Grenze International Journal of Engineering and Technology, July Issue

Pothole Detection and Prediction using Sensors and Machine Learning

Nikhil K. Khaneja[1], Kalidindi Lakshmi Sanjana[2], Raunak Bhupal[3] and Dr. Kala Bharatan[4]
[1-3]Department of Computer Science and Engineering, PES University Electronic City Campus, Bengaluru, India
Email: nkhaneja09@gmail.com, sanjanakl5595@gmail.com, raunakbhupal7@gmail.com
[4]Department of Electronics and Communication and Engineering, PES University Electronic City Campus, Bengaluru, India
Email: kala.b.anna@gmail.com

***Abstract*—In India, in metropolitan cities, a common problem seen is traffic jams or slowdown on traffic due to potholes. Most of the road accidents are due to potholes. In this paper, the authors have tried to ensure that the road surface quality and the potholes should be monitored continuously and must be repaired as necessary. The best possible distribution of resources could be attained when the real-time data is gathered from the mounted sensory system and displayed and visualized in a way (Google Maps) which is comprehensive and empirical. Different sensors would be mounted onto the public transport vehicles. The method mentioned in this paper uses the ultrasonic sensor and the accelerometer to send the data regarding the intensity of pothole straightaway to the cloud data storage, along with the location coordinates of the potholes and the time stamp. Different databases such as traffic density of the city, rain intensity of different areas is used to train different machine learning models to predict the intensity of potholes at the particular locations and get the best possible accuracy. The best-case accuracy attained was 35.5%. The pothole intensity data is visualized through a heat map application. This system helps in reducing the frequency of traffic jams in any city of India. The method suggested in the paper would help in reducing the number of accidents happening in India due to the potholes**.

***Index Terms*— Sensors, Machine Learning, Cloud Data Storage, Heat map, Real time data.**

I. INTRODUCTION

In today's world roadways are the most important means of travel and commute, but the most common problem faced in travel is the condition of the roads. The condition of the road degrades as a result of many factors such as high rain, unevenness of the road, low-quality building material, high traffic density and so on. As a result of this, the potholes are formed. According to the report by The Guardian, almost 4000 people were killed and almost 25000 people were injured leading to almost 10 deaths per day because of the potholes in 2017 which is 50% more than 2016 and it has been only rising ever since [1]. This problem is alarming and needs to be taken care of as soon as possible. The potholes are formed due to the expansion and contraction of the groundwater after the water has entered into the ground through pavements. When water freezes, it expands. This happens when water enters the pavements, it freezes and expands and thus leading the pavements to expand, bend and crack which leads to weakening of the road infrastructure and when it

Grenze ID: 01.GIJET.6.2.1_1

melts it leaves a vacant space in the pavement and when the vehicles passes over the weakened spot, the roadway material deteriorates which causes material to broke or displace which leads to formation of potholes. There has been plenty of research going on nowadays for supervising and monitoring the road conditions using smartphones [2] and image processing [3].

In our paper, we study the use of the prototype built using an ultrasonic sensor and an accelerometer while trying to minimalize the cost.

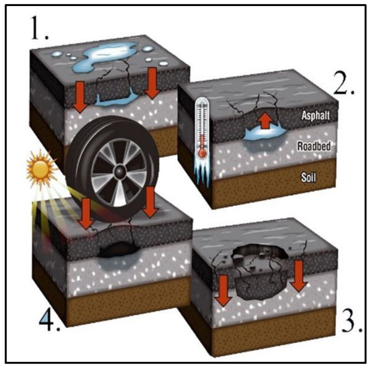


Fig 1: Formation of Potholes [4]

Figure 1 shows the formation of a pothole. It’s been divided into four phases. Seeping of water into cracks (1), Expansion of water due to high temperature (2), Evaporation of water and creation of vacant spaces (3), Due to pressure and load damage at vacant spaces (4).

The main objective of this paper is to develop an app model that helps in detecting potholes throughout any city in India which predicts the intensity of the potholes based on the datasets of the features contributing towards the formation of pothole.

The paper is been divided into four sections namely the proposed system, prediction analysis using Machine Learning, test vector model and user interface.

II. Proposed System

The Sensing subsystem gathers the data through the sensors deployed that consist of an Accelerometer and an Ultrasonic sensor along with a NodeMCU. The microcontroller is used to control all the sensors and is responsible for data gathering and sending it to the database. The system uses Google geolocation API to gather the latitude and longitude of the pothole. After the connected sensor senses, the pothole and the location data are received, this is processed and is sent to Google Cloud which is used to maintain a cloud database [5]. The subparts of the proposed system are mentioned below.

