

Optimization and Visualization of Opinion Mining and Sentiments In Tourism Dashboard: A Case of Statics Centre Abu-Dhabi (SCAD)

Hamed Saif Albusaidi¹, Prakash Kumar Udipi², Vishal Dattana³

¹ Student, Dept. of PG Studies, Middle East College, Muscat, Oman

² Department of Computing, Middle East College, Muscat, Oman

³ Department of Computing, Middle East College, Muscat, Oman

Abstract— During the last few years, many techniques have been developed [1] which not only optimizes data, but also re-structures data, that stored in traditional databases. These data which is very complex, variated, huge and fast generation can be used for potential decision making for business purposes. To deal with this, data must be optimized first, organized and then extraction of useful data must be done. Several modern techniques are dealing with this data have been emerged. So many new techniques appear in data mining and data analysis as well. These data should be used to support the world economy, and help the decisions makers to make appropriate decisions. This paper focuses on how to use twitter API streaming for analysis. Abu-Dhabi tourism places visitor's sentiments using twitter hashtag analysis to support the tourism industry, which has become one of the most important sectors to support the global economy [2]. The aim is to increase the income of many countries and help these countries to create many jobs. In-addition this sector in many countries has become a focus of attention of investors and businessmen as well as researchers and academicians of various universities across the world [3].

Keywords: Opinion Mining, Twitter API, Twitter sentiments analysis, Data analytic, Dash Board, Smart Tourism, Big data, Twitter hashtag.

I. INTRODUCTION

Information and communication technology (ICT) influence results into generation of various kinds of data and with devices, that run across the advent of the Internet. These data multiply alarmingly and increasing more with the emergence of the Internet of Things, as well as social media. [4]

This data has become difficult to deal with, being so large, complex and diverse data to the point where it has become difficult to use. Hence it has become essential to adjust optimize it, and process it to be usable.



Figure 1. data mining and analysis of tourism

The revolution of big data technology in the world of information and communications technology (ICT), has directly influenced in the business sector so that became the world's largest companies rely on it in the decision-making and improve the performance and products. [5]

The tourism sector is one of the fastest growing sectors in the world, and became the most countries even those countries that do not have natural resources, rely on it to support its economy. Some of these countries have begun to analyse the big data in order to improve tourism services, as well as rely on this technology in their decision-making for the development of the tourism sector, as it provides real-time data.

The Big data analysis may contribute to make very precise decisions, and if the firms done it tap the potential and opportunities well before the others. The mining of this data, analyse it and then present it as a graphics in dashboard, may facilitate to make appropriate decisions at the right time. In order to display the analysed data appropriately, some firms seek to use the dashboard for easy data management.

The social media is platform helps to freely express about views and feelings. Therefore, social media content analysis helps to get accurate data and feedback about all business sectors, including tourism industry. On the other side, Twitter platform is most famous and popular social media platforms. Therefore, Twitter is a rich mine of data consists of large volume of complex information architecture. By analysing this data, one can obtain accurate data and thereby can make various business decisions.

1.1 Social Media users:

Even the month of April 2016 showcased that with the worldwide population of 7.3 billion, the number of Internet users are 3.17 billion and the number of active users in social media were 2.3 billion [6]. Below table showcase the contribution of social media in terms of users.

Social Media	Users	Purpose
Facebook	1.5+ Billion	Building Relationships
WhatsApp	1+ Billion	Building Relationship Conversation
Twitter	320 Million	News and Articles
Pinterest	100 Million	Scrapbooking
WeChat	697 Million	Building Relationship Conversation
Instagram	400 Million	Building Relationship Conversation
LinkedIn	600 Million	News & articles conversation
G+	200 Million	News and Articles

1.2 Why Twitter?

Regarding twitter usage, a report issued by twitter company states that twitter has around 313 million active uses monthly, 1 billion unique visits monthly to sites with embedded in-addition, twitter supported more than 40 languages. [7]

Most institutions around the world adopted Twitter as a window to cross into the world, the fact that this network used by the educated classes around the world, as students of universities and academics and businessmen, press, media and politicians around the world.

