INTERNET OF THINGS-GROUP 1
TRAFFIC MANAGEMENT SYSTEM

STUDENT NAME : S MALINI

REGISTER NUMBER: 822621106008

NM ID : au822621106008

COLLEGE CODE : 8226

COLLEGE NAME : ARIFA INSTITUTE OF TECHNOLOGY

EMAIL ID : malinivani03@gmail.com

A Review of Traffic Congestion Prediction Using Artificial Intelligence

Introduction

Artificial intelligence (AI) is the most important branch of computer science in this era of big data. AI was born 50 years ago and came a long way, making encouraging progress, especially in machine learning, data mining, computer vision, expert systems, natural language processing, robotics, and related applications [1]. Machine learning is the most popular branch of AI. Other classes of AI include probabilistic models, deep learning, artificial neural network systems, and game theory. +ese classes are developed and applied in a wide range of sectors. Recently, it has been the leading research area in transportation engineering, especially in traffic congestion prediction. Traffic congestion has a direct and indirect impact on a country's economy and its dwellers' health. According to Ali et al. [2], traffic congestion causes Pak Rs.1 million every day in terms of opportunity cost and fuel consumption due to traffic congestion. Traffic congestion affects on individual level as well. Time loss, especially during peak hours, mental stress, and the added pollution to the global warming are also some important factors caused due to traffic congestion.

Ensuring economic growth and the road users' comfort are the two requirements for the development of a country, which is impossible without smooth traffic flow. With the development in the transportation sector by collecting traffic information, authorities are putting more attention on traffic congestion monitoring. Traffic congestion prediction provides the authorities with the required time to plan in the allocation of resources to make the journey smooth for travellers. Traffic congestion prediction problem discussed in this paper can be defined as an estimation of parameters related to traffic congestion into the short-term future, e.g., 15 minutes to a few hours by applying different AI methodologies by using collected traffic data. +ere are usually five parameters to evaluate, including traffic volume, traffic density, occupancy, traffic congestion index, and travel time while monitoring and predicting traffic congestions.

Depending on the nature of the collected data, a variety of AI approaches are applied to evaluate the congestion parameters. +is article systematically discusses the models and their advantage and disadvantages. +e primary motivation of this review is to gather the articles focusing solely on traffic congestion prediction models. +e keywords used in the search process included "traffic congestion prediction" OR "traffic congestion estimation" OR "congestion prediction modelling" OR "prediction of traffic congestion" OR "road congestion forecast" OR "traffic congestion forecast." For efficient screening, research paper search was done according to year using search engines like Scopus, Google Scholar, and Science Direct. After collecting all the peerreviewed journal and conference papers written in the English language, 48 articles were found for review. Any studies focusing on the cause of traffic congestion, traffic congestion control, traffic congestion impact, traffic congestion propagation, traffic congestion prevention, etc. were excluded from this manuscript. A general layout of the prediction approaches is provided in Section 2. +e data collection sources and congestion forecasting models are explained in Sections 3–6 and they provide the overall discussion and concluding remarks.

General Layout

Traffic congestion forecasting has two basic steps of data collection and prediction model development. Every step of the methodology is important and may affect the results if not done correctly. After data collection, data processing plays a vital role to prepare the training and testing datasets. Case area differs for different research. After developing the model, it is validated with other base models and ground true results. Figure 1 shows the general components of traffic congestion prediction studies. +ese branches were further divided into more specific sub-branches and are discussed in the following sections.

