A CREDIBILITY ANALYSIS SYSTEM FOR ASSESSING INFORMATION ON TWITTER

V.Sarala [1], E.Jetendra Gnana Sai Malleshwar[2],

[1]**Assistant professor, MCA DEPT,** Dantuluri Narayana Raju College**, Bhimavaram, Andharapradesh**
Email:- vedalasarala21@gmail.com
[2]**PG Student of MCA,** Dantuluri Narayana Raju College**, Bhimavaram, Andharapradesh**
Email:- jithendraeeli@gmail.com

ABSTRACT

Information credibility on Twitter has been a topic of interest among researchers in the fields of both computer and social sciences, primarily because of the recent growth of this platform as a tool for information dissemination. Twitter has made it increasingly possible to offer near-real-time transfer of information in a very cost-effective manner. It is now being used as a source of news among a wide array of users around the globe.

The beauty of this platform is that it delivers timely content in a tailored manner that makes it possible for users to obtain news regarding their topics of interest. Consequently, the development of techniques that can verify information obtained from Twitter has become a challenging and necessary task. In this paper, we propose a new credibility analysis system for assessing information credibility on Twitter to prevent the proliferation of fake or malicious information.

The proposed system consists of four integrated components: a reputation-based component, a credibility classifier engine, a user experience component, and a feature-ranking algorithm. The components operate together in an algorithmic form to analyze and assess the credibility of Twitter tweets and users. We tested the performance of our system on two different datasets from 489,330 unique Twitter accounts. We applied 10-fold cross-validation over four machine learning algorithms. The results reveal that a significant balance between recall and precision was achieved for the tested dataset.

1 INTRODUCTION

Generally, data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases.

2 LITERATURE SURVEY

A new model for classifying social media users according to their behaviors:

User generated content in online social media is growing rapidly, which makes it hard to be validated and verified. Face book and Twitter are the most popular social media that are being used as a means of social communication and sharing thoughts, knowledge and even news. Information in these social networks can be generated by anyone from anywhere in anytime. Classifying such huge information using traditional data mining classification algorithms is time consuming task which needs huge processing and memory space.

3 IMPLEMENTATION STUDY

EXISTING SYSTEM:

Currently, researchers have employed various methodologies in studies on information credibility . Some of them consider the problem to be one of classification that should be solved in an automated fashion using machine learning or graph-based algorithms. Others view it as a cognitive problem requiring human-centric verification.

Disadvantages:

OSNs by their very nature evolve dynamically over time and become very large in size, with various structures that make it difficult to obtain the information needed to discern the credibility of users.

PROPOSED SYSTEM

We propose a novel credibility assessment system that maintains complete entity-awareness (tweet, user) in reaching a precise information credibility judgment. This model comprises four integrated components, namely, a reputation-based model, a feature ranking algorithm, a credibility assessment classifiers engine, and a user expertise mod-el. All of these components operate in an algorithmic form to analyze and assess the credibility of the tweets on Twitter.

Advantages:

Our results show that the system that employed a reputation-based filter approach provide a significant and accurate credibility assessment.

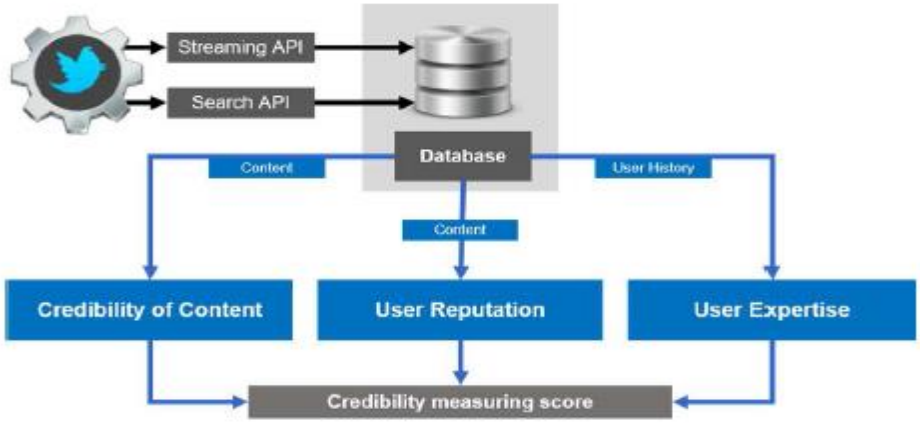


Fig:3.1 System Architecture

4.IMPLEMENTATION

MODULES:

1. Admin Module:
   Calculate credibility of content
   Calculate user reputation
   Calculate user expertise
   Find Credibility measuring score
   Graphical Representation
2. User Module:

OSN System Construction Module

Admin Module:

In this module, the Admin has to login by using valid user name and password. After login successful he can do some operations such as view all user and their details and authorize them.

Calculate credibility of content:

This module is part of admin module who will calculate each user credibility by selecting post and user name. In these calculation users credibility is measured through tweet size and @ mentions in the post. This analyzed data is sent to calculate user reputation.

Calculate user reputation:

In this user reputation calculation method total number of friends , and total number of tweets are calculated for that respective user on that specific topic. If user has more friends and more tweets his reputation is more.

Calculate user expertise:

Expertise of each user is calculated by finding users history of messages he posted on respective topic. If user has more messages on his history than he was activity participating in discussion.

