www.ijsrnsc.org

Available online at www.ijsrnsc.org

IJSRNSC

Volume-5, Issue-4, August 2017
Research Paper

Int. J. Sc. Res. in Network Security and Communication

ISSN: 2321-3256

Big Data Analytical Architecture for Real-Time Applications

Rakesh. S.Shirsath[1], Vaibhav A.Desale[2*], Amol. D.Potgantwar[3]

[1]Department of Computer Engineering, SITRC, Pune University, Nashik, India
[2*]Department of Computer Engineering, SITRC, Pune University, Nashik, India
[3]Department of Computer Engineering, SITRC, Pune University, Nashik, India

**Corresponding Author: vaibhav123desale@gmail.com*

Received 09th Jun 2017, Revised 21st Jun 2017, Accepted 28th Jul 2017, Online 30th Aug 2017

Abstract— Now a days due to the enhancement in the substantial of the real-time data coming from many social sites or any live feed of data which is continuously generating tremendous volume of information so in several fields they have got massive attention to the data or information collected by them. Also these days the usage of social media or other applications of social media has increased and data collected through this systems is of different types. Data collected by the real-time applications along with by these social media application is in massive amount. These data can be referred as term ”Big Data”. If we gather this data we can observe that data has great significance and aggregation of this data can be done very effectively. Importance to this data is increased but gathered data is in massive amount so it is an challenge to analyze, aggregate and store where data is remotely located. Existing system with the conventional techniques are not able to collect, aggregate and analyze such huge data. Results obtained are not much accurate and decision creation is also not much effective by using the previous systems and methods. Considering attention to data analysis and need of effective framework which will welcome both real-time along with offline data. Therefore in this dissertation topic proposed a framework which is capable of processing huge volume of remote data collected. By using Cloud Computing environment proposed system is more effective on Real-time applications live data feed. This system can efficiently process on the real-time data and can have effective analysis and also decision making.

Keywords— Big Data, Cloud Computing, Data Analysis, Real-Time Applications

I. INTRODUCTION

In today’s era there is a great deal of interest in the field of the Big Data[1][2][3]and it’s analysis which had lead many research challenges close to reveal applications, like as modelling, mining and Distributing large scale repositories. The term ”Big Data” classifies some specific formless data which will reside in information layer of technical computing. Some of systems and the web-apps has data stored at underlying layer of scientific computing have some individuality in common like large scale which is referring to data repositories or size of available data, other one is an issue of scalability which refer to applications which are running large data, also extraction transformation method which will extract from low or unprocessed data into well thought data also in the end advancement of simplified interpretable i.e. analytical over data repositories with a view to give a smart and momentous for them.

Big-Data analysis is very difficult and challenging task rather than any other applications on data like locating, identifying or citing data[4].Massive producer of Big-Data can be said as many transaction process of any payment or bank system, videos or audios collection of any format, email delivering systems, data generated by using the mobile phones(specially smart phones) and their apps and also a big and popular source is Social media data. So it become complicated to confine, form, store, manage, share, process, analyze, filter and visualize via typical conventional database software tools. Certain alternatives were given by the traditional database vendors but still they are not sufficient [5].Advancement in Big-Data sensing and computer technology specifies the method to gather remote data , process, analyze, and manage. The platforms for these system comes in various kind of types from software in third party hosted environment. In terms of the live data producing applications are generating big Ease of volume of information. At the very first step we prefer to do data acquiring in which we are filtering the information which is of not in any use. With the use of Data-Acquisition, we discard useless data and can have only good information to analyse. The second challenge is generation of perfect and correct metadata which will be describing composition of data and way it is stored and analysed. In this proposed system, we used the high speed live stream of real-time information or high amount of big stored information to Big-Data which is taking us to a distinct world of challenge of new problems and opportunities for its analysis. Such

problems of transforming of remotely located data to the scientific understanding are a critical task[6][7][8].Hence the speed at which amount of the remote access data is expanding and also the use and significance of this information for many purpose like research , security, health/medical issues is in demand. Therefore number of users and organizations and researchers are now asking an efficient method to gather process, and analyze, and store these information and its resources. Cloud Computing concept has many computers intertwined through a real-time network like internet. Cloud computing is distributed computing. Cloud computing enables easy to use, on demand, dynamic and steady use of distributed computing resources. It has five main characteristics- on-demand service, very large network access, resource reservation, reliability measured service.

