Transport technic and technology
Volume XX Issue 2 Year 2024
DOI: 10.2478/ttt-2024-0005

ROAD TRAFFIC PREDICTION SYSTEM USING DEEP NEURAL MULTILAYER PERCEPTRON (MLP)

Abstract. The development of a road traffic model has numerous benefits, including the reduction of congestion in cities and precise estimated travel times. This study proposes a travel time prediction model based on a deep neural multilayer perceptron trained on GPS data collected in the city of Bechar, Algeria. GPS data was collected using an Android application and a learning model was built using Artificial Intelligence and deep learning techniques. The model was integrated into a web application that allows users to choose a route and make traffic and time predictions. Tests conducted in comparison with Google Maps and real travel experiences in Bechar showed promising results with high accuracy values for most tested routes. However, the accuracy was low in some tests due to the short duration of the data collection process. To improve the project, a data collection phase should be renewed to obtain high accuracy for all destinations and time periods.
Key words: Artificial Intelligence, Traffic Congestion, Deep Learning, Neural Networks, Intelligent Transportation System, Travel Time Estimation, Machine Learning.

Youssef ELMIR[1]
[1]Laboratoire LITAN, Ecole supérieure en Sciences et Technologies de l'Informatique et du Numérique RN 75, Amizour 06300, Bejaia, Algérie; elmir@estin.dz
Hassen ELMIR[2]
[2]University Tahri Mohamed of Bechar Bechar, Algeria
Younes KHAIR[3]
[3]SGRE-Lab, University Tahri Mohamed of Bechar Bechar, Algeria

Introduction

In our modern era, time has become one of the most precious commodities, as aptly stated by Benjamin Franklin that "time is money" (Franklin, 1748). The issue of losing time, especially during city commuting, is a significant concern for people. As reported by Echourouk Online, Algerians lose an estimated three (03) hours per day in traffic congestion, equivalent to 45 days per year. This staggering statistic serves as a compelling motivation to search for solutions.

To better understand the gravity of this problem, a survey was conducted among citizens in the city of Bechar. Here's the corrected version:

A total of 78 participants were surveyed, and their responses are shown in Fig. 1, Fig. 2, and Fig. 3. When asked about their perception of traffic congestion, the results were as presented. "Do you arrive at your appointments on time? If the answer is 'No,' please mention the reason for the delay." In this survey, we investigated the correlation between travel time and traffic congestion, and their impact on intelligent transportation systems (ITS). Furthermore, we propose an intelligent model using deep learning and simple multilayer neural network to predict travel time.