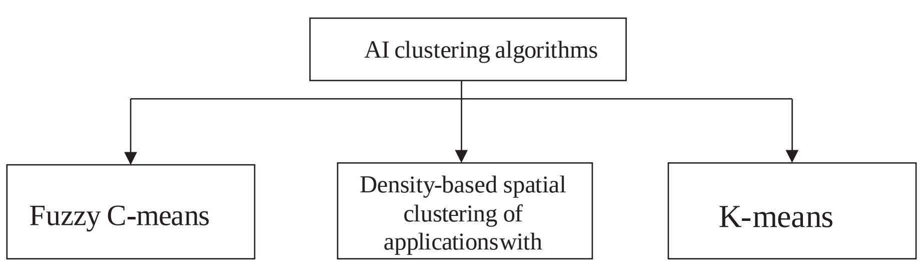
Data Source

Traffic datasets used in different studies can be mainly divided into two classes, including stationary and probe data. Stationary data can be further divided into sensor data and fixed cameras. On the other hand, probe data that were used in the studies were GPS data mounted on vehicles. Stationary sensors continuously capture spatiotemporal data of traffic. However, sensor operation may interrupt anytime. Authorities should always consider this temporary failure of the sensor while planning by using this data. +e advantage of the sensor data is that there is no confusion on the location of the vehicles. +e most used dataset was Performance Measurement System (PeMS) that collects highway data across all major metropolitan areas of the State of California of traffic flow, sensor occupancy, and travel speed in real-time. Most of the studies used dataset from the I-5 highway, in San Diego, California, every 5 minutes [3–6]. Other systems included the Genetec blufaxcloud travel-time system engine (GBTTSE) [7] and the Topologically Integrated Geographic Encoding and Referencing (TIGER) line graph [8].

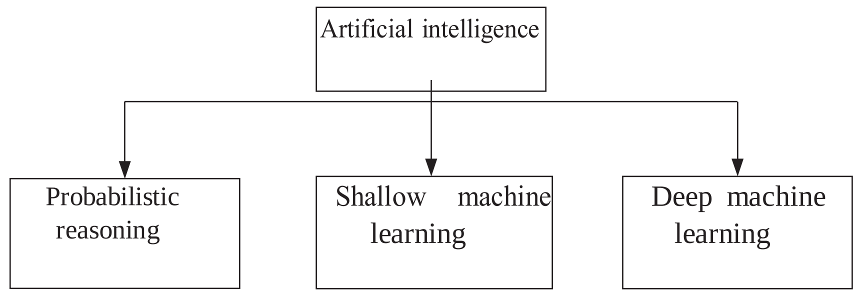
Clustering Algorithms

Some studies use clustering the acquired data before applying the main congestion models of prediction. +is hybrid modelling technique is applied to fine-tune the input values and to use them in the training phase. Figure 2 shows the commonly used AI clustering models in this field of research. +e models are described briefly in this section