Find Credibility measuring score:

Based on these 3 factors calculation credibility score is given to each user for that topic. If score is high user has high credibility.

5. RESULTS AND DISCUSSION

SCREEN SHOTS

5.3.1 HOME PAGE:

Fig: 5.1 Home Page

5.3.2 USER REGISTRATION:

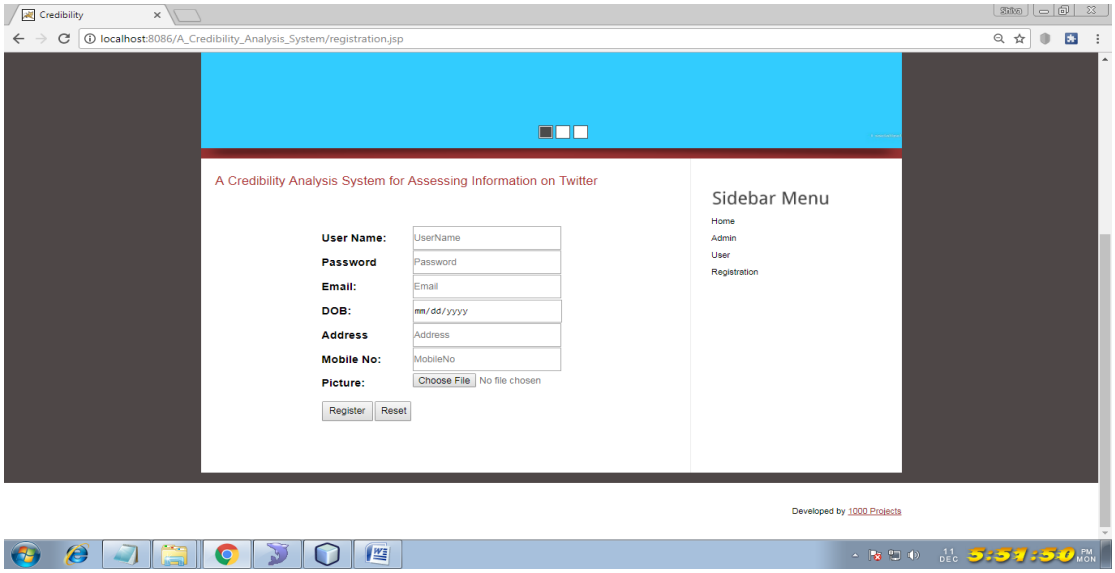


Fig:5.2 User Registration

5.3.3 USER LOGIN:

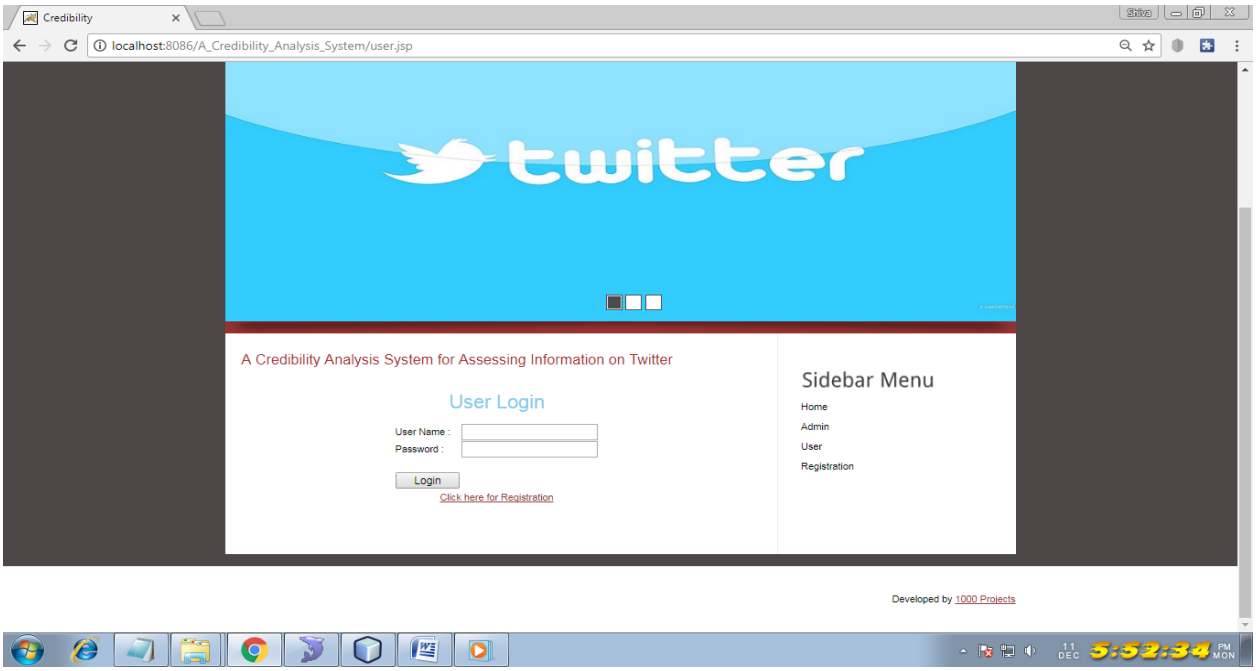


Fig:5.3 User login

