

Opinion Mining on Tourism

S. Kavitha

Assistant Professor

Department of Information Technology

Kumaraguru college of Technology, Coimbatore-India

S. Sathyavathi

Assistant Professor

Department of Information Technology

Kumaraguru college of Technology, Coimbatore-India

S. Prabhakaran

Student

Department of Information Technology

Kumaraguru college of Technology, Coimbatore-India

S. Swathi

Student

Department of Information Technology

Kumaraguru college of Technology, Coimbatore-India

R. Rajkumar

Student

Department of Information Technology

Kumaraguru college of Technology, Coimbatore-India

Abstract

Online tourism forums and social networks have become the most popular platform for sharing tourism related information, with enormous numbers of reviews posted daily. This paper proposes a platform for extraction and summarizing of opinions expressed by users in tourism related online platforms. Extracting opinions from user generated reviews are useful for clients looking for accommodation with the suitable climate and the affordable money. The proposed system extract the reviews from internet mainly focused on twitter and classify them, using an opinion mining technique. Platform is evaluated using a manually pre-classified dataset of user reviews. The proposed system retrieves a collection of reviews about tourist locations posted as tweets in twitter. The client enters the location that he wants to search and the month as an input through the application. The opinion of the people about that particular tourist spot is analysed using rapid miner and their sentiments are classified as positive, negative, neutral and suggested to the people. For classification, the classification algorithm named support vector machine(SVM) have been used. Finally, the recommended locations for the specified month have also been suggested to the client.

Keywords: Opinion Mining, Opinion Mining on Tourism

I. INTRODUCTION

Interest in Opinion Mining has been growing steadily in the last years, mainly because of its great number of applications and the scientific challenge it poses. Opinion mining is an interesting area of research because of its applications in various fields. The opinions of users are helpful for the public and for stakeholders when making certain decisions. Opinion mining is a way to retrieve information through search engines, Web blogs and social networks. Because of the huge number of reviews in the form of unstructured text, it is impossible to summarize the information manually. Accordingly, the resources and techniques to help tackle the problem are many, and most of the latest work fuses them at some stage of the process. Sentiment mining is a field of text mining to determine the attitude of people mining. One important issue is that opinions could be in different languages. The important thing in opinion mining is to extract and analyse the feedback of people in order to discover their sentiments. Growing availability of opinion-rich resources like online blogs, social media, review sites; raised new opportunities and challenges. People now can actively use information technologies to search the opinions of others. The main objective behind this system is that tourist can easily extract subjective and useful information based on the reviews available in the web services.

II. EXISTING SYSTEM

In the existing system, the opinions have been suggested to the clients only based on the rating rate not on the reviews posted by the people on the social media. The main drawback of the existing system is that the opinion of the people about particular location cannot be predicted exactly based on the rating. The positive, negative and neutral reviews of the people have not been analysed properly and hence it could not be more efficient.

A simple number on a rating system is not providing enough information, but neither along review in which users express opinions about more than tourist location, climate etc... There are a lot of reviews problems, which make them difficult to evaluate.

Some of them are

- Reviews are not concise

- Scalar reviews make difficult to compare hotels with different services offered
- Reviews refer to more than simple hotel accommodation
- Totally different opinions from one user to another
- Some aspects are more important so overall rating is not objective but more influenced on that aspect.

III. PROPOSED SYSTEM

Due to rapid growth in tourism, tourists like to find the information regarding different places. With a rapid growth in web, tourists generally share their reviews on social website. These websites have become major source of information for tourism but due to huge data and mixed emotions it becomes difficult for a tourist to make a decision regarding the travelling. This paper proposes a set of summarization methods to help the users digest the vast availability of opinions in an easy manner. This system also reduces the time required for searching and easily supports growth of tourism. This information available from social network is beneficial for business analyst for mining the user opinion about their products and considers these opinions as feedback to improve their policies, planning and process for product development.

Our system that could summarize the reviews, extracting the opinions from all this information, offering an overall perspective, would save a lot of time and ease the decision process for consumer.

