SHORT TERM TRAFFIC FLOW PREDICTION USING MACHINE LEARNING – KNN, SVM AND ANN WITH WEATHER INFORMATION

Faysal Ibna Rahman[1]

[1] *Civil and Environmental Engineering, University of Yamanashi, Yamanashi, 400-8511, Japan*

Abstract: Short term traffic prediction is one of the attractive topics in current Intelligent Transport System (ITS) research and practice. The rapid progress in machine learning and the emergence of new data sources makes it possible to observe and predict traffic conditions in cities more accurately than ever. In this study, three different sets of weather information or condition related parameters are incorporated with the different algorithms of machine learning for better traffic flow prediction. Three methods of machine learning- k-nearest neighbor (KNN), support vector machine (SVM), artificial neural network (ANN) are used in this research. However, it is hard to select the most appropriate machine learning traffic flow prediction model considering particular observation data. This paper shows the effect of the selection of each fundamental component of three machine learning algorithms and their effect on prediction accuracy. Five months of historical traffic flow data are trained with the weather condition. Then considering weather conditions, the traffic flow of one month is predicted. One hour interval traffic flow prediction considering weather information KNN gives more accurate results than SVM and ANN with 14.384% mean absolute percentage error and 0.948 R-square value.

Keywords: machine learning, prediction, traffic flow, traffic management, weather information.

1. Introduction

Due to the rapid progress of economic and technology, the mobility of city dwellers has increased day by day in urban areas, which also brought remarkable challenges like traffic congestion and causes extra hours of sitting in traffic jams. When traffic demand exceeds road capacity, congestion occurs. It has been proven to be challenging for highway agencies to respond with the present situation of congestion. Increasing road capacity to cope with present traffic demand is becoming less an option for these agencies due to the combination of financial, environmental and social issues to solve congestion. On the other hand, due to physical conditions in many cities, it is difficult to construct more roads. Therefore, the key attention has been on improving the performance of current facilities through continuous monitoring and spreading of traffic information using the intelligent traffic management system. Advanced Traffic Management Systems (ATMS) and Intelligent Transportation Systems (ITS) can help overcome or significantly reduce the impact of such negative effects on city-dwellers by providing information about

---

[1] Corresponding author: g17tc015@yamanashi.ac.jp

traffic conditions before and during their trips. Additionally, this information can be applied to provide alternatives to users so that they may make an informed decision about their trips.

In ATMSs and ITSs, it is a fundamental challenge to predict the next possible states of traffic with high precision. Machine learning methods are quite popular in traffic prediction because of their structural flexibility, accuracy, and reliability (Vlahogianni, 2009). By providing better prediction performance considering complex and congestion conditions, they gain reliability among the researches and engineers (Smith *et al.*, 2002). Moreover, such data-driven machine learning tools do not require extensive expertise in physical rules in traffic flow modeling (Van Lint *et al.*, 2004).

Machine learning is an application that provides systems the ability to automatically learn and improve from experience without being explicit. The process of learning begins with observation or data. Machine learning-based methods used in traffic prediction by capturing the patterns or relationships between data series without modeling the physical traffic process such as artificial neural networks (ANN), support vector machine (SVM) and k-nearest neighbor (KNN). Among other factors, weather condition has a great impact on traffic flow characteristics and driving behaviors (Nagy *et al.*, 2018). Three predominant variable categories are affected by inclement weather such as traffic demand, traffic safety, and traffic flow relationship (Maza *et al.*, 2006). Study on free flow traffic condition, it has been found that 6% of traffic speed reduction occurred due to rain and visibility reduce up 12% (Maza *et al.*, 2006). The research

carried by Datla and Sharma (Datla *et al.*, 2008) found that traffic flow varies with the category of trip and hour of the day due to cold and snowy weather conditions.

In 2006, a project is completed under the Road Weather Management Program where data of three cities- Seattle, Minneapolis, and Baltimore are used to develop statistical models to quantify the impacts of weather on traffic by adjusting the weather factors (Hranac *et al.*, 2006). By motivating Rahman's study (2019) where Naïve Bayes Classifier is used for forecasting traffic flow using weather conditions but this study cannot provide satisfying results. Other traffic flow prediction models, such as time series analysis (Hasnat *et al.*, 2019) cannot consider the weather condition. To overcome those problems and utilizing the Artificial Intelligent in the transportation field. The machine learning algorithms are used such as K-nearest neighbor (KNN), support vector machine (SVM) and artificial neural networks (ANN) to predict traffic flow incorporating weather information. It is also important to determine the best parameters set for machine learning algorithms considering weather-related parameters to predict traffic flow. This research can give the appropriate answer to fix the best parameter set for predicting traffic flow.

KNN approach has been applied for forecasting traffic. Some researchers have applied KNN to forecast traffic flow rates (Clark, 2003, Smith *et al.*, 2002). Similarly, other researches applied KNN for forecasting travel time (Bajaw *et al.*, 2003, Robinson *et al.*, 2005). Moreover, since the studies were conducted under different settings such as traffic parameters considered, the nature of datasets and their sources, aggregation

intervals, forecast step durations, production horizons, and performance measure considered.

A novel supervised learning method, Support vector machine used for classification and regression. SVM shows very resistance to the overfitting problem, achieving high generalization performance in solving various time series forecasting problems that have been applied in the prediction of time series (Cao *et al.*, 2003). Yeng and Yu (2009) proposed a freeway dynamic traffic flow forecasting model bases on the SMO support vector machine. Yu *et al.* (2011) develop SVM-based models to predict bus arrival time.