5.3.4 USER HOME:

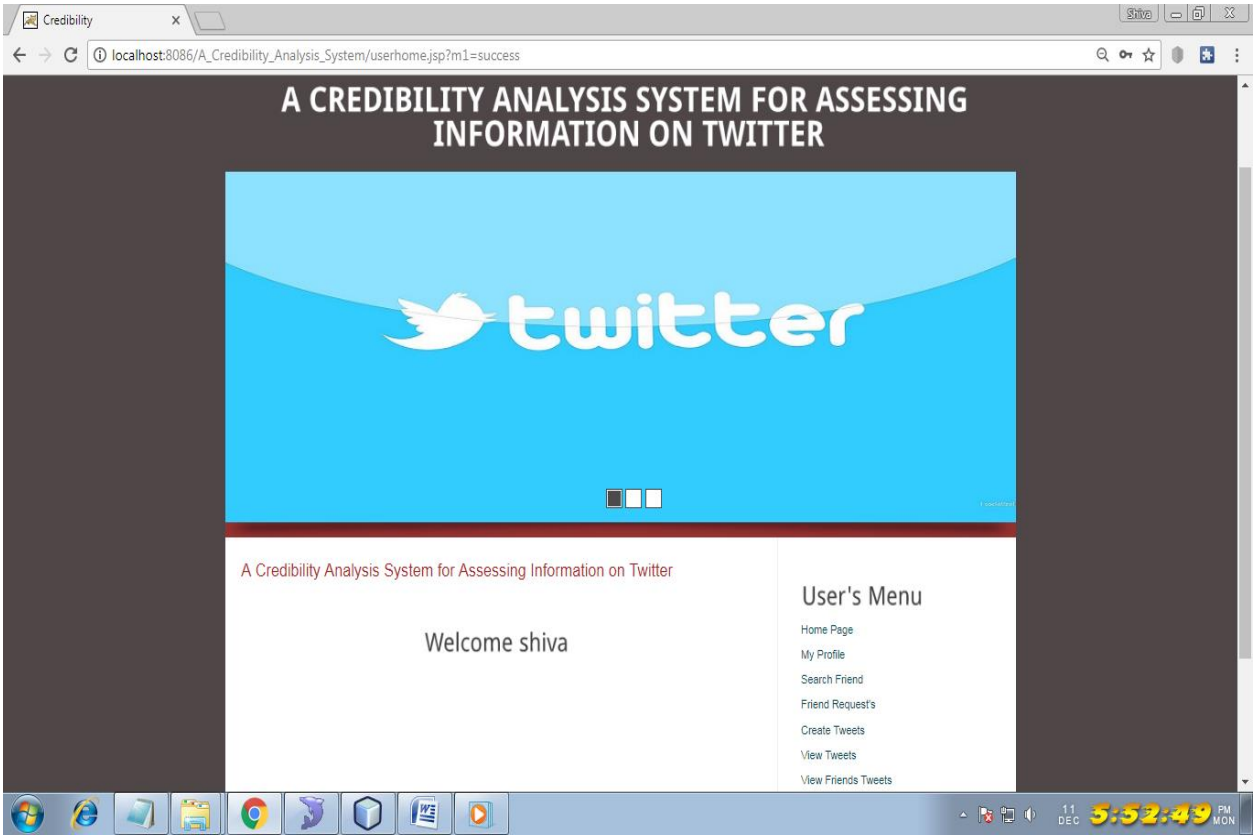


Fig:5.4 User home

5.3.5 PROFILE:

Fig:5.5 Profile

5.3.6 SEARCH FRIENDS:

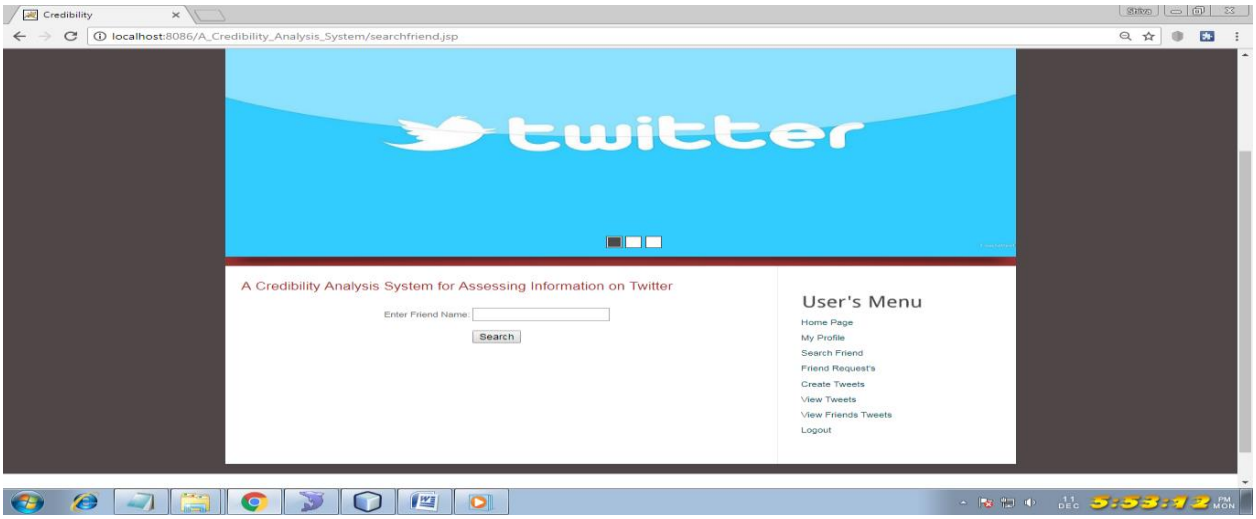
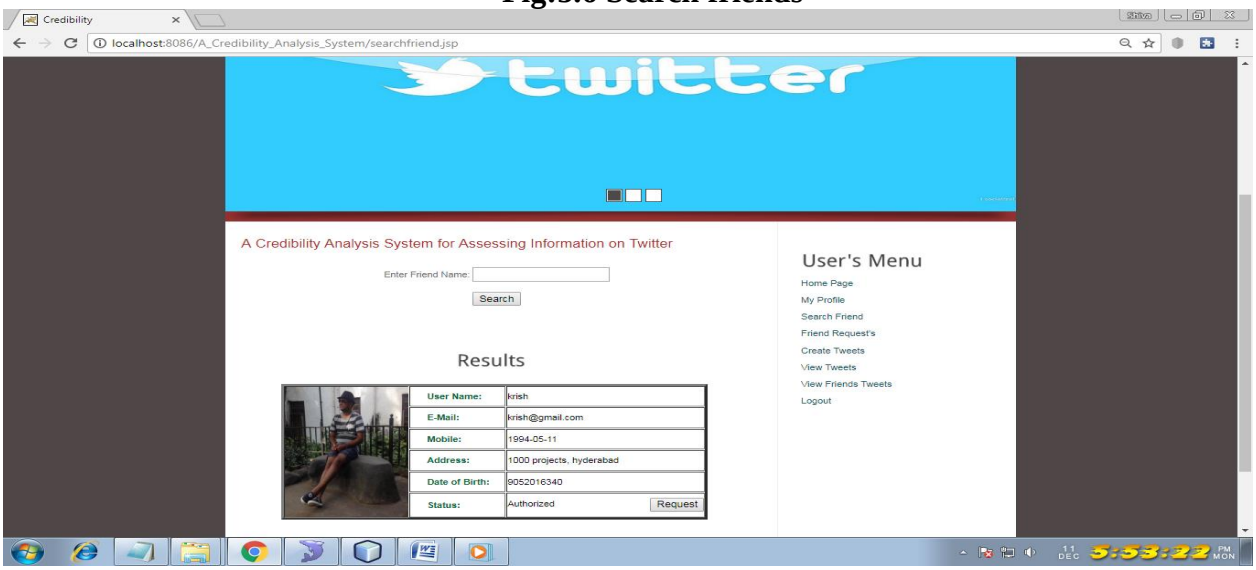


