Proceedings of the 2nd International Conference on Trends in Electronics and Informatics (ICOEI 2018)
IEEE Conference Record: # 42666; IEEE Xplore ISBN:978-1-5386-3570-4

Sentiment Analysis on top five Cloud Service Providers in the Market

Anupriya Koneru[*], Nerella Bala Naga Sai Rajani Bhavani[+], K. Purushottama Rao[*], Garikipati Sai Prakash[+], Immadisetty Pavan Kumar[+] and Velimala Venkat Kumar[+]
Department of Information Technology
Lakireddy Bali Reddy College of Engineering,
Mylavaram, India
[*]Assistant Professor
[+]Student of UG
[*]{smartykoneru, k.purushottamaa}@gmail.com

***Abstract*— A wide range of data is available from social networking sites. The collected data can be used to extract the opinion of the users on products, services or policies. In the Cloud Market, there are number of Cloud Service Providers who offer different kinds of services. It could be difficult for the Cloud Customer to decide the best provider. In this paper, we extracted the Cloud customers' opinion on the services provided by the cloud using Twitter. On this extracted data, Naive Bayes algorithm is applied to evaluate the polarity of the Cloud Service Provider. We applied this work on top Five Cloud Service providers in the market namely Amazon, Microsoft Azure, Sales force, IBM Cloud and Google cloud platform. As a result, Microsoft got more positivity than other providers.**

Keywords*— *Cloud Service Providers, Cloud market, Cloud Customers, Sentiment Analysis.

I. INTRODUCTION (*HEADING 1*)

Sentiment analysis is the way toward recognizing whether the feeling communicated in the content is positive or negative about a given subject. It is incredibly profitable in web-based social network checking as it causes us to know the review of the more broad general opinion behind particular subjects. It is to assess spoken language to decide whether the expression is good, bad or neutral. Now researchers are using algorithm based sentiment analysis to analyze the large volumes of customer feedback. Customer feedback is taken from any social media like Twitter, Face book etc. Be that as it may, it isn't sufficient to comprehend what clients are discussing. You should likewise know how they feel. Sentiment analysis is one approach to reveal those emotions. It is also known as "opinion mining".

Cloud is just like a virtual warehouse which contains a lot of data that can be used by IT resources through the internet. The league that provides cloud services such as IaaS, SaaS, and PaaS, are said to be cloud providers. All these three services like SaaS, PaaS, and IaaS are provided using virtualization technologies. The benefits of using cloud service providers are due to of its efficiency and economies of scale. Top five cloud service providers available in the marketplace are:

- Amazon Web Services (AWS)
- Sales force
- IBM
- Microsoft Azure
- Google cloud platform

A. Amazon web Services (AWS):

In 2006, Amazon Web Services (AWS) [1] was impelled. It is a standout amongst other cloud master communities in the market. Similarly, Amazon was the key cloud supplier to offer Infrastructure as a Service (IaaS), enabling people and relationship to lease virtual PCs. It furnishes infrastructure with minimal effort which is utilized for a huge number of organizations in 190 nations around the globe.

B. Microsoft Azure

Microsoft is a moment driving Provider after Amazon. Microsoft Azure [2], is otherwise called Windows Azure, which is Microsoft's open distributed computing platform. It was first discharged in February 2010. It serves IaaS and PaaS delivery models. By adopting these services the customers could build the applications, deploy the applications and manage the applications.

C. Sales force

Sales force [3] is a third leading cloud service provider after Microsoft Azure. Marc Benioff founded sales force in March 1999 Sales force and it is a top leading CRM software (Customer Relationship Management) cloud serves all its data. It handles more than 750 applications which are known for their support in various features like generating new leads, acquiring new leads, increasing sales. It was designed for the management of organization's data focused on customer and sales details. It offers features that customize its inbuilt data structures and GUI which suits the specific needs of a business. In recent days, it also started offering the IOT (internet of things) connectivity to the CRM platform.

