ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 02 (10), October 2023, pp. 12-23

Article Title: Integrating Whale Optimization Algorithm and Machine Learning for Efficient Traffic Management

Integrating Whale Optimization Algorithm and Machine Learning for Efficient Traffic Management

A. T. R. Krishna Priya, A. Yazhini

PG Scholar, Department of Computer Science and Engineering, Rohini College of Engineering and Technology, Tamilnadu, India.

Research Scholar, Department of Computer Science and Engineering, College of Engineering, Guindy Campus, Anna University, Chennai -25.

ABSTRACT

In order to protect extra retort teams, building labours, and the common community, road needs managing vehicle and pedestrian to manage traffic flows and offer guidance regarding traffic congestion, municipal or state roadway authorities may also utilize CCTV and other methods of traffic surveillance. One optimization framework is used to provide prompt and dependable decision-making by combining the dependable Whale Optimization Algorithms (WOA) and the quick-running Machine Learning (ML) algorithms. First, a standard WOA algorithm is used as a decision variable, with the network's total journey time serving as the unbiased role. In order to select the good performing model for additional hyper-tuning, multiple regression models are trained to estimate the overall journey interval in the traffic network under study. Finally, a good performance regression model, the extreme-gradient decision tree (XGBT), and WOA algorithm are clustered into a lone optimization outline.

Keywords: Traffic Control (TC), Wireless Sensor Network (WSN), Intelligent Transportation system (ITS), K-means clustering, Whale Optimization Algorithm (WOA) and eXtretme Gradient Boosting Algorithm (XGBoost).

1 Introduction

Effective traffic control (TC) has advantages for the environment, economy, society, and person. The rise in traffic volume through city roadways brought about by urban growth, traffic speeds have been steadily declining [1]. In the network of metropolitan roadways, there are dual classes of junctures: signal with and without lights. When a junction has signal lights, active signal scheduling depend on exact traffic flow forecast is essential to manage traffic; in the absence of signals [2]. In order to gather data on traffic flow in real-time and pinpoint areas of congestion, this system makes use of cameras and sensors that are installed on roads and if the data is gathered, machine learning algorithms are utilized to process and evaluate it in order to produce insights that may be applied to improve traffic flow [3]. Due to the Internet's rapid global development into almost every area of life, especially science and technology, are currently tasked with finding the best localization techniques for wireless sensor network (WSN) optimization, they are composed a lesser number of low-cost sensor nodes that interact wirelessly but are constrained in terms of memory, energy, and computational capability. [4]. Without taking into account real-time traffic data, static extent control traffic light divide the signal duration in time and phase to control system. Other algorithms for controlling traffic lights use sensors, such as inductive loop sensors, to collect data for traffic and then use that data as input [5]. Because of the promise of the underlying technology, the use of Intelligent Transportation Systems (ITS) has raised expectations for both individual travellers and society as a whole. In order to achieve traffic efficiency, key aim of ITS is to judge, generate, analyze, and add new sensor, facts and statement skills, and algorithms [6]. Data preparation is a crucial stage in the data mining process. Data preprocessing tries to improve the quality of the data and customize it to the specific data mining objective. Examples of these stages include cleaning, converting, and integrating the data [7]. An intriguing method for analyzing and forecasting traffic patterns is traffic flow prediction using K-means clustering. Data points can be grouped into clusters according to their similarity using the machine learning technique K-means clustering. [8]. Data normalization is used as a pre-processing step on the gathered data to scale the diverse collection of numerical data. Following pre-processing of the data, WOA is utilized to choose the best features. The characteristics that were essential for categorizing the stages of road traffic congestion are specified in the WOA output [9]. XGBoost is commonly employed as a predictive modelling tool in various applications because to its capacity to manage enormous datasets, comprehend intricate relationships, and generate precise forecasts. It's crucial to remember that traffic control and management are difficult jobs that frequently call for the integration of real-time data processing, machine learning, and domain knowledge in order to make wise judgments and maximize traffic flow. In this larger perspective, XGBoost can be a useful tool [10]. For making judgments quickly and with high reliability, the dependable WOA and the blazing-fast Machine Learning algorithms are integrated into a unified optimization framework. The list below demonstrates how the paper is organised: Section I gives a description of the introduction. The recent works are described in Section II. The proposed model and description are explained in Part III. The results are presented in Section IV. In Section V, the conclusion is presented.

