ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

PREDICTIVE ANALYTICS FOR OPTIMIZING URBAN AMENITIES IN SMART CITIES: A MACHINE LEARNING APPROACH

Mahesh Kumar Thota*
Research Scholar Department of CSE
UVCE, Bangalore University Bengaluru, India
thotamahi1262@gmail.com

Prathibhavani P M
Assistant Professor Department of CSE
UVCE, Bangalore University Bengaluru, India
prathibhavani.pm@uvce.ac.in

Venugopal K R
IEEE Fellow Retired Professor Department of CSE
UVCE, Bangalore University
venugopalkr@gmail.com

ABSTRACT:

The rapid urbanization and growing population in cities have led to an increased demand for efficient and sustainable management of urban infrastructure and amenities. Smart cities aim to leverage advanced technologies to enhance the quality of life for citizens by providing intelligent and optimized urban services. One crucial aspect of smart city development is the prediction and planning of amenities to meet the evolving needs of residents. This paper presents a machine learning-based approach for predicting smart city amenities. The proposed system utilizes historical data from various sources, such as citizen feedback, sensor networks, and government records, to train predictive models. By analyzing these datasets, the system can identify patterns and correlations to forecast the demand and utilization of different amenities in the future. The machine learning models employed in this study include regression, classification, and clustering algorithms. Regression models are used to predict the future usage patterns and resource requirements of existing amenities. Classification models enable the identification of potential areas for the development of new amenities based on demographic, socio- economic, and geographical factors. Clustering algorithms aid in identifying groups of similar cities or neighborhoods that exhibit similar amenity needs, facilitating targeted resource allocation and planning. To evaluate the effectiveness of the proposed approach, extensive experiments and case studies are conducted using real-world data from several cities. The results demonstrate the ability of the models to accurately forecast the demand for amenities, enabling city planners and administrators to make informed decisions regarding infrastructure investments, resource allocation, and policy-making. The contributions of this research extend beyond the realm of academia, as the developed system provides a practical tool for city planners, urban administrators, and policymakers to optimize the provision of amenities in smart cities. By accurately predicting future requirements, cities can proactively address the needs of their residents, enhance quality of life, and promote sustainable urban development.

Keywords: Smart cities, amenities prediction, machine learning, regression, classification, clustering, urban infrastructure, resource allocation, sustainable development.

I. INTRODUCTION:

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

The recent global development in urban-population and the evolvement of BDA (Big Data Analytics), there exists growing fascination in studies on the way in which cities could actively rebuild themselves into sustainable and smart city. This cities intends to afford value driven services to enhance the QoL (Quality of Life) of citizens [1]. However, in recent years, the city planners are facing insufficient contextual data on the smart city dimensions that are needed to accomplish sustainable society. Therefore, the stakeholders are in a need to take strategic decisions on the ways to execute initiatives for smart city

development. Moreover, it has also been needed to state the dimensions of smart city to be selected for making smart cities to attain sustainability. Hence, this article incorporates CBR (Case Based Reasoning) as an AI (Artificial Intelligence) method for developing RS (Recommender Systems) to promote the plans for developing smart cities. Outcomes from exploratory and descriptive analysis recommend that the introduced system has the applicability in assisting the adoption of smart cities. Analysis has also been carried out in this context.

Accordingly, the paper [2] analysed 55 papers. It has been found that twelve positive and four negative outcomes have been frequently explored. Among the twelve positive outcomes, six were entirely hypothetical namely improvising the involvement of citizens, social development facilitation, fostering innovation, environment protection, and sustainable development facilitation and enhancing the social capital. Only little focus has been provided to negative outcomes where 2 out of 4 were completely hypothetical such as security and privacy challenges, lessening democracy and speech freedom. Future research has been required in finding the evidence to pure hypothetical impacts as well as comparing the development of smart cities in emerging and advanced economies. IoT (Internet of Things) has been the frequently utilized techniques to create smart city [3]. The various layers of IoT in this context are shown in figure.1.

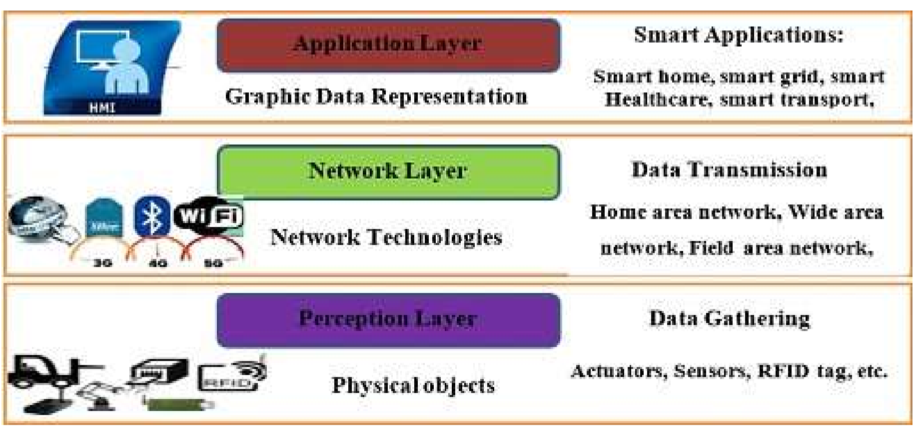


Figure.1. Layers of IoT [3]

