

TREND ANALYSIS ON TWITTER

Mr. Rahul Neve¹, Abhishek Singh², Sundaresh Narayanan³

¹Mr. Rahul Neve, Dept of IT, Thakur College of Engineering and Technology, Maharashtra, India

²Abhishek Singh, Dept of IT, Thakur College of Engineering and Technology, Maharashtra, India

³Sundaresh Narayanan, Dept of IT, Thakur College of Engineering and Technology, Maharashtra, India

ABSTRACT- In social media forums, it is important for public's perception that an organization is well notify and able to provide services that truly meet their needs and interests. Social media activity by citizens is growing exponentially and there is some evidence that certain government agencies are embracing these engagement, and communication tools, albeit at a far slower pace than other societal sectors. The on-going provocation is that the government generally lacks behind when it comes to adaption of technology tools. Social media tools and social media analytics are no anomaly. While this paper focuses on government organizations, it's fair to say all federal organization are playing catch up with commonly employed social media communication platforms. The government is even further behind in the area of social media analytics, the cluster of data from online communication tools and the interpretation of this data to form actionable and relevant information. Solutions to the provocation of this knowledge discovery from the magnitude of social media data must be automated and sophisticated enough to detect sentiment, trends, influencers, and real time issues as they arise. The basis of any social media analytics solution will be the power of its search and discovery knowledge engine.

Keywords: LDA algorithm, Clustering, Trend Analysis, Social Media Analytics

1. INTRODUCTION

Despite the talks in the market place regarding what social media is, or how it should, impact government operations, the use of information technology to change the way government works and engages with citizens is not new. President Bill Clinton, with the assistance of Vice President Al Gore, initiated what was then called The National Performance Review, which has evolved over time into The National Partnership for generating Government. That endeavor focused on generating greater operational efficiencies, as well as identifying how government could grasp the technology to work more effectively. The overall goal was to cut government costs. This undermines efforts to increase efficiency, effectiveness, and responsiveness. The proliferation of social media throughout American society has put the spotlight on a long standing troubling perception of government – that it is often slow to react and hesitant to move to new solutions and technologies that the general public has already adopted, vetted, and use in a ubiquitous manner. This slow adaptation is often a result of common issues, investigation of potential risks and security concerns, and/or compatibility analysis. Many argue this scrutiny and delay is overbearing and slows the services government agencies provide to the public; others argue it is simply a result of years of entrenched bureaucracy and regulations that lead to slow adaptation to a changing world and dynamic operating environment There are, of course, those who argue that this late-adopter role is one that is necessary to mitigate the risks to the mission of any given agency. In any case, it is clear that while a major part of society has made social media a permanent and complete aspect of their communication strategy, the government, with only a few exceptions, has not yet fully adopted by social media and analytics, and integrated them into how the government conducts its business. What is even unclear is just how government can grasp the myriad of social media tools and channels in order to provide superior service to a better informed and more engaged citizenry.

1.1 MATERIALS AND MANAGEMENT

Requirements

- Web browser
- Windows Operating System
- Windows XP, Windows 7,8
- MySQL Server

Methodology Used:

For developing a project or software it is important to know the flow or which type of software development lifecycle model the system is going to use. Out of various SDLC models there is a need to choose the model which will be perfect for the project and therefore Iterative model has been selected.

As the name suggest Iterative model is basically an iterative process where the input is processed through various phases and if it doesn't satisfy it is repeated until it becomes satisfactory.

The model consists of five phases namely:

PHASE I:

1. **Planning:** Applying agile methodology for event system event planning for better usage of resources and time with great performance.
2. **Analysis:** Prepare a detailed analysis on present event system and overcome its limitation in our upcoming system.
3. **Design:** Integration of data and designing of system event app.
4. **Coding:** Writing the whole system app code and taking help of an open source.

PHASE II:

1. **Testing:** Doing various tests on system like unit testing, smoke testing, stress testing, integration testing, regression testing, system testing etc. to check whether the integrated system works as the client desires.
2. **Deployment:** Conduct beta testing for identifying bugs and errors left out and improvements that can be made. After the testing and approval, deploy the proposed system.