Regarding to the questionnaire about using of social media, based on a questionnaire which was published, a total of 151 respondents it shows that 45.7% are prefer to using twitter social network than others of social networks for several reasons, all government and private enterprises are considered twitter accounts as an official and favourite place to publish news and Receive the customer feedback, as well as Twitter more prevalent, and can extract information easily, in addition its ease to use.

While the other are favourite Facebook, Instagram and LinkedIn as shown in the following figure:

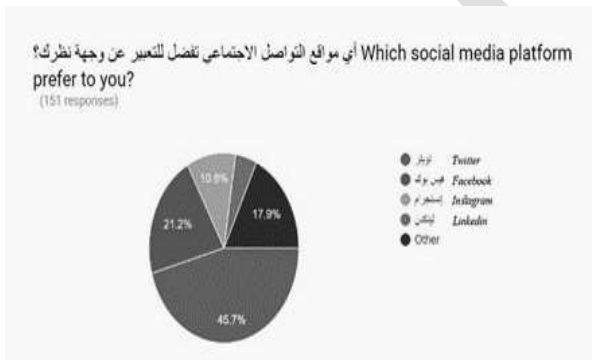


Figure 2 Simple Questionnaire about using of social media

II. REVIEW OF EXISTING SYSTEM

Many studies talked about the Opinion mining and sentiments analysis. This paper discusses how Statistics centre – Abu Dhabi (SCAD), United Arab Emirates functions. This centre is one of the most important entity in the government of Abu Dhabi, where it is the official source for statistical data in the

Emirate of Abu Dhabi. This centre has been created in the twenty-eighth of April 2008 in accordance with law No. (7) in 2008, and under this law, the centre considers an official centre for the development and regulation of statistical work in the Emirate of Abu Dhabi and the production, analysis+ and dissemination of all official statistics on the Emirate. It coordinates the process of integration and harmony based on the statistics produced by the departments and other local agencies. It also coordinates its work with these parties in respect of available statistical information [8].

2.1 The feedback methodologies collection

The centre also responsible for gathering the beneficiary's feedback of various government services provided by state institutions to do this, it resorted into several methods including the surveys carried out by the centre using traditional methods, such as paper questionnaires [9].

Although these methods are ideal at the moment but this method remains is a traditional and during the short-term will not be feasible to use.

III. PROPOSED SYSTEM FOR TWITTER DATA ANALYSIS FOR ABU-DHABI STATICS CENTRE

Everyday a huge amount of data is generated by electronic devices, especially those devices which connected to the internet like smartphones, monitoring cameras, sensors, and all devices, which connected to the internet of things. [10]

During the last few years, social media are appearing to the internet world, to change a traditional way, that is using sentiments expression like traditional discussion forums, blogging, chat rooms and others.

Statistics Centre in Abu Dhabi (SCAD), is responsible to make surveys to find out the views of the tourists of Abu-Dhabi (UAE) citizens or foreigners, residents or visitors, so they are responsible for collecting and analysing their sentiments using several methodologies.

This proposed system recommends the exploration and analysis of tourist's sentiments in Abu Dhabi with Twitter, where twitter platform is a free and very rich platform with data that contains the twitter user's sentiments. It can be through the feature which offered by Twitter to classify twitter user's sentiments to three types, positive feelings, negative feelings, and neutral. After that one can analyse those sentiments to gather feedback about the tourist services in the Emirate of Abu Dhabi.

The recommended system will also describe data mining, cleaning, and then analysing data using Twitter API streaming, which allow to extract data from twitter in real time and then stored in a database before it displayed in the dashboard.

3.1 Creating Twitter Application using API

To creating new twitter application is the first step of development called as new sentiments analysis system. During this stage, the developer should enter all the application details include application name, which is used to attribute the source of a tweet and user-facing authorization, application description and Callback URL [11]. After completing all application details, the Consumer Key (API Key), Consumer Secret (API Secret) will be created.

There are three types of twitter API authentication, which are Search API, Reset API and Streaming API.

In the following table, comparison between Search API and streaming are tabulated.

