

# Use of LSTM for Short-Term and Long-Term Travel Time Prediction

Irem Islek

Department Of Computer Engineering  
Istanbul Technical University  
Istanbul, Turkey  
isleki@itu.edu.tr

Sule Gunduz Oguducu

Department Of Computer Engineering  
Istanbul Technical University  
Istanbul, Turkey  
sgunduz@itu.edu.tr

## Abstract

Travel time prediction is an important component in intelligent transportation systems, and plays a key role in daily life. Predicting travel time for a trip is quite challenging and has been studied by many researcher. However, most of the studies focus on short term travel time prediction. In this study, LSTM (Long-Short Term Memory) neural network models are constructed to predict travel time for both long term and short term using real world data of New York city. Results of this study show that, LSTM provides satisfying results for long term travel time prediction as well as short term.

## 1 Introduction

Traffic is a common problem of urban life and ITS (Intelligent Transportation System) [1] which is an integrated system of different IoT (Internet Of Things) data sensors, cameras, computers can provide a solution for this problem. One of the most challenging part of ITS is travel time prediction because travel time is affected by numerous factors such as day of the week, time of the day, weather conditions, road conditions etc. Predicting travel time accurately for a future trip can help people to plan their route more efficiently.

In recent years, there has been an increasing interest in travel time prediction. For this reason, many researchers focus on travel time prediction. In some

of these studies, time series models such as traditional ARIMA [2] or seasonal ARIMA [3] are applied for prediction. These models use historical travel time data to fit the model and next step travel time is predicted using the fitted time series. Another common approach for travel time prediction is applying  $k$ -nearest neighbors model [4, 5]. In this model, most similar historical  $k$  days are found by the model and the mean of these travel time values is accepted as the prediction result. In addition to these models, Kalman filtering [6, 7], Support Vector Regression [8], Support Vector Machines [9], Bayesian combination of models [10] are also applied for travel time prediction. These studies use current values of several features such as speed, traffic flow, weather condition to predict future travel time.

In the forthcoming years, neural networks come in the use for travel time prediction. There are several researches which use artificial neural network models for this problem [11, 12, 13, 14]. In addition to that, some deep learning models are also applied for travel time prediction such as Deep Belief Network [15].

LSTM is also used for travel time prediction [16]. In this study, LSTM model is applied for only short term travel time prediction. Travel time data of this study is obtained from Highways England. They emphasize that deep learning models which take into account the sequence relation are quite promising in travel time prediction domain.

Although a research has been carried out on short term travel time prediction using LSTM, no single study exists which uses LSTM for both short term and long term, in details. Another contribution of this paper is that, this is the first time LSTM is applied for long term travel time prediction problem, to the best of our knowledge. As it can be guessed, long term travel time prediction is quite difficult than short term travel time prediction. In long term travel time prediction,

the main objective is predicting travel time value for a specific day and hour, at least one week ago. In summary, we can say that main objectives of this study are applying LSTM models for long term and short term travel time prediction using real world data of New York.

The paper is organized as follows. Section 2 presents the approach and details of the dataset. Section 3 describes the experiments and the results obtained. Finally, in Section 4 we conclude and discuss some possible future works.

## 2 Prediction Of Travel Time

In this study, we aim to predict travel time for short term and long term using a real world travel time data. Several different tests are done to investigate behavior of LSTM for travel time prediction problem. In this section, we first describe the data set used and then give the details of the method we applied to solve the problem of travel time prediction.

### 2.1 Dataset

The data set used in this study is obtained from New York City Department Of Transportation (NYC DOT) which provides real-time traffic information from major arterials and highways within New York city using numerous IoT sensors [17]. These IoT sensors are distributed within the five boroughs of New York City: Brooklyn, Bronx, Manhattan, Queens, Staten Island. Using a free service, this information can be collected by users for use in application development. In this dataset, "link" represents a given street section.

