Smart Traffic Management System for smart cities using IWCF Algorithm

D Venkata Siva Reddy,
Research Scholar,
Dept. of Computer Science & Engineering,
SCSVMV University,
Kanchipuram- 631561Tamil Nadu
lionshivareddy@gmail.com.

Dr. R.Vasanth Kumar Mehta,
Associate Professor ,
Dept. of Computer Science & Engineering,
SCSVMV University,
Kanchipuram- 631561Tamil Nadu
Vasanth.mehta@kanchiuniv.ac.in.

Abstract: Recently system used to manage huge traffic is assumed to be working automatically but the actual fact is that to run this system properly at one or the other point manual control is required. There is an urgent need to introduce a system whic is automated fully so that traffic control can be done with even more ease and betterment. Already a lot of work has been done by using internet of things for developing a system which is more advanced and efficient in traffic management.

The present technique is fully automated and is helpful in controlling traffic during emergencies. It is obligatory to initiate certain significant features like identifying parking space, mechanism that provides security against theft, set up sensors which help in identifying or send message regarding the traffic situation, receive certain important message during emergency and pollution control. The proposed architecture includes big data analytics and hadoop. The supervised learning methodology that helps in calculating the average speed and also analyze the passage of the vehicle.

Keywords: Feature selection, Data mining, Filter, Hadoop and IWCF

I.   INTRODUCTION

The traffic controlling system can be considered as one of the complicated system that needs to be improvised and also decrease intricacy lying in organizing road transportation appropriately. The system although appears small but there are a lot of mechanisms needed to make it function properly. There are a lot of features to be considered to attribute a perfect traffic control mechanism in urban / municipalities. With the fast development in technology the pressure on smart transport system is also increasing simultaneously because there is terrific increase in means of transportation day by day which results in substantial trouble in traffic control.

The major problem in the traffic management system lies in finding parking space in busy areas, anti-theft, reducing risk of accidents, control traffic jam during busy hours, using efficient traffic light control, estimate traffic flow, provide road maps etc. At present, GPS system is in wide use for providing necessary information or assistance to the travelers in emergency. When we consider data mining alone, most of its algorithms do not work properly in handling huge data. Here our main objective is to provide a complete automated system that can be applied to improve model performance. For that reason feature selection procedure may be carried out before applying preferred data mining algorithm.

The main challenging feature is to develop a system which is cost effective and faster in performance.

Feature selection method which is also termed as attribute or variable subset selection method is used in opting subset with appropriate predictors or features(1). This technique generally helps in interpreting features uncomplicated and also escapes trouble of dimensionality. This method improves generalization so it is generally used when there are few samples and possesses many features. Attribution or Feature selection helps us in intending new feature subsets (search method) also scores the diverse feature subsets(2). This algorithm checks possible feature subset which lessens the errors that occur. Based on the evaluation search space feature selection algorithm is classified into the Filter method, Wrapper method and Embedded method respectively.

Filter method is used as proxy to score feature subset. This technique is computationally less precise. It works more rapidly and confines significant features of the data. It omits the features with low score. After calculating the feature relevant scores, the residual feature subsets are accessible to compute algorithm (3). This algorithm is simple, works comparatively rapid and does not subject on data mining algorithm. The filter process does not pay attention to the relation with the classifiers. Since it does not regard the dependencies, the categorization may sometime deteriorate. Generally filter method is applied prior to wrapper method so that the wrapper method can work on bigger problems.

The Wrapper method describes analytical type in order to get feature subsets. This method basically works by considering data mining algorithm in selecting attributes. Each and every novel subset is employed to prepare a model which is tested on various sets(4). Quality of the feature subset is measured directly based on data mining algorithm.

The wrapper method is intensively slow as it train new model for every subset but offer finest performance feature set for the testing model. It has the potential to regard attribute dependencies.

In addition to the above two methods there is another method that includes group of practices which executes attribute selection as part of the model building process called the Embedded method. It takes in the classification model (5) . It is less rigorous in calculation than wrapper method.