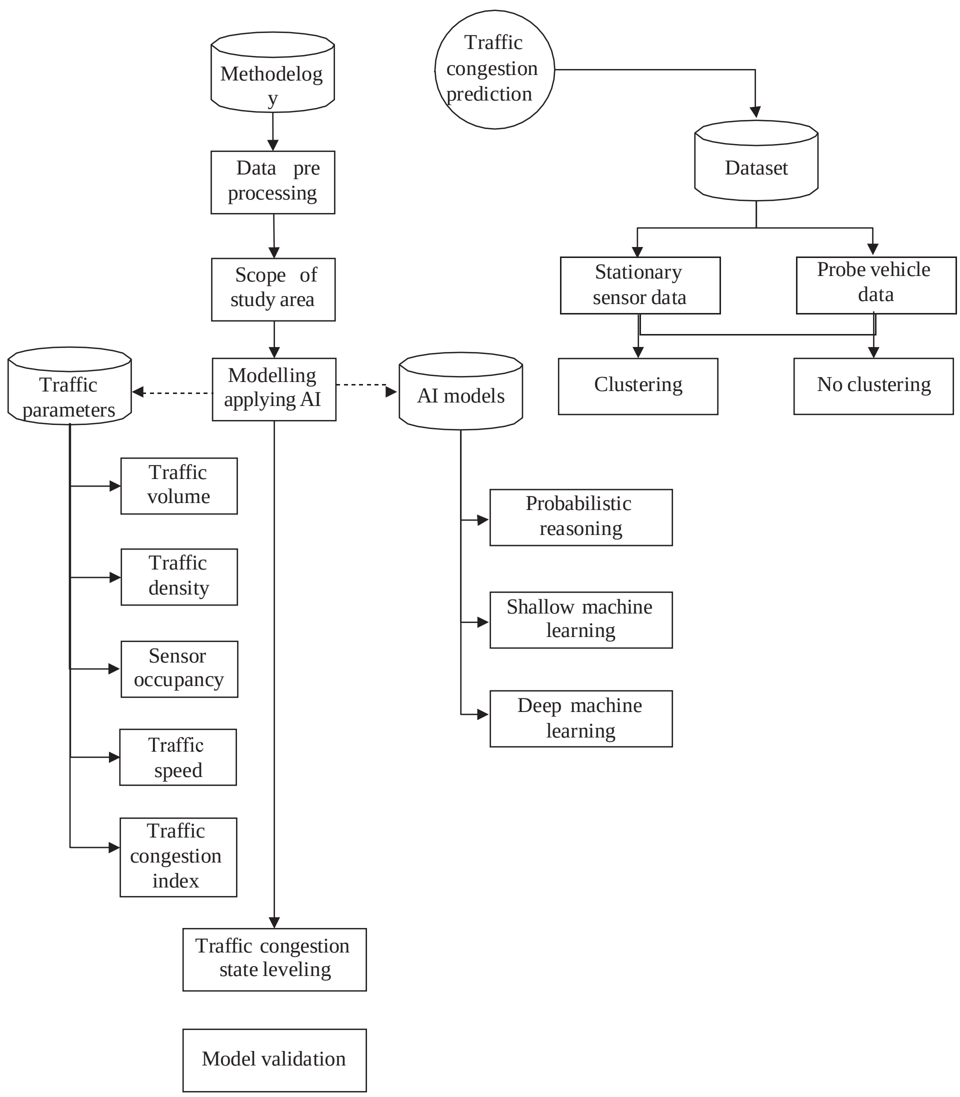
Fuzzy C-Means (FCM) is a popular nondeterministic clustering technique in data mining. In traffic engineering researches, traffic pattern recognition plays an important role. Besides, these studies often face the limitation of missing or incomplete data. To deal with these constraints, FCM has become a commonly applied clustering technique. +e advantage of this approach is, unlike original C-means clustering methods, it can overcome the issue of getting trapped in the local optimum [14]. However, FCM requires setting a predefined cluster number, which is not always possible while dealing with massive data without any prior knowledge of the data dimension. Besides, this model becomes computationally expensive with data size increment. Different studies have applied FCM successfully by improving its limitations. Some studies changed the fuzzy index value for each FCM algorithm execution [15], some calculated the Davies-Bouldin (DB) index [10], while others applied the K-means clustering algorithm [16, 17].



Commonly used AI clustering algorithms



Branches of artificial intelligence in this article.



The layout of traffic congestion prediction system.

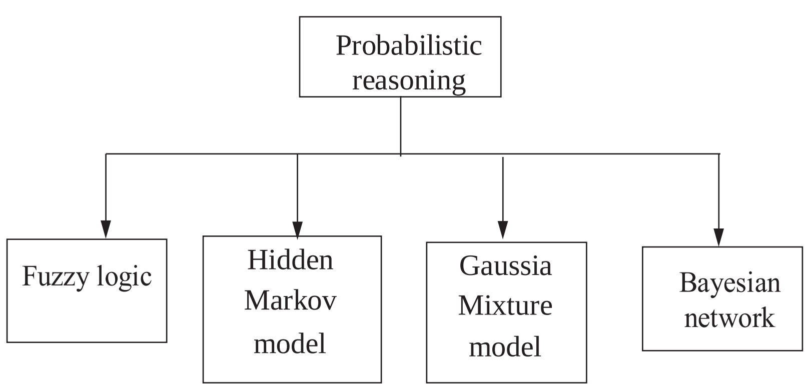
Applied Methodology

Traffic flow is a complex amalgamation of heterogenous traffic fleet. Thus, traffic pattern prediction modelling could be an easy and efficient congestion prediction approach. However, depending on the data characteristics and quality,different classes of AI are applied in various studies. Figure 3 shows the main branches—probabilistic reasoning and machine learning (ML). Machine learning comprised ofboth shallow and deep learning algorithms. However, with the progress of this article, these sections were subdivided into detailed algorithms.

