ISSN 2278-3091
Volume 9, No.4, July – August 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse195942020.pdf
https://doi.org/10.30534/ijatcse/2020/195942020

Cloud Architecture for Web Based Smart City Transport System

Revati M. Wahul [1], Dr. Santosh S. Lomate[2]
[1]Research Scholar, BAMU, India, reva.wahul@gmail.com
[2]Research Supervisor, BAMU, India, drsantoshlomte@gmail.com

ABSTRACT

The developments in cloud technology and the web of things (IoT) have produced a promising opportunity to solve the challenges caused by the increasing problems of transport. We aim to deliver a specific multi-layered communication information technology platform using cloud computing and IoT technologies to address the challenges raised by growing transportation difficulties. In this system we proposed transportation data management using cloud computing for smart cities. The entire system has described in three different phases. First using TfL application programming interface extract the real time data from London Transportation system and store it into open Cloud database. The proposed cloud database consist a mongo DB platform which is used for or data insertion as well as extraction with the minimum cost. Before inserting data into cloud database system deals with some data mining phases like ok data preprocessing, filtration, data acquisition normalization etc. During this phase system can eliminate search instances which hold null values for classified instances. The normalized data should be a balance data, and it provides good accuracy for custom as well as classical classifiers. The data extraction from our own cloud platform, the data is extracted in various manners like number of bus stations, list of routes, list of affected stations, number of buses, number of roots etc. After data extraction reply cross fold validation (e. g. 10 fold and 15 fold cross validation) for classification. For a training system uses Naive Bayes classification algorithm to build the background knowledge according to training set. Once training has done it create train policies which are basically used for classification in test set. Similarly the classification has done with NB algorithm on remaining data set. In the third phase the accuracy of propose system, and evaluate with other classification algorithms. Various experimental analysis shows how propose system is better than other classification algorithms. Finally we conclude proposed system introduces around 95% classification accuracy for different cross validation on entire data set.

Key words : API , Intelligent transportation systems, Machine learning, NLP, Public cloud, Service oriented architecture (SOA), Transportation systems.

1. INTRODUCTION

Traditional vehicles are increasingly fitted with increasingly large quantities of sensors, actuators, and communication equipment (mobile devices, GPS tracking, and integrated computers). Specifically, different vehicles have efficient monitoring, networking, communication, and processing capabilities, and can interact with range of potential or share data with external regions over a wide range of prices. The aim of the study was to research the ideas of transport management and good cities, the study focus on to describing the world efforts towards achieving transference and moral cities, further as gender problems like athletics in relation to totally different elements of the planet, jointly of the samples of transport organization in good cities. Best transport exists jointly of the first gears driving growth, a trend that has existed for extended amount of your time. Transportation organization is taken into account to be a worldwide idea. In fact, most developed countries still invest within the physical setting like infrastructure and quality, that square measure thought-about to be core drivers for transport management. Like never before, completing development in good cities needs the mixing of reliable transport systems. The driving issue of good cities is to confirm the affiliation of human wealth, organization and social capital in a very bid to realize a lot of property commercial development further as a more robust mode for residents in these cities. In different words, property transport is one among the driving components as way as achieving good cities is bothered. Through this case, the remaining steady movement of resources, be it human labor or product across given location; end up in domestic and global success. This paper we describes a smart transportation management in metropolitan cities using cloud architecture. This research also demonstrates how to handle massive generated data from Vehicular as hoc Network (VANET) or transportation systems.