Features	Search API	Streaming API
Getting real time data	x	√
Getting old tweets	√	x
Getting data from hashtags	x	√
Limitation of getting tweets	√	x

Data mining is to extract data from twitter database in real time for twitter API streaming by Microsoft azure Event Hubs. [12]

The proposed system will use the hashtag #AbuDhabi_Tourism to analyse Abu Dhabi tourist’s sentiments or opinion analysis in real time from twitter when hashtag active.

The following figure showing the processing steps of this application.

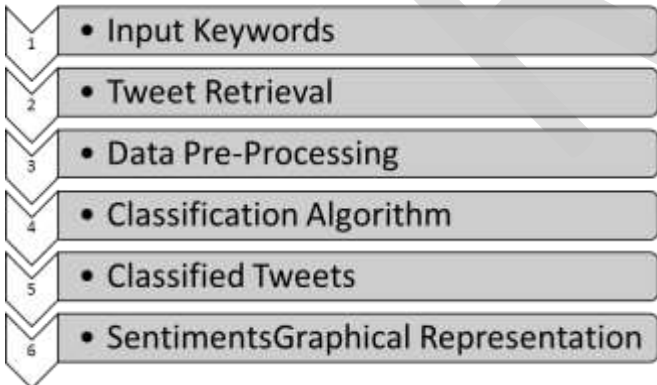


Figure 3 Twitter Sentiments Analysis Processing

3.2 Project Design and Methodology

The best methodology for the development this project is to use of Agile methodology because this methodology allows to make prototypes for the project during the design phase and thus can be easily editing done on the prototype during designing and development phases. [13]

This method is very flexible to add, modify and delete any part in the project during designing and development phases.

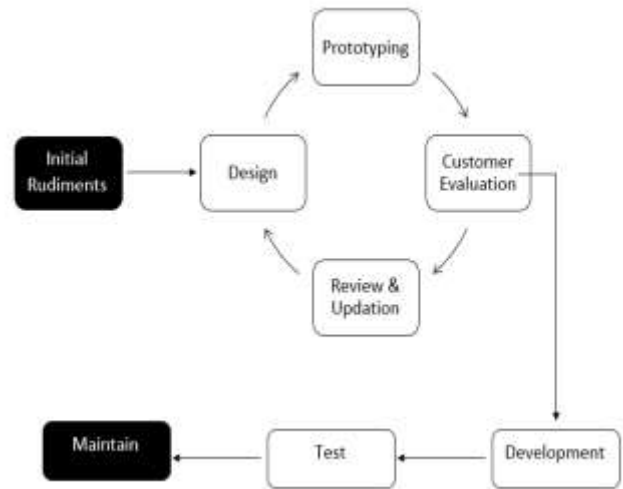


Figure 4 Prototyping Moodle

The proposed application is to mine data then do the analysis using function as shown in the following workflow diagram.

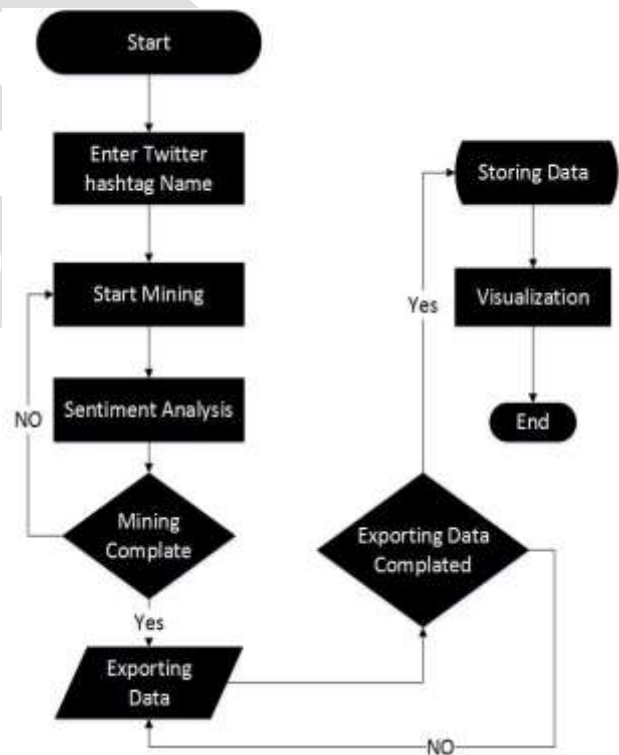


Figure 5 Workflow for analytic of twitter sentiments

where this application has two actors, one is administrator and other is normal user. Starting with administrator, first enter the hashtag name then press enter button. The application now starts to mine the hashtag data, which posts in the real time, and analysis data in three categories (Negative, Positive and neutral tweets). After that the system exports data and store it in database.