IV. IMPLEMENTATION

Rapid Miner is a software platform developed by the company of the same name that provides an integrated environment for machine learning, data mining, text mining, predictive analytics and business analytics. It is used for business and commercial applications as well as for research, education, training, rapid prototyping, and application development and supports all steps of the data mining process including data preparation, results visualization, validation and optimization.

The data have been collected from twitter using authentication key in rapid miner as shown.

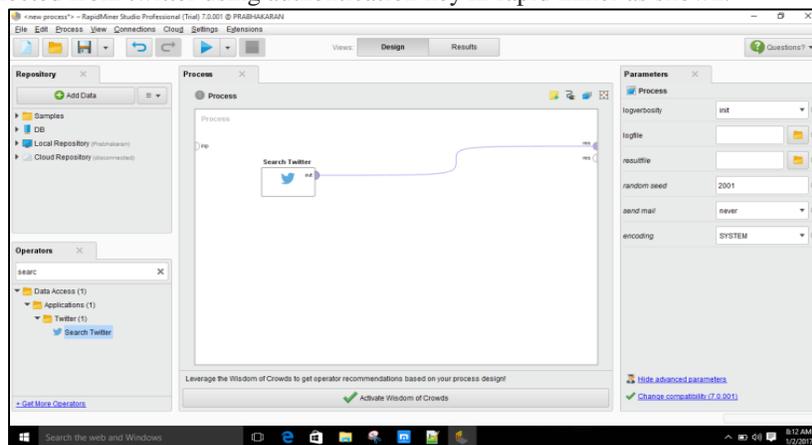


Fig. 1: Data Collection

The collected data appears as shown

Row	ID	Created-At	From-User	From-User...	To-ID	To-User...	Language	Source	Text	Geo.	Geo-L.	Retweet-Co...
1	81574	Jan 2, 2017...	muhammad...	1177878342	?	-1	en	<a href...	Toda...	?	?	0
2	81574	Jan 2, 2017...	Michael Vas...	106783540	?	-1	en	<a href...	Toda...	?	?	0
3	81574	Jan 2, 2017...	Olma Cox	244478205	?	-1	en	<a href...	#oty...	?	?	0
4	81574	Jan 2, 2017...	King Julian	2209802913	?	-1	en	<a href...	RT @...	?	?	289
5	81573	Jan 2, 2017...	Kay	2203757476	?	-1	en	<a href...	Kaa...	?	?	0
6	81572	Jan 2, 2017...	RyderGabri...	1217554484	?	-1	en	<a href...	Ody...	?	?	0
7	81570	Jan 2, 2017...	Yasir Moha...	2346809186	?	-1	en	<a href...	The...	?	?	0
8	81567	Jan 2, 2017...	LeahBerja...	1238540569	?	-1	en	<a href...	Ody...	?	?	0
9	81563	Jan 2, 2017...	LoriW	27709848	?	-1	en	<a href...	RT...	?	?	1
10	81562	Jan 2, 2017...	Charleigh	3013118513	?	-1	en	<a href...	Acc...	?	?	1
11	81562	Jan 1, 2017...	BrodyCynthia	1217443232	?	-1	en	<a href...	Cab...	?	?	0
12	81561	Jan 1, 2017...	Samuel Lun...	4010874514	Rate L...	1773288	en	<a href...	@Stak...	?	?	0
13	81561	Jan 1, 2017...	Hogne'sh	286289256	?	-1	en	<a href...	#That...	11.4...	76.649	0
14	81561	Jan 1, 2017...	Lewis Mont...	2727385620	?	-1	en	<a href...	RT @...	?	?	1
15	81560	Jan 1, 2017...	Raj	180351949	?	-1	en	<a href...	Amaz...	?	?	0
16	81560	Jan 1, 2017...	Bhanssem mt	833365255	?	-1	en	<a href...	CC...	?	?	0

Fig. 2: Output for data collection

A. Sentimental Analysis

For analyzing the sentiment from the collected data, the following function in rapid miner has been used.

