Accentuating Oman Traffic Using Big Data Analytics by Excavating Weblog & Social Networking Data

Dr Mohammed Faisal[1], Mallak Al-Abri[2], Shurooq Al Yahyai[3]

[1]Lecturer, Information Technology Department, Nizwa College of Technology, Nizwa, Sultanate of Oman

[2,3]Undergraduate-IT scholars, Information Technology Department, Nizwa College of Technology, Nizwa, Sultanate of Oman

mohammedfaisal@gmail.com

Abstract:
This research has been executed into two similar verticals to extract the best possible accurate result, first vertical has been utilized Hadoop ecosystem to perform natural text processing on tweeter data to conclude the opinion about Oman traffic another vertical has been utilized the RapidMiner capabilities to perform web mining to get insights about Oman traffic. Web mining has been chosen alternatively instead of web log analytics due to unavailability of the web log data. Oman Traffic Sentiment Analysis by using Big Data Analytics project will collect many tweets on Twitter. This project use twitter data for social purpose according to our data requirement and processing the data. Different organizations are using this huge structured and unstructured data for extracting the people's views towards their industrial and business purpose for growing the company. We are taking the same idea to analyze the big data for the betterment of Oman traffic and accident. Oman Traffic Sentiment Analysis by using Big Data Analytics project will take Tweets for people opinion or feedback about Oman traffic as input. Then, Twitter will store tweets in JSON format. Then, Tweets will be collect, aggregate and moving using Flume into Hadoop in which pre-processing is done. After that, use Hive to classify data into positive and for negative opinions generating report. Finally, report generated from Hive output. The intended user will be a member of the public who is interested in the sentiment of the Twitter population with respect to traffic in Oman topic. Users are not expected to have a very high level of technical expertise. In addition, user should be familiar with using social media program such as twitter.

I. INTRODUCTION

The Social locales have a remarkable measure of data. The social media platform are as Twitter and Facebook is the most popular and trendy and definitely we can't ignore the social media posts because people reveals their opinion and sentiments mostly genuinely. The social locales utilized for correspondence reason. Utilizing it the client can share a thought, considerations, feeling, recommendations, and individual circumstances[1]. Twitter is best a result of least word it communicates learned data. It pursues news columnist, political pioneer, motion pictures stars, and specialist. Traffic is a significant issue in numerous urban areas. Online life is a dynamic site which has numerous supporters, utilizing the traffic related tweets attempt to control traffic in OMAN. To execute the real-time traffic location and analysis of the Twitter tweets originating from those regions in the city. Android application to appear and recommend graphical course arrangement of the traffic zone. Utilizing text mining and characteristic language processing to arrange traffic related tweets in order to get new insight which leads to understand the reason of accident, apply tokenization, stop word separating, steaming and steam separating [2]. On another

hand RapidMiner application which is the well-known data science platform has been used to perform web mining from the different web sources[3].

II. PROPOSED INSIGHTS BASED ON TWITTER DATA AND HADOOP ARCHITECHTURE

Oman Traffic Sentiment Analysis by using Big Data Analytics project will take Tweets for people opinion or feedback about Oman traffic as input. Then, Twitter will stores tweets in json format. Then, Tweets will be collect, aggregate and moving using Flume into Hadoop in which pre-processing is done. After that, use Hive to classify data into positive and for negative opinions generating report. Finally, report generated from Hive output [4].

Research pursued in two different vertical, apart from the above analytics has been taken place by using web data by performing the web mining using the data science platform called Rapid Miner to find new insights to find the current reasons of road accidents in OMAN and predict the corrective measure in order to reduce the accident [5].
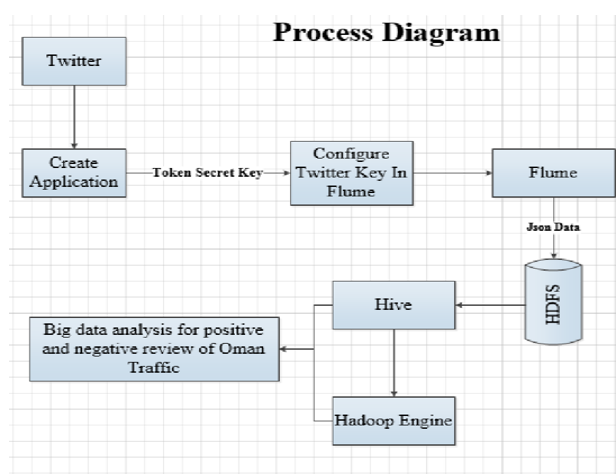