2 Recent Works

Neetesh Kumar *et al* [2021] [11] have developed an ITS with (DITLCS) Dynamic and Intelligent Traffic Light Control System, to report these problems and progress the effectiveness of the light control traffic system. However, a certain amount of transportation is entering at a relatively high rate, these traffic signal control systems will become paralyzed. Because of this, traffic cops frequently have to oversee traffic light junctions personally by waving signals by hand.

Xiaoshuang Li *et al* [2022] [12] have described to replicate traffic engineers' problem-solving abilities, a data-augmented deep behavioural cloning (DADBC) technique is put forth. Under the parallel learning (PL) conceptual framework, machine learning techniques are integrated to address complex system decision-making challenges. The DADBC approach builds upon a

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 02 (10), October 2023, pp. 12-23

hybrid demonstration by taking advantage of generative adversarial networks (GANs) and learning traffic engineers' control strategies through the use of DBC models. Advancements in technology, including reinforcement learning, help to improve the policy.

Penghao Sun *et al* [2021] [13] have proposed a machine learning, to create a model-free TE scheme with the help of a Deep Reinforcement Learning (DRL) to choose a subset of the network's links and identify the key links, Scale DRL uses the concept from pinning control theory. Make dynamic adjustments to the tie loads of the essential associates using a DRL algorithm based on the traffic distribution data. By achieving a part device on the network through the concept of pinning control, Scale DRL solves the expletive of dimensionality issue in large-scale networks by reducing action space of DRL algorithm.

Hongsuk Yi *et al* [2021] [14] have examined an auto machine learning which emerged as a popular method for learning applications, such as intelligent transportation. The application of auto-machine knowledge for hyperparameter alteration to traffic datasets in key highway system states is the major emphasis of this study. Moreover, LSTM networks can perform well in traffic data time series forecasting. Experience has shown that, in particular, search algorithms should not regard the hyper parameters equally, as this can complicate the process.

Bingzhao Gao *et al* [2020] [15] have developed an customized Adaptive Cruise Control (ACC) system that ensures car-following, comfort, and fuel efficiency while recognizing various driving styles through the use of mode-specific predictive control (MPC) and driving style identification. 66 randomly selected drivers participate in a set of real-world car trials to gather driving data, which is subsequently clustered using an unsupervised machine learning technique to determine the organiser limits defining various styles of driving. Multi-objective optimization is used to tackle the problem of customized ACC system, and the MPC approach is one approach to do so.

3 Proposed Work Explanation

The boosted whale optimization technique, which uses machine learning for traffic control, is used in this suggested work. To find groups of data objects in a dataset, k-means clustering is applied as a preprocessing step to the input dataset. The noise-removed input is supplied into WOA for optimization problem solving after pre-processing. Finally, the processed input is passed into the XGBoost, which combines the estimates of several weaker, simpler models to accurately forecast a target variable.

ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 02 (10), October 2023, pp. 12-23

Article Title: Integrating Whale Optimization Algorithm and Machine Learning for Efficient Traffic Management

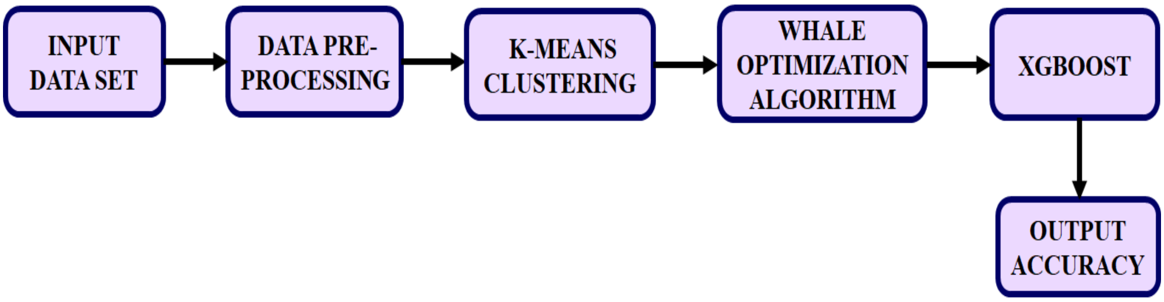


Figure 1: *Proposed Block diagram*

3.1 Data Pre-Processing

As part of the data mining and analysis process, the input traffic dataset undergoes pre-processing of data to transform unprocessed data into a format understandable and analyzed by computers and machine learning. The established steps need to go through to make sure data is successfully preprocessed.