This connectivity has been recognized to be the city's heart of what creates the smart cities namely wearable, smart grids, mobile applications and sensor networks that have been introduced to employ the innovative technology of the city to afford service thereby control their citizens in an effective way. This research showed that IoT proved to be the largest driver and enabler of initiating smart cities. It has been the stage in the ongoing transfer of conventional facilities of public into smart services and develop new things [4]. Similarly, the study intended to study a systematic analysis of ML (Machine Learning) and DM (Data Mining) methodologies selected in promoting smart cities. Methodi Ordinatio has been utilized to discover the related studies and VOS viewer software has been performed to

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

analyse the network. 39 important articles have been found where ML and DM methods have been analysed in addition to the areas engaged in smart city promotion. Predictive analytics has been the common method and the research intended to focus on fields of smart environment and smart mobility. This study pursues to motivate the methods that could be utilized by the government companies and agencies for developing smart cities has been vital to support goals for sustainable development [5]. Thus, this study aims to predict the amenities on smart city data using data mining techniques for efficient development of smart cities.

II. LITERATURE REVIEW:

This section discusses various conventional methods and approaches used by the existing systems to support smart city development.

Smart cities completely utilize cloud computing, IoT and other IT (Information Technology) generations to support deep incorporation of industrialization, urbanization and informationization in several industries of the cities. There exist several areas in developing smart cities like smart transportation, smart health care, smart energy systems, smart grid etc as the smart city enablers. The present age of digital technologies, particularly the incorporation of IoT-AI has been probable to be the main drivers of the roadmap. But digitally assisted manufacturing wouldn't be able to satisfy several or all the key indices defining smart city. It is also significant to study the way in which smart manufacturing could solve the indices like health, citizen centric, urban planning as well as environment that have been integral to the smart cities [6]. Hence, in this analysis, smart cities have been the prominent application field of IoT that afford range of smart cities with high quality to satisfy the requirements of citizens. This study [7] introduced a semi-supervised DRL (Deep Reinforcement Learning) model to fit the smart city applications that consumes unlabelled and labelled data for enhancing the accuracy and performance of the LA (Learning Agent).

Semi-supervised DRL has been an efficient solution to several IoT applications as it needs only little supervision through rewarding feedback. This learns an efficient policy for selecting among several alternative actions. Several studies focussed on one particular area of smart cities. In accordance with this, the paper [8] primarily concentrated on the traffic classification of smart cities which is significant to perform feature extraction and classification. In recent years, this topic has gained more attention as various aspects are associated with it such as intrusion detection, network security, QoS (Quality of Service) and management of network traffic. Correctly identifying internet traffic in huge network environment has been essential for IP (Internet Protocol) network management, QoS etc. But, due to classification of network traffic, ISP (Internet Service Providers) or network operator has to handle as well as control all the performances like blocking all the unwanted network traffic flow thereby managing resources. Besides, these methods have led the NO (Network Operator) and ISP to easily find the network application development and subsequently manage few resources for the preferred applications that grow every day [9]. On the other hand, to construct smart city with sustainable development, monitoring technique and fracturing framework has been established [10]. Fracturing technology has received more attention in recent days.

The enhancement in various new technologies associated with oil field has made the analysis of fracturing technology as well as remote monitoring system to be highly perfect. In addition, this paper [11] focussed on planning smart cities to estimate and control the HL (Heating Load) of buildings through a hybrid approach named PSO-XGBoost (Particle Swarm Optimisation-XGBoost). Outcomes explored that the introduced model has been the robust method to estimate HL of buildings. Though the outcomes have been perfect it assessing and predicting building's HL, further research has to be undertaken like enhancing other model's accuracy comprising of SVM, CART, RF, GP and XGBoost or attempt to construct various new hybrid AI systems relying on