D. IBM cloud

International Business Machines which is also called as IBM [4] ranks fourth provider after sales force among the world it provides a wide range of offerings that contains software, hardware and services. Global Business Services,

Global Technology Services, networking services, business continuity services and outsourcing services are the services included. IBM cloud came into force to purchase companies that are offering cloud consulting services and implementation services. IBM is well known for customer relationship management which contains a collection of different clouds like Sales Cloud, health cloud etc.

E. Google Cloud Platform

Google Cloud Platform [5] provides services that is offered by Google such as public cloud computing. In order to build and host applications, store data, websites, and analyze data on Google's scalable infrastructure. It offers of Iaas, PaaS and SaaS services. Software developers, cloud administrators and other IT professionals can access the services provided by the google through internet or any secured network connection.

In the following section we discuss the Problem definition. Section 3 provides the related work on the sentiment analysis. Section 4 describes the methodology adopted to extract the opinions. The final results are shown in Section5. Section 6 outlines the discussion points. Finally Section 7 conveys the conclusion remarks.

II. PROBLEM STATEMENT

There are number of Cloud Service Providers in the Market. But selecting the best provider is really a challenging task. To select the best one out, Customers contact the Broker [6] to decide. But it generates a layer among Provider and Customer which increases the complexity. So, there should be a direct way to decide the best provider.

III. LITERATURE REVIEW

Harshali P. Patil, Mohammad Atique [7] describes about sentiment analysis as opinion mining, which is a method used to find the opinions incarnated in text automatically, and this is becoming a challenge in many research areas, particularly in data mining field. It gives an over view of latest updates of sentiment analysis and classification methods and also discusses about the challenges of sentiment analysis. K. Mouthami et.al [8] discussed that Sentiment analysis is a process which is used to determine the attitude, emotions and opinions that are revealed by any person about a particular topic. It is done on movie reviews and implemented using sentiment fuzzy classification. M. S. Neethu, R. Rajasree [9] deals mainly on techniques of machine learning and those techniques use a test set and training set for classification. This classification can be done by using naïve bayes, Maximum Entropy, Support Vector Machine classifiers. And these Machine Learning techniques are simple and efficient compared to many other techniques. These techniques can also be applied for twitter sentiment analysis.

In [10] this paper, Author presents a mechanism to predict the overall sentiments or opinions of Indian people towards political situation and issue. This process of sentiment analysis uses some classification techniques like Random Forest, baysNet, Naive Bayesian and shows that which classifier gives good accuracy. In this paper [11], they considered two top leading cloud service providers namely Microsoft Azure and Amazon Web Services. To that they calculate the sentiments of tweets by using naïve bayes classifier for both polarities and emotions and analyze which cloud service provider is best.

Jyoti Ramteke, Samarth Shah,Darshan Godhia, Aadil Shaikh[12] collected twitter data set by two frameworks i.e., hash tag clustering and vadar. Two algorithms are applied to calculate sentiment analysis that is Support Vector Machine, Multinomial Naive Bayes to determine the polarity. Ajay Deshwal, Sudhir Kumar Sharma [13] deal with many feature extraction techniques like unigrams, emotions and word gazetteer, question mark symbol, exclamation are used to design exact sentiment classification system. This work helped to compare the result of six supervised classification algorithms and tells which algorithm gives effective performance.

Deepali Arora, et.al [14] applied sentiment analysis on Twitter data to rate users' response on popular smart phone brands and their operating systems. To analyze these results, lexicon based sentiment analysis is used.

The performance of machine learning Text classification using machine learning algorithms is well-versed [15]. Pang and Lee [16] investigated on the performance of various machine learning techniques in a specific domain of movie reviews. It is used to provide best accuracy on Naive Bayes classifier.

According to the work done by M. Vadivukarassi, N. Puviarasan and P. Aruna [17], explained about sentiment analysis which is a current research area in text mining and stem of natural language processing or machine learning techniques. It is important for decision making and can be removed, identified, examined from the online sentiment reviews. It tells about how to connect on Twitter and search for the tweets that contain a particular keyword and then calculate the polarity of the tweets as positive, neutral land negative. Pre-processing was done using Natural Language Toolkit techniques. In order to select the best feature Chi Square test is used and Naïve Bayes classifier is also used for training and testing the features.