The below mentioned Figure 2 is the flow diagram of the proposed system. Data is collected from both the sensors and Google Geolocation API is sent to the database via Node-MCU microcontroller which controls all the sensors and pre-processes the data.

A. Accelerometer (ADXL335)

Accelerometer Sensor is an electromechanical device that is used to sense the proper acceleration in all the possible directions. The acceleration measured can be gravitational or due to a moving body. Accelerometer measures the net external force applied on it and gives the corresponding readings which are analog voltage values corresponding to the acceleration. For example, if the system is stationary the reading given by the accelerometer would correspond to the earth’s gravitational force [6]. Accelerometer that is used in this project is ADXL335. It is a small, thin, a 3-axis accelerometer with a minimum measurement range of ±3g. It has a 5-pin set up out of which one is the ground pin, one is Vcc pin and the rest three are the output pins which provide the acceleration of all the 3 axes.

Figure 2 shows how all the three axes are calibrated for the accelerometer ADXL335 and acceleration of 1g acts downwards. The accelerometer detects the vibrations as soon as the vehicle passes through the pothole. As the vehicle passes, there is a gradual decrease in the values detected and if the decrease is substantial and the net value is greater than the threshold value, the data is accepted and recorded in the database.

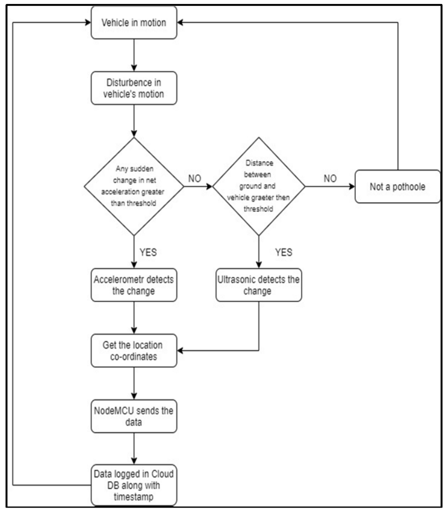


Fig 2: System Workflow

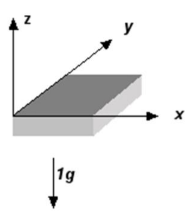


Fig 3: The Axis Calibration of ADXL335 [7]

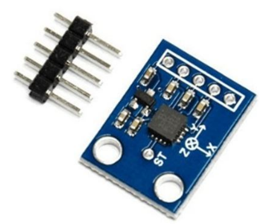


Fig 4: ADXL335 used in the system

B. Ultrasonic Sensor (HC-SR04)

An ultrasonic sensor is a non-contact distance measuring device. It is also known as an ultrasonic transducer as it uses a transmitter and a receiver to determine the distance. The transmitter sends the ultrasonic waves in one direction and waits for it to bounce back off the surface of the object. It would calculate the time till wave reaches back to the receiver and as the ultrasonic velocity in air is 340 m/s the distance is calculated as

DISTANCE = SPEED * TIME

The ultrasonic sensor used here is HC-SR04, which provides a range of 2cm - 450cm contactless measurement system which has the ranging accuracy up to 3mm and works at a frequency of 40Hz [8]. The sensor incorporates a 4-pin setup consisting of ground connection, Vcc, trigger pin and an echo pin.

Trigger pin has a fundamental role of emitting the ultrasonic waves to initialize measurement and It must be kept in a high state while the echo pin goes into the high state for a minuscule period based on the time required by the ultrasonic waves to bounce back to the sensor from the receiving edge.

Fig 5: Ultrasonic Sensor HC-SR04 used in the system

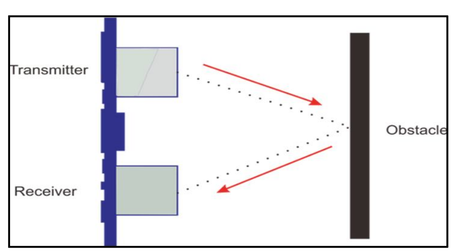


Fig 6: Working of the HC-SR04 [9]

Figure 6 shows the working of HC-SR04. The transmitter emits a signal, it bounces back the obstacle and the receiver receive the signal and calculates time.