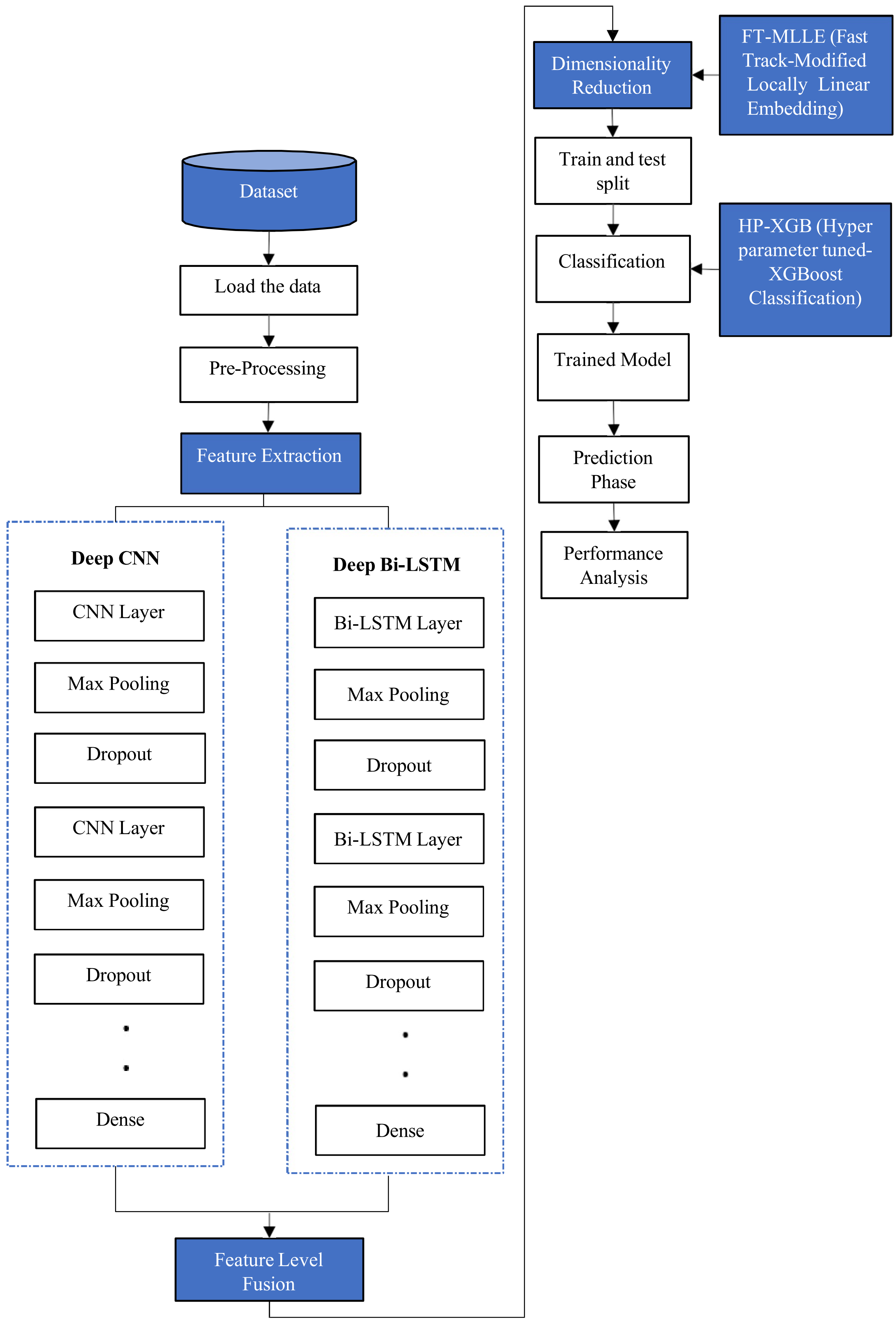
ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

optimization algorithms. It is also challenging to optimize the building design with respect to energy efficacy. Additionally, the study [12] constructed two traditional methods for predict the occupant counts at a specified location and time in smart environments. The proposed methods include ARIMA and LSTM (Long Short Term Memory) relying on Wi-Fi networks. Outcomes revealed that the LSTM had the ability to minimize the neuron counts required for building models for each time scale in comparison to other time-scale models. On contrast, social media contents have also been utilised for SCA (Smart City Application) through the proposed DS-DWR (Domain Specific-Distributed Word Representation) in the paper [13] that showed efficient outcomes [14] [15].

III. PROPOSEDE METHODOLOGY:

The study proposes methods based on data mining (ML and DL) to predict the amenities of smart city data. Various processes are involved in it to accomplish this. The overall view of the proposed system is shown in figure.2.

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)



ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

Figure.2. Overall view of the proposed system

Initially, the dataset is loaded to perform pre-processing which is the significant step to enable flexibility in further processing. After this, feature extraction is performed by the proposed DCNN-Bi-LSTM (Deep Convolutional Neural Network-Bi-directional Long Short-Term Memory) for selecting only the relevant features and ignore the irrelevant features. This process helps to improvise the classification accuracy. Followed by this, dimensionality reduction is performed through the introduced FT-MLLE (Fast Track- Modified Locally Linear Embedding) for minimizing the overfitting issue and complexity of ML model. This is then fed into train and test split. Classification is performed by HP- XGB (Hyper parameter-XGBoost). Finally, the trained model is fed into the prediction phase for predicting the amenities of smart city data. Performance analysis of the proposed system helps to evaluate the efficacy of the proposed system. For this analysis, accuracy, recall, precision and F-measure are the metrics to be taken into account.