To generalise traffic congestion forecasting studies using different models is not straight forward. The common factors of all the articles include the study area, data col- lection horizon, predicted parameter, prediction intervals, and validation procedure. Most of the articles took studied corridor segment as the study area [5, 27–30]. Other study areas included the traffic network [31, 32], ring road [9], andarterial road [33]. Data collection horizon varied from 2 years [34] to less than a day [35] in the studies. Congestion estimation is done predicting traffic flow parameters, e.g., traffic speed [4], density, speed [5], and congestion index [31], to mention a few. The Congestion Index (CI) approach is suitable to monitor the congestion level continuously in aspatiotemporal dimension. Studies those compared their results with the ground truth value or with other models used mean absolute error (MAE) (equation (1)), symmetric mean absolute percentage error (SMAPE) (equation (1)), MAPE, root-mean-squared error (RMSE) (equation (3)), false positive rate (FPR) (equation (4)), and detection rate (DR) (equation (5)). Many studies used SUMO to validate their models:

Probabilistic Reasoning. Probabilistic reasoning is a significant section of AI. It is applied to deal with the field ofuncertain knowledge and reasoning. A variety of these algorithms are commonly used in traffic congestion predictionstudies. The studies discussed hereunder probabilistic reasoning is shown in subdivision.

Fuzzy Logic. Zade is a commonly applied model in dynamic traffic congestion prediction as it allows vagueness instead of binary outcomes. In this method, several membership functions are developed those represent the degree of truth. With the vastness with time, traffic data are be- coming complex and nonlinear. Due to its ability to deal with uncertainty in the dataset, fuzzy logic has becomepopular in traffic congestion prediction studies.



Subdivision of probabilistic reasoning models

As discussed before, with the development of optimi-sation algorithms, optimisation of the fuzzy logic system'smembership functions is becoming diverse. With time, thesimplest form of FRBS-TSK has become popular due to itsgood interpretability. Some other sectors of transportationwhere fuzzy logic models are popular include traffic light/signal control [39, 40], traffic flow prediction (Zhang and Ye[41]), traffic accident prediction [42], and modified fuzzylogic for freeway travel time estimation.

Hidden Markov Model. The hidden Markov model (HMM) is a combination of stochastic characteristics of Markov process and discrete characteristics of Markov chains. It is a stochastic, time-series event recognition technique. Some studies have applied Markov chain model for traffic pattern recognition during congestion prediction [21, 25, 44]. Pearson correlation coefficient (PCC) is commonly applied among the parameters during pattern con- struction. Zaki et al. [32] applied HMM to select the appropriate prediction model from several models they developed applying the adaptive neuro fuzzy inference sys- tem (ANFIS). They obtained optimal state transition by fourprocessing steps: initialization, recursion, termination, and backtracking. The last step analysed the previous step to determine the probability of the current state by using the Viterbi algorithm. Based on the log-likelihood of the initial model parameter, defined by expectation maximization (EM) algorithm, of HMM with the traffic pattern, a suitable congestion model was selected for prediction

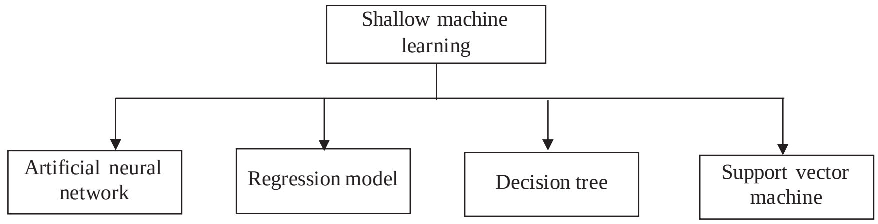
Other than the models mentioned above, the Kalman Filter (KF) is also a popular probabilistic algorithm.With the increment of available data, data fusion methods are becoming popular. The fusion of historical and real-time traffic data can achieve a higher level of traffic congestion prediction accuracy. In this regard, KF is commonly applied.Extended KF (EKF) is an extension of KF, which can be used to stochastically filter the nonlinear noises to improve the mean and covariance of an estimated state. Therefore, after data fusion, it updated the estimated covariance error by removing outliers [7].

