Sentimental Analysis Using Twitter Live Dashboard

D.I. HIMAN[1], Smt. M. PADMA[2]

[1]PG Scholar, Dept. Of CSE, G. Pulla Reddy Engineering College Kurnool, Andhra Pradesh, India.

[2]Assistant Professor, Dept. Of CSE, G. Pulla Reddy Engineering College Kurnool, Andhra Pradesh, India.

Abstract— Social networks are the most resources to collect data regarding people's opinion and sentiments towards totally different topics as they pay hours daily on social medias and share their opinion. Twitter is currently became one in every of the most effective social media platforms that's utilized by billions of individuals to share their expressions and opinions. Twitter offers organizations a quick and effective way to analyze customers' views toward the important to success within the market place. In this paper, we tend to show the application of sentimental analysis and the way to connect to Twitter and run sentimental analysis queries using IBM Watson Twitter Live Dashboard. We tend to run experiments on totally different queries from politics to humanity and show the fascinating results. We tend to realized that the neutral sentiment for tweets are considerably high that clearly shows the restrictions of the present works. We've got conducted a case study on tweets regarding the happiness based on the reactions of individuals round the world with analysis on the time and origin of the tweets.

Keywords: Streaming processing, Big Data, Twitter, Real-time, IBM Watson,PubNub

I. INTRODUCTION

Opinion and sentimental mining is a very important analysis areas as a result of due to the large range of daily posts on social networks, extracting people's opinion may be a tough task. concerning 90 percent of today's information has been provided throughout the last 2 years and obtaining insight into this massive scale data is not trivial [17, 18].

Sentimental analysis has several applications for various domains for instance in businesses to induce feedbacks for merchandise by that corporations will learn users' feedback and reviews on social medias.

Opinion and sentimental mining has been well studied during this reference and every one completely different approaches and research fields are mentioned [10]. There are some works are done on Facebook [19-23] sentimental analysis but during this paper we tend to largely target the Twitter sentimental analysis.

With the huge amounts of information being accumulated from numerous sources, analysis of massive information is immensely vital for deciding of really any kind—whether it's for businesses, scientific study, or the development of technology as some examples. Moreover, period of time applications depend on fast input and quick analysis to make a decision or action within a brief and extremely specific timeline [1]. Originally, information analytics are performed when storing information on hard disks, that eventually have a good quantity of access latency. handling great amount of structured and unstructured data in real-time makes laborious disks undesirable, as a result, there has been a recent transition from disk drive storage to memory storage. In-memory process considerably decreases the quantity of access latency, which is able to have an important role when real-time analytics is performed.
Analyzing data in real-time needs information ingestion and processing of the stream of data before the data storage step [2].a

number of the applications of the real-time information analytics of surveillance, environment, health care, business intelligence, marketing, visualisation, cybersecurity, and social media. This study presents a real-time data analytics framework for analyzing Twitter data. the essential distinction between this study and alternative researches is that the proposed framework can't solely perform the essential process tasks, however makes associate degree infrastructure for acting additional refined and sophisticated analytics on the streaming data. Present real-time methodologies use various algorithms and technologies to process Twitter data which are using event processing and one-message-at-a-time analysis. This makes it attainable to achieve real-time result, however lacks the flexibility of doing something more than plain processing. Reviewing related works during this field showed that there's a gap of capability of performing additional complicated analytical tasks like machine learning algorithms. The proposed framework offers an infrastructure for real - time processing with the flexibility of extending the analytical capability.

II. RELATED WORK

Social network analysis is that the study of people's interactions and communications on completely different topics and today it's received additional attention. Million of individuals offer their opinion of various topics on a day after day on social medias like Facebook and Twitter. it's several applications in several areas of analysis from scientific discipline to business [3].