Fig. 1. Prcess diagram

A. PROPOSED INSIGHTS BASED ON WEB MINING USING RAPID MINER

In this context researchers are proposing a

model developed by using RapidMiner where insights can be find to know the reason of the accidents and after knowing the facts recommendation can be done to take the required corrective measures

III. EXPERIMENT AND RESULT

The implementation takes place by using the big data application Hadoop, Hive and Flume. Flume capabilities have been utilized to fetch the twitter data through tweeter application and hive tool have been utilized to generate the queries in order to know the accident reasons in OMAN. Hive and Flume both has been furnished on the top of Hadoop architecture, positive and negative tweets separated to understand about the people opinion. In extend to that analytics have been executed using Hive to know the reasons of accidents.

On another hand to prove authentication of the result we have opted web mining implementation using most popular data science platform called Rapid Miner. In Rapid Miner during data preparation phase data will be acquired from the different web resources and model constructed to get the accident insights.

We note that, the percentage of negative opinion of people about Oman traffic is very near to the percentage of positive opinion of people.
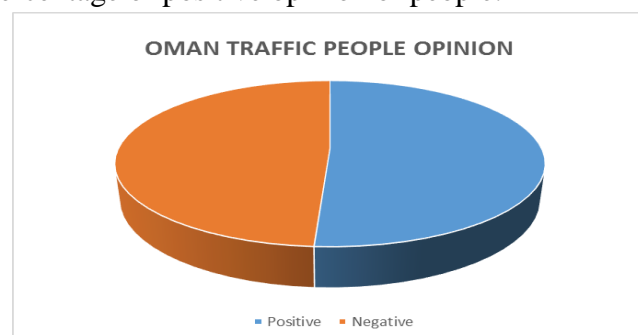

Fig .2. Oman accident and traffic Opinion

Oman should take the possible solutions to solve this traffic problem in order to reduce the negative opinion more and increase positive opinion of people about Oman traffic.

As we note from our fetched data that most tweets contain phone, drink, alcohol, speed and fast words. So because of that we can say that theses the most important reasons for traffic in Oman. As shown on the following:

a. Don not use the Mobile phone while driving:

There is no rhyme or reason to chat on a mobile while you are driving. This is very diverting. Practically when you are driving the vehicle and trying to dial a particular number or picking an incoming call you need to engage your one hand and your mind will be completely diverted which can cause the accident. Indeed, even with a without hands, you will talk, which will divert you from your driving. Set the telephone away, and based on statistics many people being killed we should realize why it is happening very frequently, it is definitely the ignorance of drivers. On the off chance that you do need to make a call, hold up until you are stopped your vehicle.

Fig. 3. Oman accident and
traffic Opinion

b. Never go for drive after consuming the alcohol:

If you are found driving alcoholic in OMAN, you will finish on the expense. Not only will you have a colossal fine to pay, you could even end up going to jail and having your vehicle appropriated.

Your driving license can be suspended for certain period. Exactly when you drunk, you don't have comparative if you are driving within the speed limit you will be save else these activities fuse the ability to drive safely. Such countless people are unwantedly injured or casualties every year in OMAN. Make an effort not to be an estimation. On the off chance that you have wanted to drink infrequently request that somebody pick you, it won't just spare your life yet it will spare others life too.

