thus recommend materials that the user is most likely to enjoy. In addition to enhancing the credibility of the results, this beneficially enhances users' awareness of new dining places they would have otherwise not come across.To ensure that the recommendation system is both scalable and efficient, PySpark, a powerful tool for processing large datasets, is utilized. PySpark allows the system to handle vast amounts of restaurant and review data quickly, making it an ideal choice for processing and analysing big data in real-time. With PySpark's ability to manage large-scale data processing, users can receive fast, reliable recommendations, even when dealing with large and complex datasets.

The system also uses machine learning algorithms that improve as more user data and feedback are fed to the system. These algorithms work by making predictions of what users would like and adapt with time so that as the system gets more information regarding the users' preferences, the suggested results are more appropriate. This makes the system more effective in giving accurate and relevant suggestions that make users experiences more personalized and beneficial as they proceed in their use.

Therefore, the restaurant recommendation system seeks to aid users in making the best choices when selecting the dining base on their past and present interactions. the system provides customized recommendations from scratch using large scale data processing which escalates through machine learning as well as PySpark. This means that users are always served with the best dining options which improve their experience and enable them make proper decisions for themselves.

II. LITERATURE REVIEW

[1] Amanuel Melese et al. This paper focus on recommender systems which are systems that supply specific information to user. There is a discussion about collaborative filtering, it operates with user feedback but has the cold start problem and content-based filtering, which is based on item characteristics. Thus, flows in between these approaches are stressed for improvement personalization into continuous hybrid methods. Some specialisations are in recommending restaurants, recipes and even developing entire menus while using machine learning algorithms to enhance the analysis ability of the systems. An overview of potential challenges is also described, these include issues of personalization and incorporation of nutritional preferences that make up the foundation on which the ensuing hybrid system is developed. The survey identifies some of the progress and lacunae in developing recommendation systems for foods.

[2] Achmad Arif Munaji et al. This paper discusses the importance of recommendation systems that help the users in decision-making in numerous terrains such as, e-commerce and social networking. Collaborative filtering is identified to be an important algorithm in the issue that uses user similarity for customized suggestions for large data sets. The evaluation of user closeness is important because users with similar preferences are likely to rate items similarly and hence improve on relevancy of

recommendations. Different approaches to calculate similarity, Including Pearson Correlation, are described, Weighted Pearson Correlation proves to be highly accurate, but data scarcity is an issue. Strengths, opportunities, weaknesses and threats in this paper are Ability to implement improvements in similarity calculations, and Self-admitted technical debt is not accurately measured by Pearson Correlation in certain environment, Foster the need to research how to improve the recommendations. Still, the survey advocates for the significance of rating systems and the collaborative approach to building the right recommendation systems for restaurants.

[3] Ketan Mahajan et al. This paper presents the aware of various methodologies for recommendation systems, with specific regard for collaborative and content-based filtering limitations. It describes challenges like overspecialization and the new user problem, which reduces the effectiveness of the recommendations. The authors insist on the use of both filtering techniques with a view of improving the results and making users more satisfied. Unlike the restaurant ratings that the study seeks to avoid, the work uses the Yelp dataset to rank restaurants according to the users' preferences. The paper also pays special attention to the criterion of the Area Under the Curve (AUC) being a more realistic measure in understanding user preferences. Altogether, the survey of the related literature proves the significance of the use of the hybrid systems to increase the accuracy of recommendations, as well as to enhance the customer's experience.

[4] T. Choenyi et al. this paper discusses on recommendation systems, filtering methods have been described in several approaches, mainly, but not limited to, collaborative filtering (CF) and content-based filtering (CBF). Memory-based and model-based facilities are of the two types of collaborative filtering; the memory-based CF employs user preferences to find similar users or items, whereas model-based CF uses methods such as Markov decision process to enhance prediction capability. Content filtering, recommend items based on their, features and user profiles, which bowed one of the sayings of the recommendation system: showing the user more of what he or she likes. Furthermore, demographic filtering uses user demographics data to improve recommendation, although no much history which solves cold-start problem. Hybrid filtering involves the use of several techniques that make it capable of overcoming most of the drawbacks like sparsity and overspecialization, making the current systems to employ the hybrid filtering as the most useful one. Combined, the examination of these techniques demonstrates that there is an increased utilization of the hybrid filter approach that depends on the known merits of each of the filter methods in order to deliver a higher level of precision and relevancy with the recommendations being provided.