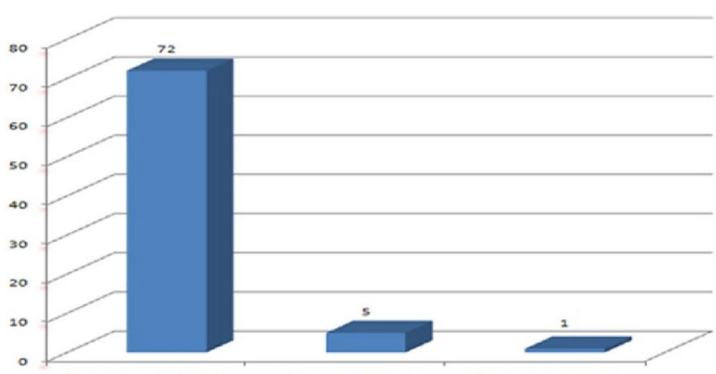

Fig. 1. Histogram of the answers for the question about traffic congestion

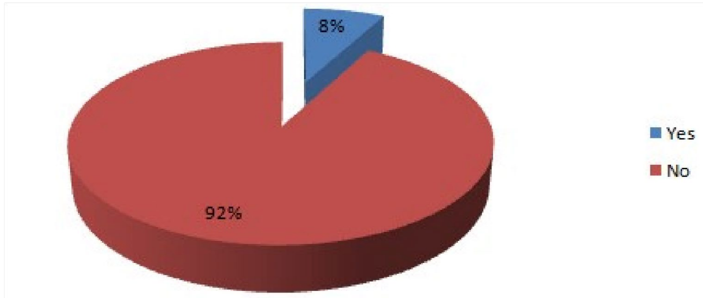

Fig. 2. Histogram of the answers for the question of appointments

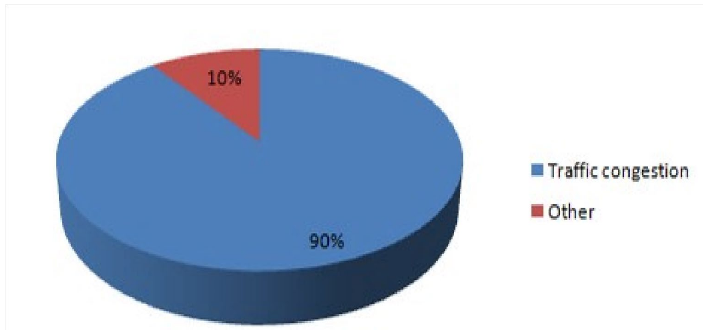

Fig. 3. Histogram of the answers for the question about the reason of delay

To address this problem, it is crucial to estimate travel time to empower individuals to better manage their movements based on time predictions. However, before proposing solutions, a comprehensive study is needed to identify the main criteria that affect travel time estimation, and as widely acknowledged, traffic congestion is one of the key factors. Based on scientific methods, potential solutions can be suggested after conducting experiments and analysing the results.

The structure of the rest of this paper is as follows:

In Section 2, the state of the art is presented, highlighting previous works related to the topic.

In Section 3, the design of the proposed idea is described.

The last section provides detailed information about the design and implementation process of the proposed model, which includes three phases: data collection, visualization and filtering of collected data, training the neural network for time prediction, and linking it with a web application. Additionally, a testing phase using real data and live experiments is conducted to evaluate the accuracy and usability of the proposed solution.

1. State of the Art

The integration of GPS technology with road mapping for traffic data collection and analysis has seen significant advancements.

Obuhuma et l. (Obuhuma, J. I. and Moturi, C. A., NOVEMBER 2012) adopted the use of GPS technology incorporated with road mapping focused on traffic data in order to collect data and analyse traffic scenarios on the roads. The adopted system has an additional ability to trigger alerts on speeding vehicles. The study was conducted in 2 major cities, Nairobi, Kenya. The results showed that there is a great need for real-time traffic information analysis. The researchers concluded that the application of GPS on traffic analysis was proven to be the most effective compared to other solutions, and it also provided a detailed study of traffic conditions.

Hoong et al. (Hoong, 2012), presented an approach to provide potentially unlimited coverage of road traffic conditions by using Bayesian Networks. The researchers trained and tested the prediction model using a dataset comprising traffic flow data collected in London for 6 months. As a conclusion, the authors identified the variables that can infer traffic conditions, such as accidents, weather, road work, roadblock, and speed of vehicles. In addition to this, the model showed promising results, but the accuracy of road traffic prediction was much lower. They intend to improve the BN model for more accurate results in the future.

Garg et al. (Neha Garg, APRIL 2017), developed a model to predict bus arrival time. This information will really help people make travel decisions. The researchers developed a 3-module GPS-based system in order to collect data. The first module was the bus module, which is an android application installed on the driver's phone, and it sends bus coordinates every 6 seconds. The second module is the passenger module, which gets the coordinates of the passenger who wants to track the location of the bus, and the third module is the server module where most of the processing is done. Finally, the system was successfully realized, and the system tracks the bus location using client-server technology. Specific bus location details are provided to the user with the bus number to easily make their decision. The implementation of this idea needs many resources, and in addition to this, the user can track the bus location easily, but they will not get information about road traffic. Therefore, the time that the user will wait until the bus arrival will not be determined correctly. Another thing is that the bus driver always needs to start the application every time they want to begin their working journey, which raises the question of what if the bus driver forgets to launch the application?