As studied by (Vlahogianni *et al.*, 2014), traffic forecasting methods have been gradually shifting from traditional statistical models to computational intelligence approaches. Computational intelligence approaches are often used to depict high dimensional and non-linear relationships. Among the computational intelligence approaches, Artificial Neural Network (ANN) receives numerous success in the field of transportation (Vlahogianni *et al.*, 2014). The precursory study of utilizing neural networks into traffic prediction can be traced back to the 1990s, when (Hua *et al.*, 1994) introduced the concept of the neural network into freeway travel time estimation.

2. Objectives

This paper focuses on the one-hour interval of the traffic flow prediction considering weather conditions in machine learning (KNN, SVM, and ANN). Which type of weather-related parameters are more accurate for parameter selection to predict is also studied. Parameters are considered

in this study such as day, day type (holiday, working, weekend), time, weather conditions (sunny, not sunny, raining, not raining) precipitation value and temperature value. Three sets of parameters are considered in this study to find the best set for traffic prediction.

Five months of traffic flow data are used for training purposes in machine learning (KNN, SVM, and ANN) algorithms considering parameters and next one month of traffic flow are predicted considering observation parameters which are then compared with observation flow to find the suitable machine learning algorithm for predicting traffic flow incorporating best weather-related parameters. This paper also shows the effect of the selection of each fundamental component of three machine learning algorithms and their effect on prediction accuracy.

3. Study Location

From January 2017 to June 2017, six months one hour interval traffic flow sensor data are collected from the link between junction-1 and junction-2 of Dublin Airport route, Ireland. The study location is shown in Fig. 1. Those data are collected from the Transportation Infrastructure Ireland (TII, 2019).

First five months traffic flow data are used as training purpose in three different machine learning algorithms then one month of data are used in accuracy test or predict traffic flow purpose. Six months of traffic flow data are shown in Fig. 2. Weather data are collected from www.accuweather and www. wunderground.com. Then weather data are combined with traffic flow for training and blue data indicates test data.

Fig. 1.
Study Location, Link between Junction-1 and Junction-2 of Dublin Airport Route
Source: (TII, 2019)

4. Machine Learning Algorithms

4.1. K-Nearest Neighbor (KNN)

The KNN algorithm is a simple, easy to appliance supervised machine learning algorithms that can be used to solve both classification and regression problems. The KNN strategy is more powerful than the parametric time series models on account of its inclination which does not require a comprehension of the idea of the connection between the features and outcomes.
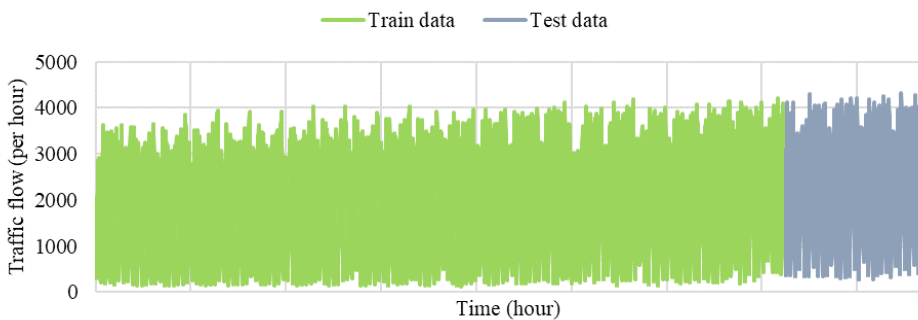

Fig. 2.
An Hour Interval Observed Traffic Flow from January to June, 2017

The basic presumption of the KNN method is "that observations which are close together in feature-space are likely to belong to the same class or to have the same a posterior distribution of their respective classes" (Devijver *et al.*, 1982). In its application of data

prediction, the KNN method can consider several numbers of observations (also termed as nearest neighbors) from a historical dataset and then predict future variables based on the nearest neighbor set. The nearest neighbor set can reflect the historical traffic flow data that are similar to the current traffic flow considering observation.

The KNN based prediction method can be deconstructed into three fundamental components: a database of observations, a neighborhood search procedure, and a prediction process. Fig. 3 depicts the general flow of data through the KNN based prediction method.
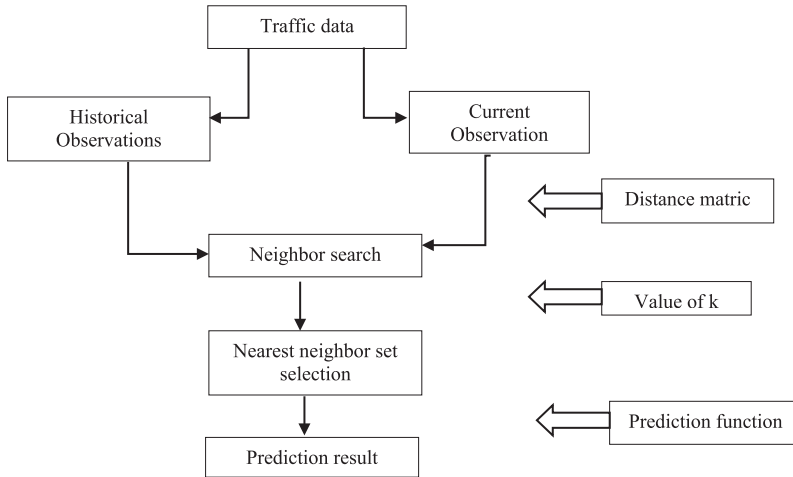

Fig. 3.
General Structure of KNN Based Prediction Method

The search procedure finds the nearest neighbors, which are the historical observations that are most similar to the current condition. The nearest neighbors then become the inputs to the prediction step so that it may calculate a predictive value. During these three procedures, key design parameters are the definition of a distance metric to determine the nearness of historical data to the current conditions, the choice of and the selection of a prediction function given a collection of nearest neighbors:

- Distance metric: This is used to determine the distance between the current input feature vector and

historical observations. The most commonly used metrics include Euclidean distance. Let $dist_{p,q}$ be the distance between two feature vectors and with dimension. The equation of the Euclidean distance is given below.

$$dist_{p,q}^2 = (x_{pn} - x_{qn}) + (x_{q(n-1)} - x_{q(n-1)}) + \cdots + (x_{pn} - x_{qn}) \qquad (1)$$

- Choice of k: Determines how many nearest neighbors are chosen from the historical dataset. For example, if is chosen to be ten, then the ten historical observations that have the nearest distances to the input feature vector will be used in the prediction process (Robinson, 2005).