Rows of obtained data contain these fields: id, speed, travel time, data as of, link points, borough and link name. Definition of these fields can be seen in Table 1. The data are updated every five minutes and contain 135 different links within the five boroughs of New York City. The dataset contains real time traffic data of links from April 2015 up to the current date but it is updated regularly by NYC DOT. Data row count of each month is nearly 1,150,000. For each link, there are nearly 275 rows (travel time values) in one day. The time intervals between each travel time values of a link are 5 minutes.

In this study, "DataAsOf" column which represents timestamp of the row and "TravelTime" column which represents the travel time value of the link are used for time series modeling. For each link, travel time values are ordered by "DataAsOf" column which represents date and time of the measurement.

### 2.2 LSTM For Travel Time Prediction

Our experimental system design consists of four steps. The first step is Data Preparation step. Afterwards,

Table 1: Definition Of Fields

Field Name	Definition
Speed	Average speed a vehicle traveled between end points on the link in the most recent interval
TravelTime	Time the average vehicle took to traverse the link
DataAsOf	The time of data was received from link
LinkId	unique id of the link (a given street section)
LinkPoints	Sequence of Lat/ Long points, describes locations of the sensor links Google compatible polyline from
Borough	NYC borough (Brooklyn, Bronx, Manhattan, Queens, Staten Island)
LinkName	Description of the link location and end points

outlier elimination step helps us to discard the out of ordinary data rows. Then, LSTM model is applied to the dataset in Prediction Model step. At the end of the workflow, the Evaluation step calculates the error rates based on selected evaluation metric. The details of the whole process is given below.

#### 2.2.1 Data Preparation Step

The aim of this study is predicting travel time for a selected link (a given street section) using previous travel time data. In order to achieve this aim, a time series is constructed using previous travel time information, for each link. This dataset contains travel time values of each link for every 5 minutes. As can be seen in Fig. 1, for predicting travel time of 5 minutes later using this dataset, the link's previous travel time values measured at 5 minute intervals should be used. On the other hand, for predicting travel time of 15 minutes later using this dataset, previous travel time values measured at 15 minute intervals should be used. In case of long term travel time prediction, for predicting the travel time value of a link at 8 o'clock on Monday, the data set should be generated using travel time values of previous weeks' Monday at 8 o'clock.

Also, data preparation step transforms data rows to a sequential structure. Using sequential output of this step, a time series forecasting model can be applied easily for predicting next step travel time value.

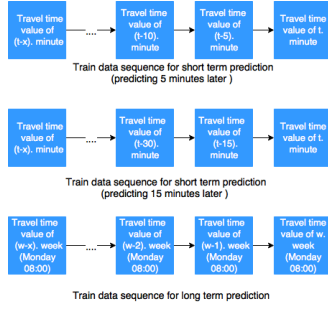


Figure 1: Datasets for Short and Long Term Prediction

### 2.2.2 Outlier Elimination Step

Traffic prediction is quite difficult because of the fact that traffic is affected by numerous different factors. Some of these factors can be handled with special solutions but some other factors such as weather conditions, special events, traffic accidents can be only handled using additional information. Because of the fact that we do not use additional dataset, these special situations are eliminated using an outlier detection method.

In this study, outliers are detected using the standard deviation of the last  $N$  travel time values. In our case,  $N$  is selected as 4. If a travel time value is quite different from the previous 4 travel time values, it is considered to be an outlier in our approach. According to this moving average approach, the outliers are eliminated from the dataset.

### 2.2.3 Prediction Step

In the prediction step of the methodology, LSTM (Long-Short Term Memories) [18] neural network is selected as prediction model of time series. LSTM network which is a special version of RNN (Recurrent Neural Networks), has chain of repeating modules in just the same way as RNN. In LSTM networks, each module which is called memory cell contains 3 different gates: an input gate, an output gate and a forget gate. This memory cell is in the hidden layer of the LSTM network. A figure of LSTM memory cell can be seen in Fig 2.

The basic job of a cell is remembering the temporal state of the network. The input gate, output gate and forget gate can be thought as a neuron of a neural network. The input gate is responsible for deciding whether a new value flows into the memory, or not. On the other hand, the output gate decides whether the memory cell state is going to have an effect on other neurons or not. The forget gate allows the cell to remember previous state or forget it.