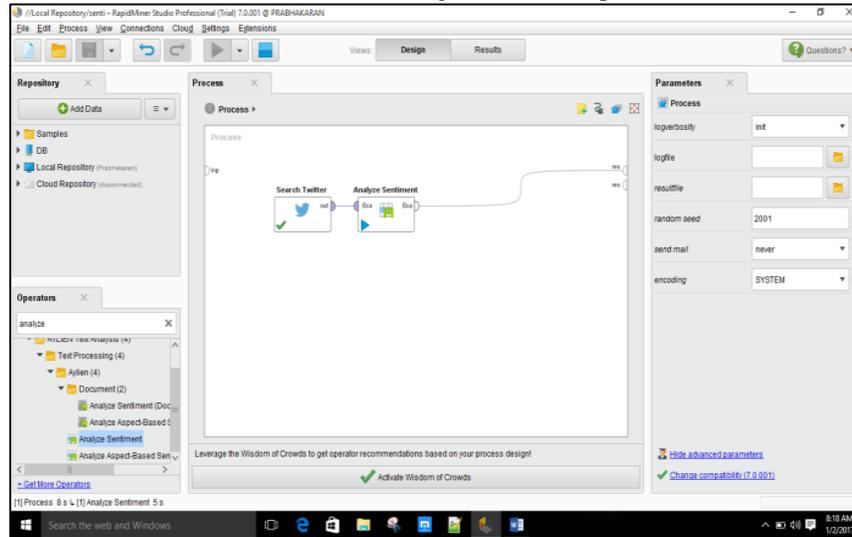


Fig. 3: Sentimental Analysis

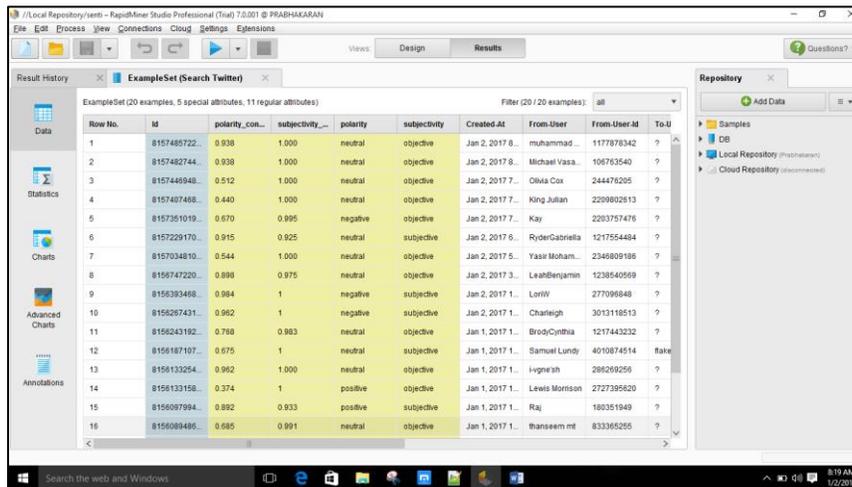


Fig. 4: Polarity Classification

Convert data to documents and then categorize the sentiments and then again convert the document back into data.