- Prediction function: It is the core of the KNN based method. Let l_y represent the set of k nearest neighbors corresponding to the current input x_c. The predictive output y_c is given by:

$$y_c = f(l_y) \qquad (2)$$

The function of $f(\ )$ is called the prediction function (Smith *et al.,* 2002) or the local estimation method (Robinson, 2005). For the first time in the early 1990s, Davis & Nihan (1991) used KNN based method to predict traffic flow.

4.2. Support Vector Machine (SVM)

A Support Vector Machine (SVM) is a discriminative classifier formally defined by a separating hyperplane. In other words, given labeled training data (supervised learning), the algorithm outputs an optimal hyperplane which categorizes new examples. In two-dimensional space, this hyperplane is a line dividing a plane into two parts wherein each class lay on either side. Along with the performing linear classification, using the kernel, SVM can effectively use a non-linear classification.

Support Vector Regression (SVR) was introduced by Vapnik (1995) based on statistical learning theory and implements the structural risk minimization principle from computational learning theory. The basic idea of SVR is to "map the data x into a high dimensional feature space F via a nonlinear mapping ϕ and to do linear regression in this space; thus, linear regression in a high dimensional (feature) space corresponds to non-linear regression in the low dimensional input space" (Müller *et al.*, 1997). The regression function can be written as:

$$f(x) = \{w. \Phi(x)\} + b \text{ with } \phi : R^n \to F, w \in F \qquad (3)$$

Where w is a vector in the feature space, $\Phi(x)$ is a function that maps the input x to a vector in the feature space and b is a threshold.

The dot product in equation 2.3 can be replaced using a kernel function. The advantage of using a kernel function is that this can enable the dot product to be calculated in a higher-dimensional feature space without explicitly mapping $\Phi(x)$ into the feature space (Al-Anazi *et al.*, 2010). There are several kernel functions, such as a linear function, polynomial function, radial basis function, and multi-layer perception functions, which are introduced by Gunn (1998). Among these, the Radial Basis Function (RBF) is the most popular for use in non-linear classification problems (Sapankevych *et al.*, 2009) and is defined as:

$$k(x, y) = \exp\left(-\gamma \|x - y\|\right)^2 \qquad (4)$$

The parameter γ is the bandwidth of the Gaussian kernel.

The goal is to find an optimal value of b and weight w. When the mapping function $\Phi(x)$ is fixed, two parts should be considered to determine w. One is the flatness of the weights the other is the error generated by the estimation process of the value, also known as the empirical risk (Sapankevych *et al.*, 2009). The value w should be determined by minimizing the sum of empirical risk $R_{emp}(f)$ and complexity $\|w\|^2$:

$$R_{reg}(f) = R_{emp}(f) + \frac{\lambda}{2}\|w\|^2 = C \sum_{i=1}^{N} \Gamma(f(x_i) - y_i) + \lambda \|w\|^2 \qquad (5)$$

Where, C is a pre-specified value, N is the sample size, Γ is a cost function and the scale factor λ is a regularization constant.

SVM can deal with traffic prediction problems in non-linear systems using kernel function by fitting a curve to a data point.

4.3. Artificial Neural Network (ANN)

Non-parametric methods such as Artificial Neural Network (ANN), which has its basis in Artificial Intelligence, have also been used for short-term traffic prediction. Neural network methods were originally motivated by the goal of having machines that can mimic the brain. As stated by Mitchell (1997), "neural network learning methods provide a robust approach to approximating real-valued, discrete-valued and vector-valued target functions". Complex non-linear relationships between multiple inputs and outputs can be modeled by neural networks to capture or learn patterns within data.

A basic framework of a neural network model includes four main elements, namely nodes, connection, layers, and transfer function. Nodes, also known as neurons, are simple processing units. Two interconnected nodes are connected by weighted connections that represent the nature of their interaction. Optimal weights for each connection can be calculated during the training process which is used to calibrate the model using patterns in data. Layers are the topology of a neural network, where nodes and connections are assigned. The transfer function determines the state of each neuron. The mathematical process at a single neuron is shown in Fig. 4. A single neuron includes a set of synapses connected to the inputs. Each of them is characterized by weight. This process includes two steps:

- Calculate a linear combination of inputs; and

- Transfer the weighted sum into output using an activation function.
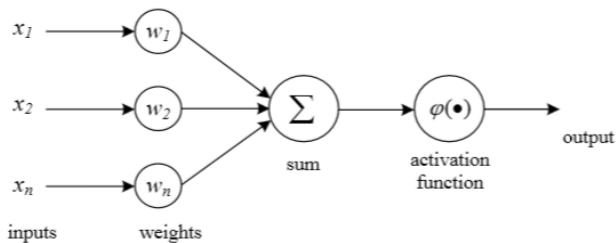

Fig. 4.
Process of a Single Neuron
Source: (Rokach, 2010)

Let x_i be the ith input and w_i be the corresponding weight. The sum of a linear combination of inputs is given by $Z = \Sigma w_i x_i$. Then a nonlinear activation function $\varphi(.)$ is applied to the weighted sum. The output of this neuron is $y = \varphi(.)$. Generally, the most commonly used activation functions include ReLU function, sign functions, piecewise-linear functions, and sigmoid functions.