III.I Deep Convolutional Neural Network (Deep CNN)

The CNN is a kind of Feed-Forward Artificial Neural Network (FFANN) that is biologically stimulated by a visual cortex organization. They are widely applicable in numerous areas like video and image recognition, Natural Language Processing (NLP) and Recommender Systems (RS). Typically, CNN comprises two major parts, namely convolutional (CNN's core building block) and pooling layers. The convolutional layer affords feature maps as output by calculating a dot product of filter and input feature map's local region. Followed by this, the non-linear function approximates complex functions, thereby squashing the NN output. Moreover, the pooling layer achieves down-sampling to feature maps through the computation of average or maximum value on the sub-region. The fully connected layers pursue stacked convolutional as well as pooling layers. Softmax is the last fully connected layer that computes scores for an individual class. Deep CNN is similar to CNN. But, deep CNN consists of various layers (ten convolutional layers, one fully connected layer and four max-pooling layers). A CNN is comprised of fundamental parts for feature extraction as well as classification. This study uses deep CNN for feature extraction that comprised of numerous convolutional layers, activation function and max- pooling. As CNN is computationally effective and has numerous layers, it learns many parameters than actually needed to solve an issue that ultimately increases the efficiency.

III.II Deep Bi-LSTM (Deep Bi-Directional Long Short Term Memory)-Neural Network

Generally, Bi-LSTM is an extended traditional LSTMs that have high capacity to improvise a model's performance on issues associated with sequence classification. These Bi-LSTMs train a combination of LSTMs instead of a single LSTM. The main motive behind this methodology is also straightforward that replicates the initial recurrent layer of the NN.

Later, this is fed as input to the corresponding input sequence and also reversing this sequence to the equivalent replicated layer. This makes the system advantageous as it resolves the issues of the conventional techniques and the Bi-LSTM can also be trained via the usage of all the available input information in past as well as future in a specific time step. Besides, in a general RNN, the state neuron split is responsible for a kind of backward state. This denotes the negative time direction. It is also liable for forward states that indicates the positive time direction. The Bi-LSTM architecture and LSTM architecture is shown in figure.3.

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

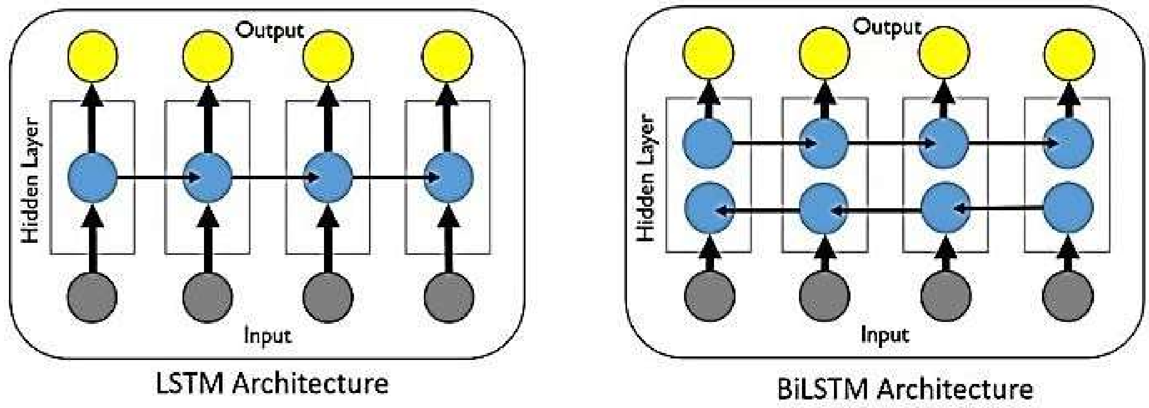


Figure.3. Architecture of Bi-LSTM and LSTM

As said earlier, Bi-LSTM is LSTM extension that comprise of input, output and hidden layer (shown in figure.3). The present study employs Deep Bi-LSTM (has several NN layers) which makes it suitable for efficient learning. Thus, this study uses both Deep CNN- Bi-LSTM for efficient feature extraction.

III.III Fast Track-Modified Locally Linear Embedding (FT-MLLE)