[5] Salu Khadka et al.This paper focuses Various techniques and algorithms used in restaurant recommendation systems are identified by different authors in the existing literature. For instance, Yelp Food Recommendation System employs collaborative and content-based filtering to make estimation in order to predict the user's preference based on the reviews and ratings they give. Furthermore, clustering approaches such as K-NN and weighted BPG has been incorporated with a view to enhancing recommendations' accuracy. This domain relies heavily on text mining, the major task of which is the classification of the restaurant reviews into relevant and important genres such as Food, Service, and Ambience. The process of tokenization is critical in text handling steps, the modern means which can also be used for cleaning to lemmatizing the words. Additionally, cold start problems for new users and the necessity of using more and more different data incoming from users are the major concerns in the development of recommendation systems. In general, the combination

of opinion mining and preference modelling is crucial to designing effective recommendations for restaurant selection.

[6] selva Kumar S et al. This paper discusses the Yelp dataset and surveys other papers that used the dataset focusing on evaluating the performance of businesses and customers. content analysed how review counts are related to business ratings while, Luca looked at how Yelp ratings affect the demand for restaurants. Moreover, other authors such as proposed to improve the results of collaborative filtering algorithms by incorporating various types of data into one or another sort of recommender systems' classification. Regression analysis have been widely used and more specific, Zeng and Li used logistic regression model to model restaurant success, while Alam et al combined regression and classification methods for rating forecast. Ye et al. have also used clustering techniques in order to classify restaurants based on reviews. These works also stress issues of search for optimal hyperparameters as the key to better performance of the model, which is presented as the ongoing process of maximizing the potential of machine learning methods.

[7] Elham Asan et al. The paper is focused on the review of the development of recommender systems, and more specifically, on the restaurant domain, where users' preferences are stated by different means such as queries and star ratings. It increases an awareness towards identification of user preferences from comments collected in contrast of fixed questionnaires by including an approach towards sentiment analysis. Natural language processing (NLP) is another important component stressed as being responsible for processing users' comments for implicit sentiments. Moreover, the proposed system is context sensitive and takes into account the comments and the features of the restaurants to make an appropriate recommendation; this considerate a research gap that is usually instantiated with static information. A qualitative analysis is performed using precision, recall and f-measure to show that the proposed system is superior to previous research. In total, this literature survey indicates the movement towards the dynamic, sentiment-based approaches for improved user preference extraction from the recommender systems.

[8] Taufiq Ahmed et al. The study on the restaurant recommendation system presents the literature as follows, though they are scarce concerning Dhaka city. Some examples of using collaborative filtering were also given where it is used with platforms like Yelp to recommend restaurants that the user should try. A number of location-based systems based on user preferences and Cosine Similarity which have been considered regarding right recommendations. Using data on psychographics and demographics, recommendation systems involve collaborative scores. Information obtained from applications for instance, foursquares has been used to recommend restaurants in view of previous check-ins and friends. Moreover, concepts such as k-means and k-NN shows the effectiveness of recommendation algorithms Being.

[9] DANIEL RIANDY et al. this paper discusses the Kaggle restaurant data sets have been used to provide solution on recommendation system by employing collaborative filtering method which analyse the user history and rating using Pearson correlation. Logistic Regression models along with Support Vector Machine for customer rating has been anticipated whereby liner SVM yielded the best outcome. Applying a combination of content-based and collaborative filtering systems has revealed high efficiency in the context of improving recommendation accuracy. NLP technologies in sentiment analysis of customer reviews also enhance the ratings to afford better

recommendations. It should, however, be noted that clustering techniques improve recommendation processes by affording less computational time and high precision. In Location-Based Filtering, the Cold Start Problem is solved by making recommendations to users that have just started using the application based on their location.

[10] Ankita Singh et al. this paper discusses various strategies in the field of restaurant recommendation systems. Customer feedback was used by for the preference-based recommender system and topic modelling and Ramzan et al. applied the sentiment analysis combined with the collaborative filtering using Hadoop for massive datasets. they supported the expansion of hybrid over conventional single filtering techniques, Jiang employed machine learning algorithms such as Slope One and KNN for rating forecasts. And focused on the geographical information with SVM classifier and proposed an approach based on food and service rating using The paper suggests an improvement to the current recommendation system by adding the collaborative and content-based filtering approaches.