A variety of different neural network models have been applied to traffic engineering

applications (Dougherty, 1995). Feed-Forward neural network (FFNN) (Kriesel, 2007) is a simple and widely used network for traffic flow prediction. In Fig. 5, the structure of FFNN which has an input layer, some hidden layers, and an output layer. In FFNN, each neuron in one layer strictly feeds forward to the output units of the next layer.
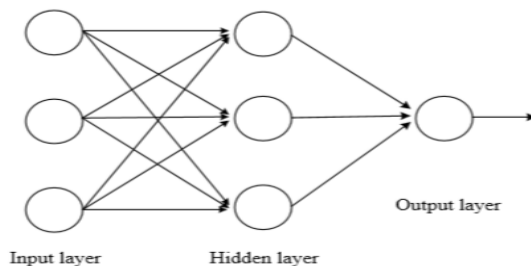

Fig. 5.
General Architectures of The Feed-Forward Network
Source: (Mitchell, 1997)

Huang and Ran (2002) used a FFNN model to forecast future traffic speed using both weather data gathered each hour and traffic speed data gathered every five minutes. In this study back propagation neural network is used for traffic flow prediction considering the weather condition. The advantage of the back-propagation neural network is that it speeds up the training process as well as reduce the error in training data sets compare with FFNN. The back-propagation algorithm searches for weather values that minimize the total error of the network over the set of training examples.

The back-propagation Neural Network algorithm consists of two passes. In the forward pass, the network is activated on one example and the error of (each neuron of) the output layer is computed. And in the backward pass, the network error is used for updating the weights. Starting at the output layer, the error is propagated backward through the network layer by layer. This is done by recursively computing the local gradient of each neuron. Structure of backward propagation neural network is shown in Fig. 6.

The advantage of neural network-based methods is their learning ability in capturing the traffic patterns from large historical traffic datasets. They can make full use of historical traffic datasets to address complex problems where the relationships between data series are not clear and well-defined.

Fig. 6.
Backward Propagation Neural Network Structure

5. Parameters

In this study, a one-hour interval of six months of traffic flow data which is observed from an expressway, Durbin, Ireland. Traffic flow is directly affected by weather conditions as well as time of day and day type. For example, due to snowfall, the speed of vehicles is significantly reduced as well as during heavy rainfall, the speed of vehicles is also reduced.
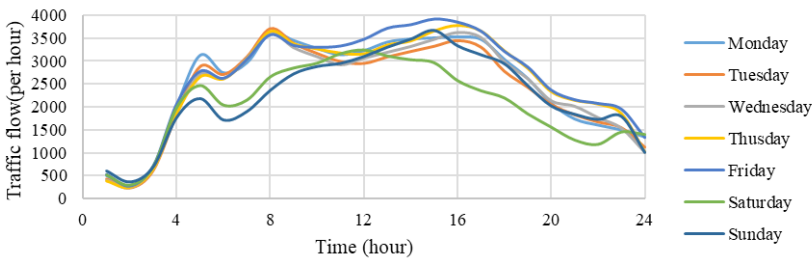

Fig. 7.
Diurnal Variation of Traffic Flow in The Various Day

Fig. 7 indicates the traffic flow variation considering day and day of time. In Saturday traffic flow is comparatively lower than other days after 12.00. Friday traffic flow is comparatively higher than other days throughout the day time.

In the morning time, Sunday traffic is comparatively lower than other days. Fig. 8 indicates the traffic flow variation considering the type of days such as working day, weekend day and national holidays. On holiday and weekend, morning time traffic flow is comparatively lowered then working day.

Traffic flow is influenced by day, time of day and weather condition. In this research, three parameters set are considered for training and prediction purposes to determine the best combination of weather parameters. Day, day type (holiday/weekend/working day) and time are considered as common parameters for all three parameters set. As those parameters have a significant effect on traffic flow variation. Then weather-related parameters raining, sunny, temperature value, precipitation values are considered in this research. Three types of parameter sets are used in this research are shown in Table 1.
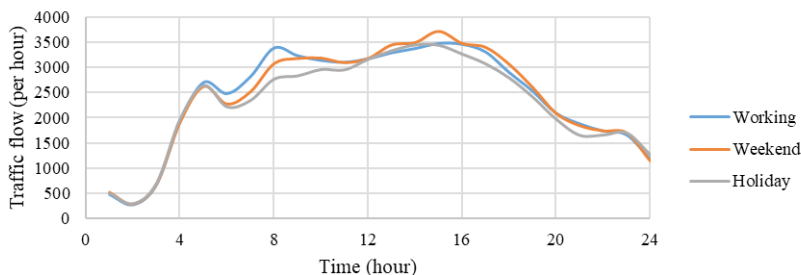
Fig. 8.
Diurnal Variation of Traffic Flow of The Various Type of Day

Table 1

Parameter Sets Used in Prediction

	Parameter set-1	**Parameter set-2**	**Parameter set-3**
Day	Monday/ Tuesday/ Wednesday/ Thursday/ Friday/ Saturday/ Sunday	Monday/ Tuesday/ Wednesday/ Thursday/ Friday/ Saturday/ Sunday	Monday/ Tuesday/ Wednesday/ Thursday/ Friday/ Saturday/ Sunday
Day Type	Working/ Weekend/ Holiday	Working/ Weekend/ Holiday	Working/ Weekend/ Holiday
Clock Time	1/ 2/ 3/ 4/ 5/ 6/ 7/ 8/ 9/ 10/ 11/ 12/ 13/ 14/ 15/ 16/ 17/ 18/ 19/ 20/ 21/ 22/ 23/ 24	1/ 2/ 3/ 4/ 5/ 6/ 7/ 8/ 9/ 10/ 11/ 12/ 13/ 14/ 15/ 16/ 17/ 18/ 19/ 20/ 21/ 22/ 23/ 24	1/ 2/ 3/ 4/ 5/ 6/ 7/ 8/ 9/ 10/ 11/ 12/ 13/ 14/ 15/ 16/ 17/ 18/ 19/ 20/ 21/ 22/ 23/ 24
Weather Condition	*sunny + not rain* *sunny + rain* *not sunny + not rain* *not sunny + rain*	*precipitation value* *temperature value*	*sunny + not rain* *sunny + rain* *not sunny + not rain* *not sunny + rain* *precipitation value* *temperature value*