This paper is organization as: In chapter II this includes literature survey. We have also discussed the Problem Statement, Proposed System, Algorithm, Mathematical Model, Analysis and Discussion of results that are obtained in experiment and conlusion.

II. LITERATURE SURVEY

Digital real data on digital world is growing exponentially on large scale. Use of current techniques and traditional methods on such form of information will not give an correct solutions because these methods does not extract perfect data sets. Therefore there is need for framework that will effectively analyse both remote access real-time data and off-line data. When a business enterprise can get off all the good information obtainable in the Big-Data in place of a sample of its data set, in that case, it has an influential benefit over the market competitors. Big-Data analytics helps us to gain insight and get better decisions. hence, with the purpose of using Big-Data, modifications in model are at utmost. To support our motivations, we have described some areas where Big Data can play an important role.

In healthcare examples, medical practitioners gather huge volume of data about patients, medical history, medications, and other details. The above-mentioned data are accumulated in drug-manufacturing companies. The nature of these data is very complex, and sometimes the practitioners are unable to show a relationship with other information, which gives result in missing of important information. With a view in employing advance analytic methods for arranging and getting useful information from Big-Data results in personalized medication, the advance Big-Data analytic techniques give insight into hereditarily causes of the disease.

In the Securable Machine-Learning Online Services for analysis of real-time huge data they have implemented system by using RESTful API on the Hadoop framework. In this system Real-Time data can be effectively accessed by using the RESTful API. RESTful API permits offering Real-Time analysis tools as a favour, so that specific analysis could be performed online. Process is done in two ways that are system Batch processing for collected big collection of the information and a continuous flow of data processing for Real-Time data. RESTful API which allows to bundle machine learning systems supported by this architecture as a service, such as the user can perform requests while stream Big-Data is flowing. Results in this system are calculated in the two modules one batch processing where performing query analytics, performing clustering, building machine-learning and in the stream processing for real-time data carrying out predictions, carrying segmentation, providing recommendations is done. Results obtained in this system were encouraging but still studying the enhancement of the recommender system with real users and comparing how the system is able to scale out horizontally is left as future work resulting in the design and execution of a more complete evaluation, and a side-by-side comparison of the game prediction system using public datasets.

Map-Reduce: Simplified Data Processing on Large Clusters they have studied implementation of Map-Reduce goes on a large cluster of machines and is more scalable: a typical Map- Reduce computation processes many terabytes of data on thousands of machines. Programmers and the system easy to handle: hundreds of Map-Reduce programs have been executed and upwards of one thousand Map-Reduce tasks are executed on Googles clusters every day. But limitations for this study are network bandwidth is the scarce of network. Slow machines or machine failure may occur failure or loss of data[5,9,10]

A cloud based grid system was used for Big-data analysis. It was used for efficient planning of power production and distribution. The execution and usage of cloud grid have more demand as they provide high amount of parallelism. This grids were acting on base of Hadoop , Cassandra and hive. Limitations came for this system like security and privacy in cloud. Also slow system for failure in system may cause huge data loss.

In GPU processing for remote sensing data they offer effective parallelism for effective Data-Processing and analysis. In this system many image processing algorithms were implemented and proven it also rapids development of GPU adoptions. In this system several cases like from simplex to complex were implemented and observed that efficiency was improved up to 400 times. Two modules were used in this system first image size scalability: streaming and second Rise to multi-core GPUs. Calculation or results were extracted by Pixel based processing, Neighborhood Processing and irregular and non-local processing. But limitations came like there was need to improved or better approach for seamlessly transform from CPU to GPU version of algorithms.[6]

RS-SWBM (remote sensing driven soil moisture balance model) was used for soil moisture monitoring model. Results in this system were collected on the parameters of co-relational analysis and co-relational coefficients. Results in this model indicated vegetation types in different seasonal type, vertical patterns and consumption efficiencies with strong correlation between these parameters and land use as well as precipitation. System had two models first soil moisture balance module and vegetation interpretation module. Results were collected by detecting position, observation content and equipment, information collection and processing and result verification. Limitations came for this system are satellite images have some limitations and we do not have enough observation errors for interpolation of data for multiple days.[11-17]

A survey on Big Data in real-time, they studied Big data is that whose characteristics are beyond traditional methods that are prevalent at the time. It is analyses , collect , store of data. Limitations for their study was Question of unity, also lack of structure, error handling, privacy control and visualization of data analysis from acquisition [11,14, 16,17]

III. PROBLEM STATEMENT

Design an architecture which will be useful for analysis of both offline and online Real-Time Data. Designed Architecture would be efficient to process, analyse big quantity of data very easily and effectively. Architecture is responsible for effective collection, analysis, aggregation and for decision making. Use of various techniques like Data-Acquisition and data analysis, Data-Processing unit and decision making unit and also various algorithms which will be useful for development of system.