There may be times when the vehicle dodges the pothole and passes over the pothole. In this case, the ultrasonic sensor comes into the picture. When the vehicle dodges the obstacle, the vehicle allows the pothole to pass in between the tires. A threshold value is set to calculate the distance between the base of the vehicle and the road. The sensor continuously measures the distance and as soon as the value rises above the threshold, the pothole is detected and by subtracting the threshold distance, the depth of the pothole is calculated. The depth is recorded in the database.

C. Location System (Google Geoocation API)

An API is an acronym for Application Programming Interface. To get the desired location of the device, the google geolocation API is used. Google's Geolocation API uses radio towers and Wi-Fi nodes in the system's vicinity that the system can detect and returns the location and accuracy radius of the pothole. Communication and data transfer is done over HTTPS using POST and both request and response are formatted as JSON, and share a common resemblance in their message types, that being *application/JSON* [10] The geolocation request is sent in a JSON format using a POST request to a geolocation URL which processes the request and returns the response and the location and the radius also in the JSON format.


```
{
  "location": {
     "lat": 12.7916
     "lng": 77.5946
  }
  "accuracy": 30
}
```


Fig 7: Response from Geolocation API

The *lat* parameter specifies the latitude coordinate of the location while the *lng* means the longitude coordinate of the location and *accuracy* specifies the radius of the location

The sensors together along with the geolocation API makes the system whole. The sensors detect the presence of the pothole and API retrieves the location of the pothole and all the data is sent to the database.

The database is maintained over the google sheets. The system uses a third-party intermediate client to send a log the data to the google sheets. An HTTP web server client is used where a non-persistent connection is established to the google sheets and the data is sent. In the connection type used, the client establishes a connection, sends the data packet and closes the connection and if the sending of the packet is successful an acknowledgment is received by the sender station the data was received successfully and the receiver is ready to receive another data packet. As soon as the acknowledgment is received, another data packet is sent.

III. Prediction Analysis Using Machine Learning

"*Prevention is always better than cure*"

Stating the above mentioned, to predict the occurrence of the pothole at a location before it appears is of much help for the people who commute daily via that particular location. The machine learning model helps the maintaining bodies to a larger extent as it will reduce both, the cost and the human effort in repairment of the potholes. The road can be amended at an early stage where the signs of damage are present.

Many machine learning models and algorithms have been presented and used which are trained on certain selected datasets like the traffic density on that particular road and rainfall patterns and after processing and data filtering, the data was fed to the models, and as a result the occurrence of the pothole at a certain location where the road traffic density and rainfall pattern was defined. The results of both the models used were tabulated. The models that were used are mentioned below.

A. Model-1: Linear Discriminant Analysis (LDA)

The Linear discriminant analysis (LDA) is an unsupervised learning method and is preferred whenever there are two or more datasets involved in the training process. LDA estimates the probability of the new sets of the input data for each class thus helping in making the predictions. The class that acquires the highest probability in the output dataset for the particular input dataset, that dataset is selected as the final predicted output.

$$P_k(X) = \frac{\pi_k f_k(x)}{\sum_{l=1}^{K} \pi_l f_l(x)}$$

This model is based on above mentioned Bayes' theorem which states the probability of the event can be calculated based on the prior information regarding the event that has had happened and is known [11].
Π_k is the overall probability that an observation is associated to the *kth* class. $f_k(x)$ denotes the density function of *x* for an observation that comes from the *kth* class.

B. Model-2: Artificial Neural Networks

Artificial neural networks (ANN) are thoroughly interconnected and multi-tier neural nets. Each node in a layer is associated with every other node in the next layer. By increasing the number of hidden layers, the neural network is made deep and profound. Each layer of the ANN does a non-linear transformation of the input from one vector space to another. They enhance existing data analysis technologies and helps in producing meaningful outputs [12].

$$z = f(b + x \cdot w) = f\left(b + \sum_{i=1}^{n} x_i w_i\right)$$

$$x \in d_{1\times n},\ w \in d_{n\times 1},\ b \in d_{1\times 1},\ z \in d_{1\times 1}$$

The following equation is the general equation, here the weighted sum of all the inputs is passed to a non-linear activation function and can be expressed as a vector dot product where n is the number of inputs, w is the respective weights of inputs and b is the bias which has value 1. It is used to shift the values of the activation function from left to right and helps train the model when inputs are 0 [13].