II. DATA PRE-PROCESSING

Usually the data that is available is frequently imperfect, unfinished, contradictory or lacking some theme or in some cases hold unexpected errors. Thus preprocessing the targeted data is to be done prior to the authentic method.(3) In order to process the raw matter certain methods are to be followed. Initially the data which is to be processed is observed to know whether there are any variables or facts or records lost and try to rectify or refill the missing part (attributes) using mean algorithm or learning algorithm.
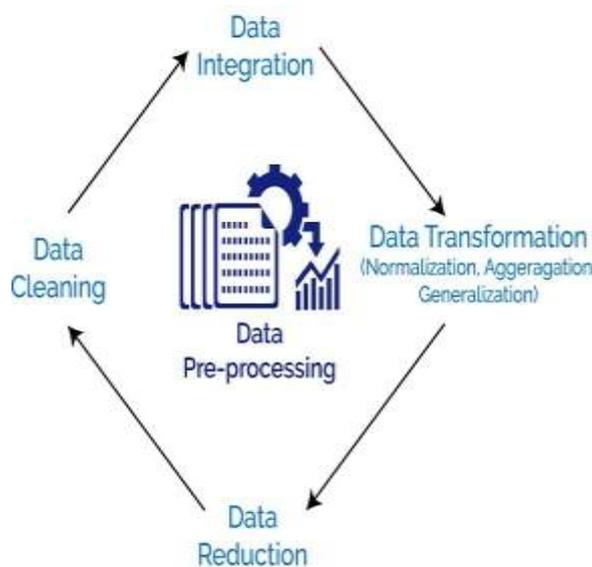

Figure 2.1: Data Pre-processing

In addition being huge, the data may contain noise or outliers or certain inconsistencies. So these are identified and reduced using regression function. This whole process of refilling, re-correcting or smoothing the data is called as data cleaning.

After the first step, then the huge data is organized properly that is rearranging the data properly so that any uncertainty present will be limited (6). This we can term as integration of data. Now the data is structured, planned, ordered and reorganized. This helps to normalize the given data. Then generalization and aggregation of data is also done. This type of preprocessing data is called as data transformation(6). Now unwanted or repeated data is removed and the data is reduced in size. Reduction of data is carried out carefully by retaining the important records
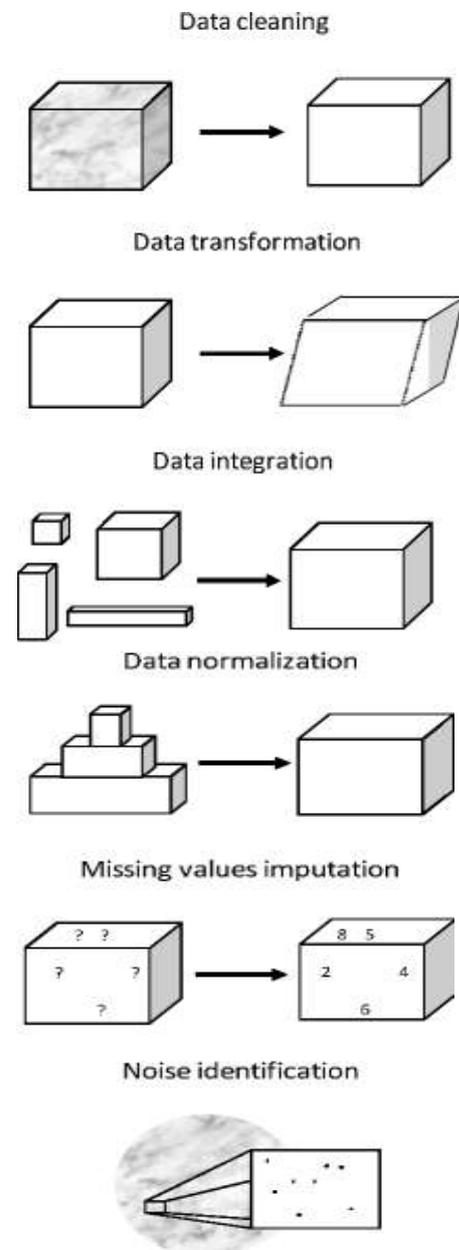

Figure 2.2: Different Steps in Data Preprocessing.