[11] Heeyoung Kim et al. This paper discuss the AI chatbots in restaurant recommendations lead to improved user experiences, in comparison with one-directional information-providing approaches. Recently there have been improvements with the collection of personalization information, but the extent of its use in restaurant industries is still very much restricted, therefore the proposed app avails itself with a well-utilized element in the demographic information of the users. There are no necessarily developed services as the AI chatbots themselves can give the information of the reservation status and menu details. Other efficient methods in usage of the collaborative filtering methods increase the recommendation process where people can find out new spot among the restaurants. On balance, therefore, deployed AI chatbots are accurate recommendations in dining options since they enhance user satisfaction due to their real-time services. It is suggested that future studies should be done to improve the results of these technologies for finding similar terms.

[12] Shefali Goyal et al. this paper aims at surveying the literature on methodologies for constructing a recommendation system in the context of a restaurant site, based on approaches related to machine learning, specifically, user preferences as expressed in the reviews. Machine learning methods like topic modelling and regression models were taken to map relationships between users and restaurants and therefore making geographical recommendations. For collaborative filtering methods, multiple list class SVM, k-Nearest Neighbours, and Slope One methods were used, while SVM was found to be the most accurate. Some of the factors governing decisions of the users have described below, which have been extracted from the existing mobile apps including Swiggy and Zomato are type of food/ cuisine offered, location of food delivery, food prices, and the most important rating. Several self-organizing, content-filtered recommendation systems were suggested to provide conflicting solutions according to the users' profile. Python, the highly efficient libraries, and Flask were used for implementation, and the data was obtained from Kaggle for the suggestion. Statistical analysis and exploratory data techniques further improved system efficiency, and the utility provided a complete dining experience.

III. PROPOSED METHODOLOGY

The following section highlights how the Restaurant Recommendation System has been developed. The approach covers preparatory steps, creating features, implementing models and assessment, using PySpark for magnification. The system should therefore give user review, ratings and restaurant attributes for restaurant recommendation in a way that can be fine-tuned for optimum scalability through the use of machine learning.

A. Data Pre-processing

The data used in this project was collected from restaurant data obtained from publicly accessible restaurant data, which contains information about restaurant names, ratings, reviews, types of restaurants, and the restaurant prices. PySpark has been used in data preprocessing to deal with the huge data set. The data was pre-processed involving column deletion and modification such that, there was elimination of columns that contained too many missing values, unimportant or repeated columns, and even modification of numeric and qualitative columns to make them appear consistent.

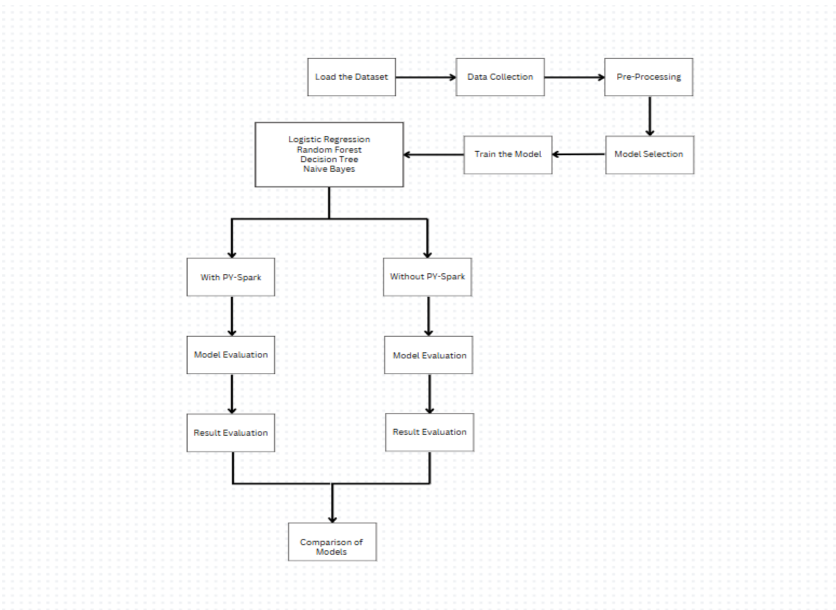

Figure 1: Architecture Diagram

1)Data Cleaning: Based on Data Frame functions in PySpark, null values and duplicated data were filtered out from this dataset. Categorical variables including ratings, reviews, and price were cleaned for aspect-level analysis where they were checked for the presence of valid values excluding nulls. To make the context clean and simple, I removed any columns that were repeated.