In [18] this, author applied the Dictionary based approach on twitter data for the extraction of opinion on public on GST in India.

IV. METHODOLOGY

In this work, customer can directly select the best provider by considering the tweets given for the provider services in the twitter. Figure1 shows the complete work flow of this method.

A. Data Collection

In this work, we have chosen top five cloud service providers that are available in the market place i.e. Amazon web services, Microsoft Azure, Sales force, IBM cloud, Google Cloud Platform. We extract the tweets of these cloud service providers using twitter API. Twitter API is connected to twitter using authentication keys and then fetches all recent tweets of those five cloud service providers. For each cloud service provider, we collected tweets by using screen name of that particular cloud service provider which is shown in Table 1.

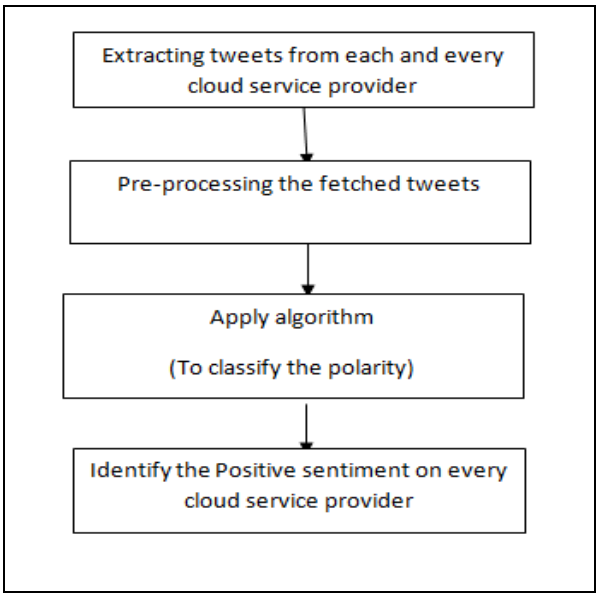

Fig. 1. Work Flow Diagram

B. Data Preprocessing

Preprocessing is a major task in data analysis. Tweets contains unnecessary details like acronyms, URLs, hash tags and emotional symbols. In order to remove them, we apply preprocessing techniques. For preprocessing of tweets, we have selected three stages for filtering of tweets: Tokenization, Stop words removal and Regex matching for special character removal.

- Tokenization

The words in the tweet are separated individually in the absence of space and irrelevant symbols like emotions.

- Stop words Removal

Words which does not specify any emotion (is, or, the) are called stop words. Words which are not helpful to analyze the sentiment are removed.

- Regex matching and special Characters removal

Regex matching is implemented in python to remove URLs in the tweet and also remove the special characters like @, #. For example @microsoft will be replaced as Microsoft. Words like goooooood! are replaced with good! After completion of these three steps, now tweets are ready for sentiment classification.

C. Classification

Every pre-processed Cloud Service Provider tweet set is given as an input to the classifier. In this, the Naïve Bayes classification algorithm is applied to categorize the positive and Negative tweets. In python, there is a package called Text blob to implement the naïve bays algorithm.

Table 1: Data size

s. no	Provider name	Provider website id	No. of tweets
1	Microsoft azure	https://azure.microsoft.com/en-us/	3335
2	Amazon	https://aws.amazon.com/	3223
3	Sales force	https://salesforce.com/	3135
4	IBM Cloud	https://www.ibm.com/cloud/	3348
5	Google cloud platform	https://cloud.google.com/	3432

V. Result Analysis

A. Environment

To implement this work, we used Python 3.5 which is installed on Personal Computer with 4GB RAM on Intel i5 processor. Pip is installed to use of different packages.

B. Results

In this step we extracted the positive, negative and neutral polarities of the five cloud service providers as shown in Table2. The polarities have been compared and selected the best provider which is having high positive polarity.