Kyaw et al. (Thura Kyaw, MARCH 2020), developed a model to predict traffic congestion of urban road networks by using machine learning with the assistance of public transport GPS data collected over 18 months in Yangon, Myanmar. The model was realized by getting the data from the buses by installing GPS trackers, building the road network, and creating a model to estimate travel speed. The next step was constructing the traffic congestion prediction model using decision trees, Random Forest classifiers, and Extra Tree classifiers algorithms, relying on the data collected and travel speed estimation. The tree-based algorithms were well used, and the researchers obtained an error ranging from 97% to 98%. The result was promising, and the traffic codes (red, orange, yellow, green) were displayed.

Abhijeet et al. (Abhijeet Tikar, MAY 2017), realised a model to predict Taxi trip travel times using machine learning and GPS collected data. The estimation of taxi trips was applied by using KNN algorithm. The system showed good results in comparison with the real examples. They also introduced the location clustering-based approach to map raw GPS co-ordinates to meaningful locations and then used them for travel time estimation.

Bai et al. (Cong Bai, 2015) have realised a model for dynamic travel time prediction for buses on road with multiple routes. The goal of this project is to better help passengers to plan their trips and avoid long periods of waiting. The results showed that the model is applicable for bus travel time prediction and has the best performance.

Kaulage et al. (A N Kaulage, MAY 2017) presented a statistical approach to predict the public bus arrival time based on GPS information and taking on consideration all the factors affecting the bus travel time like the departure time, workdays, current bus location, number of links etc. The researchers concluded that the proposed system could not make the bus tracking and passengers service more convenient.

Zaki et al. (Mohamed Zaki, MAY 2015) presented a model-based approach for travelling time prediction. The development of this model is based on DsmT (DezertSmarandache Theory) as a fusion technique and an integration of different data sources such as GPS, historical travel time data, weather conditions, and real time traffic data. The proposed model showed promising results in terms of accuracy and can be used to improve travel time prediction in intelligent transport systems.

2. Proposed Approach

The proposed solution in this study consists of several components. The first component serves as a mobile application, providing information to the user about his ID, location, and credit earned as a reward for using this application while traveling in the city. To obtain the user's location, the main activity calls the location service class, which uses predefined Java classes to check the availability of GPS service and determine the user's latitude and longitude. This information is then sent back to the main activity.

The communication between the main activity and the server involves saving collected GPS data to the server database after checking the network availability, and calculating the user's payoff, which is displayed on the screen as motivation for users to continue using the application. More details about data collection and communication between the server and the android application are shown in Fig. 4. and Fig. 5., respectively.

One of the challenges faced in designing the android application was dealing with the limitations of Android in terms of background service and foreground service. These limitations are in place to conserve the battery life of the device, as using the GPS sensor and mobile internet connection can be energy consuming modules. However, the proposed design aims to optimize the usage of these resources while tracking the user's location continuously and without interruptions.

In summary, the android application component of the proposed design consists of the main activity as the user interface, the location service class for obtaining the user's location, and communication with the server for data collection and calculating user's payoff. Efforts have been made to overcome the challenges posed by Android limitations and optimize the battery usage while ensuring continuous tracking of the user's location. Further details of the proposed design can be found in Fig. 4. and Fig. 5. Overall, the organized design of the android application is an essential part of the proposed idea in this paper, which aims to collect GPS data, train a deep neural multilayer perceptron using Python, and provide traffic and travel time predictions to users through a web application.

The Deep Neural Multilayer Perceptron, an artificial neural network used for basic deep learning modelling, was employed for model training. Four hidden layers were included in the model, each with a specific number of neurons (256, 128, 64, 32). Combined with the input and output layers, a total of six layers were used for training.