Data quality Assessment: Take a close look at the data to determine its overall quality, project relevance, and consistency.

Data cleaning: Data cleaning is the process of filling in the gaps in a data set and removing, correcting, or restoring erroneous or superfluous information.

Data transformation: Data conversion into the format(s) needed for analysis and other processes that come after.

Data reduction: Simplifies and improves the accuracy of analysis while storing less data.

3.2 K-Means Clustering Algorithm

After preprocessing, the preprocessed data is segmented to group the relevant data features. With each data point belonging to a separate group, the iterative K-Means segmentation technique aims to split K pre-defined distinct into the dataset, non-overlapping subgroups (clusters). The objective is to preserve the greatest possible distance between the clusters and the highest possible similarity between the intra-cluster data points. The method includes dividing the data points into clusters with the goal of minimizing the sum of squared distances between all of the data points in a cluster, measured from the arithmetic mean of the cluster centroid.

3.3 Whale Optimization Algorithm

To select the appropriate features the proposed work involves WOA, to extract the relevant data features. This involves:

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 02 (10), October 2023, pp. 12-23

3.3.1 Encircling Prey

Because the location of optimal design in the search space is uncertain, a priority for humpback whales, when they detect prey, is to circle around them, assuming that the current best candidate solution is either the target prey or extremely close to the optimum. Once the best search agent has been identified, the other search agents will try to reposition themselves in respect to this agent.

The succeeding equalities represent conditions:

$$\vec{D} = \left|\vec{C}.\overrightarrow{X*}(t) - \vec{X}(t)\right| \tag{1}$$

$$\vec{X}(t+1) = \overrightarrow{X^*}(t) - \vec{A}\vec{D} \tag{2}$$

Where, t defines the current iteration, $\vec{\mathbf{A}}$and $\vec{\mathbf{C}}$ are coefficient vectors, $\mathbf{X}^*$ is the position vector of the finest result got so far, $\vec{\mathbf{X}}$ is the position vector, and is a part by part growth. It mentioning here that $\mathbf{X}^*$ must be simplified if there is a best output.

The vectors $\vec{\mathbf{A}}$ and $\vec{\mathbf{C}}$ are intended as follows:

$$\vec{A} = 2\vec{a}.\vec{r} - \vec{a} \tag{3}$$

$$\vec{C} = 2.\vec{r} \tag{4}$$

Where, $\vec{\boldsymbol{a}}$ is linearly lowered from 2 to 0 over the path of repetitions and $\vec{\boldsymbol{r}}$ is a casual course in [0, 1].

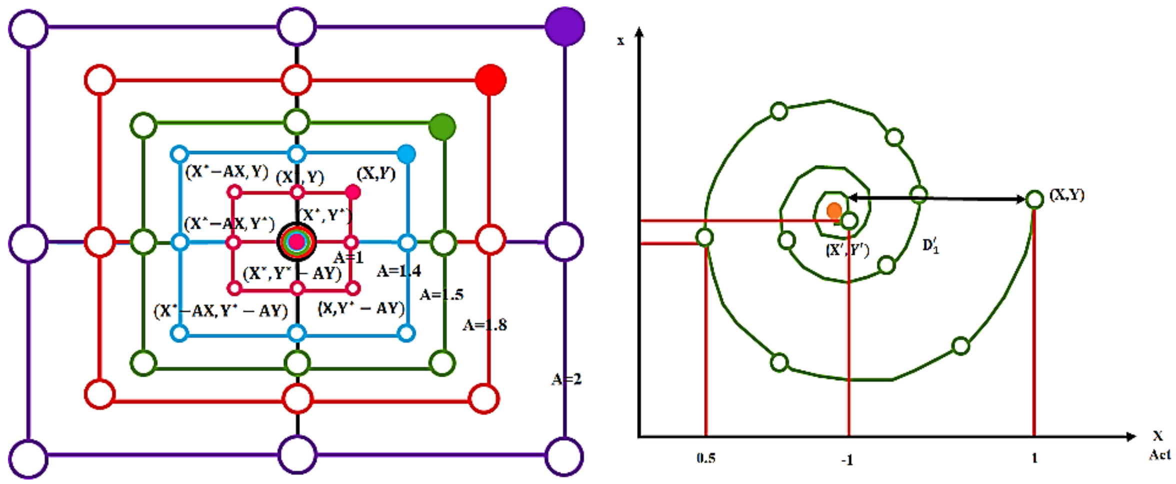
Bubble-net attacking method (exploitation phase)