IV. PROPOSED SYSTEM

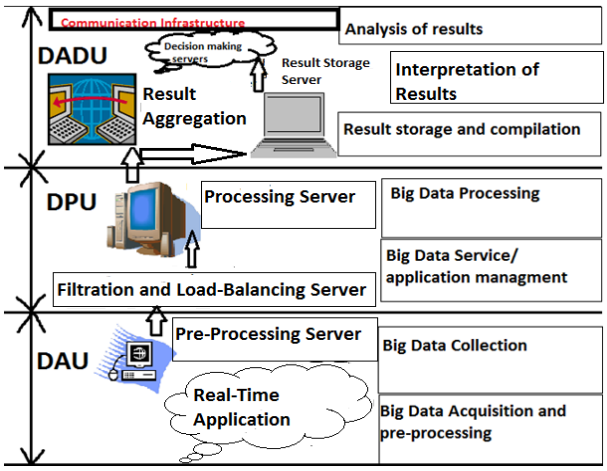

Fig1. Real Time Big Data Analysis Architecture

In our proposed system for analysis real time and off-line data for real-time applications using term Big Data we have split Real-Time Big-Data processing framework into three parts, those are namely: i) Data-Acquisition Unit ii) Data-Processing Unit and iii) Data-Analysis and Decision Unit. In next three units various algorithms or techniques will be implied on data for its analysis. The functionalities and working of three units is as explained and shown in diagram below:

A. Data Acquisition Unit :-

The requirement for collateral processing of the big amount of data was required, which could give effective analysis of the Big Data. For that reason, the proposed unit is introduced in the real-time Big-Data processing architecture that collects the big volume of data from different types of available data collecting unit around the world. We assume that the data capturing unit can correct the erroneous data. For better data analysis in the Base unit system pre-processing data under several conditions to associate the information from several locations, which not only reduces cost of storage, but also enhances analysis exactness. Some relational data pre-processing methods are data assimilation, data brushing, and redundancy elimination. The data should be accurate in several methods to eliminate distortions caused due to the motion of the platform. We splited the data processing procedure into two steps, such as real-time Big Data processing and offline Big Data processing. In the case of offline data processing, the live data repository System transmits the information to the data centre for storage. This data is then utilized for future analyses. However, in Real-Time data processing, the data are directly sent to the filtrating server and load-balancer server, since storing of coming Real-Time data decreases the performance of real-time processing.

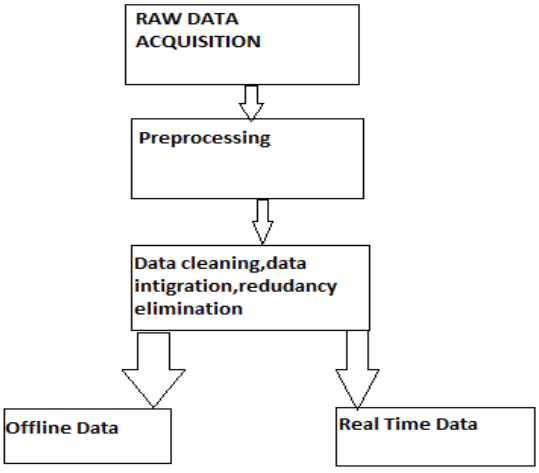

Fig2.Data Acquisition Unit

B. Data Processing Unit:-

In information stream processing unit, has two basic functionalities filtration server and load-balancer server. Filtration mainly includes elimination of information and load-balancing of processing power. This process of filtering information which is helpful to us for analysis and blocks other data. It will defiantly help to enhance performance of system as we are interested in only useful data. The filtration server and Load-Balancing server algorithm deviates from analysis to analysis. Each process acting has its own algorithm implementation for processing incoming segment of data from load balancer. Each process acting server makes statistical calculation, any measurements, and performs other mathematical or logical tasks to generate intermediate results against every segment of data. These tasks are performed parallel and independently such that the interpretation of system is increased at an extent and result segments are generated in real time. The outcomes given by each server are then given to the aggregation server for compilation, organization, and keeping for next processing.