The collected GPS data was structured to include various information, such as the latitude and longitude of point A (the coordinates of the first point of the segment), the latitude and longitude of point B (the coordinates of the second point of the segment), the distance between point A and point B, the day of the week when the travel operation occurred (denoted as ”Day”), and the period of time when the travel took place. The period was divided into three categories: 7:00 am to 4:00 pm (denoted as 1), 4:00 pm to 9:00 pm (denoted as 2), and 9:00 pm to 7:00 am (denoted as 3). This structured data was used for training the model.

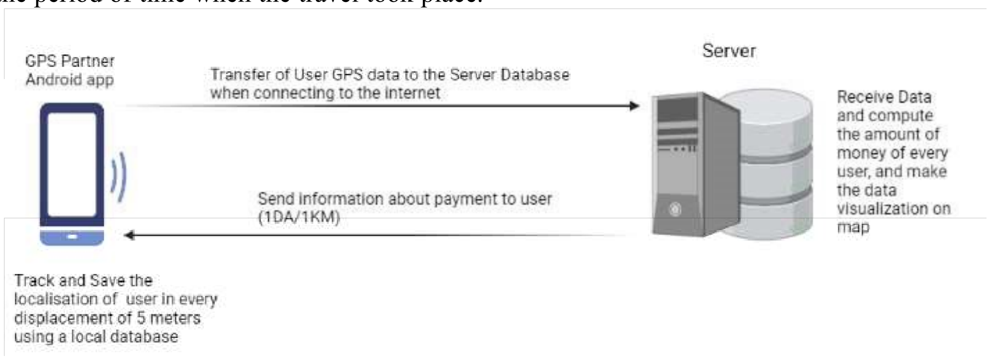

Fig. 4. Communication between GPS Partner and the server

Fig. 5. User interface of GPS Partner.

Traffic state: determined by the speed of the traveller passing the segment, with values assigned based on speed thresholds (e.g., very slow, slow, normal, fast) ranging from 10 km/h to 50 km/h or greater.

Understanding the problem and collecting sufficient data for machine learning was not the only challenge in this study. A well-designed web application was also crucial. The concept is to allow users to select their route, segment it, and then send each segment to the prediction module to estimate traffic state and travel time for each segment, ultimately providing an estimate for the entire route. Overall, a systematic design is essential for the successful implementation of the proposed idea, which involves an android application for data collection, model training using a deep neural multilayer perceptron, and a web application for traffic prediction and travel time estimation. The collected GPS data is structured to include information such as latitude, longitude, distance between points, day of the week, and time period during which travel occurred, as well as traffic state based on speed thresholds.

The implementation of the proposed idea required a well-organized plan to effectively collect and utilize the data for building the prediction model. This section is divided into four parts:

1. Development of the Android application using Android Studio editor and Java programming language to collect data from the inhabitants of the city of Bechar and store it in a remote database. PHP was used to create the server-side web application that manages the collected data and assigns user IDs. This stage involved significant efforts and support from volunteers, resulting in a prolonged timeline.
2. Treatment of the collected data, including visualization on a map using Google Maps services and filtering to retain only relevant data collected on known roads.

3. Building the deep neural multilayer perceptron model using Python programming language and Scikit-learn machine learning library for machine learning.
4. Evaluation and accuracy measurement of the deployed model are crucial steps after model deployment.