or without changing the quality of the given data. Finally the data is further reduced by substituting some concepts or attributes numerical form by splitting or by using intervals. This method of reduction of data is called as data discretization.
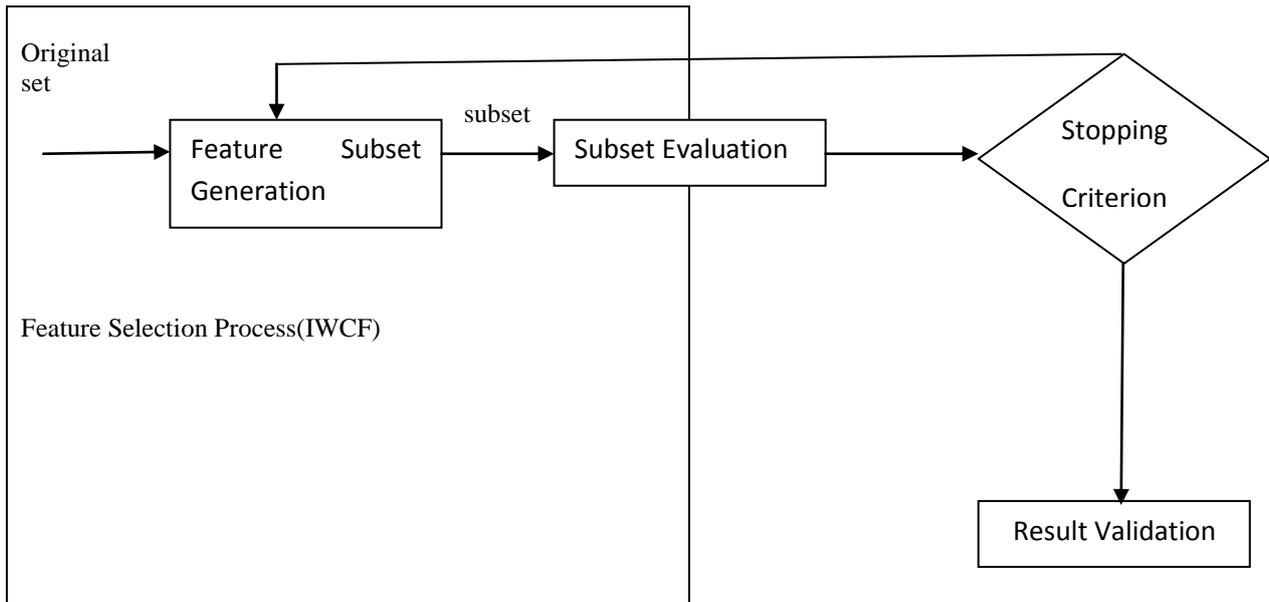
Figure 3.1: Hybrid model

III.   INCREASE OF WEIGHT FOR ALL FEATURES BASED ON CONTINUES FEATURE (IWCF)

Attribute or Feature selection is categorized into three techniques depending on selection approach. First one is the filter method which mainly works to pick the quality or elements or aspects of the subset feature by means of the essential or fundamental characteristics of the given data. Because of this quality this filter method is functional with high aspects. Filter method is comprehensive because it has ability to work with simplicity and with more efficiency. The wrapper technique is the second method which can work more successful and well-organized in selecting the best characteristic in the data.(7) This practice depends on the ideal mining algorithm to execute its action. It is considered to be more costly to apply on massive data. Hybrid method is applied which is the combination of filter and wrapper approach. This helps to grab the virtues that are underlying in both filter and wrapper approach. Selection process is done independently and also by considering mining algorithm. Thus helps in evaluating the virtue of recently formed subset. Filter method is not so precise in producing best class of subset it helps in revealing the actual interaction among the subsets. This method is first applied prior to the wrapper method since it has the capacity to confine the valuable aspects of the subset feature. This helps the wrapper method to concentrate on the key problems prevailing in the subset.

Thus, initially filter method is applied which is followed by wrapper method. Filter approach generally minimize the search space. This results in accomplishing finest subset feature. Selection of attribute or feature, need

training data for studying. The training data can be present either in the form of labeled or as unlabeled data. The big data available may be labeled or unlabeled so the algorithm applied are also divided into supervised which generally uses the labeled form of data and unsupervised form of algorithm that depends on unlabelled data.

IV.   WORKING METHOD OF IWCF

The main aim of this method is to necessitate required information that ensure safe and smooth journey to the tourist. The most important information or data which people want during their travel are data regarding the distance and road map to their destination, traffic conditions, any alarms regarding road signals, toll gates and warnings near accident prone areas etc. The vehicle to infra structure interaction will help the tourist to assemble appropriate information essential for their safe journey. To evade mishaps or any other sort of inconvenience, vehicle to vehicle interaction may be considered. This helps us to know the information regarding vehicles moving in front in advance and also facilitate in adjusting the speed of the vehicle. This vehicle to vehicle communication may be one way that is only receiving the message or may also be capable of retransmit the information. This can be referred as one hop or multi chip sort of communication system.

I.   Characteristics of IWCF
- The model sustains guarantee to predict the location and actions based on the progress of the particular vehicle
- If there is irregular connectivity in the network then the connection should be set to other node so that the connectivity will be established.

- The path and speed of the vehicle helps to define the particular dynamic pattern.
- The well-organized network can be planned depending on the mobility model and node calculation.
- The important application of IWCF is to send message in time to the appropriate nodes during emergencies.
- The conventional algorithm has to get acquainted based on the various mobility circumstances.