Binary methods were developed in demand to exactly simulate the bubble-net characters of humpback whales:

Shrinking encircling mechanism

This behavior is achieved by lowering the value of $\vec{a}$ in the Equation 4.5. Note that the difference range of $\vec{A}$ is a casual value in the recess [−a, a] where a is lessened from 2 to 0 over the path of recurrences. Set chance prices for $\vec{A}$ in [−1, 1], the new site of a study cause can be defined anywhere in between the creative place of the cause and the spot of the recent finest cause. Fig. 2 (a) shows the likely positions from (X, Y) towards (X^*,Y^*) that can be reached by $0 \leq\ A\ \leq 1$ in a 2D space.

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 02 (10), October 2023, pp. 12-23



(a) (b)

Figure 2: *The best solution found thus far is the bubble-net search mechanism used in WOA (X∗): (a) shrinking encircling mechanism and (b) spiral updating position.*

Spiral updating position

As can be seen in Fig. 4.2 (b), this technique first calculates the space between the whale situated at (X, Y) and prey found at (X^*,Y^*). A curved calculation is then made among the place of whale and prey to satirist the helix-shaped effort of humpback whales as follows:

$$\vec{X}(t+1) = \overrightarrow{D'}.e^{bl}.cos(2\pi l) + \overrightarrow{X^*}(t) \quad (5)$$

Where $\overrightarrow{D'}= |\overrightarrow{X^*}(t) - \vec{X}(t)|$ and shows the space of ith whale to the prey (best solution got so far), b is a continuous major for shape of the logarithmic bent, l is a arbitrary sum in [−1, 1], and. is an element-by-element increase. The calculated ideal is as follows:

$$\vec{D} = |\vec{C}.\overrightarrow{X_{rand}} - \vec{X}| \quad (6)$$

$$\vec{X}(t+1) = \overrightarrow{X_{rand}} - \vec{A}.\vec{D} \quad (7)$$

Where $\overrightarrow{X_{rand}}$ is a casual point direction (a random whale) select from present inhabitants.

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 02 (10), October 2023, pp. 12-23

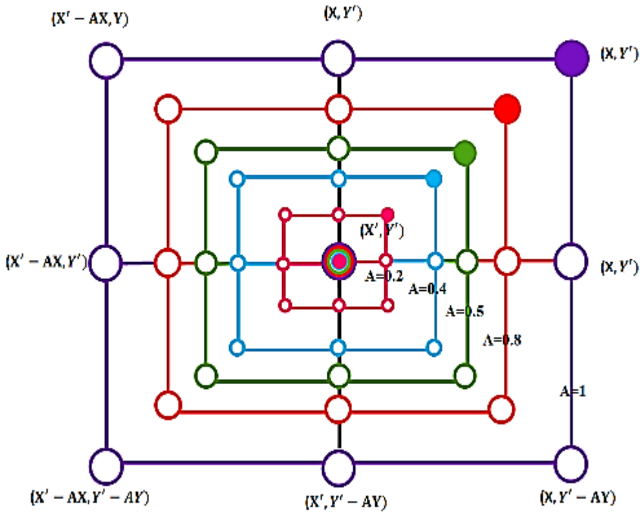


Figure 3: *WOA's exploration mechanism ($\boldsymbol{X}^*$ is a randomly selected quest cause)*

3.5.2 Search for prey (exploration phase)

To find prey, the same technique based on the variation of the $\vec{A}$vector is applied (exploration). Actually, humpback whales casually search based on one another's positions. Therefore, to push the search agent to go far away from a reference whale, use $\vec{A}$with random values larger than 1 or less than −1.