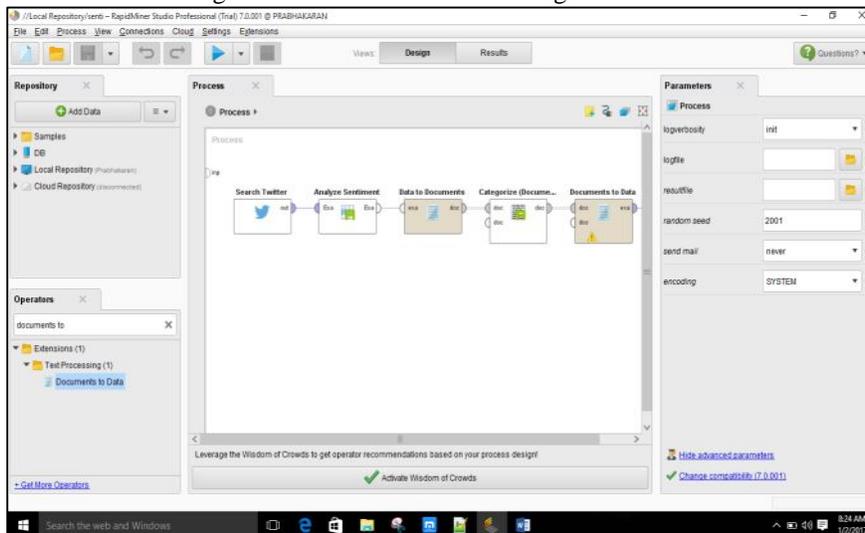


Fig. 5: Document Categorization

We can view results in statistical and graphical formats like histogram, pie chart , bar chart etc.,

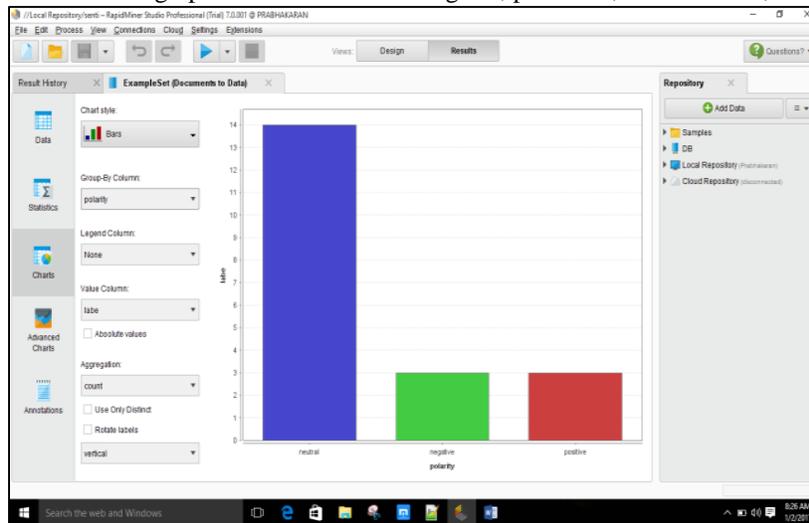


Fig. 6: Output for classification

V. CONCLUSION

In this paper it was presented an opinion mining platform for extracting and classifying the reviews posted by users on twitter. Here, we have classify and segregate the reviews based on certain categories so that the user can easily extract useful and necessary information. This also enhances the growth of tourism.

REFERENCES

- [1] Aarti Potdar, Pranav Patil, Raunak Bagla, Rohitashwa Pandey, Prof. Nagesh Jadhav “Sentiment Based Product Review Analysis System” 11-12 December 2015
- [2] Aditya Bhardwaja, Yogendra Narayanb , Vanrajc , Pawana , Maitreyee Duttaa “Sentiment Analysis for Indian Stock Market Prediction Using Sensex and Nifty” 2015
- [3] Chetan Mahajana, Preeti Mulay “Effective Emoticon Extractor for Behavior Analysis from Social Media” 2nd International Symposium on Big Data and Cloud Computing (ISBCC’15)
- [4] Emma Haddia , Xiaohui Liua , Yong Shi “The Role of Text Pre-processing in Sentiment Analysis” ITQM2013
- [5] Lin Zhanga,,Kun Huac ,Honggang Wangd ,Guanqun Qiane ,Li Zhanga “Sentiment Analysis on Reviews of Mobile Users” MobiSPC-2014
- [6] Utkarsh Srivastavaa , Santosh Gopalkrishnan “Impact of Big Data Analytics on Banking Sector” ISBCC’15
- [7] Abinash Tripathya, Ankit Agrawalb, Santanu Kumar Rathc “Classification of Sentimental Reviews Using Machine Learning Techniques” ICRTC-2015