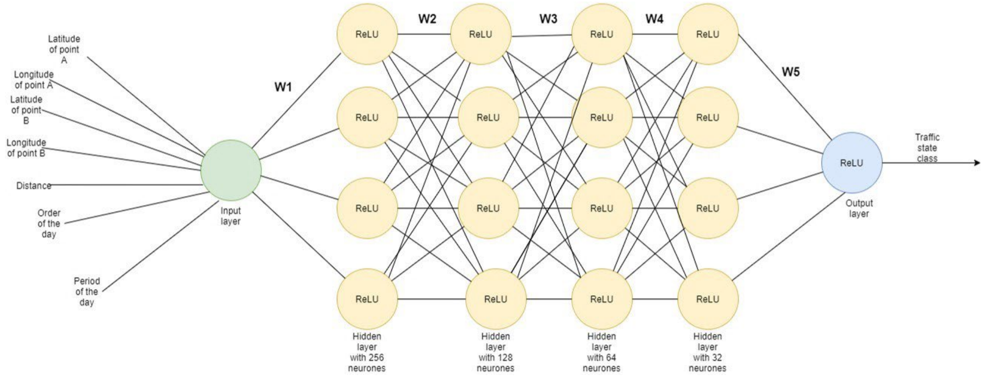

Fig. 6. Architecture of the proposed deep neural multilayer perceptron

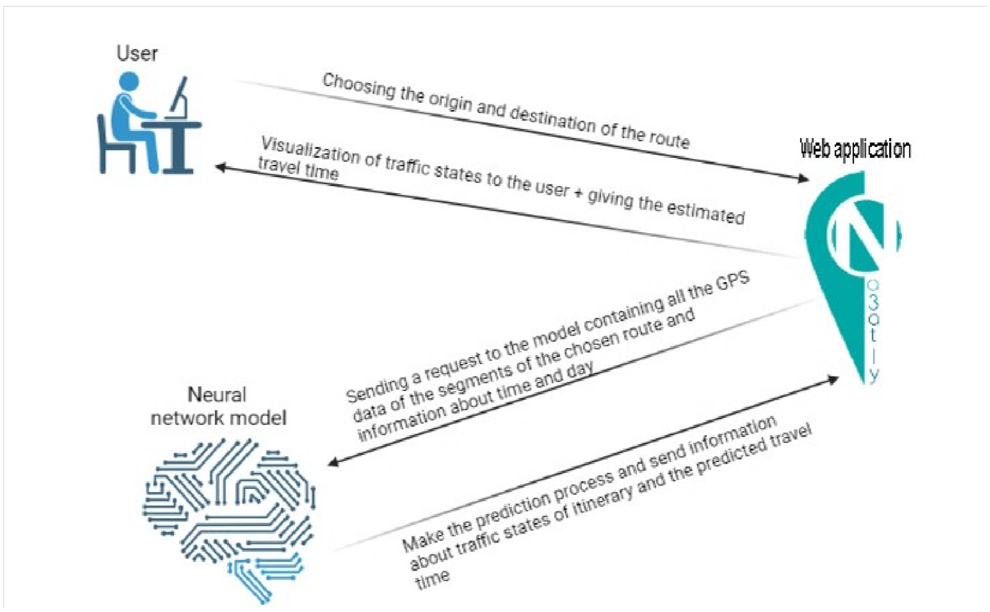

Fig. 7. A description of how the proposed web application works

2.1 Data Collection Process

To effectively study traffic congestion and develop an accurate model for estimating travel time based on this condition, a significant amount of data is required. To collect this data, an android application named "GPS Partner" was developed. This application tracks the user's location using GPS services on their Android phone, recording their coordinates every 5 meters during their travel within a 10-kilometer radius from the city centre of Bechar, Algeria.

To encourage user participation, users were paid 1 Algerian dinar for every kilometre they travelled with the application running and GPS sensor enabled. A total of 52 participants from various parts of Bechar city contributed to this phase, covering over 2 220 kilometres and collecting a total of 297 009 GPS positions[1].

After the data collection phase, the collected data needed to be visualized and filtered to retain only relevant data, as the focus of the project is on addressing traffic congestion. The data was visualized using Google Maps API for better visualization and filtering purposes (see Fig. 8.). This step was crucial in preparing the data for further analysis and model building.

[1] https://labrier-univ-bechar.dz/na3atly/

After generating a CSV file containing all the required, cleaned, and ready GPS data for the training process, approximately 239 346 segments of data were used for machine learning. Given the high complexity of the problem, which involves classifying segments based on the state of traffic flow, Artificial Neural Networks were chosen as the learning algorithm.