Fig. 4. Alcohol finding in
tweets using hive query

c. Driver should mind the speed limit:

Exactly when you are driving too much quick, you are genuinely placing yourself in peril. For certain something, in case you have to stop out of the sudden, chances are that you won't have the choice to. Imagine a situation wherein a youth was to run out before you and you were going exorbitantly quick. If you are driving inside the posted speed limit, you won't have this problem, since driving within the limit will give you the ease to stop in emergency. It isn't just individuals by walking you need to fear over either. Another driver may not be centering and run a stop sign,

3230

and you ought to go steadily enough to have the alternative to react so as to control your vehicle [10].

Fig. 5. Speed finding in tweets
using hive query

Oman Traffic Sentiment Analysis by using Big Data Analytics project will take Tweets for people opinion or feedback about Oman traffic as input. Then, Twitter will store tweets in JSON format. Then, Tweets will be collect, aggregate and moving using Flume into Hadoop in which pre-processing is done. After that, use Hive to classify data into positive and for negative opinions generating report. Finally, report generated from Hive output [6].

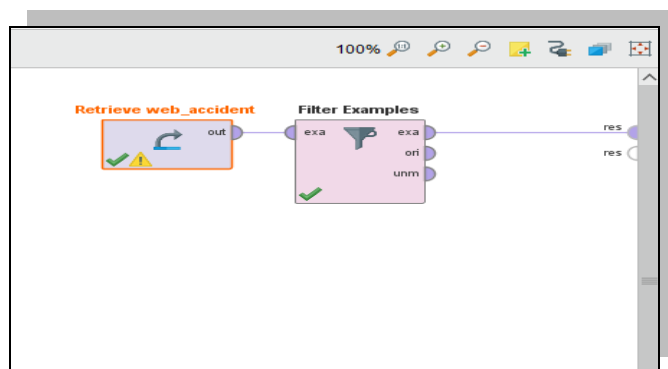*Web Mining using RapidMiner to get accident insights*
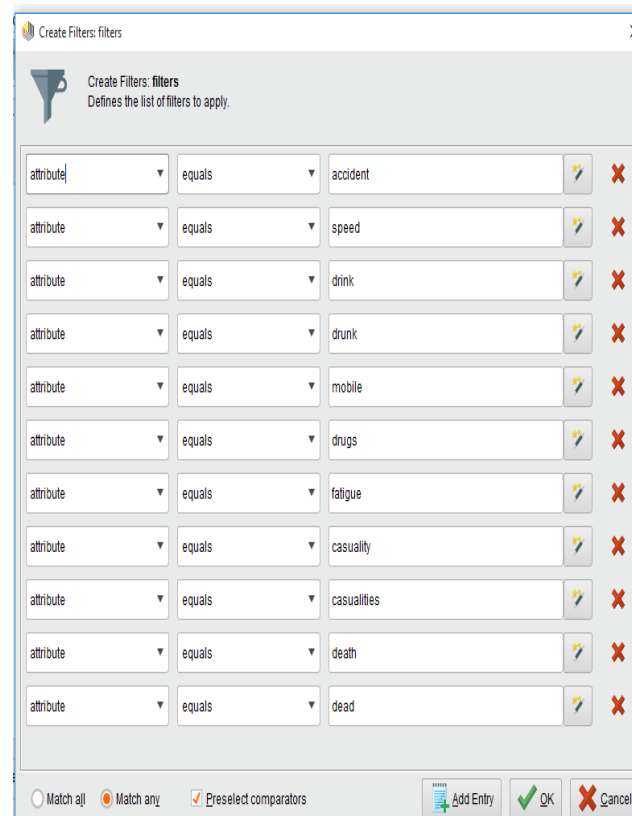

Fig. 6. Filtering the web data

Fig. 7. Filter keyword as accident reason

The above fig 5 and fig 6 illustrated the result of web mining, data has been collected from various web sites and mined the accident reason by using Rapid miner insights capabilities. To get the above insights we have gone through all the phases of knowledge discovery to make the report more user friendly we have presented our result in summarized form [8] [9].