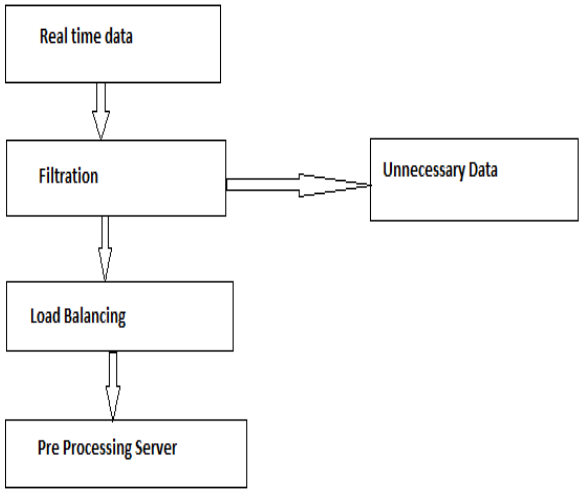

Fig3.Data Processing Unit

C. Data Analysis and Decision Unit:-

This unit includes three major functions, such as aggregating and compiling server, results storage server, and decision making server. When results are to be send to the compilation the data is not in aggregated form. So it is required to prepare the given data in aggregated form for proper storage and processing. In this unit many aggregating algorithms are implied so that organized results are stored into the storage. The aggregation server also gives the dummy copy of that result to the decision-making server to process that result for making decision

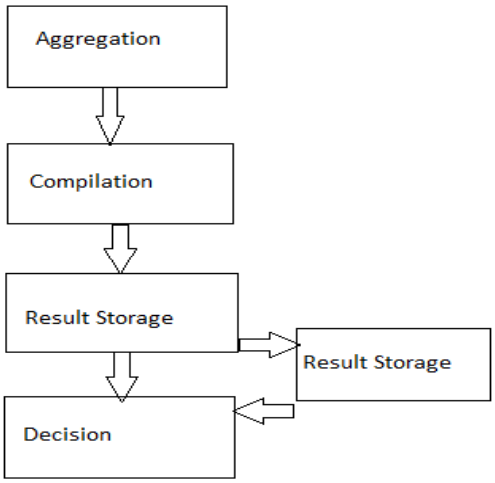

Fig4.Data Analysis and Decision Unit

V. ALGORITHM

Algorithm I. Filtration and Load Balancing Algorithm
Input: Collect Real-Time Data Feed and Process it
Output: Filtered data in fixed size block and send each block to processing Mechanism
Steps:
1. Filter related data i.e. Processed data. All other unnecessary data will be discarded.
2. Divide the Data into Appropriate Key Value Pair.
3. Transmit Unprocessed data directly to aggregation step without processing.
4. Assign and transmit each distinct data block of Processed data to processing steps in Data Processing Unit.

This algorithm takes live data and then filters and splits them into segments and perform load-balancing algorithm. In step 1, related data is filtered out. In step 2, filtered information is the association of diverse key value pairs and each pair is different numbers of sample, which concludes in forming a data block. In Next steps, these blocks are forwarded to processed by Data Processing Unit

Algorithm II. Processing and Calculation Algorithm
Input: Filtered Data
Output: Normalized Disruption data.
Steps:
1. For each event data, Categorical Data like G for good, A for average is extracted.
2. Normalize the disruption data for all the live feed.
3. Persist the data into collection and forward it.

The processing algorithm generates results for different parameters against each incoming filtered data and send them to the next level. In step 1, the calculation of Good

and Average along with trend Furthermore, in the next step, the results are transmitted to the aggregation mechanism.

Algorithm III. Multi Modal Summarization Algorithm
Input: Normalized Disruption Data.
Output: Final Result Summary
1. Collect the data from data store in normalized format.
2. Apply Summarization for Individual modal pie from the total disruption data capture.
3. Persist the final disruption summary into data store.

Here the data is collected and the results from each modal is processed against all and then combines, organizes, and stores these results in NoSQL database.

Below figure it shows the process flow algorithm. In the first step data is collected from the given API. After receiving data collected data is fetched subject wise and the clusters of stored data are formed. Classifiers are applied on the collected data for getting accurate or data cleaning process. Fake and truth values are calculated and truth values are also calculated after that it is compared with threshold value. When the threshold value is matched then results are collected in the given result storage server and reviews are made according to the given results.