Twitter nowadays is one in every of the popular social media which according to the statistician [4] presently has over three hundred millions accounts. Twitter is that the rich source to learn about people's opinion and sentimental analysis [2]. for every tweet it's vital to see the sentiment of the tweet whether or not is it positive, negative, or neutral. Another challenge with twitter is simply one hundred forty characters is that the limitation of every tweet that cause folks to use phrases and works that aren't in language process. Recently twitter has extended the text limitations to 280 characters per every tweet. addressing social media information, as well as many alternative data varieties equivalent to text messages, photos, and videos that are incoming in a very massive volume in each second, desires a correct framework that doesn't rely on storing data on hard disks and can process data in memory, as it arrives [3].

A. TWITTER SENTIMENTAL ANALYSIS

Social networks is a rich platform to find out about people's opinion and sentiment regarding different topics as they will communicate and share their opinion actively on social medias as well as Facebook and Twitter. There are completely different opinion-oriented operation systems, that aim to extract people's opinion regarding different topics. The sentiment-aware systems these days have several applications from business to social sciences.

Since social networks, particularly Twitter contains tiny texts, people might use completely different words, and abbreviations, that are troublesome to extract their sentiment by current natural language process systems easily, thus some researchers have used deep learning and machine learning techniques to extract and mine the polarity of

the text [15]. a number of the top abbreviations are FB for Facebook, B4 for before, OMG for oh my God.

There are studies showing the ability of social media particularly Twitter data in monitoring the impacts of earthquakes. The authors [17] analyzed the flexibility of Twitter in contribution of data in terms of location and time of the earthquake happened on the geographic area of the U.S. on August 23, 2011. They compared their results with the information gathered by U.S. geologic Survey discovered that social media information might complement different sources of information and should facilitate to boost our understanding of this kind of events.

This study presents a framework for analyzing the streaming data in real-time. using a framework with a robust engine that has the potential of process large streaming information beside its advanced analytic options (for instance, machine learning) makes it attainable to offer sophisticated analysis in real-time. Sentiment analysis is one in every of the samples of implementing machine learning algorithms in real-time analysis. However, building the machine learning model is generally supported the historical data, during this method the created model are going to be utilized in production on live events. However, the engine used will give some streaming machine learning algorithms, however still usually there's a necessity to do an analysis of historical data.

III. METHODS
IBM WATSAN FRAMEWORK

Watson is a question-answering computer system capable of responsive queries posed in natural language, developed in IBM's DeepQA project by a research team. Watson uses the Apache UIMA (Unstructured info Management Architecture) framework. The system was written in numerous languages, as well as Java, C++, and Prolog, and runs on the SUSE UNIX operating system Enterprise Server 11 operating system using the Apache Hadoop framework to provide distributed computing.
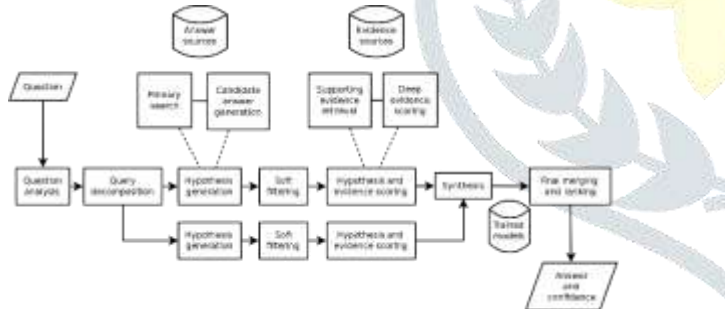