2. LITERATURE SURVEY

The Intelligent transportation networks in multimodal procurement: a case of role and contribution by means of wireless vehicle networks in a maritime port location [1]. The development of a multifaceted nature of teamwork and its importance as a noteworthy monetary activity has raised the profile of data and information and communication

technology (ICT) as it seeks to increase components. With the utilization of remote vehicular systems, Intelligent Transport Systems (ITS) can possibly shape the eventual fate of multimodal coordination. Without complicated System and enable, the possible function and dedication of ITS and specifically remote system components in collaboration is studied in a multi - modal instances of a container terminals that takes care of ocean-transported load data that is unloaded into vehicles for transportation. Occasional stream monitoring and investigation demonstration systems are used to evaluate the possibility for ITS to assist the persistent network data associated with the exchanging of messages, which demonstrate the stream of occasions occurring in multimodal communication and which may be connected to high-impact capabilities with monetary implications, such as monitoring and follows entire system. Moreover the remote vehicular systems guarantee to majorly affect how transportation and coordination's tasks are run. So as to accomplish the job and commitment of ITS the recognizable proof of data and material courses through mapping is an essential advance towards characterizing a system to actualize remote vehicular systems DSRC. The job and commitment of ITS through remote vehicular systems may incorporate the capacity to give moment continuous following and following, which can uncover if merchandise are conveyed/gathered to/from the correct spot, just as constant updates to corporate data frameworks, expanded security, robbery avoidance, expanded vehicle use, driver/administrator checking, and so forth speaking to a difference in worldview on how the store network can be overseen. Information traffic is vital to help the majority of the above mentioned. A huge scale sending of ITS through remote vehicular systems, for example, DSRC would enable universal access to data. Be that as it may, preceding substantial scale organizations occurring, test beds should be conveyed to run various preliminaries and to realize what kind of results will be accomplished. In the specific instance of remote vehicular systems, for example, DSRC, a testing situation to send a proving ground is spoken to by multimodal coordination's involving street and ocean transport. A similar system utilized in the investigation of mass materials can be connected to different tasks including holders just as mass fluids.

A Descriptive Bayesian Approach to Modeling and Calibrating Drivers' En Route Diversion Behavior described in [2] [3], This method provides a probabilistic framework for predicting and calibrating drivers and determination change of route with usage patterns obtained from laboratory driving simulation models and field Bluetooth sensors. The models of action are not founded on assumptions of absolute rationality. Alternatively, a new concise approach is suggested and illustrated, based on the principles of naive Bayes. The concept of en-route intervention is first determined with conduct data from a simulation environment. While [3] describes multi label classification arrival time estimation using ANN for bus transportations network. To classify the passenger data using Naive Bayes [4] system used terrorist dataset for classification purpose. Forecasting the urban conditions based on historical data which is also defined in [5] , classification approach to predict the traffic conditions based on collected web data on traffic from web maps. Naive Bayes, j48 and LibSVM modules has used for prediction of system.

Naive Bayes classification method has used for prediction for purchasing methods, defined in [6] which carried out the vehicle purchase prediction system based on various considerable parameters. The Naive Bayes shows around 75% prediction accuracy on large synthetic dataset. Meanwhile publication transportations reconstruction is define in [7]. It propose a state-based Bayesian techniques to the reconstruction of the transport system state from limited available background knowledge. The system is efficient also in the presence of multiple real-world type of noise, such as data blackouts. The NB algorithm defined for classification model with some local attribute weighted based on KNN approach has used in [8], it optimized the public transportations parameters resources, It is also find the rules of the bus lines various people taking, and predict which bus lines different people choosing. The Naïve Bayes (NB) Classifier Based Traffic Detection System using Cloud environment [9] illustrated on identifying the traffic condition by analyzing the behavior of vehicle primarily based on GPS as well as device history data. Entire system has define into the two parts, client as well as Cloud Server. In the Client side, the system distinguishes whether a phone carrier is taking a vehicle or walking.

This study geared to showcasing the world efforts towards achieving transference and moral cities, further as gender problems like athletics in relation to totally different elements of the planet, jointly of the samples of transport organization in good cities. Best transport exists jointly of the first gears driving growth, a trend that has existed for a extended amount of your time [10]. Transportation organization is taken into account to be a worldwide idea. In fact, most developed countries still invest within the physical setting like infrastructure and quality, that square measure thought-about to be core drivers for transport management. Like ne'er before, completing development in good cities needs the mixing of reliable transport systems. The driving issue of good cities is to confirm the affiliation of human wealth, organization and social capital in a very bid to realize a lot of property commercial development further as a more robust mode for residents in these cities. In different words, property transport is one among the driving components as way as achieving good cities is bothered.