In a given time  $t$ , the input of cell is

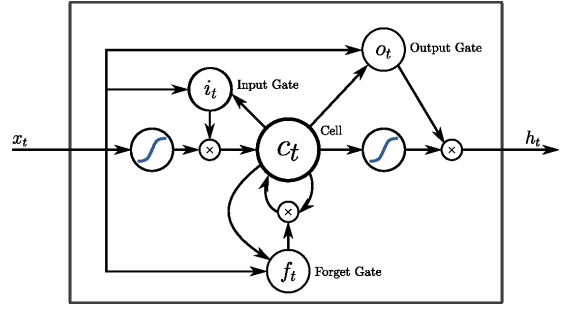


Figure 2: Long-Short Term Memories

$x_t$ ,  $h_t$  shows the output of the gate and  $W_i, W_f, W_c, W_o, U_i, U_f, U_c, U_o, V_o$  correspond to weight matrices.  $b_i, b_f, b_c, b_o$  are bias vectors in the model equations of LSTM.

First of all, input gate values ( $i_t$ ) and possible memory cell state values  $\widetilde{C}_t$  are calculated using Eq. 1 and Eq. 2.

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (1)$$

$$\widetilde{C}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (2)$$

After that, the forget gate activation can be seen in Eq. 3.

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (3)$$

Using the  $\widetilde{C}_t$  (candidate state value),  $i_t$  (input gate value) and  $f_t$  (forget gate value),  $C_t$  which is the memory cell value can be calculated using Eq. 4.

$$C_t = i_t * \widetilde{C}_t + f_t * C_{t-1} \quad (4)$$

After state of the memory cell is calculated, the value of output gate  $o_t$  and output  $h_t$  can be calculated using Eq. 5 and Eq. 6.

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + V_o C_t + b_o) \quad (5)$$

$$h_t = o_t \tanh(C_t) \quad (6)$$

LSTM networks can be applied to several different problems which contain prediction on sequential data. In this study, LSTM is applied for travel time prediction which is a time series prediction problem. Chain like structure of LSTM network which is used for this study, can be seen in Fig. 3. Reason of choosing LSTM as travel time prediction model is that LSTM neural network model is a special kind of neural network which considers sequential relation.

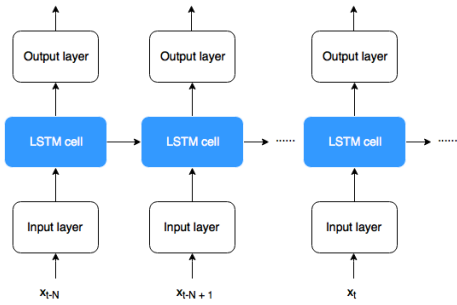


Figure 3: LSTM Time Series Model

### 2.2.4 Evaluation Step

The error rates of the models are estimated using Mean Average Percentage Error (MAPE). In Eq. 7 which shows calculation formula of MAPE,  $A_t$  means actual value and  $F_t$  means forecasting value.

Comparison of all experiments can be seen in the next section, Experiments and Results.

$$MAPE = \frac{100}{n} \sum_{t=1}^n \frac{|A_t - F_t|}{|A_t|} \quad (7)$$

## 3 Experiments and Results

In this section, we describe the details of our experiments and present the results obtained by our experiments for travel time prediction.

In the first group of experiments, LSTM network is applied to long term travel time prediction problem. In long term travel time prediction, the aim is predicting travel time for a specific day and specific time of a week. In long term travel time prediction approach, the datasets are constructed using travel time values on the same day and time of previous weeks on a particular day's previous weeks' travel time values. For instance, if we want to predict travel time value at 2 p.m. on Wednesday, we construct the related time series dataset using travel time values on the same day and time of previous weeks. Afterwards, the outlier values are eliminated from the datasets because of the fact that these outlier values originate from unusual situations such as traffic accident, bad weather conditions, special occasions, etc. These cases cannot be handled without any extra information from other resources; therefore outliers are ignored during time series construction.