Table 2: Polarity Comparison for five cloud service providers

Cloud Provider	Positive polarity	Negative polarity	Neutral polarity
Amazon	71.50%	5.50%	23.00%
Microsoft Azure	72.00%	4.50%	23.50%
Salesforce	64.00%	6.00%	24.00%
IBM Cloud	55.50%	9.50%	35.00%
Google Cloud Platform	43.50%	5.50%	51.00%

Figure 2 shows Positive polarity comparison among five cloud service providers. Microsoft Azure has the highest positive polarity which is of 72% and Amazon has 71.5% where as Sales force has 64% and IBM Cloud has 55.5% and Google cloud platform has 43.5%.

Figure 3 shows Negative polarity comparison among five cloud service providers. IBM Cloud has the highest negative polarity which is of 9.5%, while Amazon and Google cloud platform have the same negative polarity that is 5.50%, Sales force has 6% of negative polarity and Microsoft Azure has 4.5% of negative polarity.

Figure 4 shows the comparison of neutral polarity among five cloud service providers. Google cloud platform has the highest neutral polarity which is of 51% and IBM Cloud has 35%, Sales force has 24%, Amazon has 23% where as Microsoft Azure has 23.5% of neutral polarity.

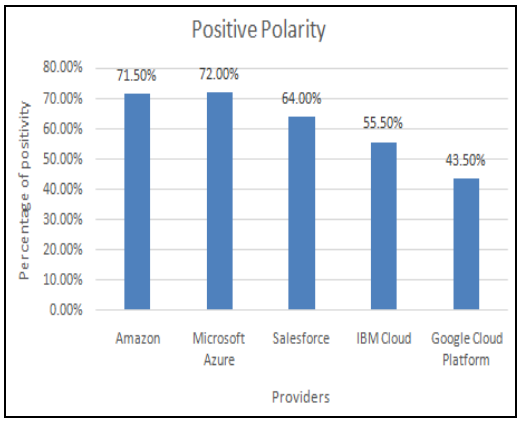

Fig. 2. Comparison of Positive Polarity

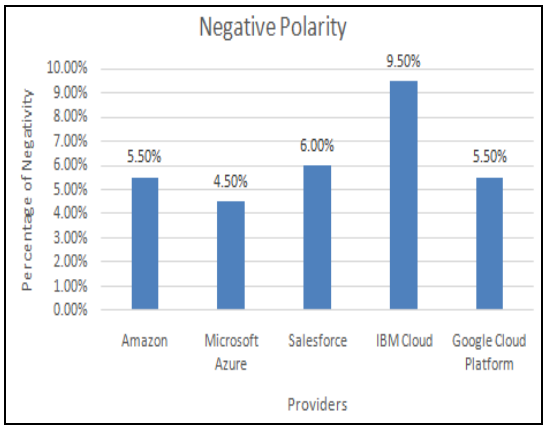

Fig. 3. Comparison of Negative Polarity

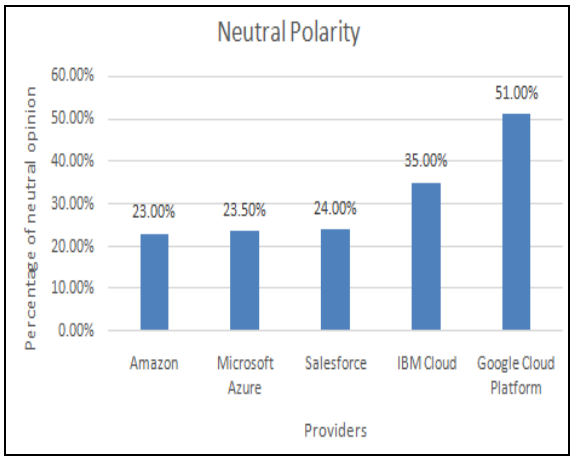

Fig. 4. Comparison of Neutral Polarity

VI. DISCUSSIONS

Our object is to select the best cloud service provider among the available providers without using any third party like Broker. So we used Naive Bayes classifier to do that task.

We didn't compare the performance of naive Bayes with any other classifier because our work is to find the best cloud service provider.