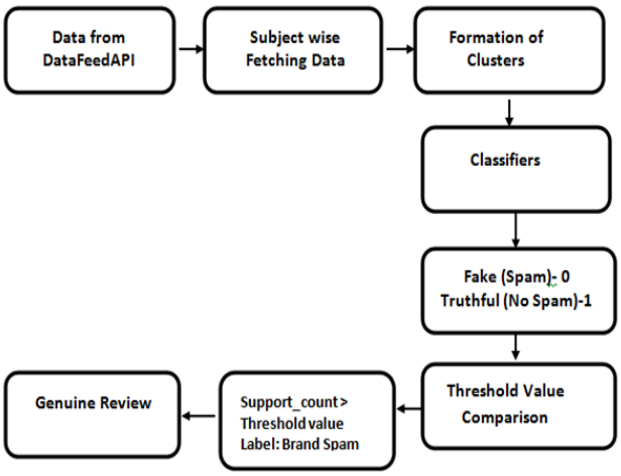

Fig5.Process Flow Algorithm

Below Diagram represents the data flow in the system. At first data is collected and then filters are applied on the given data to discard meaningless data. After filtered data is classified according to subject and stored in the form of clusters. After storing it into form of clusters fake values are calculated by which only useful data is remained into the clusters. Based on these values results are calculated and they are stored into the server from which decisions are made. Based on these decisions and computations a genuine review is made

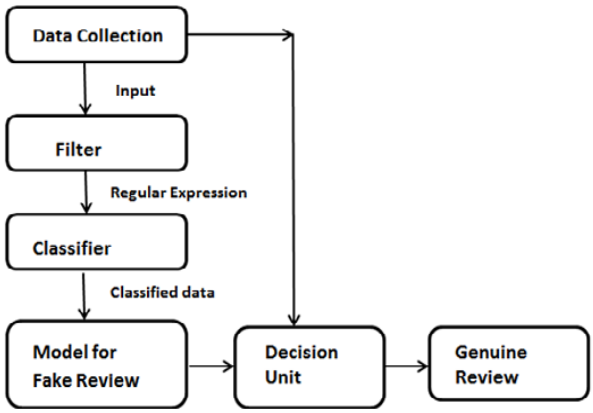

Fig6.Data Flow Diagram

VI. MATHEMATICAL MODELLING

Set Theory Model:

Let 'S' be the | Error detection in Big-Data as the final set
S = {……}
Identify the inputs as D
S = {D, …}
D = {D1, D2, D3, D4| 'D' given Data files}
Identify the outputs as O
S = {D, L, A…}
D = {D1, D2, D3, D4| 'D' gives data files }
L = {L1, L2 …| 'L' gives the log files for upload and download and repair }
A = {A1, A2, A3, … | 'A' gives alert }
Identify the functions as 'F'

S = {D, L, A, F…}
F = { F1(), F2(), F3(), F4(), F5(), F6() }
F1(V) :: Upload
F2 (V) :: integrity check
F3 (V) :: Log generation
F4 (T) :: Alert the system
F4 (D) :: Restore the file
F6 (V) :: Download the data file

Hence the functionality can be shown as,

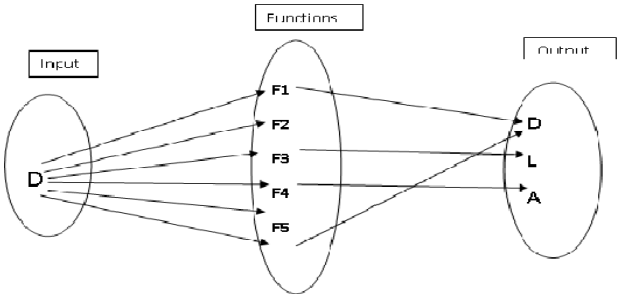

Fig7.Functional Dependency

VII. ANALYSIS AND DISCUSSION

Using the proposed framework for Real Time traffic, we perform a simple analysis on Real-Time Data. We assume that the real-time data from an application is big in nature and difficult to handle for a single server. The data is continuously coming from a Real-Time application with high speed. Hence, special algorithms are needed to process, analyse, and make a decision from that Big Data. Here, in this section, we analyse real-time data for finding semantic analysis, genuine reviews, or fake reviews. We have used the proposed architecture to perform analysis and proposed an algorithm for making decision. First, we take Real-Time Big Data samples from Twitter API Tweepy to analyse semantic analysis, genuine reviews, and fake reviews separately. On the basis of these analyses, we proposed a set of algorithms for handling, processing, analysing, and decision-making (sentiment analysis, genuine reviews and fake reviews) for Real Time Big Data using our proposed architecture.