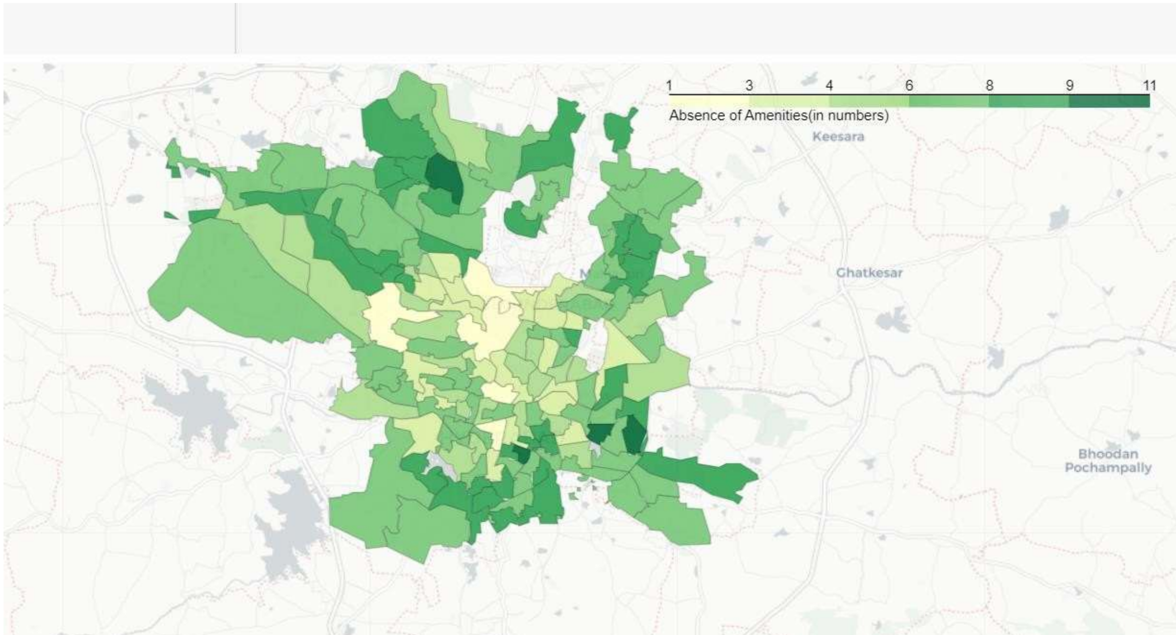
FT-MLLE attempts to minimize n-dimensions thereby attempt to reserve the actual feature structure's geometric features. This technique has gained high attention to investigators as it has the capability to handle many data of high dimension. It also follows a non-iterative path to find the embedding. It is an unsupervised technique employed for the process of dimensionality reduction that minimizes the ML model's complexity thereby reduce the overfitting issue. This occurs due to the use of several features. As the model gets complex, this might make the model fit the specific data leading to overfitting. On the other hand, features which don't decide output label might be utilized. But, this might not be helpful in real-life. Hence, dimensionality reduction is a significant step that is performed by the proposed FT-MLLE quickly.

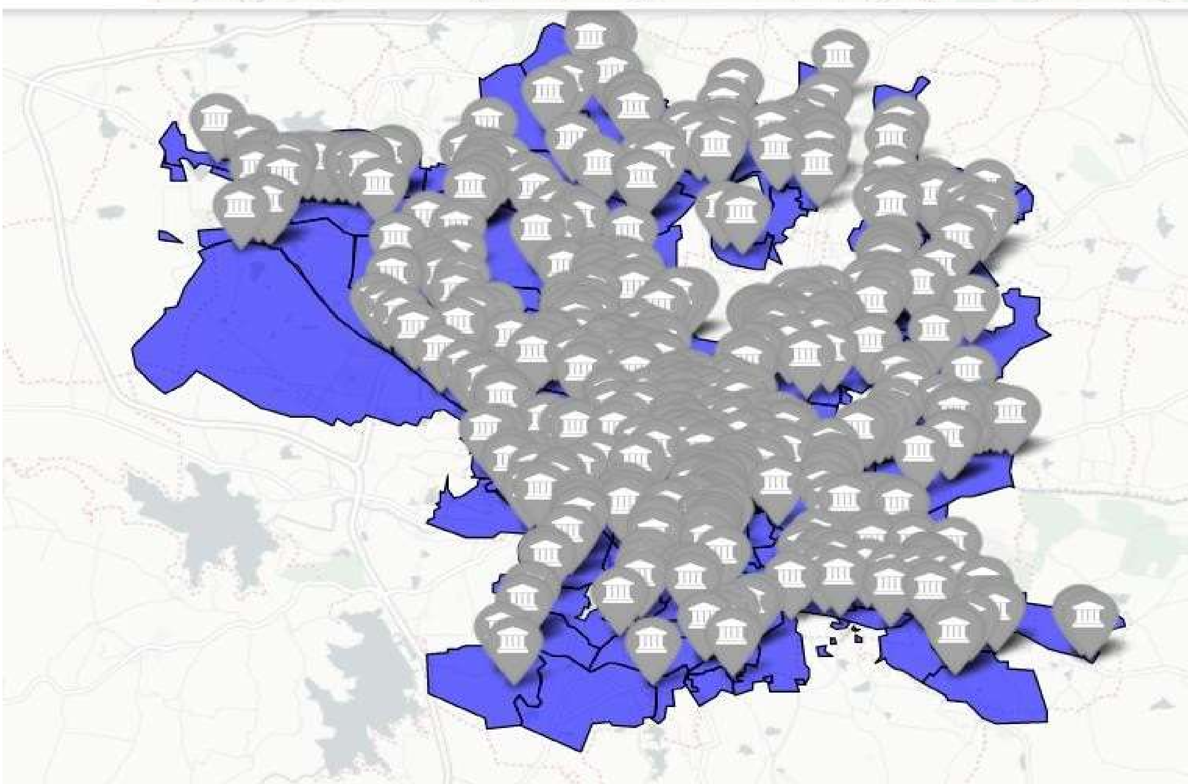
III.IV Hyper-parameter tuned XGBoost (HP-XGB)

The study proposes HP-XGB for classification. Typically, XGBoost is an efficient algorithm that has various advantages such as high flexibility, supports regularization, and works effectively in all datasets, ability to perform parallel processing, ability to deal with missing data, works quickly in comparison to gradient boosting etc. On the other hand, hyper-parameters (HP) can be defined as weights/certain values that find an algorithm's learning process. It operates by running several trials in a training job. Individual trial indicates the overall implementation of the proposed training model with values of the selected HP that is set within specific limits. XGBoost affords large HP range. Maximum count of XGBoost could be leveraged by HP tuning. Hence, the proposed HP-XGBoost is expected to perform efficient classification due to the above-mentioned advantages.