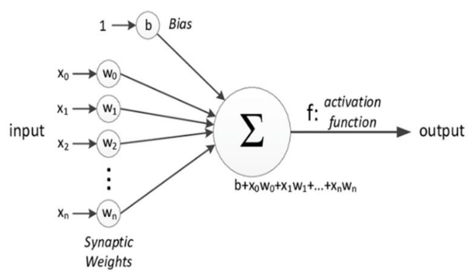


Fig 8: Representation of a one particular input

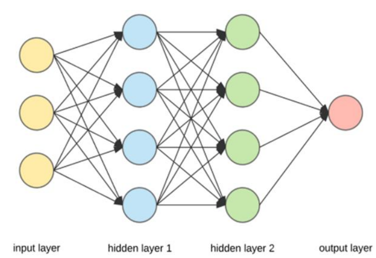


Fig 9: Structure of a developed Neural Net [14]

Figure 8 and 9 shows the basic structures of Artificial Neural Networks. It has hidden layers between the input and output layers for better results.

IV. Observations

TABLE I: COMPUTATION TAKEN FOR EACH MODEL

Model / Algorithm	Accuracy (%)
Linear Discriminant Analysis (LDA)	35.5
Artificial Neural Networks (ANN)	28.9

TABLE II: ACCURACY RECORDED BY EACH MODEL

Model/Algorithm	Time Taken (sec)
Linear Discriminant Analysis (LDA)	5.959
Artificial Neural Networks (ANN)	2.990

Table 1 shows the time taken by each algorithmic model to produce the output and Figure 10 depicts the graphical representation of the same. Table 2 shows the time taken by each algorithmic model to produce the output and Figure 11 depicts the graphical representation of the same.

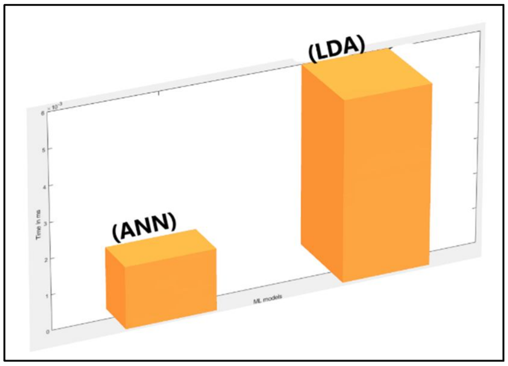


Fig 10: Computational Time Plot of Models

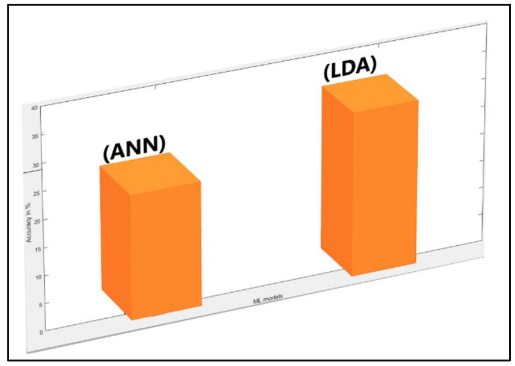


Fig 11: Accuracy Plot of Models

The results of LDA and ANN were considered because the pothole formation at a given particular location depends on a lot of factors and data sets of some of the factors were not accessible.

V. Test Vector Model

For the testing purposes of the model, the raw data used is gathered from the below-mentioned sources:
The raw data of the Annual Precipitation used to train and test the models that were considered is Monthly Precipitation Probabilities of New York State” [14] .
The raw data of the road traffic density used to train and test the models that were considered is “Annual Average Daily Traffic of 2015 of New York City” [15] .

VI. User Interface

Navigation is an important integral of any vehicle. Potholes are one of the major reasons for traffic jams and road accidents in India. The intimation of a pothole while navigating would be helpful.
The location of the pothole is dynamically updated to the application and the user can get to see it while navigating. The proposed system helps any developing authorities (like in Bangalore BDA) in the direction of filling up of potholes for road safety.
The user can click on the point of the pothole which will direct them to show them the location along with the intensity of it. Since intensity is known well in advance, they can take the materials sufficient for filling up that pothole and go to the directed location using the navigation provided by this mobile application and fill up the pothole.
The application developed uses a heat map to visualize the depth and the intensity of pothole by placing a specific coloured marker.
The heat map is a graphical representation of data that is to be visualized where the individual values are represented by particular colours that represent the category the values lie in.