In this section, we generalize the data sets and tools that are used to perform analysis. Furthermore, we described the analysis findings of the Data Sets.

A. .Tools, Data Set and Implementation Enviroment

In this framework we used Twitter API Tweepy for simple analysis of the Real-Time application data sets. Twitter API Tweepy gives an easy way to understand the composition of textual real-time application data sets. It also useful for simple sentiment analysis. Cloud environment provides efficiency, flexibility and performance increment for real-time applications, since they have capacity of parallel processing and high performance processing power using huge volume of data. Therefore it is suitable for large of data set. The real-time data collected by using Tweepy will be very large in amount and nature of data will be full of variety therefore it is necessary to analyse such data effectively for this we used Data using Deep Learning NLP API for finding the relevance of the data collected. We use meaning cloud API for finding the relevance of data collected. Data collected from twitter application is in Real-Time so we are storing data which coming from live feed will be stored in database at cloud environment. These data will be of various type, nature and form and also from very huge numbers of users who are using twitter application. So we are collecting, storing, analysing this data. We analyse tweets as per categories Ex. If we want tweets related to latest mobile phones which are new in market by using them we want to calculate number of genuine and fake reviews related to products we can classify them by using Deep Learning API

B. Findings and Discussion

We collected data from live feed of twitter by using twitter API tweepy. At first by using Data Acquisition first task is to collect data and store this real-time of data. After that we apply pre-processing techniques and then only textual data will be remaining for processing. Next task is to process that pre-processed data in the first step we will process the data as per requirement. Now here we are using classification on collected tweets by using latest mobile phones to calculate trend that which one is affecting market more. In the processing step we apply deep learning API i.e. we use meaning cloud API will classify data and after that classified data will be processed as per requirement. Data-Processing unit will process data and find out how many tweets are related to the mobile phones. Then next major task is decision making now we have to find how many among them are fake and genuine reviews because a single user can give multiple reviews and that are of different type might be sometime positive tweet or negative tweet also. Task of decision making unit is to calculate such genuine and fake reviews related to it. And in the last aggregate the collected results. Following graph and table shows the total Number of Tweets received for every product and genuine review related to each product. This can be done and it will be always useful for getting business insights for getting market trends and making better decision making.

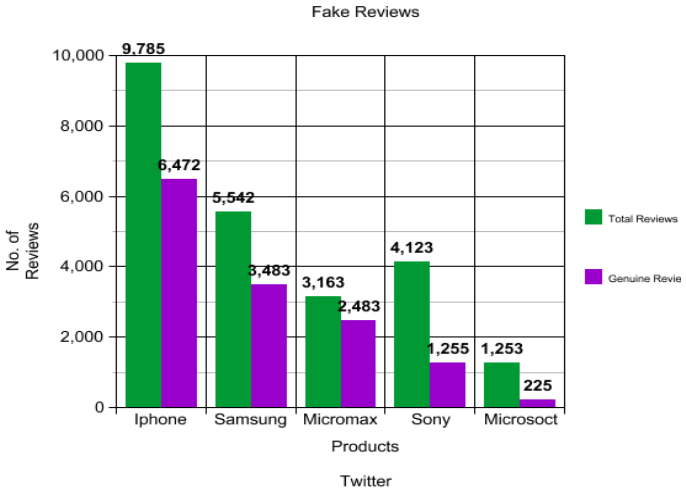

Fig8. Total Number of Tweets

Table 1. Total Number of tweets

Product	Total Number Reviews	Genuine Reviews
Iphone	9785	6472
Samsung	5542	3483
Micromax	3163	2483
Sony	4123	1255
Microsoft	1253	255

C. Performance Evaluation

When it comes to the performance calculation accuracy and processing time plays important role. We use deep learning API (meaning cloud API) for performing classification. It

provides highly scalable environment against complex functionalities and performs better with less time. Following graph and table gives the processing time required for the 100 tweets. In our system time required for processing 100 tweets is 146 seconds and for other rosette API it is 174 seconds per hundred tweets so our system proves better in time. It is comparatively faster.