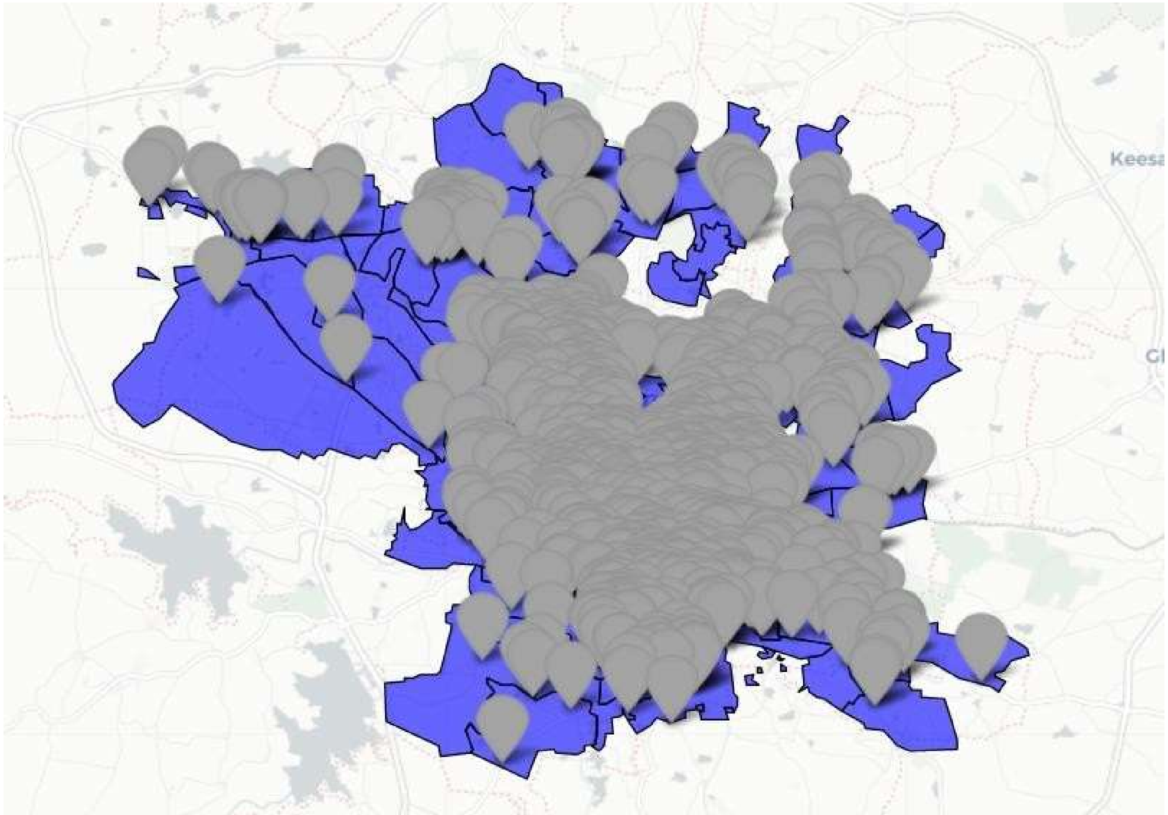
IV. EXPERIMENT AND RESULTS:

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)