Figure 1: IBM Watson Framework

Watson parses queries into different keywords and sentence fragments so as to search out statistically connected phrases. Watson's main innovation wasn't within the creation of a new algorithm for this operation however rather its ability to quickly execute many proven language analysis algorithms at the same time. The additional algorithms that find the same answer independently the additional likely Watson is to be correct. Once Watson includes a tiny range of potential solutions, it's ready to check against its info to determine whether or not the answer is sensible or not.
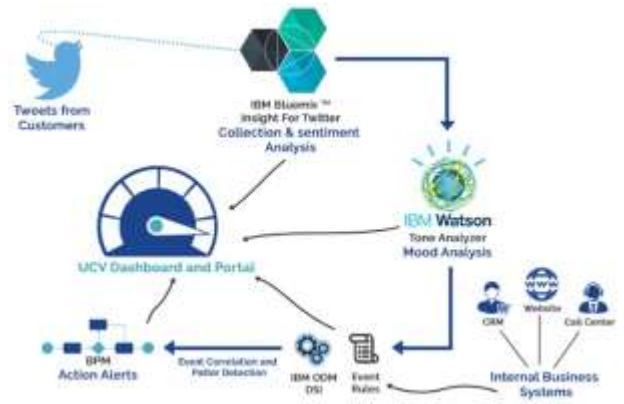

Figure 2: IBM Watson process Flow

Data process work flow typically connects computing resources to automate a sequence of tasks by processing massive volumes of data, where {different|totally totally different|completely different} resources are connected for automating different tasks. in the case of streaming data processing, a scalable and distributed platform is needed for combining massive volumes of historic and streaming information at the same time. The framework includes a core element referred to as Brandceptions.

Brandceptions consists 3 components: i) Twitter API; ii) IBM Watson Tone analyser Service; iii) PubNub BLOCK.
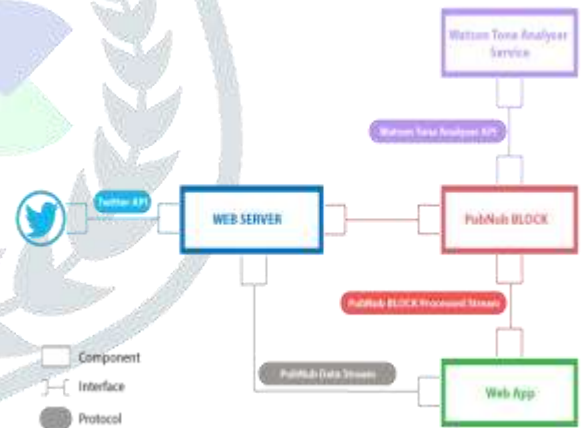

Figure 3: IBM Watson elements and information Flow

A. Twitter API

Twitter API fetches the most recent tweet stream mentioning the hash code (data type) that needed to be compared. This uses the twitter key to attach to twitter account and streams the information unceasingly. To buy Tweets from a selected handle or hashtag during this application, it's needed to create a Twitter account and a Twitter application. The Twitter account are going to be used because the account that receives the messages from different Twitter users furthermore because the owner of the application, required by Twitter, to receive Tweets.

A normal Twitter account will be created in Twitter or will use an existing account. it's needed to provide a unique email id that isn't already related to an existing Twitter account furthermore as a telephone number to verify the account.

Once we've the Twitter account created and verified, log in to Twitter Dev and build an application.

Select the Keys and Access Tokens tab and generate a consumer Key and Secret. Keep this page open as you may ought to use these tokens into setup procedure within the application later on.

B. IBM Watson Tone analyser Service

IBM Watson Tone analyser Service runs the sentiment analysis and extract the emotion score. The IBM Watson™ Tone analyser service uses linguistic analysis to discover emotional and language tones in transcription. The service will analyze tone at each the document and sentence levels. This service will be accustomed perceive however the written communications are perceived then to boost the tone of the communications. Businesses will use the service to be told the tone of their clients' communications and to reply appropriately to every customer, or to know and improve their client conversations generally.