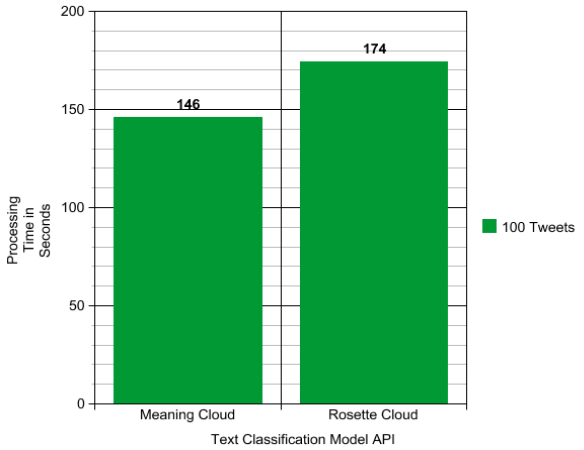

Fig 9. Processing Time

Table 2: Processing Time

API	Time Taken in Secs	Number of Tweets Processed
Meaning Cloud	146	100
Rostte Cloud	174	100

We observe the accuracy of our proposed system comparatively better. Following graph and table shows that the total number of relevant jobs per 100 tweets. We found that in our system number of relevant job out of 100 are 84 while in comparative API it is only 77 out of 100 which lower than our system. These shows that accuracy of system is more than the other system.

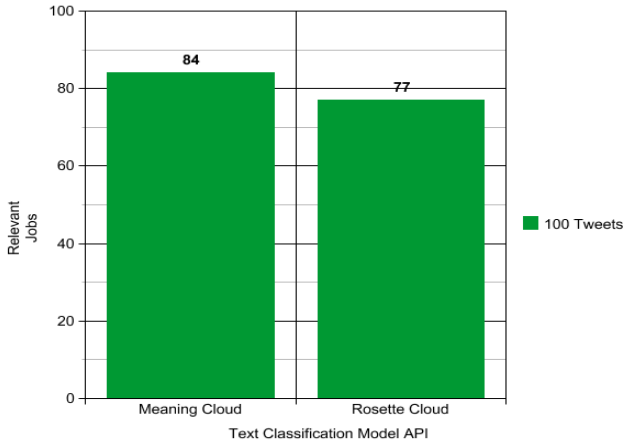

Fig10. Accuracy Comparison

Table 3. Accuracy Comparision

API	Number Tweets	Relevant Tweets
Meaning Cloud	100	84
Rosette Cloud	100	77

We have taken another requirement instead of using mobile phones instead now we have taken news item starting from '#'. The same process will be repeated and tweets will be collected out of them we have to retain only those tweets as per our requirement. Data-Acquisition unit will collect, store and pre-process data. After that processing unit will process the pre-processed data and then results will be given to the aggregation and decision server. Following table describes that it has taken a search of five hot news topics which are trending and then all related tweets related to them are processed. So it has got certain number of tweets related to each topic. Next is to aggregate results such that we want to calculate how many are genuine and fake review. So Decision making unit finds out the number fake and genuine reviews and aggregates the results.

Table 4. Aggregated Results

Sr No	Hash tags (#)	No of Tweets	Positive tweets (%)	Negative tweets (%)
	#narendramodi	100	78	22
	#kumbh	100	69	31
	#drought	100	35	65
	#madhyapradesh	100	75	25
	#pd12345	10	6	4

VIII. CONCLUSION

We have proposed a framework for processing of Real-Time Big-Data. Framework is designed in such a way that it can be effectively used for processing of Real-Time Data and offline data and analysed for effective decision making. Big data is independent application which is also used for any type processing for huge volume of data. It has more capability for filtering, dividing and parallel processing of data. It can discard all that extra data and process only good data. Techniques which we have used in this paper are more capable of processing parallel and analyses data sets for better understanding of the data.

REFERENCES

[1] Tanuja A, Swetha Ramana D, "*Processing and Analyzing Big data using Hadoop*", International Journal of Computer Sciences and Engineering, Vol.4, Issue.4, pp.91-94, 2016.