Fig 12: Screenshot of the heat map of pothole

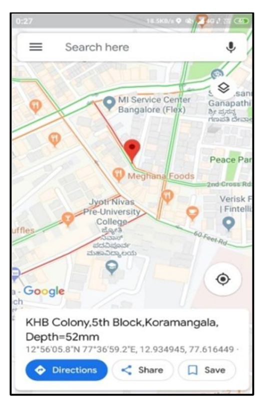


Fig 13: Screenshot of the navigation route to the selected pothole on the heat map

Figure 10 shows the screenshot of the android application developed to display the heatmap of the data collected.

A benchmark of 50mm is set for the pothole to be of high intensity. Thus, if the intensity is greater than 50mm then it is denoted by a red colour. An intensity varying between 25mm to 50mm is denoted by green colour and that lesser than that is denoted by yellow colour.

The developed product model is supported by a mobile application. The application is developed using Android Studio software using Java as the programming language with the SDK version 23. Figure 9 shows the directions to the location of the pothole located on heat map.

VII. Results and Figures

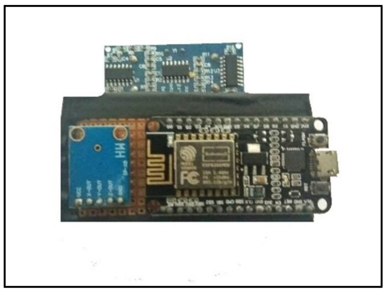


Fig 14: Side view of the Sensing Subsystem

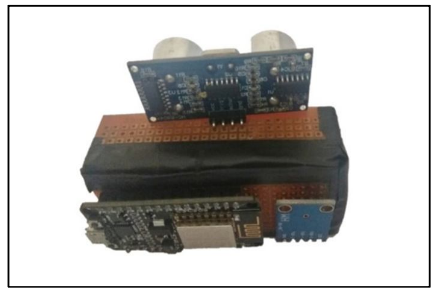


Fig 15: Top View of the Sensing Subsystem

The figures show the different components (the sensors and the micro-controller from the side view (14) and top view (15).

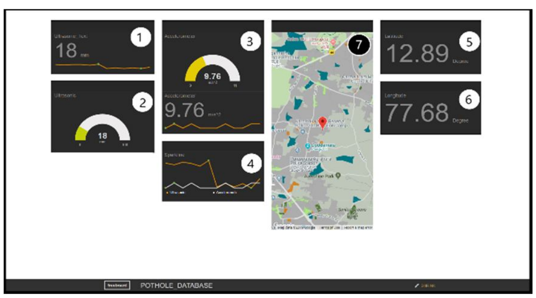
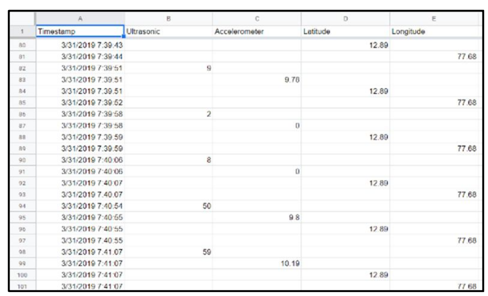


Fig 16: Freeboard.io data visualization dashboard



	A	B	C	D	E
1	Timestamp	Ultrasonic	Accelerometer	Latitude	Longitude
80	3/31/2019 7:39:43			12.89	
81	3/31/2019 7:39:44				77.68
82	3/31/2019 7:39:51	9			
83	3/31/2019 7:39:51		9.76		
84	3/31/2019 7:39:51			12.89	
85	3/31/2019 7:39:52				77.68
86	3/31/2019 7:39:58	2			
87	3/31/2019 7:39:58		0		
88	3/31/2019 7:39:59			12.89	
89	3/31/2019 7:39:59				77.68
90	3/31/2019 7:40:06	8			
91	3/31/2019 7:40:06		0		
92	3/31/2019 7:40:07			12.89	
93	3/31/2019 7:40:07				77.68
94	3/31/2019 7:40:54	50			
95	3/31/2019 7:40:55		9.8		
96	3/31/2019 7:40:55			12.89	
97	3/31/2019 7:40:55				77.68
98	3/31/2019 7:41:07	59			
99	3/31/2019 7:41:07		10.19		
100	3/31/2019 7:41:07			12.89	
101	3/31/2019 7:41:07				77.68

Fig 17: Database maintained on Google Sheets

In figure 16, labels 1 and 2 represent the data gathered by the ultrasonic sensor, label 3 represents the data gathered by accelerometer, labels 5 and 6 represent the latitudes and longitudes of the location and label 7 represent the visualized map of the aforementioned location co-ordinates.