We submit JSON, plain text, or hypertext markup language input that contains the written content to the service. The service accepts up to 128 kilobyte of text, that is concerning one thousand sentences. The service returns JSON results that report the tone of the input. we will use these results to boost the perception and effectiveness of your communications, making certain that your writing conveys the tone and style that we want for the intended audience. the subsequent diagram shows the fundamental flow of calls to the service.
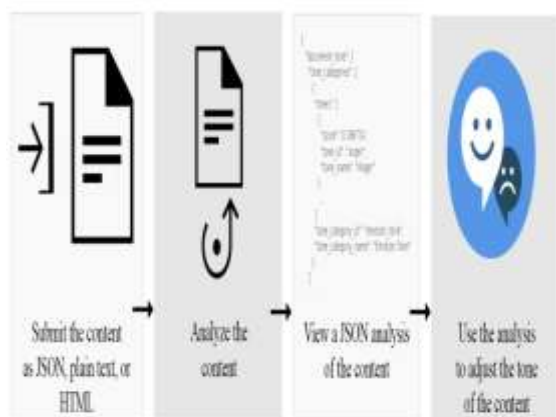

Figure 4: The Tone analyser Service

C. PubNub BLOCK

PubNub Functions enable the creation and execution of business logic on messages at the edge of the network. Quickly build microservices and incorporate logic like re-routing, augmenting, filtering, transforming, and aggregating information. PubNub Functions are serverless, there's no ought to worry about deploying, maintaining, or scaling server infrastructure. PubNub processes the user requests and orchestrate the information stream containing the response from Watson Tone analyser.

IV. CASE STUDY

The purpose of this study is to develop a framework for analyzing Twitter data in real-time using IBM Watson. This framework has some characteristics, that distinguish it from traditional data analytics approaches. the main idea here is that there's a need for methods to analyze thousands of tweets coming every second, in a very short amount of your time. Also, the framework ought to be independent of imported data volume; this can be vital as a result of the volume of tweets is growing at a noticeable rate. Figure 1 shows a schematic of a scalable stream process. The concept here is to gather event streams

by totally different nodes and let multiple process nodes to investigate information in parallel. So, the challenge here is the way to manage streaming information and the way to investigate it over the clusters.

In this case study, we've used PubNub to access IBM Watson framework to research twitter data. The high-level steps are given below.

a) creating PubNub

Firstly, a PubNub account have to be compelled to be created with the giving details. Once you're in and on the main page, scroll down to create an app. Apps are treated like projects, therefore all has distinctive keys and blocks.

b) IBM Bluemix

IBM Bluemix provides us access to the Watson genus Apis - most significantly, the sentiment analysis. initial you need to signup for it. when a successful registration, you'll need to verify your account via email. Once back on the main page, navigate to products & Services -> Launch Bluemix. you will be prompted to name your "organization" and "space" - I simply named mine "Twitter Stream" and "dev"

c) API key

Next we'd like to provision the natural language service. Click on "Apps" or "Catalog" and either look for "Natural language service'or look under Services -> Watson. Clicking on the natural language service can take you to a page describing it that additionally permits you to alter the service and credentials names. If you scroll down you'll see the pricing plans - "Free" should already be chosen.

d) test the Streamer

Go to the IBM Sentiment Analysis block inside your PubNub app. confirm that the block is Running which you added your Watson API username and password to the code. Copy-paste the subsequent into the "Test Payload" box: one, "text": "I am happy!" }

e) data visualization

For data visualization we tend to use Initial State that we tend to create an account for it.

After creating an account we tend to create a bucket(to show our twitter data) and that we get an API key for initial state account.

In PubNub initial state app we've to add the api key to that So we tend to use Twitter API to urge the twitter tweets and that we sentiment analysis with IBM Watson api and that we showcase with the Initial State.

V. RESULTS

The ASCII text file is split into 2 components:

App: The Node.js based mostly web SERVER that hosts the web App. the web server interfaces with the Twitter stream to extract the tweets relevant to the brands that the user desires to compare.
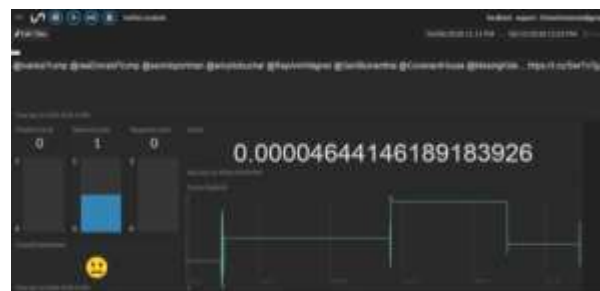

Figure 5: data Plot by IBM Watson

Block: The routine that runs on the PubNub BLOCK. this is often wherever the action happens. All tweets are sent to the Watson Tone analyser service, and also the emotion score is processed to present to the user.