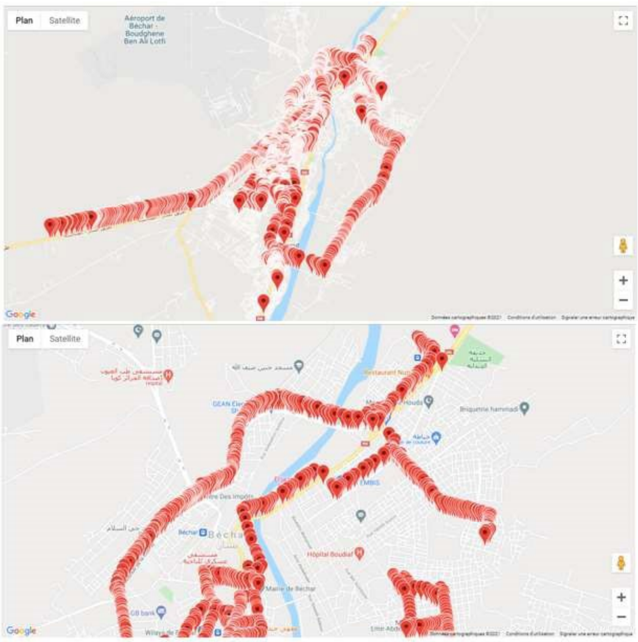

Fig. 8. Visualization of the last 10 000 collected GPS points

The training phase commenced with the data being split into 80% (191 476) for training purposes and 20% (47 869) for testing purposes. To ensure that the input data was normally distributed and cantered around zero, with variance in the same order, a scaler was created. The Rectified Linear Activation Function (ReLU) was used as the activation function, which is a piecewise linear function that outputs the input directly if it is positive, and zero otherwise. ReLU has become the default activation function for many types of neural networks due to its ability to facilitate easier learning and often achieve better performance. The maximum number of iterations for training was set to 200 iterations.

$$\text{Accuracy} = \frac{\text{Number of well predicted values}}{\text{Total number of real values}} \times 100 \qquad (1)$$

After conducting a comparison between the predicted values and the real values of traffic states, the algorithm calculated the accuracy of the model according to Equation (1), which was approximately 99.47% as per the confusion matrix (refer to Fig. 9.). This level of training performance can be considered as highly commendable for a neural network model.

As Google Maps does not currently offer travel time estimation based on real-time traffic conditions in the city of Bechar, due to limited data availability in the area, the proposed service that allows inhabitants of Bechar to share their experiences about traffic conditions has the potential to significantly enhance the accuracy of travel time predictions. By leveraging local data from Bechar city's residents, the proposed service can provide more precise and up-to-date information about traffic conditions, leading to improved travel time estimates for users. This would address the limitations of relying solely on Google Maps data and enhance the accuracy of the service.

3. Experiments and Obtained Results

To evaluate the accuracy of the proposed solution, an evaluation test was conducted by comparing it with Google Maps, which provides time estimations based on average speeds without taking into consideration the traffic conditions in the city of Bechar. Real-time experiments were conducted in Bechar, Algeria, using various routes to different locations.

For example, in an experiment shown in Fig. 10., the proposed application estimated the travel time from Taouba mosque to Culture house as 17 minutes and 56 seconds, while Google Maps estimated it as 16 minutes. The actual travel time recorded during the experiment was 18 minutes and 5 seconds. The proposed application had a minimum error of -6 seconds, whereas Google Maps had an error of -1 minute and 55 seconds. This demonstrates that the proposed application provided more accurate estimations, taking into consideration the real time traffic conditions in Bechar city.