And in parameter set-1, weather conditions such as sunny and rain, create four combinations (sunny + not rain, sunny + rain, not sunny + not rain, not sunny + rain) to training and prediction purposes. In parameter set-2, precipitation value and temperature value are considered as weather parameters. Combining all those parameters used in set-1 and 2 are used in parameter set-

3. As traffic flow is predicted by considering weather conditions, here, three types of parameter sets are considered and which parameter set is more convenient for accurate prediction with day (Monday/ Tuesday/ Wednesday/ Tuesday/ Friday/ Saturday/ Sunday), day type (Working/ Weekend/ Holiday), day type (Working/ Weekend/ Holiday) and clock time are studied.

6. Results

6.1. K-Nearest Neighbor (KNN)

Considering the closest feature traffic prediction is done in the k-Nearest Neighbor method. The number of k and distance matrices are important for predicting or classification in the KNN method. For calculating distance, the Euclidian distance method is used and the number of k is selected by optimization. Equation 6 shows the Euclidian distance equation.

$$dist^2_{p,q} = (x_{pn} - x_{qn}) +$$
$$+ (x_{q(n-1)} - x_{q(n-1)}) + \cdots + (x_{pn} - x_{qn}) \quad (6)$$

One-hour interval traffic flow data are trained with the base on weather information and temporal. Five months of traffic flow data are used for training purposes to predict the next one-month traffic flow. Training is done individually considering three different parameters set shown in Table 1. Then the prediction is done to find out which parameter set is more accurate to predict traffic flow.

Considering all parameter sets and k-numbers prediction results are shown in Table 2. From the analysis, it is found that for parameter set-1, in K-Nearest Neighbor method gives more accuracy in one-hour interval traffic flow. For parameter set-1, with the variation of the number of k, the prediction result is slightly different with R-square value 0.94 to 0.95 and Mean Average Percentage Error (MAPE) about 14.00.

Table 2

Prediction Result in KNN by Considering Three Parameters Set

K-value	Parameter set-1		Parameter set-2		Parameter set-3	
	R-Square	MAPE	R-Square	MAPE	R-Square	MAPE
5	0.948	14.399	0.852	15.507	0.857	15.714
7	0.948	14.423	0.857	16.071	0.857	15.994
9	0.948	14.384	0.862	16.015	0.862	15.813
15	0.954	14.537	0.859	16.724	0.859	16.500
19	0.955	14.692	0.852	17.421	0.851	17.456
25	0.953	14.535	0.845	18.260	0.848	18.098

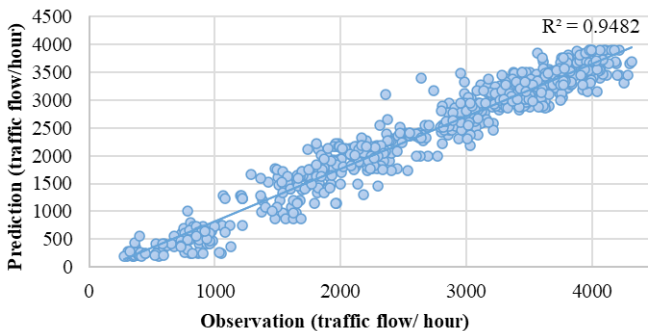**Fig. 9.**

R-Square Value in KNN for Parameter Set-1 (k=9)

For parameter set-1 with 9 k-number R-square value and Mean Average Percentage Error (MAPE) are 0.948 and 14.384 respectively and number R-square value and Mean Average Percentage Error (MAPE) are shown in Fig. 9 and 10 respectively.
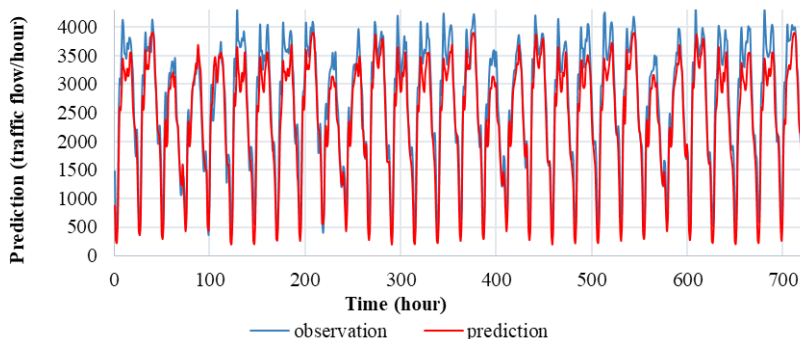

Fig. 10.
Compare One Hour Interval Observation and Prediction Traffic Flow in KNN Considering Parameter set-1 (k=9)

From the analysis, it is found that parameter set when weather information is considered only weather state or condition rather than precipitation and temperature value, prediction accuracy is significantly increase shown in Table 2. It is also found that selecting the number of k has no significant in prediction results.

In summary, it is found by KNN method that:
- Best parameter set: parameter set-1;
- R-square value: 0.948 (k=9);
- Mean Average Percentage Error (MAPE): 14.384.

6.2. Support Vector Machine (SVM)

In Support Vector Machine (SVM), traffic flow prediction is done mainly by classification and pattern recognition with creating hyperplane. Creating a hyperplane kernel such as Radial Basis Function (RBF) is used. The advantage of using kernel is that if n number of parameters is used in the system, the kernel creates n-1 number of dimensions in the system. So that classification is done properly by using kernel.