Within this study is described the concept of traffic precision-replication. The model field is to spot critical connections between the intermodal transportation systems of the large framework. To this end, one link at a time is investigated by examining the sources of inspiration a suspension has on the representatives of the traffic that move this link unusually. Thus, first, this same model replicates the intermodal transport network as well as its real traffic 8before; secondly, one in each of the network's links is

semi-continuous. A newly revised of traffic is then used to traverse traffic supporters via the network, which is deformity-continuous. Due to the replicated disruptions, the criticality is mainly managed since before the delay time of all traffic representatives inside the network [11].

The multi class classification problem has solved using various machines learning algorithm in [16] [17]. Both system deals with large unstructured data and proceed using supervised learning approach. The achieved results demonstrated they got better classification accuracy than traditional machine learning algorithms.

3. PROPOSED SYSTEM DESIGN

The system carried out in Figure 1 demonstrates transportations system management and interruption detection using supervised learning algorithm on real time dataset which is collaborated form Transport for London (TfL) API's [18]. The entire work predicts interruption of road as well as affected stations using Naïve Bayes algorithm[19], below we define the proposed methodology description which we followed to achieve proposed results.

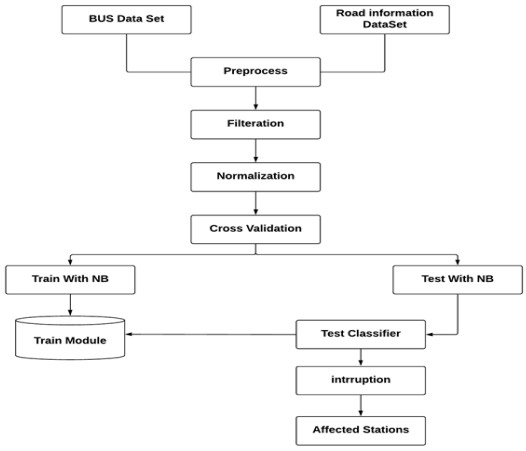

Figure 1: -Proposed System Architecture Design

3.1 Data Extraction form TfL

Using TfL (provided by London Government www.tfl.gov.uk) [20] API's is used to collect the real time transportation dataset. Various transportation dataset are available on given link, we extract Bus and road dataset with custom programming.

3.2 Data Preprocessing

The basic data mining processes has done in second phase, Basically real time data sometimes hold some misclassified instances and null values, using data acquisition, data normalization, feature extraction etc. The normalize data does contains actual values of routes and bus information.

3.3 Disruption classification using Naïve Bayes

Once data preprocessing has done system deals with classification module of system, according the supervised learning approach we apply 10 fold and 15 fold cross validation on entire dataset for training and testing respectively. The classification results will shows actual disruption on roads based on real time data.

3.4 Station affected List

After completion of classification phase, we apply data reduction and own optimization technique to generate the station affected list show as final classification result of system.

3.5 Analysis

In final phase we shows the predicted classification accuracy of system as well as comparative analysis of proposed system with various machine learning algorithm show display the effectiveness of proposed system.

The entire search describes the smart transportation management which is deal with high dimensional large data set in a smart City. This system is also effective to identify the distribution on different roads as well as bus stations. The proposed research basically deals with supervised learning classification with data summarization using naive Bayes classifier. In the preprocessing phase the data reduction whenever it contains miss-classified instances and normalization has done to reduce the dimensionality of feature extraction as well as feature selection. The TfL platform provide the live streaming data and our cloud has intermediate data storage to handle the entire work which provides efficiency and 24*7 data accessibility over the internet. The proposed classification algorithm also provides good classification accuracy for entire system.

4. ALGORITHM DESIGN

4.1 Training using updated NB

Input: Training dataset TrainData[], Various activation functions[], Threshold Th

Output: Extracted Features Feature_set[] for completed trained module.