Fig:5.6 Search friends



5.3.7 FRIENDS REQUEST:

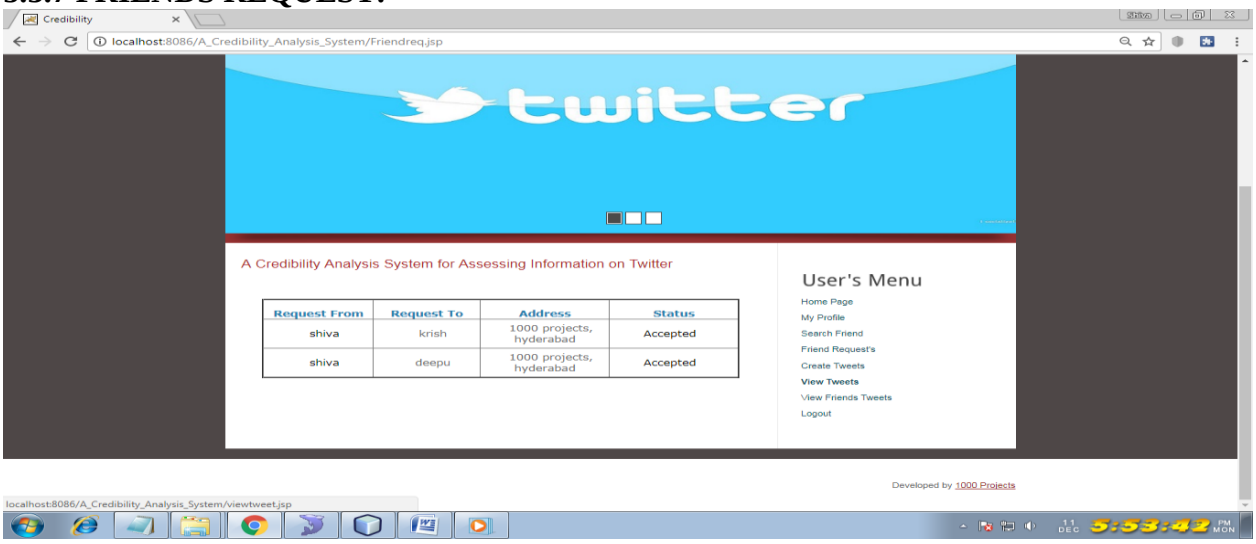


Fig:5.7 Friends request

5.3.8 CREATE TWEET:

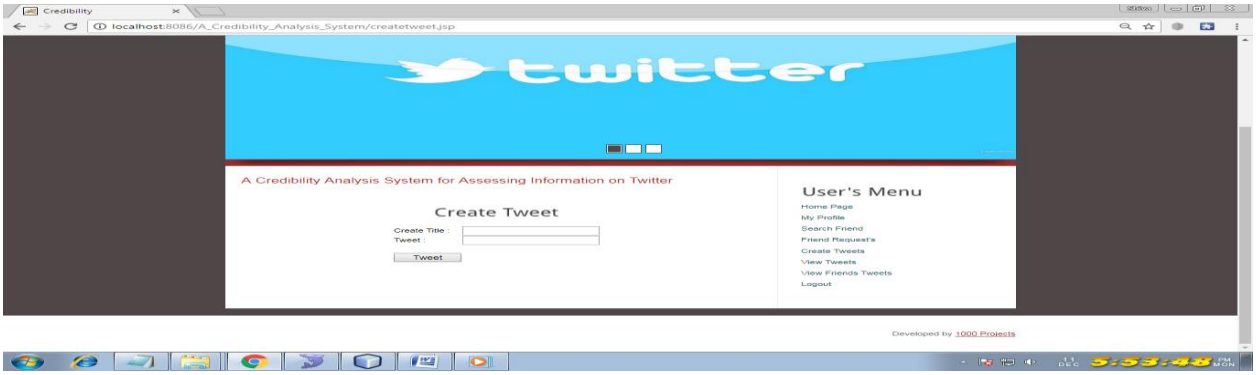


Fig:5.8 Create tweet

5.3.9 VIEW TWEET:

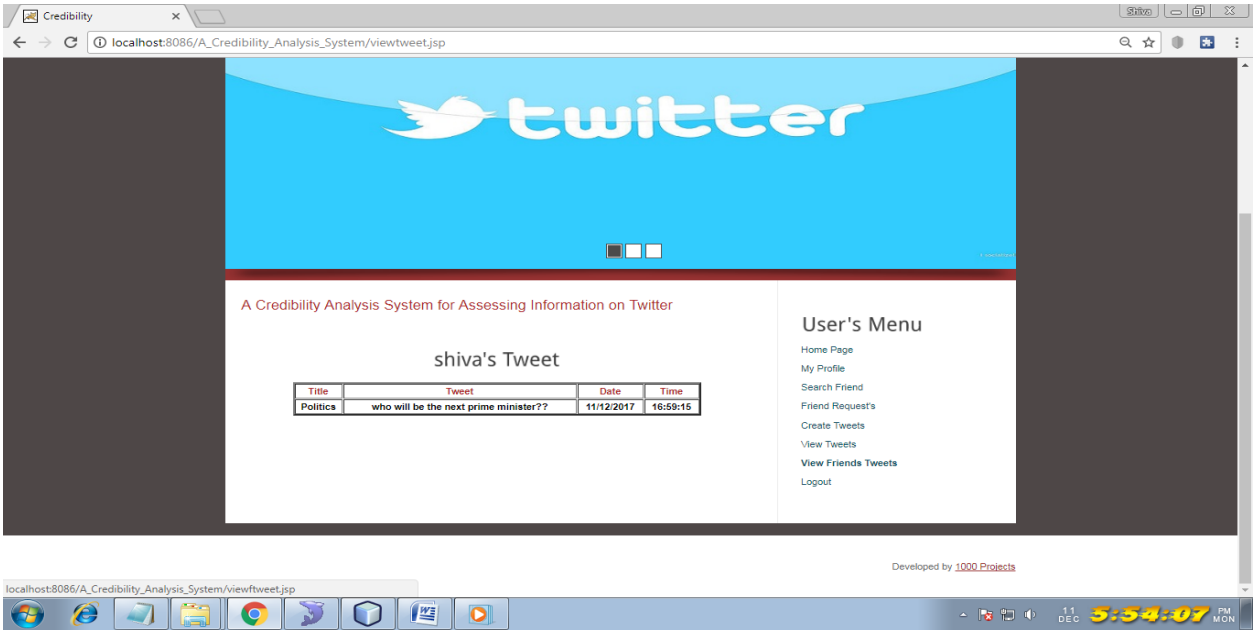