The long term travel time series datasets are constructed separately for each link, day and hour. For each link, there are 2016 different test instances which corresponds to data of one week. For each test instances, related previous travel time values are selected for training the related model. There 153 different links in the dataset.

In our LSTM experiments, there are a visible input layer, a visible output layer and 3 hidden layers with 4 LSTM neurons. Adam algorithm is used for stochastic optimization [19]. The batch size is 1 in our experiments.

In order to compare performance of the LSTM neural network on long term travel time prediction, we conduct the a prediction model using ARIMA. The results of these experiments can be seen in Table II.

Table 2: Long Term Travel Time Prediction Model

Model	Error Rate % (MAPE)
ARIMA	20.9%
LSTM	11.2%

According to our tests, it can be said that LSTM provides satisfying results for long term travel time prediction. Reason of this situation is that, LSTM is quite suitable for prediction on sequential data.

Rest of the experiments are focus on short term travel time prediction. In short term travel time prediction case, we complete 2 different test cases: predicting 5 minutes later and predicting 15 minutes later.

For predicting travel time value of 5 minutes later, the train datasets obtained using travel times of every 5th minutes for previous 2 hours. For each link, this experiment is done 12 times. You can see the comparison of LSTM and ARIMA models for predicting travel time of 5 minutes later, in Table III.

As we mentioned before, there is another paper which applies LSTM for short term travel time prediction [16]. In their study, travel time value of 15 minutes later is predicted. In order to apply same tests for our dataset, train datasets are obtained using travel times of every 15th minutes for previous 2 hours. For each link, this experiment is repeated 12 times. In addition to LSTM network model, ARIMA model is also applied for predicting travel time of 15 minutes later. The results of these tests can be seen in Table III.

Table 3: Short Term Travel Time Prediction Model

Test case	Model	Error Rate (MAPE)
5 minutes later	ARIMA	19.4%
5 minutes later	LSTM	9.8%
15 minutes later	ARIMA	23.2%
15 minutes later	LSTM	12.7%

It is apparent from Table III that LSTM models provide lower error rates than ARIMA models for both cases. In addition to that, it can be said that predicting 15 minutes later with LSTM network gives higher

error rates than predicting 5 minutes later with LSTM network. Reason of this situation is that, previous 5 minutes gives more accurate information about travel time value of 5 minutes later.

Finally, some multi-step ahead predictions are performed using LSTM and obtained results are shown in Table IV.

From the Table IV, it can be seen that the lowest error rates are provided by 1 step ahead predictions. As the number of steps increases, the error rates are increases for both prediction models (model for 5 minutes step interval and model for 15 minutes step interval). What is interesting in Table IV is that 3 step ahead prediction with 5 minutes interval model and 1 step ahead prediction with 15 minutes interval model are both aim to predict 15 minutes later. 1 step ahead prediction with 15 minutes interval model gives lower error rate than the other. Because of the fact that multi-step ahead prediction uses its own prediction results for predicting the next step, errors of each step are cumulates.

Table 4: Short Term Travel Time Prediction - Multi step Ahead LSTM

Step ahead	Step interval	Error Rate (MAPE)
1	5 minutes	9.8%
2	5 minutes	16.2%
3	5 minutes	21.4%
1	15 minutes	12.7%
2	15 minutes	18.4%
3	15 minutes	22.3%

## 4 Conclusion

The main goal of the current study was to investigate the performance of LSTM models for both short term and long term travel time prediction. As far as we know, this is the first time LSTM is applied for long term travel time prediction and obtained results show that it provides satisfying results for this problem. In addition to that, LSTM network model is applied for short term prediction and it is shown that LSTM model for predicting 5 minutes later overcomes LSTM model for predicting 15 minutes later. Also, multi-step ahead prediction performances are measured for short term predictions. Evaluation results show that the 1-step ahead predictions give better results than multi-step ahead prediction. Taken together, these results suggest that LSTM network model is quite suitable for travel time prediction problem. As a feature work, we are going to improve these results by involving other data resources to predict outlier situations

such as bad weather conditions, traffic jam which is originated from social events, etc.