Artificial Neural Network. Artificial neural network (ANN) was developed, mimicking the function of the human brain to solve different nonlinear problems. It is a first-order mathematical or computational model that consists ofa set of interconnected processors or neurons. Figure 7shows a simple ANN structure. Due to its easy implementation and efficient forecasting ability, ANN has becomepopular in the field of traffic congestion prediction research.Hopfield network, feedforward network, and back- propagation are the examples of ANN. Feedforward neural network (FNN) is the simplest NN, where the input data go to the hidden layer and from there to the output layer.Backpropagation neural network (BPNN) consists of feed- forward and weight adjustment of the layers and is the mostcommonly applied ANN in transportation management. Xuet al. [31] applied BPNN to predict traffic flow, thus to evaluate congestion factor in their study. They proposed occupancy-based congestion factor (CRO) evaluation method with three other evaluated congestion factors based on mileage ratio of congestion (CMRC), road speed (CRS),and vehicle density (CVD). They also evaluated the effect ofdata-size on real-time rendering of road congestion. Complex road network with higher interconnections showed higher complication in simulation and rendering. The advantage of the proposed model was that it took little processing time for high sampling data rendering. The model can be used as a general congestion prediction modelfor different road networks. Some used hybrid NN for congestion prediction. Nadeem and Fowdur [11] predicted congestion in spatial space, applying the combination of one of six SML algorithms with NN. Six SML algorithms included moving average (MA), autoregressive integrated moving average (ARIMA), linear regression, second- and third-degree polynomial regression, and k-nearest neighbour (KNN). The model showing the least RMSE value wascombined with BPNN to form hybrid NN. The hidden layerhad seven neurons, which was determined by trial and error.However, it was a very preliminary level work. It did not show the effect of data increment in the accuracy.

On the other hand, Jain et al. [33] developed both linear and exponential regression model using IBM SPSS softwareto find the relevant variables. The authors converted het- erogenous vehicles into passenger car unit (PCU) for sim- plification. Three independent variables were considered to estimation origin-destination- (O-D-) based congestion measures. They used PCC to evaluate the correlation amongthe parameters. However, simply averaging O-D node pa- rameters may not provide the actual situation of dynamic traffic patterns.

Table 1: Traffic congestion prediction studies in probabilistic reasoning.

Methodology	Road type	Data source	Input parameters	Target domain	No. of congestion state levels*	Reference
Hierarchical fuzzy rule-based system	Highway corridor	Sensor	Occupancy	Speed	2	Zhang et al. [30]
			Speed	Speed		Lopez-garcia et al. [37]
Evolutionary fuzzy rule learning			Traffic flow	Traffic density		Onieva et al. [28]
Mamdani-type fuzzy logic inference	Highway, trunk road, branch road	—	Speed	Congestion Index	4	Cao and Wang [3]
			Density Travel time			Wang et al. [58]
Fuzzy inference	Highway corridor	Camera	Traffic flow Speed			
Fuzzy comprehensive evaluation	Highway corridor	Probe	Traffic volume Speed	Saturation Density speed	5	Kong et al. [4] Yang et al. [5]
Hidden Markov model	Highway network	Sensor	Emission matrix	Traffic pattern selection	—	Zaki et al. [32]
			Emission matrix Transition matrix	Traffic pattern determination	—	Zaki et al. [25]
	Main road	Probe	Observation probability Transition probability	Mapping GPS data	—	Sun et al. [45]
Gaussian distribution	Highway corridor	Sensor	Traffic volume	Optimal feature selection	—	Yang [29]
Bayesian network	Build-up area	Simulation	Road and bus increment	Congestion probability	—	Yi Liu et al. [59]
	Bridge	Sensor	Intensity Occupation Average speed Average distance Network direction			Asencio-Cortés et al. [54]
	Highway network	Sensor	Day and time weather Incidents Traffic flow Occupancy Speed Level of service Congestion state	Congestion probability		Kim and Wang [34]



the time-series stationary. A lghamdietal. [67] took d as 1 asone differencing order could make the model stationary. Next, they applied the autocorrelation function (ACF) and predictability, an optimal segment length and velocity was found. However, with less available data, an increased number of segments increased the predictability. Another traffic parameter, travel time, was used to find CI by Liu andWu [73]. They applied the random forest ML algorithm to forecast traffic congestion states. At first, they extracted 100sample sets to construct 100 decision trees by using boot- strap. The number of feature attributes was determined as the square root of the total number of features. Chen et al. [16] also applied the CART method for prediction and classification of traffic congestion. The authors applied Moran's I method to analyse the spatiotemporal correlation among different road network traffic flow. The model showed effectiveness compared with SVM and K-means algorithm.