One-hour interval traffic flow data are trained with the base on weather information and temporal. Five months of traffic flow data are used for training purposes to predict the next one-month traffic flow. Training is done individually considering three different parameters set shown in Table 1. Then the prediction is done to find out which parameter set is more accurate to predict traffic flow.

In prediction result in Table 3, it is found that for parameter set-1, the Support Vector machine gives more accuracy in one hour interval traffic flow prediction with R-square vale 0.859 and Mean Average Percentage Error (MAPE) 19.338. R-square value and Mean Average Percentage Error are shown in Fig. 11 and 12 respectively.

Table 3

Prediction Result in SVM by Considering Three Parameters Set

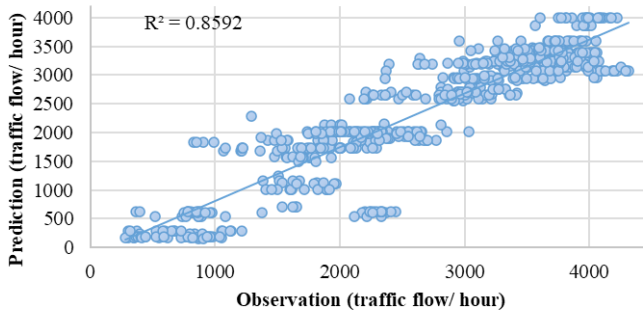|	Parameter set- 1	Parameter set- 2	Parameter set- 3
R-square value	0.859	0.331	0.416
Mean Average Percentage Error (MAPE)	19.338	47.623	42.631

Fig. 11.

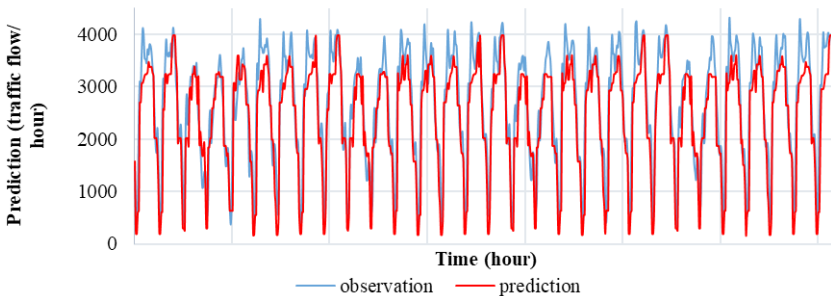*R-Square Value in SVM for Parameter Set-1*

Fig. 12.

Compare One Hour Interval Observation and Prediction Traffic Flow in SVM Considering Parameter Set-1

From the analysis, it is found that parameter set when weather information is considered only weather state or condition rather than precipitation and temperature value, prediction accuracy is significantly increase shown in Table 3.

In summary, it is found by the SVM method that :

- Best parameter set: parameter set-1;
- R-square value: 0.859;
- Mean Average Percentage Error (MAPE): 19.338.

6.3. Artificial Neural Network (ANN)

With the complex relationship between input and output data, the neural network can be used

successfully in the prediction arena. Traffic flow has irregular fluctuation characteristics due to time of day, weather and other factors. Considering the complex nature of traffic flow, the neural network can play a vital rule in traffic flow prediction. Back-propagation neural network algorithm with the number of hidden layers is used here. Hidden layers are used in the algorithm so that it is also called deep learning. In back-propagation, the error is immunized by two step and the ReLU function is used for activation function.

The activation function is used to determine the output of neural network like yes or no. it maps the resulting values in between 0 to 1 or -1 to 1 etc. depending on activation function. The ReLU is the most used activation function in the neural network. The nature of ReLU is shown in Fig. 13, where it is found that the ReLU is half rectified from the bottom. f(z) is zero when z is less than zero and f(z) is equal to z when z is above or equal to zero. Other hand range of ReLU is [0 to infinity).
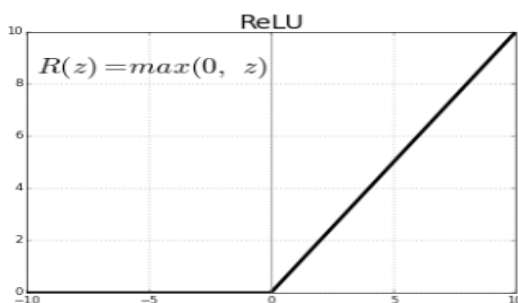

Fig. 13.
ReLU Activation Function

Iteration number and learning rate are significant in the neural network to avoid overtraining. First, find optimum Iteration number and learning rate. Then find suitable hidden layers contain the number of neurons of each hidden layer.

One-hour interval traffic flow data are trained with the base on weather information and temporal. Five months of traffic flow data are used for training purposes to predict the next one-month traffic flow. Training is done individually considering three different parameters set shown in Table 1. Then prediction is done to find out which parameter set is more accurate to predict traffic flow.

Considering Input-10-5-output structure in neural network with the variation of iteration number and learning rate variation of accuracy are shown in Table 4 and 5. And it is shown that when for 500 iteration number and 0.001 learning rate, the prediction results shown more accurately.