Fig:5.9 View tweet

5.3.10 VIEW FRIENDS TWEET:

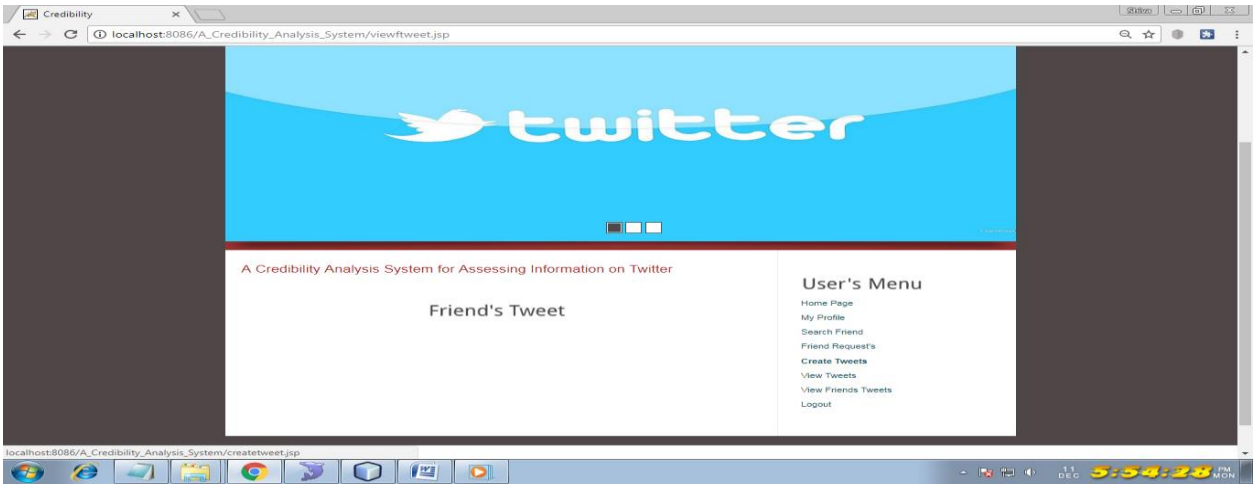


Fig:5.10 View friends tweet

5.3.11 ADMIN LOGIN:

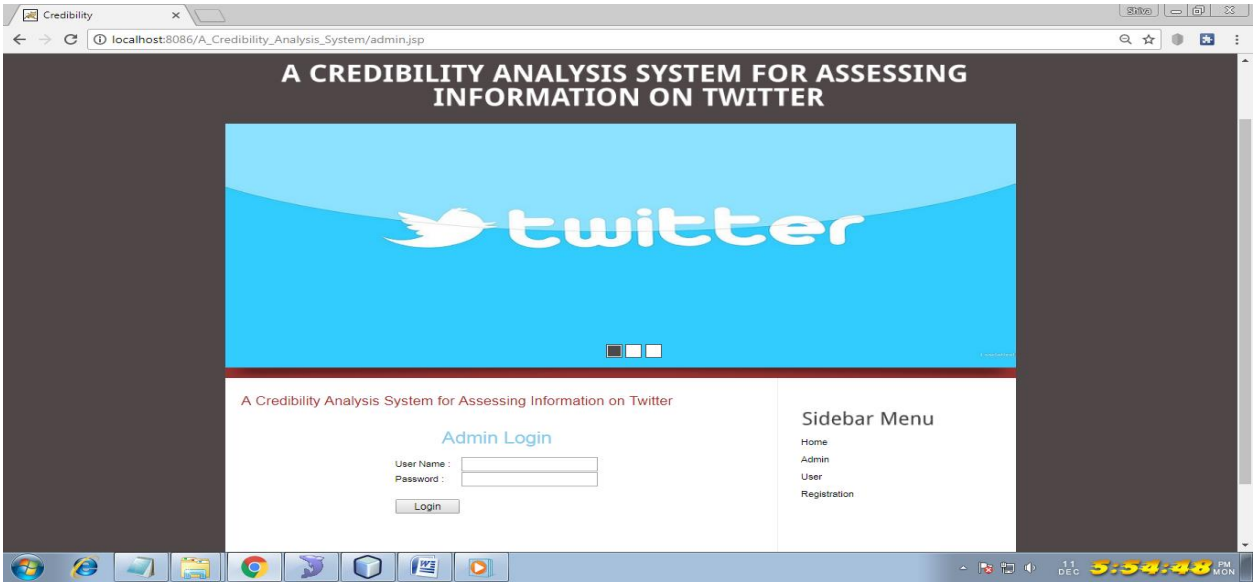


Fig:5.11 Admin login

5.3.12 ADMIN HOME:

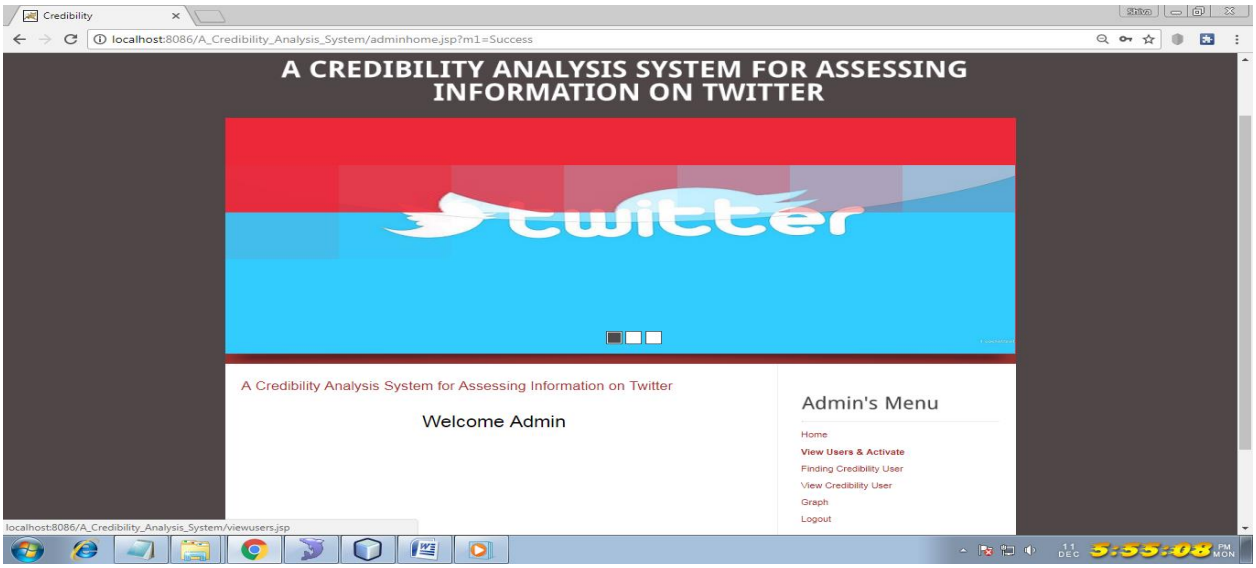


Fig:5.12 Admin home

5.3.13 VIEW USERS & ACTIVATE:

Fig:5.13 View users & activate

5.3.14 FINDING CREDIBILITY USER:

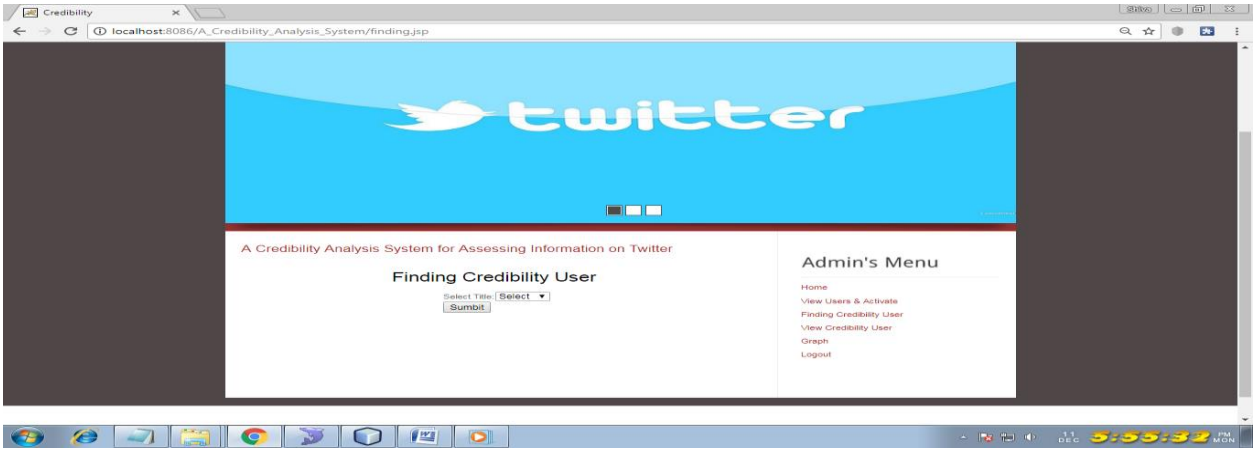
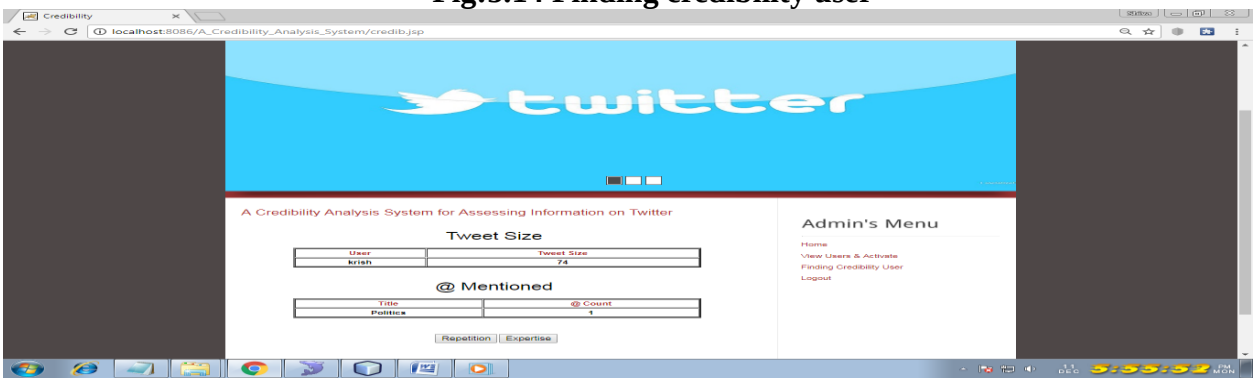