Regression models are useful to be applied for time series problems. Therefore, regression models are suitable for traffic forecasting problems. However, these models are notreliable for nonlinear, rapidly changing the multidimension dataset. The results need to be modified according to pre- diction errors.

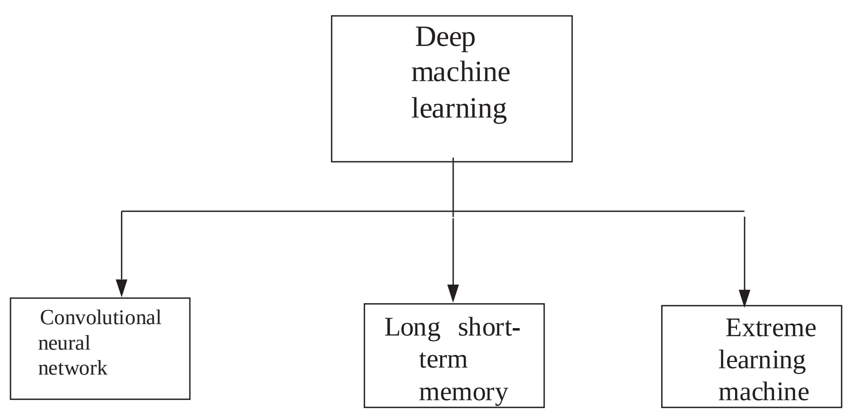
Decision tree is a simple classification problem-solving model that can be applied for multifeature data, e.g., Liuand Wu [73] applied weather condition, road condition, time period, and holiday as the input variables. This model's knowledge can be represented in the form of IF- THEN rules, making it an easily interpretable problem. It isalso needed to be kept in mind that the classification results are usually binary and therefore, not suitable where thecongestion level is required to be known. Other sectors of transport, where decision tree models applied are trafficprediction [74] and traffic signal optimisation with Fuzzy logic [75].

Support Vector Machine. The support vector machine(SVM) is a statistical machine learning method. The main idea of this model is to map the nonlinear data to a higher dimensional linear space where data can be linearly classified by hyperplane [1]. Therefore, it can be very useful in traffic flow pattern identification for traffic congestion prediction. Tseng et al. [13] determined travel speed in predicting real- time congestion applying SVM. They used Apache Storm toprocess big data using spouts and bolts. Traffic, weather sensors, and events collected from social media of close proximity were evaluated together by the system. They categorised vehicle speed into classes and referred them as labels. Speed of the previous three intervals was used to trainthe proposed model. However, the congestion level cat- egororised from 0 to 100 does not carry a specific knowledgeof the severity of the level, especially to the road users. Increment in training data raised accuracy and computa- tional time. This may ultimately make it difficult to make real-time congestion prediction.

Deep Machine Learning. DML algorithms consist of several hidden layers to process nonlinear problems. The most significant advantage of these algorithms is they can extract features from the input data without any prior knowledge. Unlike SML, feature extraction and model training are done together in these algorithms. DML can convert the vast continuous and complex traffic data with limited collection time horizon into patterns or feature vectors. From last few years, DML has become popular in traffic congestion prediction studies. Traffic congestion studies that used DML algorithms are shown in Figure 8 anddiscussed in this section.

Convolutional Neural Network. Convolutional neuralnetwork (CNN) is a commonly applied DML algorithm in traffic engineering. Due to the excellent performance ofCNN in image processing, while applying in traffic prediction, traffic flow data are converted into a 2-D matrix to process. There are five main parts of a CNN structure in transportation: the input layer, convolution layer, pool layer, full connection layer, and output layer. Both the convolution and pooling layer extracts important features. The depth of these two layers differs in different studies. Majority of the studies converted traffic flow data into an image of a 2-D matrix. In the studies performed by Ma et al.