[2] J. Cohen, B. Dolan, M. Dunlap, J. M. Hellerstein, and C. Welton, "Mad skills: New analysis practices for Big Data", PVLDB, vol. 2, no. 2,pp. 1481–1492, 2009

[3] J.Dean and S.Ghemawat , "*Mapreduce : Simplified data processing on large clusters*", Commun. ACM, vol. 51, no. 1, pp. 107–113, 2008. 5

[4] Prakash Singh , "*Efficient Deep Learning for Big Data: A Review*", International Journal of Scientific Research in Computer Science and Engineering, Vol.4, Issue.6, pp.36-41, 2016.

[5] .Renu Bhandari, Vaibhav Hans and Neelu Jyothi Ahuja, "*Big Data Security – Challenges and Recommendations*", International Journal of Computer Sciences and Engineering, Vol.4, Issue.1, pp.93-98, 2016.

[6] S. Jadav, P. Rawool, V. Shah, "*Image Merging in Transform Domain*", International Journal of Scientific Research in Network Security and Communication, Vol.5, Issue.1, pp.38-39, 2017.

[7] Y.Wang et al., "Using a remote sensing driven model to analyze effect ofland use on soil moisture in the Weihe River Basin, China", IEEE J. Sel.topics Appl. Earth Observ. Remote Sens., vol. 7, no. 9, pp. 3892–3902,Sep. 2014.

[8] M. Clerc, "The Swarm and the Queen: Towards a Deterministic and Adaptive Particle Swarm Optimization", In Proceedings of the IEEE Congress on Evolutionary Computation (CEC), pp. 1951-1957, 1999. (conference style)

[9] T. Balaji , "*Robust and Realistic Classification of Massive Gray Level Thresholding in Remote Sensing Images*", International Journal of Computer Sciences and Engineering, Vol.2, Issue.11, pp.31-38, 2014.

[10] P.Kodimalar, "*A Study on Big Data and Big Data Analytical Research and Issues*", International Journal of Computer Sciences and Engineering, Vol.3, Issue.11, pp.171-179, 2015.

[11] S. Pumpichet, N. Pissinou, X. Jin, and D. Pan, "*Belief-based cleaning in trajectory sensor streams",* 2012 IEEE International Conference on Communications (ICC), Jun. 2012.

[12] A. Paul, J. Wu, J.-F. Yang, and J. Jeong, "*Gradient-based edge detection for motion estimation in H.264/AVC",* IET Image Processing, vol. 5, no. 4, p. 323, 2011.

[11] JHS Tomar, JS Kumar, "*A Review on Big Data Mining Methods*", International Journal of Scientific Research in Network Security and Communication, Vol.4, Issue.3, pp.7-14, 2016.

[12] Antonio Plaza ,Jon Atli Benediktsson , Joseph W. Boardman , Jason Brazile , Lorenzo Bruzzone., "*Recent advances in techniques for hyperspectral image processing",* Remote Sensing of Environment, Vol.113, No.1, pp. S110-S122

[13] Y.Wang, "*Using a remote sensing driven model to analyze effect ofland use on soil moisture in the Weihe River Basin, China",* IEEE J. Sel.topics Appl. Earth Observ. Remote Sens., vol. 7, no. 9, pp. 38923902,Sep. 2014.

[14] E. Christophe, J. Michel, and J. Inglada., "*Remote sensing processing:From multicore to GPU*", IEEE J. Sel. Topics Appl. Earth Observ. Remoteens., vol. 4, no. 3, pp. 643652, Aug. 2011.

[15] R. A. Dugane and A. B. Raut., "*A survey on Big Data in real-time.; Int. J.Recent Innov*", Trends Comput. Commun., vol. 2, no. 4, pp. 794797, Apr.2014.

[16] A. Plaza et al., *"Recent advances in techniques for hyperspectral image processing*",Remote Sens. Environ., vol. 113, pp. 110122,2009.

[17] Zilong Zou, Yuequan Bao, Fodan Deng, and Hui Li., *"An Approach of Reliable Data Transmission With Random Redundancy for Wireless Sensors in Structural Health Monitoring",* IEEE Sensors journal, vol. 15, no. 2, Feb 2015.

[18] M. Mayilvaganan and M. Sabitha., "*A cloud-based architecture for Big-Data analytics in smart grid: A proposal",* in Proc. IEEE Int. Conf. Comput. Intell. Comput. Res. (ICCIC), 2013.

[19] Z. Liu, B. Jiang, and J. Heer., immense, "*Real-time visual querying of Big Data",* Comput. Graph. Forum, vol. 32. no. 3, pp. 421430, pt. 4, 2013.