2)Feature Transformation: For my feature transformation, the numerical features particular, the review_count and price_for_two for the respective restaurants were scaled using PySpark'sStandardScaler, which standardizes the data, that is, brought to a comparable scale. Categorical features such as restaurant type were transformed to engineered features using StringIndexer so that they can be processed by machine learning algorithms.

3)Data Splitting: After cleaning the data and transforming the data, the feature set was divided into training and test data, with 80% for training data and 20% for the test data. This meant that there was adequate data for the training of the models in addition to conserving a part that was intact for model validation. For split of the data, PySparkrandomSplit function has been used so that it is feasible to handle larger data.

B. Model Development and Implementation

In this project, recommendation of restaurants has been done by using PySpark-based recommendation system as well as by other traditional Machine Learning models. In order to deal with big data and keep the recommendation system scalable, PySparkwas used in the project. To enhance the recommendation accuracy, many models were used as follows.

1)PySpark-Based models: Models including Logistic Regression of which the goal was to estimate the propensity of customers to like a restaurant by the number of ratings, reviews, and type of cuisine. This had the advantage of being probabilistic, and gave more preference to films with higher ratings. The goal of using Random Forest is that this model is good at dealing with large numbers of features and quite effective even with substantial levels of overfitting, which will be suitable for large amount of data. The Decision Tree Classifier was used for classification and a model was given in order to be able to understand the user preferences. Probabilistic classifier was used, specifically Naive Bayes because it is good at dealing with categorical data as it deals with restaurant type and cuisine where features chosen usually assigned with pure independence condition.

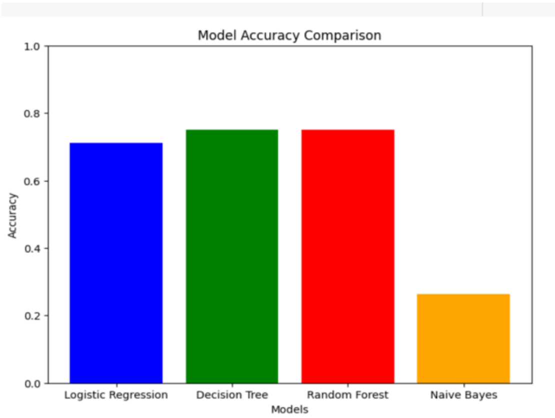

Figure 2: Model Comparsion-I

2) Traditional Models: In this study, Traditional Models were developed using technologies compatible with Python: Scikit-learn library in particular, to compare the computational relative costs and benefits and forecast models' performance. Both the feature engineering and hyperparameters tuning were performed on an XGBoost framework which is a gradient-boosted decision tree algorithm to achieve accurate recommendations of restaurants. A conventional Logistic Regression model was employed, which offered fast probabilistic predictions as a simplified baseline against which to compare with complex models. Decision Tree Classifier was also used in the traditional procedure due to its interpretability, and Naive Bayes as fast for ordinal and nominal features, quickly provide restaurant suggestions without high computational costs.

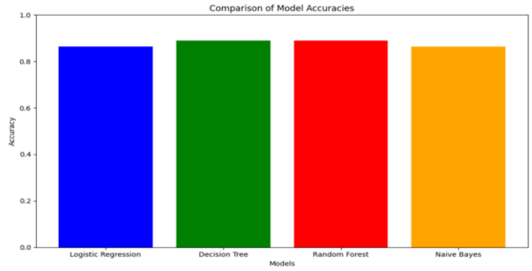

Figure *3:* Model Comparison-II

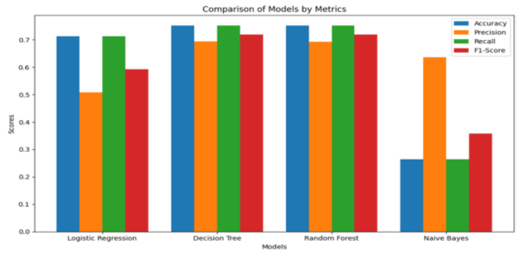

Figure 4*:* Evaluation Metrics Comparison

C. Evaluation Metrics

All models were assessed using several measures that established their separateness and relevance in making restaurant recommendations. Accuracy calculated the degree of actuality of models in identifying the preferred options of users. Other measures incorporated included precision that relied on the percentage of relevant restaurants recommended, and recall that looked into the number of relevant restaurants suggested by the algorithm as well as the F1-score that was a combination of both precision and recall. Cosine similarity was used in order to compare the similarity between restaurant feature vectors and to determine the accuracy with which the system is able to identify restaurants to recommend. These evaluation metrics ensured that PySpark-based was favourably compared with the traditional models used in identifying the most appropriate recommendation system.