Step 1: Set input block of data d[], activation function, epoch size,

Step 2 : Features.pkl ← ExtractFeatures(d[])

Step 3 : Feature_set[] ← optimized(Features.pkl)

Step 4 : Return Feature_set[]

Testing using updated NB

Input: Training dataset TestDBLits [], Train dataset TrainDBLits[] and Threshold Th.

Output: Resulset <class_name, Similarity_Weight> all set which weight is greater than Th.

Step 1: For each testing records as given below equation

$$testFeature(k) = \sum_{m=1}^{n} (. featureSet[A[i] \ldots\ldots A[n] \leftarrow TestDBLits)$$

Step 2 : Create feature vector from $testFeature(m)$ using below function.

Extracted_FeatureSet_x [t……..n] = $\sum_{x=1}^{n.}(t) \leftarrow testFeature$ (k)

Extracted_FeatureSet_x[t] holds the extracted feature of each instance for testing dataset.

Step 3: For each train instances as using below function

$$trainFeature(l) = \sum_{m=1}^{n} (., featureSet[A[i] \ldots\ldots A[n] \leftarrow TrainDBList)$$

Step 4 : Generate new feature vector from $trainFeature(m)$ using below function

.

Extracted_FeatureSet_Y[t……n] = $\sum_{x=1}^{n.}(t) \leftarrow TrainFeature$ (*l*)

Extracted_FeatureSet_Y[t] holds the extracted feature of each instance for training dataset.

Step 5 : Now evaluate each test records with entire training dataset

$$weight = calcSim\ (FeatureSetx \,||\, \sum_{i=1}^{n} FeatureSety[y])$$

Step 6 : Return Weight

5. RESULTS AND DISCUSSIONS

We evaluated that our work provides better classification accuracy for various cross fold validation, the Naïve Bayes algorithms provides better classification accuracy than other classification results. Various supervised learning algorithms has used for comparative analysis like ANN [12], Random Forest [13], J48 [14] and K-Means [15] and out of them proposed NB provides best accuracy than other classification approaches.

The Figure 2 to Figure 4 shows the similar experiment analysis of proposed system with some existing systems. Finally this graphs proved the proposed classification provides better accuracy than other machine learning based classification algorithms.

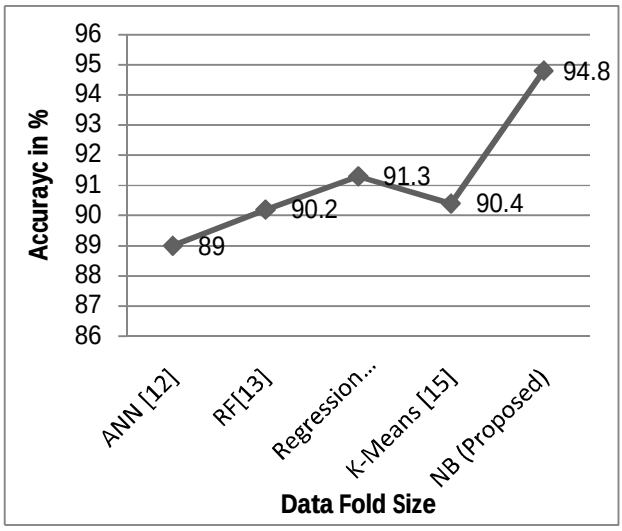

Figure 2 : Classification Accuracy of existing machine learning with proposed Naïve Bayes Algorithm when 10 Fold Cross data Validation