Fig. 9. The confusion matrix of the training process

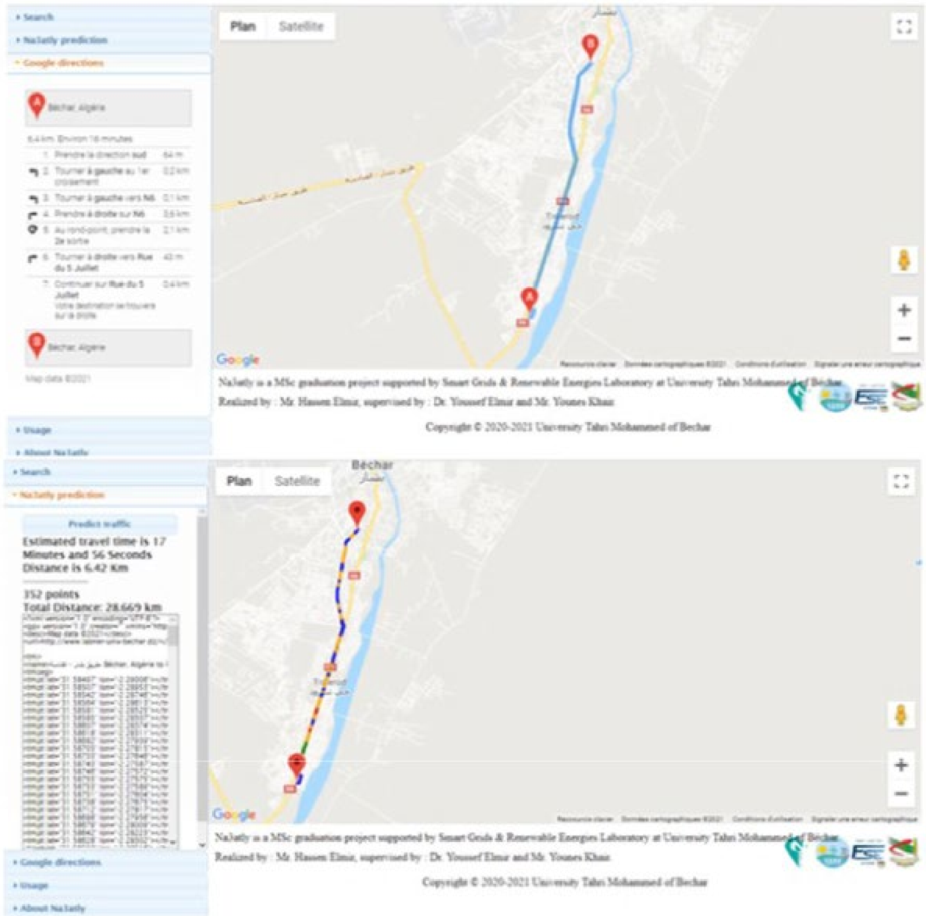

Fig. 10. Example of a real time experiment from Taouba mosque to the Culture house

3.1. Results and Discussion

The performance of the proposed web application in real-time experiments was encouraging, with an average error of 50 seconds. In most cases, it provided more accurate estimated travel times compared to Google Maps, taking into consideration the condition of traffic. However, there were four tests where the accuracy was not as good as in the initial tests. This could be attributed to the limited data available for those specific routes in the proposed web application, which resulted in less accurate estimates. Nevertheless, overall, the proposed web application outperformed Google Maps in providing precise values of travel time, considering the traffic condition in the city of Bechar.

$$\text{Average error} = \frac{\sum \text{Error in each experiment}}{\text{Total number of experiments}} x100 \tag{2}$$

Based on the findings presented in Table 1, the study's objective was successfully achieved as the proposed solution demonstrated better accuracy in estimating travel time and providing traffic state information for segments of the selected route. The performance of the solution can be further improved by incorporating additional training data and extending the data collection period. These results also support the hypothesis that traffic has a significant impact on route travel time estimation.

With the proposed solution, users can view the estimated time for their chosen route before embarking on their journey. This empowers them to make informed decisions on whether to take the suggested route or explore alternative options. Furthermore, the service of providing traffic state information enables users to avoid congested roads and choose more efficient routes, enhancing their overall travel experience.