V.   CONCLUSION

Depending on the flow of vehicles the traffic signals are controlled. In case of heavy traffic where vehicles are progressing in all the four directions there may be chances to crash. To avoid such miserable situations sensor nodes are connected to the traffic signals. The sensors located in the signals will pass the information regarding the chances of clash to the vehicle way before 5 km from the point to prevent serious mishaps. The sensors also detect the vehicle which is in emergency and show green signal towards that particular direction. So that the emergency condition of that vehicle is satisfied first. During heavy traffic preference is always given to the emergency vehicles and if emergency vehicles are moving in more than one direction then sensor will communicate with the vehicle in much need and allows it to move first. The main principle is to control the traffic properly by avoiding any sort of annoyance or failure in busy roads.

REFERENCE

1. B.M Vidhyavathi, " A New Approach to Feature Selection for Data Mining", International Journal of Computational Intelligence Research, ISSN 0973-1873 Vol.7 Number 3, pp 263 – 269, 2011.
2. Mark A. Hall and Lloyd A. Smith, "Feature Selection for Machine Learning: Comparing a Correlation-based Filter Approach to the Wrapper", Proceedings of the Twelfth International FLAIRS Conference, 1999.
3. H.Liu and H.Motoda, Feature Selection for Knowledge Discovery and Data Mining. Kluwer Academic Publishers, 1998.
4. Zilin Zeng, Hongjun Zhang, Rui Zhang, Youliang Zhang, "Hybrid Feature Selection Method based on Rough Conditional MutualInformation and Naïve Bayesian Classifier", Hindawi Publishing Corporation, ISRN Applied Mathematics, Vol 2014, Articla Id 382738,11 pages .
5. K.Dunne, Cunningham and F.Azuaje, "Solution to instability problems with sequential wrapper-based approaches to feature selection", Journal Of Machine Learning Research,2002. [6] Lei Yu, Huan Liu, "Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution", Department of Computer Science & Engineering, Arizone State University, Tempe, AZ 85287-5406, USA, 2003
6. Hwang, Young-Sup, "Wrapper-based Feature Selection Using Support Vector Machine", . Department of Computer Science and Engineering, Sun Moon University, Asan, Sunmoonro 221-70, Korea, Life Science Journal 2014;11 (7)
7. Yongkoo Han, Kisung Park and Young-koo Lee, "Confident Wrapper-type Semi-Supervised Feature Selection Using an Ensemble Classifier", IEEE, 2011
8. K.Kira and L.A Rendell, "The Feature Selection Problem: Traditional methods and A New Algorithm," Proc. 10th National Conference Artificial Intelligence, pp.129-134, 1992.
9. Kashif Javed, Haroon A.Babri and Mehreen Saeed, "Feature Selection based on Class-Dependent Densities for High Dimensional Binary Data", IEEE Transactions on Knowledge and Data Engineering, Vol 24, No 3, 2012 (www.computer.org/csdl/trans/tk/2012/03/ttk2012030465-abs.html)
10. "Feature Selection Challenge by Neural Information Processing Systems Conference (NIPS), "http://www.nipsfsc.ecs.soton.ac.uk,2003
11. H.Liu and H.Motoda, Feature Selection for Knowledge Discovery and Data Mining. Kluwer Academic Publishers, 1998.
12. Y.Kim, W.Street, and F.Menczer, "Feature Selection for Unsupervised Learning Via Evolutionary Search," Proc. Sixth ACM SIGKDD International Conference, Knowledge Discovery and Data Mining, pp 365 – 369, 2000

AUTHORS BIOGRAPHY

D.Venkata Siva Reddy is currently working as an Assistant Professor, Department of Computer Science and Applications at Besant Theosophical College, Madanapalle, Andhra Pradesh, India. He has 15 years of teaching experience. His research interests are data mining, machine intelligence and image processing. He has published many articles in Scopus, Springer and many peer-reviewed journals and international conferences. He is a life Member in ISC, professional Member in IEAE and also a Member of CSI.

R.Vasanth Kumar Mehta is currently working as an Associate Professor, Department of Computer Science and Engineering at SCSVMV Deemed University, Kanchipuram, Tamil Nadu, India. He obtained his Ph.D. in Computer Science and Engineering from SCSVMV and M.Sc (Tech) and B.Sc in CSE from BITS Pilani, Rajasthan, India. He has 15 years of teaching experience. His research interests are data mining, machine intelligence and image processing. He has acted as program chair, reviewer and program committee member for many international conferences and has published many articles in IEEE, Springer and many peer-reviewed journals and international conferences. He is currently guiding 6 scholars for Ph.D. Dr. RVK Mehta is also a Member of ACM.