The database was mentioned in Google Sheets as a third-party service of dweet.io and freeboard.io. These intermediaries were used to create a dashboard to visualize and analyze the data on the server. Dweet.io is a server that accepts the post requests from the micro-controllers such as NodeMCU and directs them to freeboard.io to visualize them.

VIII. Results and Figures

The proposed system was successful in acquiring the location of the pothole and simultaneously updating it in the cloud database and the app. This also helps in reducing the number of potholes present thus resulting in reducing the number of traffic jams in the city. The detection of the potholes helps in saving the cost of repairment of roads as they can be repaired before the condition gets worse. The proposed system is cost-effective and compact and can be reconfigured. The net value of the product is INR 800 only. The results obtained are accurate and efficient.

REFERENCES

[1] Article-The Guardian dated 25th July 2018 (Online) Available: https://www.theguardian.com/world/2018/jul/25/more-deadly-than-terrorism-potholes-responsible-for-killing-10-people-a-day-in-india

[2] Avik Ghose, Provat Biswas, Chirabrata Bhaumik, Monica Sharma, Arpan Pal, Abhinav Jha "Road Condition Monitoring and Alert Application Using In-Vehicle Smartphone as Internet-Connected Sensor" in IEEE International Conference on Pervasive Computing and Communications Workshops 2012 pp 489-491.

[3] Emir Buza, Samir Omanovic, Alvin Huseinovic (University of Sarajevo), "Pothole Detection with Image Processing and Spectral Clustering" in 2nd International Conference on Information Technology and Compter Networks-2013 pp 48-53

[4] Article in Toledo Blade dated 2nd March 2014 (Online) Available: http://jothut.com/cgi-bin/junco.pl/blogpost/32730/Toledo-is-located-in-Pothole-Alley

[5] Vijaya Shetty S., Niveditha H. N, J.Janani, T. Hrishikesh Reddy, Praveen Datt "Pothole Detection and Notification System" in International Journal of Advance Research and Innovative Ideas in Education (IJARIIE) volume 4 issue 2 2018 pp 3913-391

[6] A Pothole Detection System (M. Tech Project Report– II stage) by Sudarshan Rode Roll No: 06329001 Dept. of CSE, IIT, Bombay 2007

[7] Article on Accelerometers by Hobby Electronics (Online) Available: http://www.hobbytronics.co.uk/accelerometer-info

[8] Yoshiaki Taniguchi, Kodai Nishii, Hiroyuki Hisamatsu "Evaluation of a Bicycle-Mounted Ultrasonic Distance Sensor for Monitoring Road Surface Condition" in 7th International Conference on Computational Intelligence, Communication Systems and Networks (CICSyN) 2015 pp 31-34.

[9] Article" HC-SR04 Ultrasonic sensor with Arduino Tutorial" dated 13th September 2017 (Online) Available: http://miliohm.com/hc-sr04-arduino-tutorial/

[10] Developers Guide – Documentation of Google's Geolocation Service (Online) Available: https://developers.google.com/maps/documentation/geolocation/intro

[11] Jason Brownlee "Linear Discriminant Analysis for Machine Learning dated 6th April 2016 (Online) Available: https://machine earningmastery.com/linear-discriminant-analysis-for-machine-learning/

[12] Arden Dertat "Applied Deep Learning – Part 1: Artificial Neural Networks" dated 8th August 2017 (Online) Available: https://towardsdatascience.com/applied-deep-learning-part-1-artificial-neural-networks-d7834f67a4f6

[13] Nagesh Singh Chauhan "Introduction to Artificial Neural Networks (ANN) – Your first step in deep learning" dated 3rd October 2019 (Online) Available: https://towardsdatascience.com/introduction-to-artificial-neural-networks-ann-1aea15775ef9

[14] Arden Dertat "Applied Deep Learning – Part 1: Artificial Neural Networks" dated 8th August 2017 (Online) Available: https://towardsdatascience.com/applied-deep-learning-part-1-artificial-neural-networks-d7834f67a4f6

[15] Climatography of The United States No. 81 Supplement No. 1 "Monthly Precipitation Probabilities and Quintiles 1971-2000" by National Oceanic and Atmospheric Administration (Online) Available: https://www.ndsu.edu/fileadmin/ndsco/normals/documents/7100/Monthly_Precip_Prob_ND.pdf

[16] Traffic Volume Report of State Roads (2015) by New York State Department of Transport (Online) Available: https://www.dot.ny.gov/divisions/engineering/technical-services/hds-respository/NYSDOT_2015TrafficVolumeReport-Routes.pdf