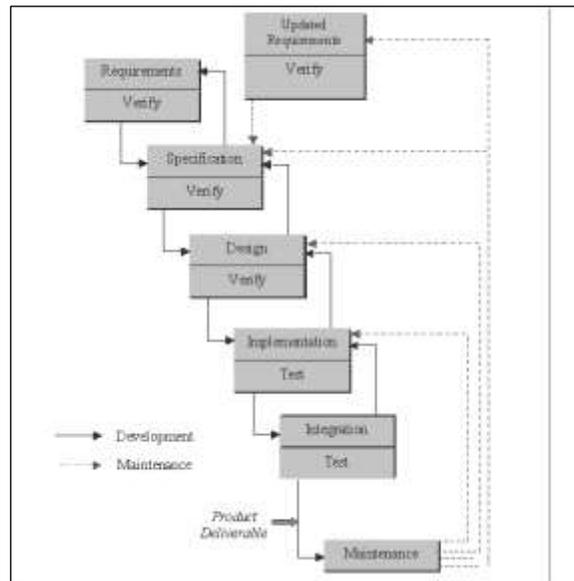


Chart 1: Iterative model

Feasibility Study:

Prior developing the product or software it is necessary to know all the factors important for developing a good product. It is important to know the resource requirements, strengths and weakness of the project, etc. All these factors ultimately revolve around two major criteria cost and value, in order to know how much feasible, it is to create a project and how much profit will it return.

The feasibility of a particular software product can be determined from the following factors:-

1. Technical Feasibility.
2. Operational Feasibility.
3. Economic Feasibility.
4. Legal Feasibility.

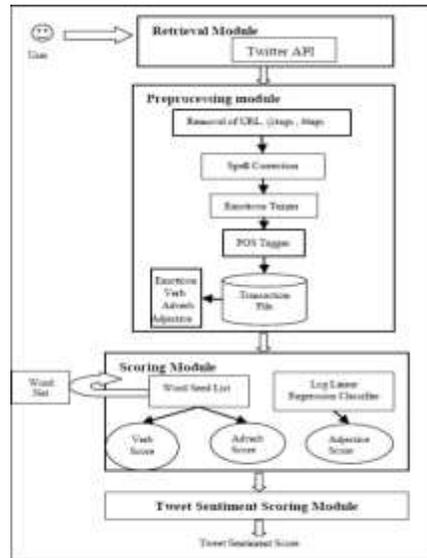


Chart 2: Flowchart of Model.

1.2 Result

Users can look for social events happening around them through our web application. They can get acknowledged of recent trending events just by entering the keywords. Tweets from all around the world can be accessed with the click of a button. Users can view the result graphically which will be segregated on the basis of time, location and legitimacy.

a) Output

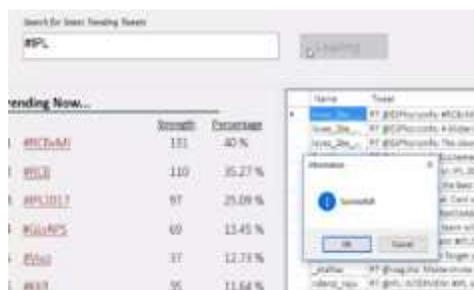


Figure 1: Query Output

b) Outcomes

The client should be able to connect to the server properly without any problems. The connection establishment between the mobile device and the server should take minimal time. The mobile device should be able receive data from the server uninterrupted. Information provided by the application should be correct and as per the user's need.

c) Discussion of the results

Connection can be established easily provided that the server is on. The connection with the server takes time as it uses Internet connection. Receiving data from the server takes time. Information coming from the database is correct

1.3 Discussion

Before the introduction of our web application users had to have an access of Television or Newspapers to get aware about the trending events happening around us. To overcome this problem, we introduced a web application that could present you with trending tweets from around the world with the click of a button. It becomes convenient for users to acknowledge themselves about the events around them through a non-complex application.

Latent Dirichlet Allocation Algorithm has been used to cluster the tweets into different categories. Providing security to authorized user using Two Factor authentication was a one of the major problem.

2. Conclusion

In this project we develop a real-time non parametric model to know about the recent trends. It successfully detects the recent trends and it also shows the trends according to the location of a user.

3. References

1. Hsinchun Chen "Exploring extremism and terrorism on the web: The dark web project" 2007
2. Amit Pratap Singh, Dr, R. C. Jain "A survey on different phases of web usage for Anomaly user Behavior Investigation" 2007
3. Floras Tsai, KapLuk Chan "Detecting cyber security threats in weblogs using probabilistic model" 2007
4. Yulei Zhang, Shuo Zeng, Li Fan, Yan Dang, Catherine A. Larson and Hsinchun Chen "Dark web forums portal: Searching and analyzing Jihadist forums" 2009.
5. The application responds quickly while making a decent decision. The application uses LDA algorithm which is efficient for clustering. Irrespective of users location he can access trending tweets around the world. Positive and negative tweets are being clustered to perform sentimental analysis.