The app will compare 2 brands and show a straightforward bar chart of their comparative scores across the 5 emotions that's predicted by the Watson Tone analyser service.

Here is a glimpse of the app under operation when we compare 2 world leader and 2 coming movies
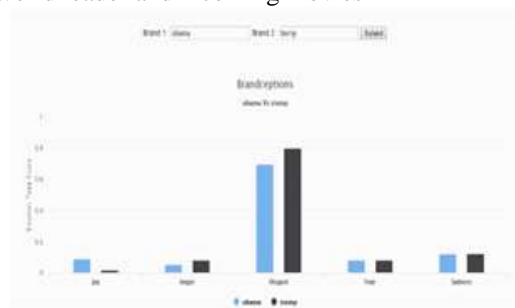

Figure 6: data Plot2 by IBM Watson

The app follows a straightforward 3 step process:

　　　　User enters the 2 complete names for comparison and web app sends request to web server.

　　　　The request from web app to web server is sent on the PubNub channel 'realtimebrandmonitor-brands'.

　　　　Web server pulls all the newest tweets mentioning the brands and forwards them to block.

　　　　The Web server publishes all the tweets for every brand on the PubNub channel 'realtimebrandmonitor'. The publish is done as 2 separate data blobs containing the tweets for every brand.

　　　　BLOCK sends all the tweets to Watson Tone analyser and collates the results of emotion perception score and sends it back to the web app.

VI. CONCLUSION

　　　　In this paper, we've used IBM Watson and PubNub for real-time analysis of Twitter data. Watson Tone analyser encompasses a heap of use cases. Not solely will it extract the emotional tone of the text, however it's additionally capable of sensing social and language tone. This framework is intended to collect, filter, and analyze streams of data and gives us an insight to what's widespread throughout a particular time and condition. The case study during this analysis aims to point out the strength and also the importance of real-time data analytics on social media streaming data.

VII. REFERENCES

[1] N. Mohamed, J. Al-jaroodi, Real-Time Big Data Analytics: Applications and Challenges. International Conference on High Performance Computing & Simulation (HPCS), 2014.

[2] S. Cha and M. Wachowicz. Developing a real-time data analytics framework using Hadoop. 2015 IEEE International Congress on Big Data, pages 657–660, June 2015.

[3] B. Yadranjiaghdam, N. Pool, N. Tabrizi, "A Survey on Real-time Big Data Analytics: Applications and Tools," in progress of International

Conference on Computational Science and Computational Intelligence, 2016.

[4] H. Hedayati, N. Tabrizi, "MRSL: Autonomous Neural Network-Based 3-D Positioning System," 2015 International Conference

on Computational Science and Computational Intelligence, Pages: 170 - 174, doi: 10.1109/CSCI.2015.88.

[5] A. Bifet, "Mining Big Data in real time," Informatica, 37(1), 2013,

Pages 15 - 20.

[6] D. T. Nguyen and J. E. Jung. Real-time event detection for online behavioral analysis of big social data. Future Generation Computer Systems, 2016.

[7] D. Preotiuc-Pietro, S. Samangooei, T. Cohn, N. Gibbins, and M. Niranjan. Trendminer: An architecture for real time analysis of social media text. Proceedings of the workshop on real-time analysis and mining of social streams, 2012.

[8] J. Zaldumbide, R. O. Sinnott, "Identification and Validation of Real-Time Health Events through Social Media," 2015 IEEE International
Conference on Data Science and Data Intensive Systems, Pages 9 – 16, doi 10.1109/DSDIS.2015.27.