3.4 XGBOOST

Finally, after selecting the optimal features, classification is performed using XGBoost, which is renowned for its potent prediction capabilities and adept handling of intricate datasets. It is frequently utilized for classification tasks in both real-world applications and machine learning contests. When using XGBoost models, it is imperative to be mindful of overfitting and hyperparameter adjustment.

4 Results and Discussion

After implementing the proposed work in a PYTHON simulation, the following outcomes were attained.

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 02 (10), October 2023, pp. 12-23

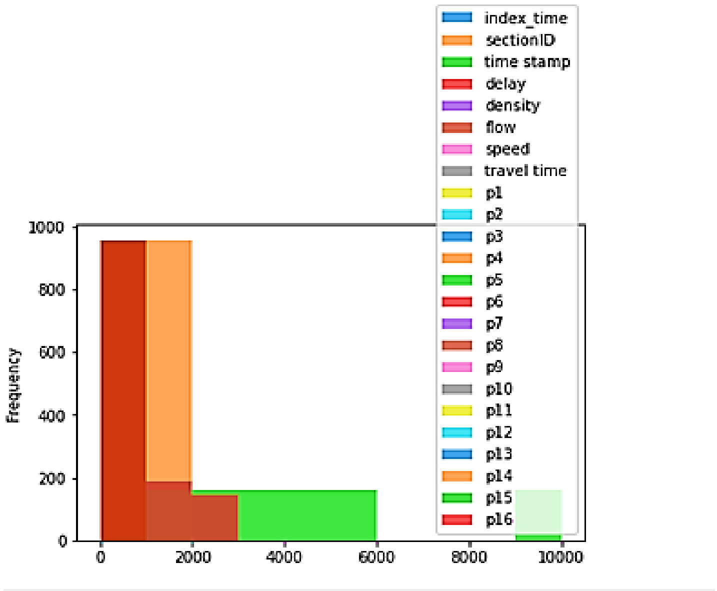


Figure 4: *Represents the boosted genetic algorithm machine learning dataset*

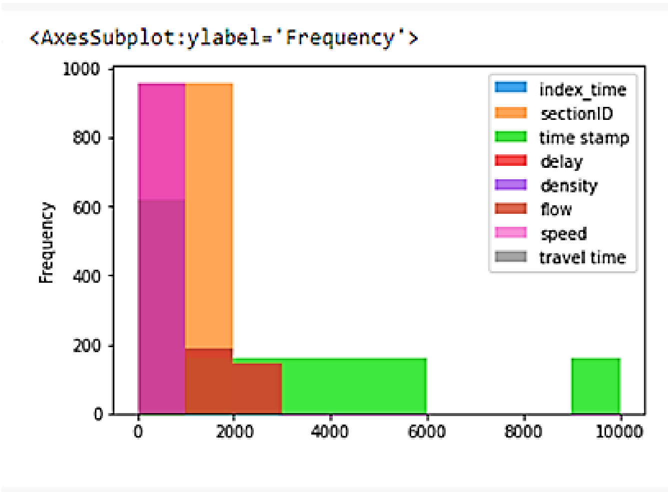


Figure 5: *Sub Plot of Main Data*

The major data subplot is displayed in Fig. 5. The primary data are the index time, section ID, time stamp, delay, density, and flow, speed, and travel time.

ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 02 (10), October 2023, pp. 12-23

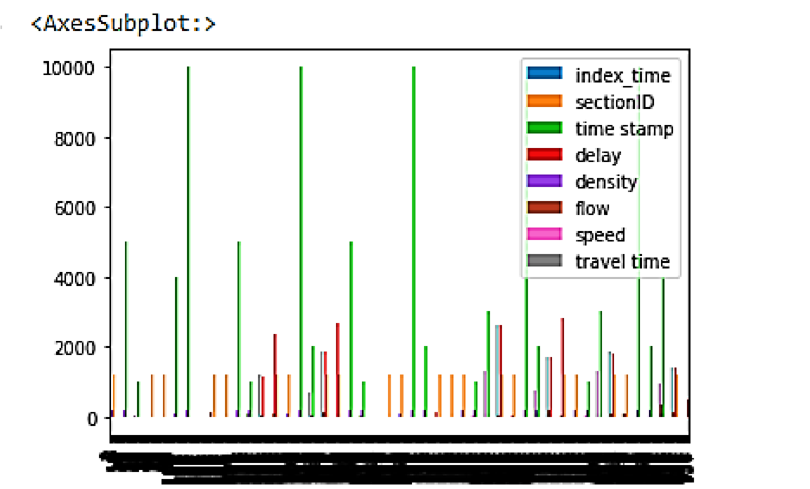


Figure 6: *Bar Graph of Main Data*

The primary data is represented graphically in Fig. 6. The time stamp indicates the actions that are now occurring in the traffic area and is higher.