Whereas Chen et al. [68] used a five-layered convolution offilter size of (2 2) without the pooling layer. The authors applied a novel method called convolution-based deep neural network modelling periodic traffic data (PCNN). The study folded the time-series to generate the input combining real-time and historical traffic data. To capture the correlation of a new time slot with the immediate past, they duplicated the congestion level of the last slot in the matrix. Zhu et al. [49] also applied five convolution- pooling layers as well as (3 3) and (2 2) sizes, respec- tively. Along with temporal and spatial data, the authors also incorporated time interval data to produce a 3-D input matrix. Unlike these studies, Zhang et al. [6] preprocessed the raw data by performing a spatiotemporal cross-cor- relation analysis of traffic flow sequence data using PCC. Then, they applied a model named spatiotemporal feature selection algorithm (STFSA) on the traffic flow sequence data to select the feature subsets as the input matrix. A 2- layered CNN with the convolutional and pooling size as same as the previous studies was used. However, STFSA considers its heuristics, biases, and trade-offs and does not guarantee optimality.



Extreme Learning Machine. In recent years, a novel learning algorithm called the extreme learning machine (ELM) is proposed for training the single layer feed-forwardneural network (SLFN). In ELM, input weights and hidden biases are assigned randomly instead of being exhaustively tuned. Therefore, ELM training is fast. Therefore, taking thisadvantage into account, Ban et al. [19] applied the ELM model for real-time traffic congestion prediction. They de- termine CI using the average travel speed. A 4-fold cross-validation was done to avoid noise in raw data.

Discussion and Research Gaps

Research in traffic congestion prediction is increasing ex- ponentially. Among the two sources, most of the studies used stationary sensor/camera data. Although sensor data cannot capture the dynamic traffic change, frequent change in source makes it complicated to evaluate the flow patternsfor probe data [95]. Data collection horizon is an important factor in traffic congestion studies. The small horizon of a few days [3] cannot capture the actual situation of the congestion as traffic is dynamic. Other studies that used datafor a few months showed the limitation of seasonality

Probabilistic reasoning algorithms were mostly applied for a part of the prediction model, e.g., map matching and optimal feature number selection. Fuzzy logic is the most widely used algorithm in this class of algorithms. From otherbranches, ANN and RNN are the mostly applied models. Mostof the studies that applied hybrid or ensembled models belongto probabilistic and shallow learning class. Only two studies applied hybrid deep learning models while predicting net- workwide congestion. Tables 4, 5–6 summarize the advantageand weaknesses of the algorithms of different branches. Among all DML models, RNN is more suitable for timeseries prediction. In a few studies, RNN performed better than CNN as the gap between the traffic speeds in different classes was very small [12, 69]. However, due to little re- search in traffic congestion field, a lot of new ML algorithmsare yet to be applied.

Future Direction

Traffic congestion is a promising area of research. There- fore, there are multiple directions to conduct in future research.

Numerous forecasting models have already been applied in road traffic congestion forecasting. However, with the newly developed forecasting models, there is more scope tomake the congestion prediction more precise. Also, in this era of information, the use of increased available traffic databy applying the newly developed forecasting models can improve the prediction accuracy.

The semisupervised model was applied only for the EML model. Other machine learning algorithms should be ex- plored for using both labelled and unlabelled data for higherprediction accuracy. Also, a limited number of studies havefocused on real-time congestion forecasting. In future, re- searches should pay attention to real-time traffic congestion estimation problem.

Another future direction can be focusing on the levelof traffic congestion. A few studies have divided the traffic congestion into a few states. However, for better traffic management, knowing the grade of congestion is essen- tial. Therefore, future researches should focus on this. Besides, most studies focused on only one traffic pa- rameter to forecast congestion for congestion prediction. This can be an excellent future direction to give attention to more than one parameter and combining the results during congestion forecasting to make the forecasting more reliable.