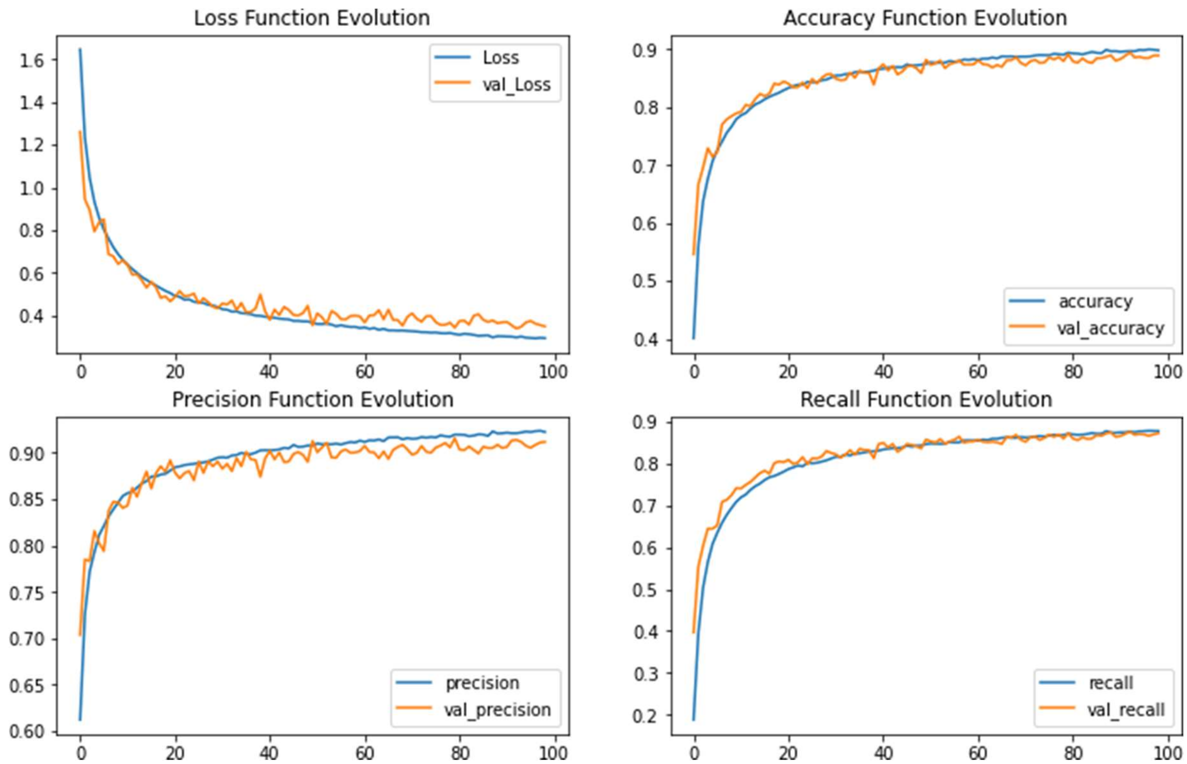
ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

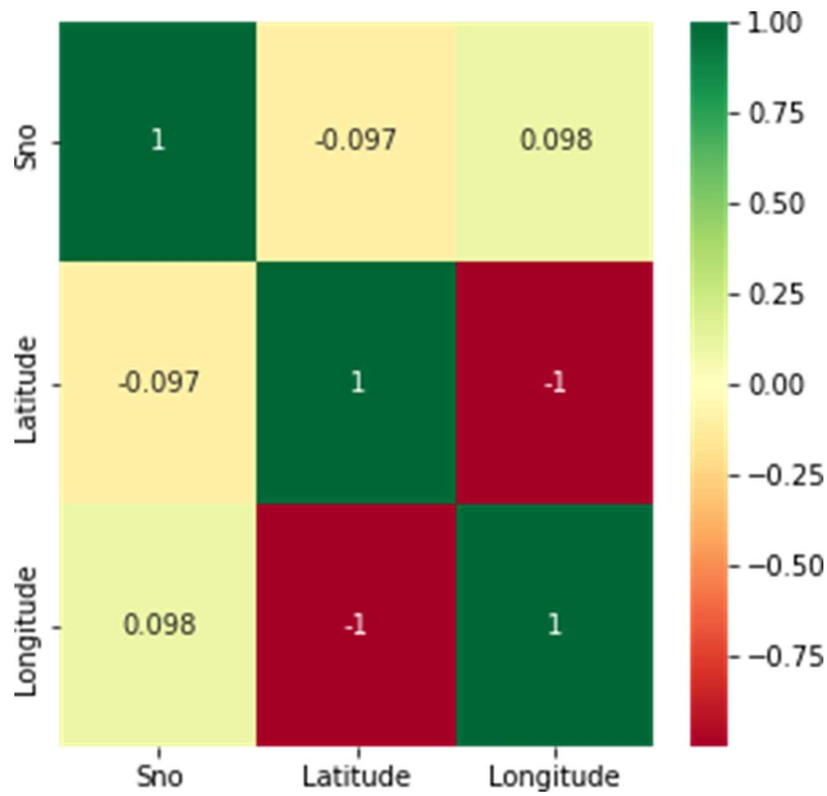



```
Total Amenities within GHMC
DUMPER BINS              2887
COMMUNITY HALLS          1374
OPEN SPACES              1355
HOSPITALS                 889
PARKS                     823
ANGANWADI CENTERS         796
TOILETS                   559
PLAY GROUNDS              310
ANNAPURNA MEALS           131
HERITAGE BUILDINGS         95
MARKET COMPLEXES           46
2BHK                       22
FIRE STATIONS              21
NIGHT SHELTERS             11
Name: Category, dtype: int64
```

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

```
313/313 [==============================] - 3s 9ms/step - loss: 0.3492 - accuracy: 0.8890 - precision: 0.9119 - recall: 0.8723
Test Accuracy : 88.90%
```





ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

V. CONCLUSION:

This study gives a strong framework to predict and manage the amenities of smart cities using advanced methodologies in machine learning and deep learning. The proposed system includes various models, such as Deep CNN, Bi-LSTM, FT-MLLE, and HP- XGBoost, for optimizing feature extraction, classification, and prediction accuracy. The results prove the potential of the system to be helpful in sustainable urban development through efficient resource allocation and improved quality of life for city dwellers. Contributions include a data-driven approach to smart city planning, which aligns with the needs of urban populations and facilitates goals for sustainable development.