Table 4

Effect of Iteration Number on Traffic Flow Prediction Accuracy

Iteration Number	R-Square	MAPE
100	0.153	65.040
500	0.478	28.285
700	0.472	33.593
1000	0.463	32.526
1500	0.473	33.267

Table 5

Effect of Learning Rate on Traffic Flow Prediction Accuracy

Learning Rate	R-Square	MAPE
0.8	0.000	79.764
0.1	0.000	85.756
0.01	0.523	30.606
0.001	0.577	26.427
0.0001	0.478	28.285
0.00001	0.025	50.408

Table 6

Prediction Result in NN by Considering Three Parameters Set

Neural per Hidden Layer	Parameter set-1		Parameter set-2		Parameter set-3	
	R-Square	MAPE	R-Square	MAPE	R-Square	MAPE
10-5	0.712	23.367	0.579	28.240	0.376	31.399
100-50-20-10-5	0.583	25.205	0.582	27.608	0.499	30.418
100-100-100-100	0.880	15.665	0.572	26.009	0.594	24.542
500-200-100-50-10	0.816	21.737	0.615	25.084	0.649	24.159
10-10-10-10-10	0.830	20.758	0.456	32.663	0.650	23.842

After fixing the iteration number and learning rate, then focus on try to find suitable for the number of hidden layers in the neural network structure. For this reason, traffic flow is predicted by considering three types of parameter sets with the variation of the number of layers and number of neurons each layer and results are shown in Table 6. From the analysis, it is found that for parameter set-1, in neural network method gives more accuracy in one-hour interval traffic flow. It is found that when considering input-100-100-100-100-output network structure in the neural network which means that consider four hidden layers and each hidden layer has a hundred neuron and for this structure, prediction results are more accurate. Considering parameter set-1 and input-100-100-100-100-output neural network structure, R-square value and Mean Average Percentage Error (MAPE) are 0.948 and 14.384 respectively and number R-square value and Mean Average Percentage Error (MAPE) are shown in Fig. 14 and 15 respectively.
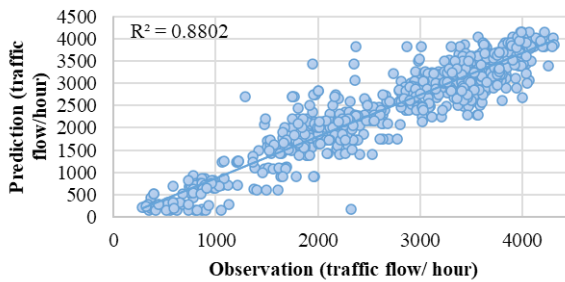
Fig. 14.
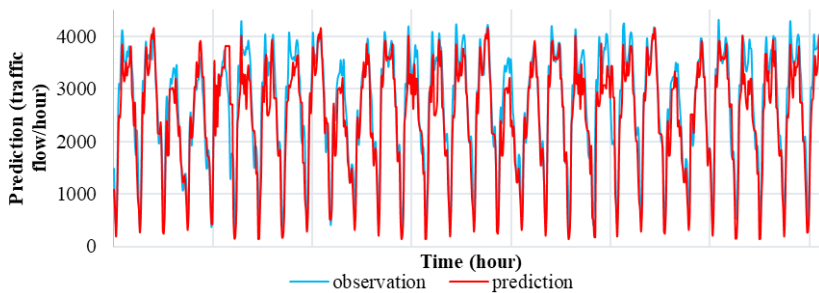*R-Square Value in NN for Parameter Set-1*

Fig. 15.
Compare One Hour Interval Observation and Prediction Traffic Flow in NN Considering Parameter Set-1

From the analysis, it is found that parameter set when weather information is considered only weather state or condition rather than precipitation and temperature value, prediction accuracy is significantly increase shown in Table 6. It is also found that the number of iteration and number of hidden layers with the number of neurons of each hidden layer has a significant effect on prediction accuracy.

In summary, it is found by the ANN method that:

- Best parameter set: parameter set-1;
- R-square value: 0.880;
- Mean Average Percentage Error (MAPE): 15.665;
- With considering neural network structure input-100-100-100-100-output with 500 iteration number and 0.001 learning rate.

6.4. Comparisons and Discussion

In machine learning, the selection of parameters is an important factor in traffic flow prediction. This study answers to select the preface parameter set for accurate prediction. It is found that weather information like as sunny and raining state

information gives more accurate prediction rather than considering precipitation value and temperature value. For parameter set-1, we get better accuracy for all three machine learning algorithms to compare to other parameter sets. From the analysis, it is also found that KNN gives the most accurate result than SVM and ANN considering the best fundamental component. The results are shown in Table 7.

Table 7

Compression of Prediction Result of Different Machine Learning Algorithm

	KNN	SVM	ANN
Best Parameter Set	Set-1	Set-1	Set-1
R-Square Value	0.948	0.859	0.880
MAPE	14.348	19.338	15.665

It is found that our algorithms can reflect the traffic flow trend considering weather conditions. Sometimes it is found that peak the point of traffic flow distribution cannot predict perfectly. It is happened due to a shortage of training data. In this stud, five months of traffic flow data are considered for training purposes. It can be overcome by increasing the number of training data. Traffic flow has dynamic property, it can be fluctuated by other parameters like traffic jams, accidents, etc. rather than weather conditions. This study only focuses on considering the parameters of weather information. Not considering other factors like accident and traffic jam density. Those factors are also responsible for accuracy. As weather information is easily available in the real world. Therefore, only weather information is considered for traffic flow prediction in KNN, SVM and ANN machine learning algorithms. Integrated weather information in traffic prediction by machine learning will be useful tools for predicting traffic conditions in normal conditions as well as adverse weather conditions.

due to the dynamic property of traffic flow. Researches are trying to incorporating real-world information for traffic flow prediction. This study can contribute to parameter selection related to weather information. Also, it opens the door to use other parameters like an accident, jam density for traffic flow prediction. Nowadays real-world data are easy to collect and machine learning acquires a vital place in the transportation arena. Researches are also trying to improve the machine learning algorithms to get an accurate result. This study answers to find the best fundamental components of each machine learning algorithms and their effect on accuracy. KNN gives a more accurate result with 14.348 mean average percentage error and 0.948 R-square value than SVM and ANN considering weather state-related parameters rather than numeric value of weather.

This study can applicable in traffic management, logistic support of transportation agencies. It is also applicable in the field of autonomous vehicle management.