VII. CONCLUSION

In this paper, we did analysis on the top five cloud service providers, by considering opinion of the customers. Based on the analysis made, we conclude which cloud service provider provides best services. As an extension to this work, analysis can be made by including context aware classification.

REFERENCES

[1] "Amazon Web Services (AWS) - Cloud Computing Services," Amazon Web Services, Inc.[Online]. Available: https://aws.amazon.com/. [Accessed: 19-Mar-2018]

[2] J"Your vision. Your cloud.," Microsoft Azure Cloud Computing Platform & Services. [Online]. Available: https://azure.microsoft.com/en-us/. [Accessed: 19-Mar-2018]

[3] "The Customer Success Platform To Grow Your Business," Salesforce.com. [Online]. Available: https://salesforce.com/. [Accessed: 19-Mar-2018]

[4] "This is not just any cloud. This is the IBM Cloud.," IBM Cloud, 14-May-2015. [Online]. Available: https://www.ibm.com/cloud/. [Accessed: 19-Mar-2018]

[5] Google Cloud Computing, Hosting Services & APIs | Google Cloud Platform. [Online]. Available: https://cloud.google.com/. [Accessed: 19-Mar-2018

[6] A. Koneru and M. S. Latha, "Three tier architecture: To select CSP through CloudServiceBroker in multicloud environment," 2016 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, 2016, pp. 1-6. doi: 10.1109/INVENTIVE.2016.7823236

[7] Atique, M., & Patil, H.P. (2015). Sentiment Analysis for Social Media: A Survey. 2015 2nd International Conference on Information Science and Security (ICISS), 1-4.

[8] Bhaskaran, V., Devi, K.N., & Mouthami, K. (2013). Sentiment analysis and classification based on textual reviews. 2013 International Conference on Information Communication and Embedded Systems (ICICES), 271-276.

[9] M. S. Neethu and R. Rajasree, "Sentiment analysis in twitter using machine learning techniques," 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, India, 2014, pp. 1-5.

[10] A. P. Jain and V. D. Katkar, "Sentiments analysis of Twitter data using data mining," 2015 International Conference on Information Processing (ICIP), Pune, 2015, pp. 807-810. doi: 10.1109/INFOP.2015.7489492

[11] L. M. Qaisi and I. Aljarah, "A twitter sentiment analysis for cloud providers: A case study of Azure vs. AWS," 2016 7th International Conference on Computer Science and Information Technology (CSIT), Amman, 2016, pp. 1-6. doi: 10.1109/CSIT.2016.7549473

[12] Ramteke, Jyoti & Shah, Samarth & Godhia, Darshan & Shaikh, Aadil. (2016). Election result prediction using Twitter sentiment analysis. 1-5. 10.1109/INVENTIVE.2016.7823280.

[13] Deshwal, A., & Sharma, S.K. (2016). Twitter sentiment analysis using various classification algorithms. 2016 5th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 251-257.

[14] D. Arora, K. F. Li and S. W. Neville, "Consumers' Sentiment Analysis of Popular Phone Brands and Operating System Preference Using Twitter Data: A Feasibility Study," 2015 IEEE 29th International

Conference on Advanced Information Networking and Applications, Gwangiu, 2015, pp. 680-686. doi: 10.1109/AINA.2015.253

[15] C. D. Manning and H. Schutze. “Foundations of statistical natural language processing”. MIT Press, 1999.

[16] B. Pang, L. Lee, and S. Vaithyanathan. “Thumbs up? Sentiment classification using machine learning techniques”. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 79–86, 2002.

[17] M. Vadivukarassi, N. Puviarasan and P. Aruna, “Sentimental Analysis of Tweets Using Naive Bayes Algorithm”2017.

[18] K. Purushottama Rao, Anupriya Koneru and Dr. Naga Raju Devarakonda, “OEFC Algorithm - Sentiment Analysis on Goods and Service Tax System in India,” Proceeding of International Conference on Cognitive Informatics & Soft Computing (CISC-2017), 20th -21st December, 2017.