IV. RESULTS AND ANALYSIS

Model	Accuracy with using Pyspark	Accuracy with not using Pyspark
Logistic Regression	0.7122	0.8637
Decision Tree	0.7515	0.8892
Naive Bayes	0.2634	0.8623
Random Forest	0.7513	0.8892

Figure 5: Model Performance Table

A. Accuracy Comparison:

The evaluation of the models in terms of accuracy showed some insights when perform a comparison among the PySpark-based development and the traditional ones. Logistic Regression provided an accuracy of 0.7122 through PySpark, however the conventional method outperformed Logistic

Regression with accuracy of 0.8637. The Decision Tree model did noticeably better with the traditional method with an accuracy of 0.8892 to 0.7515 of PySpark showing how effective traditional methods are in small data sets. The Random Forest model also indicated slightly higher points in the conventional method (89.92%) than the PySpark estimated (75.13%). Conventional Naïve Bayes which underperformed with PySpark having an accuracy of 0.2634 improved scoring an average accuracy of 0.8637 in the conventional working environment. These results shown the capability of PySpark for distributed computations though it signals that traditional approach can provide better accuracy for less demanding data sets.

B. Precision, Recall, and F1- score comparison:

The performance will also be discussed further by quantitative precision, recall, and F1-scores analysis. Logistic Regression on PySpark proved to be precise with .5073, recall level of .7122, and F1-score of .5925. A similar concept-based approach, but in its traditional form, outperformed the model with precision at 0.7460, recall at 0.8637, and the F1-score of 0.8006. The Decision Tree model was also boosted in the traditional setup, with the precision to the accuracy and the Recall of 0.8818 and the F1-score of 0.8572 as opposed to the PySpark Data set which had a Precession of 0.6935, Recall of 0.7515 and the F1-score of 0.7195. Similarly, PySpark Random Forest had an accuracy of 0.7513, precision was 0.6931, recall = 0.7513and F1-score = 0.7191 Following the above metrics, Random Forest received accuracy of 0.8892, thus has precision of 0.8818, recall of 0.8892 and F1-score of 0.8572. Although it yielded low accuracy with PySpark, the Naïve Bayes classifier had a precision of 0.6358, recall of 0.2634 and F1-score of 0.3585. The configuration gave a substantial enhancement to the ordinary setup with precision of 0.7460, recall of 0.8637, and F1-score of 0.8006. Such measures clearly showed that even though PySpark is useful for handling big data, using baseline methods offers greater accuracy and certainty for tiny data sets.

V. CONCLUSION

The work presented in this paper successfully utilizes the ML models written in PySpark and in the regular Python to make recommendations matching the user preferences, as well as the restaurant characteristics. The proposed PySpark-based models accommodated large databases effectively; thus, they were scalable. However, traditional models on average were more accurate compared to PySpark models for smaller sized datasets and most evaluation metrics used. Logistic Regression and Random Forest were found to be stable models performing significantly good and achieving high degree of accuracy with fairly satisfactory precision, recall and F1 measures. The comparison made during the work revealed the trade-off between scaling and predictive accuracy, indicating that while PySpark offers the best results for computationally heavy distributed data analysis for large, distributed data, traditional methods give the best results for relatively large but non-distributed data for big data analysis at moderate scale. This system proves the applicability of machine learning in the recommendation system for increasing the level of satisfaction of users and the overall decision-making on selection of restaurants.

References

[1] Mandava, R., Vellela, S. S., Malathi, N., Haritha, K., Gorintla, S., & Dalavai, L. (2025, May). Exploring the Role of XAI in Enhancing Predictive Model Transparency in Healthcare Risk Assessment. In *2025 International Conference on Computational Robotics, Testing and Engineering Evaluation (ICCRTEE)* (pp. 1-5). IEEE.