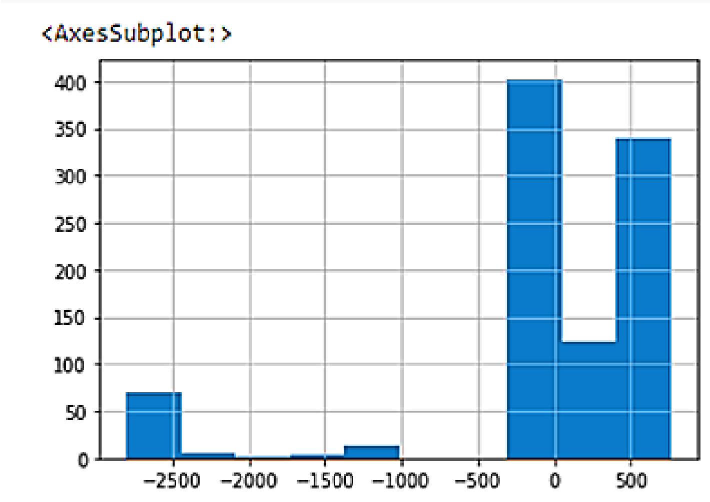


Figure 7: *Graph of Travel Time*

Figure 7, displays a graphical representation of the trip time in the traffic region.

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 02 (10), October 2023, pp. 12-23

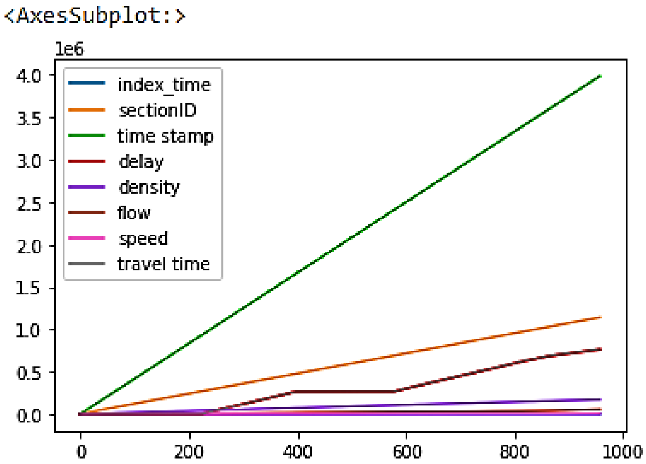


Figure 8: *Direction Range of Main Data*

The main graph, which shows the successive growth of the primary data, is seen in its entirety in Fig 8.

Table 1: *Comparison of Vehicle flow between proposed and Existing Method*

Techniques	MAE	RMSE	MAPE (%)
SVM	4.362	5.891	4.321
Proposed XGBOOST	2.315	3.215	2.023

The Table 1 explains the comparison of vehicle flow between proposed and existing method of MAE, RMSE and MAPE. The proposed method lays out lower percentage while comparing the existing method.

5 Conclusion

This study proposes an updated whale optimization method for traffic management that is based on machine learning. To identify the categories in the data that aren't clearly labeled, apply K-means clustering. By reducing error and improving prediction and classification accuracy, the WOA is used to address optimization problems. By eliminating unnecessary steps to save time, reduce errors, and avoid repetitive work, optimization strategies help you work more efficiently. Using XGBoost at last, an accurate prediction of the target variable is obtained. Large datasets and missing data are handled effectively by XGBoost. The results will be expanded in the near future by Deep Learning methods. To improve CNN's categorization accuracy even more, a more.

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 02 (10), October 2023, pp. 12-23

Reference

1. Guan Gui; Ziqi Zhou; Juan Wang; Fan Liu; Jinlong Sun, Year: 2020, "Machine Learning Aided Air Traffic Flow Analysis Based on Aviation Big Data", in IEEE Transactions on Vehicular Technology, Vol: 69, no: 5, pp. 4817 – 4826.