IV. CONCLUSION

Big data analytics and web mining is known to handle variety of data in order to through the accurate insights. This research carried out into two different methodologies using two different but relevant top-notch technologies that is Hadoop ecosystem and RapidMiner. With reference to the above findings, the first dimension of research result mined the opinion of the people on Oman accidents. We found more than 50% people having the positive opinion; the research says it is definitely nowhere concern about the rule but it's

completely result of driver ignorance, in another words we can say the reason of accidents is rash and careless driving. Its majorly happening due to using of mobile phones while driving, crossing the speed limits and driving after consuming the alcohols.

To assure the accuracy of the research finding we have executed the second dimension of the research and decided to perform web mining to know the reason of the accidents. We had performed the data acquisition from the multiple websites including the online news portals to collect the accidental data for the text mining. Based on the web mining we found the common accidents reason that is the same as we found through opinion mining using tweeter data. Based on our results we would like to give the message to the country civilians to take the corrective measure by avoiding the three major factors and authorities need to enhance monitoring capacity to control the accidents.

V.   ACKNOWLEDGEMENT

VI. REFERENCE

[1]. Chandrasekhr Rangue, Shuvojit Chatterjee, Srinivasa Rao Valluru, "Text Mining Approach for Product Quality Enhancement" in IEEE 2017.

[2]. Mr. Piman Barneghi and John G. Breslin , ", Opinion Mining and sentiment polarity on Twitter and correlation between Events & Sentiment", International Conference on Big Data Computing and Application , IEEE 2016.

[3]. Prakash, P., & AR, A. A Review on Different Approaches to Efficiently Detect and Recognize Information from Traffic Panels.

[4]. Judith Sharin Tilsa S, Sobha M.S.," A Survey on Twitter Data Analysis Techniques to Extract Public Opinion.", IJARCSE , Vol. 5 , Issue 11 , Nov 2015 , 2277128X.

[5]. Lokmnyathilak G S S," A Framework for Fast-Feedback Opinion Mining on Twitter Data Streams", IEEE 2015.

[6]. Gunavathi, N. S. R. Sentimental Analysis using Twin Extreme Learning Machine Classifier.

[7]. T. K. Das , D. Acharya & M. R. Pattra, " Opinion Mining about a product by Analyzing Public Tweets in Twitter ", ICCCI- 2014, Jan 03-05, 2014.

[8]. Ninge Zhonge, Yuefeng Li, and Sheng-Tang Wu, "Effective Pattern Discovery for Text Mining", IEEE Transactions on Knowledge And Data Engineering, Vol. 24, No.1, January 2012.

[9]. Deep, S. Intelligent Traffic Management System.

[10]. Andri Sechela, Tien Do Huu, Evangelos Zimos, and Nikos Deligiannis, "Twitter Data Clustering and Visualization", in 2016 23rd International Conference on Telecommunications (ICT), 2016 IEEE.

[11]. Kemoabe, R. K., Gower, E., Ngebani, I., Zibani, I., & Chuma, J. Correlation and Critical Path Analysis based Road Traffic Routing Algorithm.

[12]. Shrute Kohle, Himani Singal, "Data Analysis with R" in 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing.

[13]. Sheikh, S. M., Powder, L., & Ngebani, I. A Smart Microprocessor-based Four Way Stop Road Traffic Controller.

[14]. Arun Jalanela, Nirml Subrmanian, "Comparing SAS® Text Miner, Python,

R" in 2016 IEEE International Conference on Healthcare Informatics.

[15]. Tanveer, S. (2016). Application of Graph Theory in Representing and Modelling Traffic Control Problems. International Journal of Mathematics and Computer Applications Research (IJMCAR) ISSN (P), 2249-6955.

[16]. Dr. S. Vijayrani, Ms. J. Ilamethi, Ms. Nithya, "Preprocessing Techniques for Text Mining - An Overview" in International Journal of Computer Science & Communication Networks,Vol 5(1),7-16.

[17]. Logeswaran, S., & Elavarasan, S. Design of Traffic Intersection at Moolapattarai in Erode City.