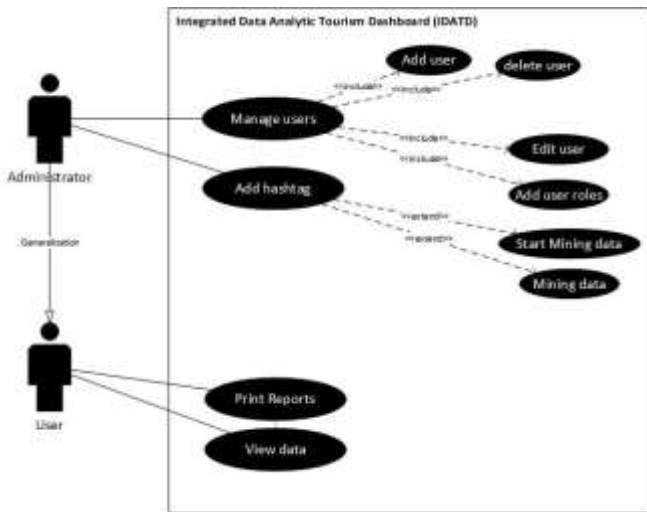


Figure 6. Use-Case Diagram

3.6 Twitter Data Analysis Result

The following figure is shows a simple result of twitter sentiments of tourism data capturing and analytics.

This console screen is providing the primary capturing data before showing on dashboard, which contains date and time of capturing tweets, the title of hashtag, tweets texts and score of tweets where 0=Negative tweets, 1=Neutral tweets and 2=Positive Tweets.

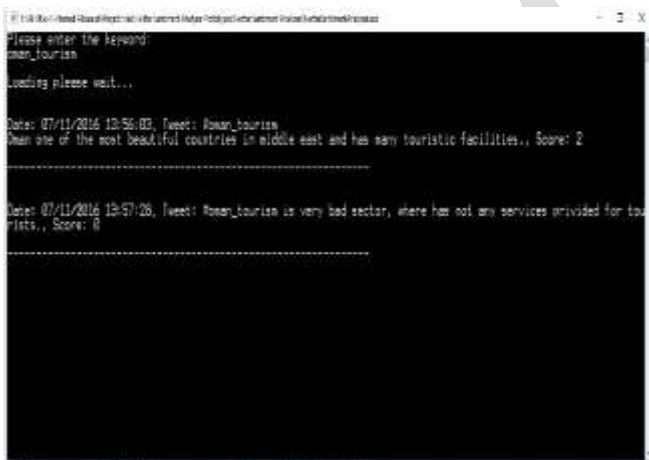


Figure 7 Simple Result of twitter sentiments of tourism data capturing and analytics

3.5 Development Dashboard

There are three types of dashboard- Strategic, Tactical and Operational.

The following table describes the functionality of these dashboards

Dash Type	Functionality	users
Strategic	Use to monitor	Executive / Board
Tactical	Use to Analysis	Manager / Analysts
Operational	Display details	Operation Staff

Dashboard application will have developed using Asp.Net with programming language C#, this application connected with SQL server database.