[2] Mandava, R., Vellela, S. S., Gorintla, S., Dalavai, L., Malathi, N., & Haritha, K. (2025, May). Evaluating the Impact of Explainable AI on User Trust in Financial Decision-Support Systems. In *2025 International Conference on Computational Robotics, Testing and Engineering Evaluation (ICCRTEE)* (pp. 1-6). IEEE.

[3] Vellela, S. S., & Balamanigandan, R. (2024). Optimized clustering routing framework to maintain the optimal energy status in the wsn mobile cloud environment. *Multimedia Tools and Applications*, *83*(3), 7919-7938.

[4] Polasi, P. K., Vellela, S. S., Narayana, J. L., Simon, J., Kapileswar, N., Prabu, R. T., & Rashed, A. N. Z. (2024). Data rates transmission, operation performance speed and figure of merit signature for various quadurature light sources under spectral and thermal effects. *Journal of Optics*, 1-11.

[5] Biyyapu, N., Veerapaneni, E. J., Surapaneni, P. P., Vellela, S. S., & Vatambeti, R. (2024). Designing a modified feature aggregation model with hybrid sampling techniques for network intrusion detection. *Cluster Computing*, *27*(5), 5913-5931.

[6] Vuyyuru, L. R., Purimetla, N. R., Reddy, K. Y., Vellela, S. S., Basha, S. K., & Vatambeti, R. (2025). Advancing automated street crime detection: a drone-based system integrating CNN models and enhanced feature selection techniques. *International Journal of Machine Learning and Cybernetics*, *16*(2), 959-981.

[7] Vullam, N. R., Geetha, G., Rao, N., Vellela, S. S., Rao, T. S., Thommandru, R., & Rao, K. N. S. (2025, February). Optimized Multitask Scheduling in Cloud Computing Using Advanced Machine Learning Techniques. In *2025 International Conference on Intelligent Control, Computing and Communications (IC3)* (pp. 410-415). IEEE.

[8] Reddy, N. V. R. S., Chitteti, C., Yesupadam, S., Desanamukula, V. S., Vellela, S. S., & Bommagani, N. J. (2023). Enhanced speckle noise reduction in breast cancer ultrasound imagery using a hybrid deep learning model. *Ingénierie des Systèmesd'Information*, *28*(4), 1063-1071.

[9] Vellela, S. S., & Balamanigandan, R. (2023). An intelligent sleep-awake energy management system for wireless sensor network. *Peer-to-Peer Networking and Applications*, *16*(6), 2714-2731.

[10] Vellela, S. S., & Balamanigandan, R. (2024). An efficient attack detection and prevention approach for secure WSN mobile cloud environment. *Soft Computing*, *28*(19), 11279-11293.

[11] Vellela, S. S., Roja, D., Purimetla, N. R., Thalakola, S., Vuyyuru, L. R., & Vatambeti, R. (2025). Cyber threat detection in industry 4.0: Leveraging GloVe and self-attention mechanisms in BiLSTM for enhanced intrusion detection. *Computers and Electrical Engineering*, *124*, 110368.

[12] Vellela, S. S., Malathi, N., Gorintla, S., Priya, K. K., Rao, T. S., Thommandru, R., & Rao, K. N. S. (2025, March). A Novel Secure and Scalable Framework for a Cloud-Based Electronic

Health Record Management System. In *2025 3rd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT)* (pp. 131-135). IEEE.

[13] Praveen, S. P., Nakka, R., Chokka, A., Thatha, V. N., Vellela, S. S., & Sirisha, U. (2023). A novel classification approach for grape leaf disease detection based on different attention deep learning techniques. *International Journal of Advanced Computer Science and Applications (IJACSA)*, *14*(6), 2023.

[14] Vellela, S. S., Rao, M. V., Mantena, S. V., Reddy, M. J., Vatambeti, R., & Rahman, S. Z. (2024). Evaluation of Tennis Teaching Effect Using Optimized DL Model with Cloud Computing System. *International Journal of Modern Education and Computer Science (IJMECS)*, *16*(2), 16-28.

[15] Vellela, S. S., Vullam, N. R., Gorintla, S., Rao, T. S., & Harinadh, T. (2025, July). Exploring the Anti-Inflammatory Potential of Green-Synthesized Pyrazolines. In 2025 6th International Conference on Data Intelligence and Cognitive Informatics (ICDICI) (pp. 814-819). IEEE.