VI. FUTURE SCOPE:

Model Enhancements: Explore hybrid techniques that combine newer optimization algorithms with traditional ones such as SVM, CART, and RF to enhance the predictive accuracy of existing models.
Real-Time Integration: Design real-time analytics frameworks through IoT devices for continuous feeding of data into the models for dynamic adaptability to the requirements of the city.
Scalability: Scale up the application to larger datasets and diverse city structures for validation and generalization of the system's applicability.
Addressing Limitations: Focus on the ethical concerns, such as data privacy and security, in the use of citizen data in predictive analytics.
Comparative Analysis: Compare the developed system in various economic contexts, such as emerging and advanced economies, to identify context-specific optimizations.

VII. REFERENCES:

[1]. B. Anthony Jnr, "A case-based reasoning recommender system for sustainable smart city development," *AI & SOCIETY,* vol. 36, pp. 159-183, 2021.
[2]. Y. Lim, J. Edelenbos, and A. Gianoli, "Identifying the results of smart city development: Findings from systematic literature review," *Cities,* vol. 95, p. 102397, 2019.
[3]. R. J. Hassan, S. R. Zeebaree, S. Y. Ameen, S. F. Kak, M. A. Sadeeq, Z. S. Ageed*, et al.*, "State of art survey for iot effects on smart city technology: challenges, opportunities, and solutions," *Asian Journal of Research in Computer Science,* pp. 32-48, 2021.
[4]. Z. S. Ageed, S. R. Zeebaree, M. M. Sadeeq, S. F. Kak, Z. N. Rashid, A. A. Salih*, et al.*, "A survey of data mining implementation in smart city applications," *Qubahan Academic Journal,* vol. 1, pp. 91-99, 2021.
[5]. J. T. d. Souza, A. C. d. Francisco, C. M. Piekarski, and G. F. d. Prado, "Data mining and machine learning to promote smart cities: A systematic review from 2000 to 2018," *Sustainability,* vol. 11, p. 1077, 2019.
[6]. M. Suvarna, L. Büth, J. Hejny, M. Mennenga, J. Li, Y. T. Ng*, et al.*, "Smart manufacturing for smart cities—overview, insights, and future directions," *Advanced Intelligent Systems,* vol. 2, p. 2000043, 2020.
[7]. M. Mohammadi, A. Al-Fuqaha, M. Guizani, and J.-S. Oh, "Semisupervised deep reinforcement learning in support of IoT and smart city services," *IEEE Internet of Things Journal,* vol. 5, pp. 624-635, 2017.

ISSN:2731-538X | E-ISSN:2731-5398
Vol. 19 No. 01 (2025)

[8]. M. Shafiq, Z. Tian, A. K. Bashir, A. Jolfaei, and X. Yu, "Data mining and machine learning methods for sustainable smart cities traffic classification: a survey," *Sustainable Cities and Society,* vol. 60, p. 102177, 2020.
[9]. N. Manju, B. Harish, and V. Prajwal, "Ensemble feature selection and classification of internet traffic using XGBoost classifier," *International Journal of Computer Network and Information Security,* vol. 11, p. 37, 2019.
[10]. H. Liang, A. Xian, M. Mao, P. Ni, and H. Wu, "A research on remote fracturing monitoring and decision-making method supporting smart city," *Sustainable Cities and Society,* vol. 62, p. 102414, 2020.
[11]. L. T. Le, H. Nguyen, J. Zhou, J. Dou, and H. Moayedi, "Estimating the heating load of buildings for smart city planning using a novel artificial intelligence technique PSO-XGBoost," *Applied Sciences,* vol. 9, p. 2714, 2019.
[12]. B. Qolomany, A. Al-Fuqaha, D. Benhaddou, and A. Gupta, "Role of deep LSTM neural networks and Wi-Fi networks in support of occupancy prediction in smart buildings," in *2017 IEEE 19th International Conference on High Performance Computing and Communications; IEEE 15th International Conference on Smart City; IEEE 3rd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, 2017, pp. 50-57.
[13]. M. Alam, F. Abid, C. Guangpei, and L. Yunrong, "Social media sentiment analysis through parallel dilated convolutional neural network for smart city applications," *Computer Communications,* vol. 154, pp. 129-137, 2020.
[14]. S.-U. Hassan, M. Shabbir, S. Iqbal, A. Said, F. Kamiran, R. Nawaz*, et al.*, "Leveraging deep learning and SNA approaches for smart city policing in the developing world," *International Journal of Information Management,* p. 102045, 2019.
[15]. M. Chong, A. Habib, N. Evangelopoulos, and H. W. Park, "Dynamic capabilities of a smart city: An innovative approach to discovering urban problems and solutions," *Government Information Quarterly,* vol. 35, pp. 682-692, 2018.