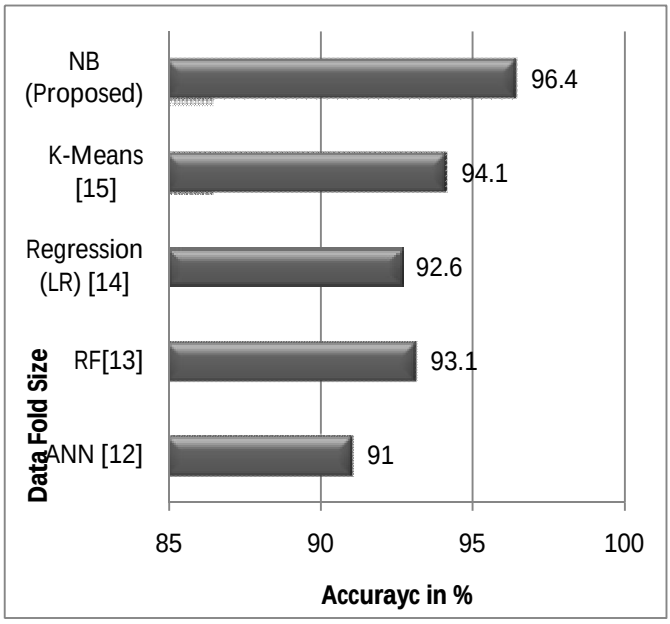

Figure 3 : Classification Accuracy of existing machine learning with proposed Naïve Bayes Algorithm when 15 Fold Cross data Validation

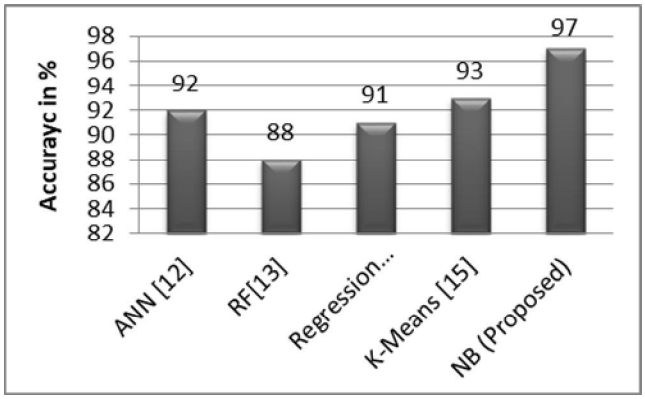

Figure 4 : Classification Accuracy of existing machine learning with proposed Naïve Bayes Algorithm when 20 Fold Cross data Validation

The above sections described experiment analysis and result analysis of system, it provided implementation execution and result calculation of system with proposed algorithm and comparison with existing algorithms. Data extraction form TfL and store into middleware Linode cloud system provides like Service Oriented Architecture (SOA) to system. Various machine learning based supervised algorithms evaluated in system and finally concluded the result.

6. CONCLUSION

The research's main objective is to build a Cloud data model for a multi - modal transportation process that integrates multiple modes with one network, enabling various modal compositions in traffic management. The optimization technique adopted for this work to achieve these objectives is mainly to provide a separate entity with each mode route and to distinguish each of these identities functionally. In various methods the distinction is fixed and variable for the various routes of the same mode. Getting connected these separate things is done by means of connectors which reflect the transferring action through one route to another. A multi - modal network model is built using this principle upon its TfL platform.

REFERENCES

[1] Adrian E. Coronado Mondragon, Chandra S. Lalwani, Etienne S. Coronado Mondragon, Christian E. Coronado Mondragon, Kulwant S. Pawar, **Intelligent transport systems in multimodal logistics: A case of role and contribution through wireless vehicular networks in a sea port location**, International Journal of Production Economics, Volume 137, Issue 1, 2012, Pages 165-175, ISSN 0925-5273
https://doi.org/10.1016/j.ijpe.2011.11.006

[2] Xiong C, Zhang L. **A Descriptive Bayesian Approach to Modeling and Calibrating Drivers' En Route Diversion Behavior**. IEEE Transactions on Intelligent Transportation Systems. 2013 Jul 15;14(4):1817-24.

[3]Kee CY, Wong LP, Khader AT, Hassan FH. **Multi-label classification of estimated time of arrival with ensemble neural networks in bus transportation network**. In2017 2nd IEEE International Conference on Intelligent Transportation Engineering (ICITE) 2017 Sep 1 (pp. 150-154). IEEE.
https://doi.org/10.1109/ICITE.2017.8056898

[4]Singh S, Verma S, Tiwari A, Tiwari A. **A novel way to classify passenger data using Naïve Bayes algorithm (A real time anti-terrorism approach)**. In2016 2nd International Conference on Next Generation Computing Technologies (NGCT) 2016 Oct 14 (pp. 312-316). IEEE.