## References

- [1] F.-Y. Wang, "Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 11, no. 3, pp.630-638, 2010.
- [2] D. Billings, J.S. Yang, "Application of the arima models to urban roadway travel time prediction-a case study," in *Systems, Man and Cybernetics, 2006. SMC06. IEEE International Conference on*, 2006, pp.2529-2534.
- [3] A. Guin, "Travel time prediction using a seasonal autoregressive integrated moving average time series model," in *IEEE International Conference on Intelligent Transportation Systems*, pp. 493-498, 2006.
- [4] J. Myung, D. K. Kim, S. Y. Kho, C. H. Park, "Travel time prediction using k nearest neighbor method with combined data from vehicle detector system and automatic toll collection system," *Transportation Research Record: Journal of the Transportation Research Board*, pp. 51-59, 2011.
- [5] B. Bustillos, Y. C. Chiu, "Real-time freeway-experienced travel time prediction using N-curve and k nearest neighbor methods," *Transportation Research Record: Journal of the Transportation Research Board*, pp. 127-137, 2011. ISO 690
- [6] J. S. Yang, "Travel time prediction using the GPS test vehicle and Kalman filtering techniques," *In American Control Conference*, pp. 2128-2133, 2005.
- [7] Y. Yuan, J. Van Lint, R. E. Wilson, V. Wageningen-Kessels, S. P. Hoogendoorn et al., "Real-time lagrangian traffic state estimator for freeways," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 59-70, Mar 2012.
- [8] C. H. Wu, J. M. Ho, D. T. Lee, "Travel-time prediction with support vector regression," *IEEE transactions on intelligent transportation systems*, vol. 5, no.4, pp. 276-281, 2004.

- [9] L. Vanajakshi, L. R. Rilett, "Support vector machine technique for the short term prediction of travel time," in *Intelligent Vehicles Symposium, 2007 IEEE*, 2007, pp. 600-605.
- [10] C. P. I. J. Van Hinsbergen, J. van Lint, "Bayesian combination of travel time prediction models," *Transportation Research Record: Journal of the Transportation Research Board*, pp. 73-80, 2008.
- [11] D. Park, L. R. Rilett, G. Han, "Spectral basis neural networks for real-time travel time forecasting," *Journal of Transportation Engineering*, vol. 125, no. 6, pp. 515-523, 1999.
- [12] J. W. C. Van Lint, S. P. Hoogendoorn, H. J. van Zuylen, "Accurate freeway travel time prediction with state-space neural networks under missing data," *Transportation Research Part C: Emerging Technologies*, vol. 13, no. 5, pp. 347-369, 2005.
- [13] J. Van Lint, S. Hoogendoorn, H. Van Zuylen, "Freeway travel time prediction with state-space neural networks: modeling state-space dynamics with recurrent neural networks," *Transportation Research Record: Journal of the Transportation Research Board*, pp. 30-39, 2002.
- [14] D. Park, L. R. Rilett, G. Han, "Travel-time prediction with support vector regression," *IEEE transactions on intelligent transportation systems*, vol. 5, no.4, pp. 276-281, 2004.
- [15] C. Siripanpornchana, S. Panichpapiboon, P. Chaovalit, "Travel-time prediction with deep learning," in *Region 10 Conference (TENCON)*, pp. 1859-1862, 2016.
- [16] Y. Duan, Y. Lv, F.Y. Wang, "Travel time prediction with LSTM neural network," in *Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on*, 2016, pp. 1053-1058.
- [17] City of New York Department of Transportation, <https://data.cityofnewyork.us/Transportation/Real-Time-Traffic-Speed-Data/xsat-x5sa/data>
- [18] S. Hochreiter, J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no.8, pp. 1735-1780, 1997.
- [19] D. P. Kingma, J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.