[16] Vellela, S. S., Manne, V. K., Trividha, G., Chaithanya, L., & Shaik, A. (2025). Intelligent Transportation Systems AI and IoT for Sustainable Urban Traffic Management. Available at SSRN 5250812.

[17] Vellela, S. S., Singu, K., Kakarla, L. S., Tadikonda, P., & Sattenapalli, S. N. R. (2025). NLP-Driven Summarization: Efficient Extraction of Key Information from Legal and Financial Documents. Available at SSRN 5250908.

[18] Vuyyuru, L. R., Purimetla, N. R., Reddy, K. Y., Vellela, S. S., Basha, S. K., & Vatambeti, R. (2025). Advancing automated street crime detection: a drone-based system integrating CNN models and enhanced feature selection techniques. International Journal of Machine Learning and Cybernetics, 16(2), 959-981.

[19] Dalavai, L., Purimetla, N. M., Vellela, S. S., SyamsundaraRao, T., Vuyyuru, L. R., & Kumar, K. K. (2024, December). Improving Deep Learning-Based Image Classification Through Noise Reduction and Feature Enhancement. In 2024 International Conference on Artificial Intelligence and Quantum Computation-Based Sensor Application (ICAIQSA) (pp. 1-7). IEEE.

[20] Haritha, K., Vellela, S. S., Vuyyuru, L. R., Malathi, N., & Dalavai, L. (2024, December). Distributed Blockchain-SDN Models for Robust Data Security in Cloud-Integrated IoT Networks. In 2024 3rd International Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 623-629). IEEE.

[21] Madhuri, A., Jyothi, V. E., Praveen, S. P., Sindhura, S., Srinivas, V. S., & Kumar, D. L. S. (2024). A new multi-level semi-supervised learning approach for network intrusion detection system based on the 'goa'. Journal of Interconnection Networks, 24(supp01), 2143047.

[22] Biyyapu, N., Veerapaneni, E. J., Surapaneni, P. P., Vellela, S. S., & Vatambeti, R. (2024). Designing a modified feature aggregation model with hybrid sampling techniques for network intrusion detection. Cluster Computing, 27(5), 5913-5931.

[23] Praveen, S. P., Sarala, P., Kumar, T. K. M., Manuri, S. G., Srinivas, V. S., & Swapna, D. (2022, November). An adaptive load balancing technique for multi SDN controllers. In 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS) (pp. 1403-1409). IEEE.

[24] Vellela, S. S., & Balamanigandan, R. (2022, December). Design of Hybrid Authentication Protocol for High Secure Applications in Cloud Environments. In 2022 International

Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 408-414). IEEE.

[25] Vellela, S. S., Balamanigandan, R., & Praveen, S. P. (2022). Strategic survey on security and privacy methods of cloud computing environment. Journal of Next Generation Technology, 2(1).

[26] SrinivasVellela, S., Praveen, S. P., Roja, D., Krishna, A. R., Purimetla, N., Rao, T., & Kumar, K. K. (2024, April). Fusion-Infused Hypnocare: Unveiling Real-Time Instantaneous Heart Rates for Remote Diagnosis of Sleep Apnea. In 2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS) (Vol. 1, pp. 1-5). IEEE.

[27] Vellela, S. S., & Krishna, A. M. (2020). On Board Artificial Intelligence With Service Aggregation for Edge Computing in Industrial Applications. Journal of Critical Reviews, 7(07).

[28] Praveen, S. P., Vellela, S. S., & Balamanigandan, R. (2024). SmartIris ML: harnessing machine learning for enhanced multi-biometric authentication. Journal of Next Generation Technology (ISSN: 2583-021X), 4(1).

[29] Vellela, S. S., Chandra, S. S., Thommandru, R., Mastan Basha, S., & Sri Ram, D. (2023). Novel Approach to Mitigate Starvation in Wireless Mesh Networks. Available at SSRN 5262254.

[30] Vellela, S. S., Singu, K., Kakarla, L. S., Tadikonda, P., & Sattenapalli, S. N. R. (2025). NLP-Driven Summarization: Efficient Extraction of Key Information from Legal and Financial Documents. Available at SSRN 5250908.