[5]Chen YY, Lv Y. **Analysis and forecasting of urban traffic condition based on categorical data**. In 2016 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI) 2016 Jul 10 (pp. 113-118). IEEE.

[6]Harahap F, Harahap AY, Ekadiansyah E, Sari RN, Adawiyah R, Harahap CB. **Implementation of Naïve Bayes Classification Method for Predicting Purchase**. In2018 6th International Conference on Cyber and IT Service Management (CITSM) 2018 Aug 7 (pp. 1-5). IEEE.

[7]Pagani A, Bruschi F, Rana V, Restelli M. **Reconstruction of public transport state.** In2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC) 2016 Nov 1 (pp. 2285-2292). IEEE.

[8]Mao X, Zhao G, Sun R. **Naive Bayesian algorithm classification model with local attribute weighted based on KNN.** In2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC) 2017 Dec 15 (pp. 904-908). IEEE
https://doi.org/10.1109/ITNEC.2017.8284867

[9]Aung SS, Naing TT. **Naïve Bayes classifier based traffic prediction system on cloud infrastructure.** In 2015 6th International Conference on Intelligent Systems, Modelling and Simulation 2015 Feb 9 (pp. 193-198). IEEE.

[10] Y.M. Bontekoning, C. Macharis, J.J. **Trip Is a new applied transportation research field emerging?—a review of intermodal rail–truck freight transport literature Transportation Research Part A: Policy and Practice**, 38 (2004), pp. 1–34 Article

[11] S.V. Ukkusuri, W.F. Yushimito **A methodology to assess the criticality of highway transportation networks** Journal of Transportation Security, 2 (2009), pp. 29–46

[12] R Development Core Team R: **A Language and Environment for Statistical Computing R Foundation for Statistical Computing**, Wien (2009)

[13] Nguyen H, Bui XN. **Predicting blast-induced air overpressure: a robust artificial intelligence system based on artificial neural networks and random forest**. Natural Resources Research. 2019 Jul 1;28(3):893-907.

[14] Samara A, Rempe F, Fastenrath U, Göttlich S. **Assessing the probability of arriving on time using historical travel time data in a road network**. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC) 2019 Oct 27 (pp. 1343-1348). IEEE.
https://doi.org/10.1109/ITSC.2019.8916919

[15] M.A.P. Taylor, G.M. **D'Este Transport network vulnerability: a model for diagnosis of critical locations in transport infrastructure systems** A.T. Murray, T.H. Grubesic (Eds.), Critical Infrastructure: Reliability and Vulnerability, Springer, Heidelberg (2007), pp. 9–30.

[16] B.Pandu Ranga Raju, B.Vijaya Lakshmi, C.V.Lakshmi Narayana, **Detection of Multi-Class Website URLs Using Machine Learning Algorithms,** International Journal of Advanced Trends in Computer Science and Engineering, pp. 1704-1712 ,Volume 9, No.2, 2020.
https://doi.org/10.30534/ijatcse/2020/122922020

[17] Roobaea Alroobaea, **An Empirical combination of Machine Learning models to Enhance author profiling performance**, International Journal of Advanced Trends in Computer Science and Engineering, pp. 2130-2137,Volume 9, No.2, 2020.
https://doi.org/10.30534/ijatcse/2020/187922020

[18] Wahul R.M., Lomte S.S. **Conflict Analysis of Smart City Services using Live Data from Cloud** , Journal of Advanced Research in Dynamical and Control Systems, pp. 709-715, 02-Special issue, ISSN: 1943-023X. 2018
[19] Wahul R.M., Lomte S.S. **Smart City Project Management System Using Cloud**. In: Iyer B., Nalbalwar S., Pathak N. (eds) Computing, Communication and Signal Processing. Advances in Intelligent Systems and Computing, vol 810. Springer, Singapore. 2019
https://doi.org/10.1007/978-981-13-1513-8_24
[20] www.tfl.gov.uk