Table 1. Time estimation and errors for various routes

Origin	Destination	Distance (km)	Real time (s)	Google map (s)	Proposed (s)	Error (Google map) (s)	Error (Proposed) (s)
Echohada mosque	AGB bank	1.4	177	240	150	+63	-27
University of Bechar	Directorate of religious affairs	2.2	235	360	228	+125	-7
Taouba mosque	Culture house	6.4	1085	960	1076	-115	-6
Hammadi station	Sogral station	3.9	347	480	329	+133	-18
Debdaba amusement park	El-Kods pharmacy	3.8	510	720	497	+210	-13
El-Kods pharmacy	Algeria Telecom Department	3.9	690	780	549	+90	-141
Hai Essalam's Post	Ophtalmology Hospital	1.7	151	180	284	+29	+133
Sogral station	Bechar grand post	3.5	445	540	345	+95	-100
Bechar's town hall	University of Bechar	3.5	535	480	529	-55	-6
AGB Bank	Bechar's town hall	1.3	270	240	230	-30	-50
Average error						**±95**	**±50**

Conclusions

The purpose of this study was to develop a system that estimates travel time while taking into consideration the impact of traffic conditions, which is not provided by Google Maps in the city of Bechar, Algeria. To achieve this, a neural network model was trained using traffic data collected from the city of Bechar through an Android application called GPS Partner.

The data collection process involved distributing the GPS Partner application to gather sufficient traffic data. Subsequently, the trained neural network was integrated into a web application to provide traffic state predictions and travel time estimations for selected routes.

The results obtained from most tested routes were accurate, but some routes showed inconsistencies. This was primarily due to the limited availability of traffic data in certain areas, as the data collection period was only one month (April 2021) and the testing phase was conducted in June. To improve the accuracy of the predictions for all routes in Bechar, a longer data collection period of at least 2 years with data collected every month is recommended.

The developed application was a valuable addition to the city of Bechar and its residents, as it provided a service that Google Maps did not offer in the city. The availability of traffic information can benefit various stakeholders such as delivery companies and hospital emergency services. Furthermore, the research findings serve as a foundation for future research in building a comprehensive Intelligent Transportation System in Bechar or similar cities. The results of this study can motivate researchers to undertake similar works and contribute towards the development of smarter transportation systems in urban areas. Overall, this research is a significant step towards addressing traffic congestion and improving transportation efficiency in Bechar and beyond.

Acknowledgments

Authors of this work would like to express their sincere gratitude to Prof. Abdelkhalek Othmane head manager of SGRE laboratory for providing valuable materials and resources for this project.

References

A N Kaulage, A. P. (MAY 2017). Bus tracking and Bus arrival time, location prediction system. *International Research Journal of Engineering and Technology (IRJET)*.

Abhijeet Tikar, N. J. (MAY 2017). Travel time estimation for taxi trips using GPS sequence. *International Education and Research (IERJ)*.

Cong Bai, Z.-R. P.-C. (2015). Dynamic Bus Travel Time Prediction Models on Road with Multiple Bus Routes. *Hindawi Publishing Corporation Computational Intelligence and Neuroscience*.

Franklin, B. (1748). *Advice to a Young Tradesman.* National Archives and Records Administration/University of Virginia Press.

Hoong, P. K.-Y. (2012). Road traffic prediction using Bayesian Networks. *Faculty of Computing and Informatics, Multimedia University, Cyverjaya, Malaysia*.

Mohamed Zaki, A. H. (MAY 2015). Travel Time Prediction under Egypt Heterogeneous Traffic Conditions using Neural Network and Data Fusion. *Egyptian Computer Science Journal*.

Neha Garg, P. V. (APRIL 2017). Bus Tracking using GPS an Real Time Prediction. *International Research Journal of Engineering and Technology (IRJET)*.

Obuhuma, J. I. and Moturi, C. A. (NOVEMBER 2012). Use of GPS With Road Mapping For Traffic Analysis. *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY VOLUME 1, ISSUE 10*.

Thura Kyaw, N. N. (MARCH 2020). Predicting On-road Traffic Congestion from Public Transport GPS Data. *International Journal of Advances in Scientific Research and Engineering (IJASRE)*.