2. Jinbao Mou, Year: 2020, "Intersection traffic control based on multi-objective optimization", IEEE Access, Vol: 8, pp. 61615 – 61620.

3. Tuo Mao; Adriana-Simona Mihăiță; Fang Chen; Hai L. Vu, Year: 2021, "Boosted genetic algorithm using machine learning for traffic control optimization", IEEE Transactions on Intelligent Transportation Systems, Vol: 23, no: 7, pp. 7112 – 7141.

4. Nagham A. Al-Madi; Adnan A. Hnaif, Year: 2022, "Optimizing Traffic Signals in Smart Cities Based on Genetic Algorithm", Computer Systems Science & Engineering, Vol: 40, no: 1.

5. Penghao Sun; Julong Lan; Junfei Li; Jianpeng Zhang; Yuxiang Hu; Zehua Guo, Year: 2021, "A Scalable Deep Reinforcement Learning Approach for Traffic Engineering Based on Link Control", in IEEE Communications Letters, Vol: 25, no: 1, pp. 171 – 175.

6. Palwasha W. Shaikh; Mohammed El-Abd; Mounib Khanafer; Kaizhou Gao, Year: 2020, "A review on swarm intelligence and evolutionary algorithms for solving the traffic signal control problem", IEEE transactions on intelligent transportation systems, Vol: 23, no: 1, pp. 48 – 63.

7. Yang Zhang; Dongrong Xin, Year: 2020, "Dynamic optimization long short-term memory model based on data preprocessing for short-term traffic flow prediction", IEEE Access, Vol: 8, pp. 91510 – 91520.

8. Lin Cao; Tao Wang; Dongfeng Wang; Kangning Du; Yunxiao Liu; Chong Fu, Year: 2020, "Lane determination of vehicles based on a novel clustering algorithm for intelligent traffic monitoring", IEEE Access, Vol: 8, pp. 63004 – 63017.

9. R. Vijayanand; D. Devaraj, Year: 2020, "A novel feature selection method using whale optimization algorithm and genetic operators for intrusion detection system in wireless mesh network", IEEE Access, Vol: 8, pp. 56847 – 56854.

10. Xuchen Dong; Ting Lei; Shangtai Jin; Zhongsheng Hou, Year: 2018, "Short-term traffic flow prediction based on XGBoost", In 2018 IEEE 7th Data Driven Control and Learning Systems Conference (DDCLS), pp. 854 – 859. IEEE.

11. Neetesh Kumar; Syed Shameerur Rahman; Navin Dhakad, Year: 2021, "Fuzzy Inference Enabled Deep Reinforcement Learning-Based Traffic Light Control for Intelligent Transportation System", in IEEE Transactions on Intelligent Transportation Systems, Vol: 22, no: 8, pp. 4919 – 4928.

International Journal of Advanced Trends in Engineering and Management (IJATEM)
Vol. 02 (10), October 2023, pp. 12-23

12. Xiaoshuang Li; Peijun Ye; Junchen Jin; Fenghua Zhu; Fei-Yue Wang, Year: 2022, "Data Augmented Deep Behavioral Cloning for Urban Traffic Control Operations Under a Parallel Learning Framework", in IEEE Transactions on Intelligent Transportation Systems, Vol: 23, no: 6, pp. 5128 – 5137.

13. Penghao Sun; Julong Lan; Junfei Li; Jianpeng Zhang; Yuxiang Hu; Zehua Guo, Year: 2021, "A Scalable Deep Reinforcement Learning Approach for Traffic Engineering Based on Link Control", in IEEE Communications Letters, Vol: 25, no: 1, pp. 171 – 175.

14. Hongsuk Yi; Khac-Hoai Nam Bui, Year: 2021, "An Automated Hyper parameter Search-Based Deep Learning Model for Highway Traffic Prediction", in IEEE Transactions on Intelligent Transportation Systems, Vol: 22, no: 9, pp. 5486 – 5495.

15. Bingzhao Gao; Kunyang Cai; Ting Qu; Yunfeng Hu; Hong Chen, Year: 2020, "Personalized Adaptive Cruise Control Based on Online Driving Style Recognition Technology and Model Predictive Control", in IEEE Transactions on Vehicular Technology, Vol: 69, no: 11, pp. 12482 – 12496.