[9] V. Ta, C. Liu, G.W. Nkabinde, "Big Data Stream Computing in Healthcare Real-Time Analytics", 2016, IEEE International Conference on Cloud Computing and Big Data Analysis, Pages: 37 -

42, doi: 10.1109/ICCCBDA.2016.7529531.

[10] M. Wachowicz, M.D. Artega, S. Cha, and Y. Bourgeois, "Developing a streaming data processing workflow for querying space–time activities from geotagged tweets" Computers, Environment and

Urban Systems Journal. 2015.

[11] M. T. Jones. Process real-time Big Data with twitter Storm. IBM Technical Library, 2013.

[12] G. Mishne, J. Dalton, Z. Li, A. Sharma, and J. Lin. Fast data in the era of Big Data: Twitter's real-time related query suggestion architecture. In Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data, pages 1147–1158. ACM, 2013.

[13] A. Toshniwal, S. Taneja, A. Shukla, K. Ramasamy, JM. Patel, S. Kulkarni, J. Jackson, K. Gade, M. Fu, J. Donham, N. Bhagat,

"Storm@ twitter", InProceedings of the 2014 ACM SIGMOD international conference on Management of data 2014 Jun 18 (pp. 147-156). ACM.

[14] A. Sotsenko, M. Jansen, M. Milrad, J. Rana, "Using a Rich Context Model for Real-Time Big Data Analytics in Twitter", 4th International Conference on Future Internet of Things and Cloud Workshops, 2014, Pages 228 -233, DOI 10.1109/W-FiCloud.2016.55.

[15] M. Z. Mosharaf Chowdhury and T. Das, "Resilient distributed datasets: a fault-tolerant abstraction for in-memory cluster computing," in NSDI'12 Proceedings of the 9th USENIX Conference on Networked Systems Design and Implementation. San Jose, CA: USENIX Association Berkeley, Apr. 2012.

[16] Y. Yan, L. Huang, L. Yi, "Is Apache Spark Scalable to Seismic Data Analytics and Computations?", IEEE International Conference on Big
Data (Big Data), 2015, Pages: 2036 - 2045, doi: 10.1109/BigData.2015.7363985.

[17] A. Crooks, A. Croitoru, A. Stefanidis, J. Radzikowski, "#Earthquake: Twitter as a Distributed Sensor System", Transaction in GIS, 8 October 2012, doi: 10.1111/j.1467-9671.2012.01359. x.

[18] P. Earle, "Earthquake Twitter", Nature Geoscience 3, 221 - 222 (2010), doi:10.1038/ngeo832.

[19] T. Sakaki, M. Okazaki, Y. Matsuo, "Earthquake shakes Twitter users: real-time event detection by social sensors", Proceedings of the 19th international conference on World wide web, Pages 851-860, 2010, doi: 10.1145/1772690.1772777.

[20] P.S. Earle, D.C. Bowden, M. Guy, "Twitter earthquake detection: earthquake monitoring in a social world", Annals of Geophisycs, vol 54, No 6, 2011, doi: 10.4401/ag-5364.

[21] M. Bilal, L. O. Oyedele, O. O. Akinade, S. O. Ajayi, H. A. Alaka, H. A. Owolabi, J. Qadir, M. Pasha, and S. A. Bello. Big Data architecture for construction waste analytics (cwa): A conceptual framework. Journal of Building Engineering, 6:144–156, 2016.

[22] http://radiostud.io/build-social-percept

[23] https://hackernoon.com/visualizing-trump

[24] https://www.ibm.com/blogs/bluemix/2017/06/build-a-social-perception-dashboard-using-watson-tone-analyzer-and-pubnub-blocks/

[25] https://github.com/initialstate/pubnub-live-twitter-dashboard/wiki/Part-2.-IBM-Bluemix-Setup