7. Conclusions

Short term traffic flow prediction is an important arena in the transportation field

References

Al-Anazi, A.; Gates, I. 2010. Support vector regression for porosity prediction in a heterogeneous reservoir:

A comparative study, *Computers & geosciences* 36(12): 1494-1503.

Bajaw, S.; Chung, E.; Kuwahara, M. 2003. A Travel Time Prediction Method Based on Pattern Matching Technique, *Presented at 11th Conference of the Road Engineering Association of Asia and Australasia and the 21st Conference of the Australian Road Research Board.* Cairns, Australia.

Cao, L. J.; Tay, F. E. H. 2003. Support vector machine with adaptive parameters in financial time series forecasting, *IEEE Transactions on Neural Networks* 14(6): 1506–1518.

Clark, S. 2003. Traffic prediction using multivariate nonparametric regression, *Journal of Transportation Engineering* 129(2): 161– 168.

Datla, S; Sharma, S. 2008. Impact of cold and snow on temporal and spatial variations of highway traffic, *Journal of Transport Geography* 16(5): 358-372.

Davis, G. A.; Nihan, N. L. 1991. Non parametric regression and short-term freeway traffic forecasting, *Journal of Transportation Engineering-ASCE* 117(2): 178-188.

Devijver, P. A.; Kittler, J. 1982. *Pattern Recognition: A statistical approach.* Prentice Hall International, ISBN 0136542360.

Dougherty, M. 1995. A review of neural networks applied to transport, *Transportation Research Part C: Emerging Technologies* 3(4): 247-260.

Gunn, S. R. 1998. *Support vector machines for classification and regression*. Technical report, School of Electronics and Computer Science, University of Southampton.

Hasnat, A.; Rahman, F. I. 2019. Short term traffic flow forecasting by time series analysis. In *Proceedings of International Conference on Planning, Architecture and Civil Engineering,* Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh.

Hranac, R.; Sterzin, E.; Krechmer, D.; Rakha, H.; Farzaneh, M. 2006. *Empirical studies on traffic flow in inclement weather.* FHWA-HOP-07-073, Federal Highway Administration, USA.

Hua, J.; Faghri, A. 1994. Applications of artificial neural networks to intelligent vehicle-highway system, *Transportation Research Record* 1453: 83-90.

Huang, S.; Ran, B. 2002. An application of neural network on traffic speed prediction under adverse weather condition. In *Proceedings of the Transportation Research Board 82nd Annual Meeting,* Washington D.C, USA.

Kriesel, D. 2007. A brief introduction to neural networks. Available from internet: <http://www.dkriesel.com>. [Accessed 28/4/2019].

Maze, T.; Agarwai, M.; Burchett, G. 2006. Whether matters to traffic demand, traffic safety, and traffic operations and flow, *Transportation Research Record* 1948: 170-176.

Mitchell, T. M. 1997. *Machine Learning.* New York, McGraw-Hill, ISBN: 0071154671.

Müller, K. R.; Smola, A. J.; Rätsch, G.; Schölkopf, B.; Kohlmorgen, J.; Vapnik, V. 1997. *Predicting time series with support vector machines*. In: Gerstner, W.; Germond, A.; Hasler, M.; Nicoud, J.-D. (eds.) Artificial Neural Networks — ICANN'97, Springer, Berlin Heidelberg, 999-1004.

Nagy, A. M.; Simon, V. 2018. Survey on traffic prediction in smart cities, *Pervasive and Mobile Computing* 50: 148-163.

Rahman, F. I.; Hasnat, A.; Lisa, A. A. 2019. Traffic flow prediction by incorporating weather information in Naïve Bayes Classifier, *Journal of Advanced Civil Engineering Practice and Research* 8: 1-16.

Robinson, S. 2005. *The development and application of an urban link travel time model using data derived from inductive loop detectors*. PhD Thesis. Centre for Transport Studies, Imperial College London.

Robinson, S.; Polak, J. W. 2005. Modeling urban link travel time with inductive loop detector data by using the k-NN method, *Transportation Research Record* 1935(1): 47–56.

Rokach, L. 2010, *Pattern Classification Using Ensemble Methods*. World Scientific Publishing Company Incorporated, ISBN: 9814271063.

Sapankevych, N. L.; Sankar, R. 2009. Time series prediction using support vector machines: A survey, *IEEE Computational Intelligence Magazine* 4(2): 24-38.

Smith, B. L.; Williams, B. M.; Oswald, R. K. 2002. Comparison of parametric and nonparametric models for traffic flow forecasting, *Transportation Research Part C: Emerging Technologies* 10(4): 303-321.

TII. 2019. Transport Infrastructure Ireland Available from internet: <https://www.tii.ie/>. [Accessed 2 January 2019].

Van Lint, J. W. C. 2004. *Reliable travel time prediction for freeways*. PhD Thesis. Delft University of Technology, Delft, Netherlands.

Vapnik, V. 1995. *The Nature of Statistical Learning Theory.* Springer-Verlag New York, ISBN 978-1-4419-3160-3.

Vlahogianni, E. I. 2009. Enhancing predictions in signalized arterials with information on short-term traffic flow dynamics, *Journal of Intelligent Transportation Systems* 13(2): 73-84.

Vlahogianni, E. I.; Karlaftis, M.G.; Golia, J.C. 2014. Short-term traffic forecasting: where we are and where we're going, *Transportation Research Part C* 43(1): 3-19.

Yang, J. H.; Yu, X. N. 2009. The support vector machines prediction model of freeway dynamic traffic flow, *Journal of Xi'an Technological University* 3: 280–284.

Yu, B.; Lam, W. H. K.; Tam, M. L. 2011. Bus arrival time prediction at bus stop with multiple routes, *Transportation Research C* 19(6): 1157–1170.