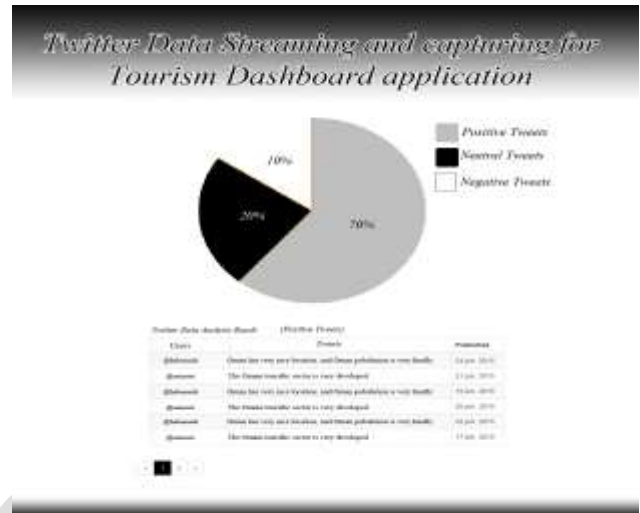


Figure 8 Twitter Data Streaming and capturing Dashboard

IV. CONCLUSION

The Outcomes of this research focuses on how to improve the result of sentiment analysis of the tourists in the Emirate of Abu Dhabi through using social media data. The proposed system specifically recommends Twitter platform using the Twitter Streaming API (Application Programming Interface) that provided by twitter company, and allow the developers to access to the twitter data, analyse data and Twitter hashtags in real time when the hashtag activated.

The result shows data extracted from twitter hashtag #AbuDhabi_Tourism that, twitter users can express their sentiments more freely than if they had filled out a questionnaire about the tourist services provided in the Emirate of Abu Dhabi. The analysis result also showed that most twitter users expressed their views more freely and some users back more than once to Twitter hashtag.

Therefore, one can conclude that Twitter is a platform with full freedom to express their sentiments. Through twitter one can get the perspectives of tourists in all sincerity and when analysed these data, better business decision can be made to enhance the tourism industry services.

REFERENCES

- [1]. Nazarian, A., 2014. *The huffington post*. [Online] Available at: www.2014.com/angella-nazarian/the-technology-revolution-b-4809786.html [Accessed 13 December 2016].
- [2]. Turner, R., 2012. *The Comparative Economic Impact of Travel & Tourism*, s.l.: American Express .
- [3]. Xhiliola Agaraj, M. M., 2016. Tourism an Important Sector of Economy Development. *ideas*.
- [4]. Schreyer, P., 2000. *THE CONTRIBUTION OF INFORMATION AND COMMUNICATION TECHNOLOGY TO OUTPUT GROWTH: A STUDY OF THE G7 COUNTRIES*, Paris: The

- Organisation for Economic Co-operation and Development (OECD).
- [5]. Jacques Bughin, M. C. a. J. M., 2007. Clouds, big data, and smart assets: Ten tech-enabled business trends to watch. *McKinsey Quarterly*.
- [6]. Chaffey, D., 2016. *Smart Insights*. [Online] Available at: <http://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/> [Accessed 13 December 2016].
- [7]. Newberry, C., 2016. *hootsuite*. [Online] Available at: <https://blog.hootsuite.com/twitter-demographics/> [Accessed 13 December 2016].
- [8]. SCAD-Authors, 2016. *Statistics centre – Abu Dhabi (SCAD)*. [Online] Available at: <https://www.scad.ae/en/pages/aboutus.aspx> [Accessed 13 December 2016].
- [9]. Alzaabi, M. M., 2016. *Data Collection methodologies in SCAD* [Interview] (9 October 2016).
- [10]. Cloudera-Authors, 2016. *Cloudera*. [Online] Available at: http://www.triforce.com.au/pdf/Cloudera_Case_Study_Network_Service_Provider_Machine-Generated_Data.pdf [Accessed 13 December 2016].
- [11]. Gray, I. A., 2016. *How to Register a Twitter App in 8 Easy Steps*. [Online] Available at: <https://iag.me/socialmedia/how-to-create-a-twitter-app-in-8-easy-steps/> [Accessed 13 December 2016].
- [12]. Contributors, 2016. *Microsoft Azure*. [Online] Available at: <https://docs.microsoft.com/en-us/azure/hdinsight/hdinsight-hbase-analyze-twitter-sentiment> [Accessed 13 Decemebr 2016].
- [13]. Warsta, J., 2002. Software development methods.

IJRIS