Fig:5.14 Finding credibility user



5.3.15 REPETITION:

Fig:5.15 Repetition

5.3.16 EXPERTISE:

Fig:5.16 Expertise

5.3.17 VIEW CREDIBILITY USER:

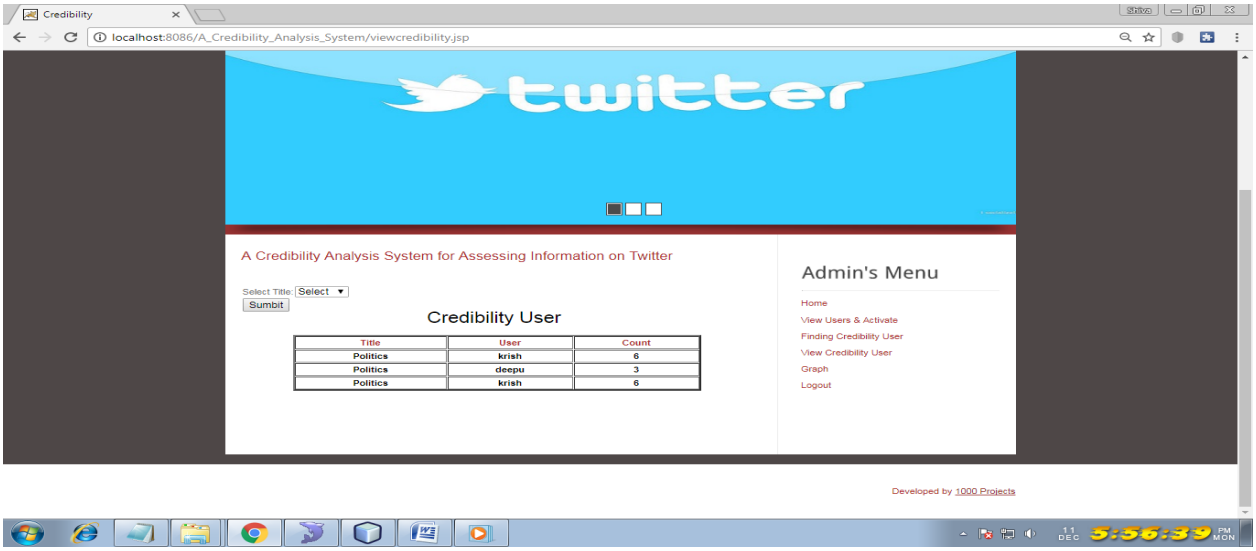


Fig:5.17 View credibility user

5.3.18 GRAPH:

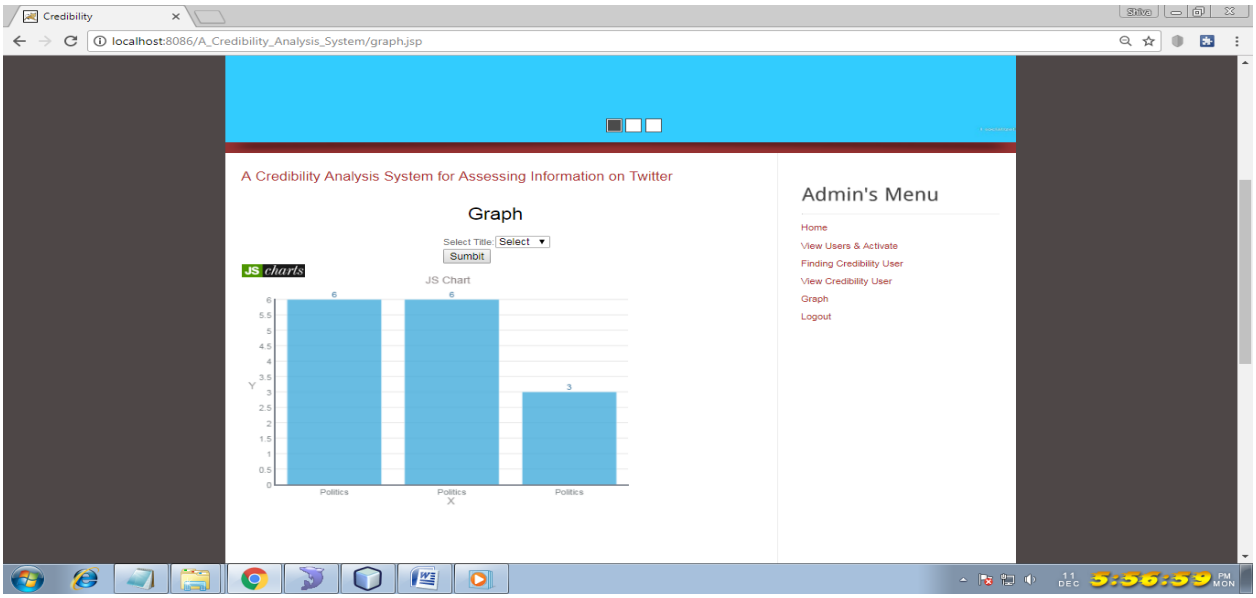


Fig:5.18 Graph

6. CONCLUSION AND FUTURE WORK

CONCLUSION

This paper presents the results of a study of the problem of assessing information credibility on Twit-ter. The issue of information credibility has come un-der scrutiny, especially in social networks that are now being used actively as first sources of information. Twitter and other social networks have be-come widely used in disaster mitigation in cases of high-impact events because they make it possible for relevant parties to obtain important information sufficiently quickly to coordinate countermeasures to such events.

To obtain a better understanding of how to assess information credibility on Twitter, we measured and characterized the content and sources of Twitter tweets. By crawling Twitter, we collected data from more than 1,416,443 tweets by 489,330 unique users. In addition, we examined data for 2,843 Twitter users with more than 7,870,549 tweets. Based on the data, we extracted the features that can be of most help in the assessment process. Based on our feature ex-traction process, we designed an automated classification system that consists of four main components: a reputation-based component, a credibility classifier engine, a user experience component, and a feature rank algorithm.

7. REFRENCES

- M. Al-Qurishi, R. Aldrees, M. AlRubaian, M. Al-Rakhami, S. M. M. Rahman, and A. Alamri, "A new model for classifying social media users according to their behaviors," in Web Applications and Networking (WSWAN), 2015 2nd World Symposium on, 2015, pp. 1-5
- M. AlRubaian, M. Al-Qurishi, M. Al-Rakhami, S. M. M. Rahman, and A. Alamri, "A Multi-stage Credibility Analysis Model for Microblogs," presented at the Proceedings of the 2015 IEE E/ACM International Conference on Advances in Social Networks Analysis and Mining 2015, Paris, France, 2015
- A. AlMansour, L. Brankovic, and C. S. Iliopoulos, "Evaluation of credibility assessment for microblogging: models and future directions," in Proceedings of the 14th International Conference on Knowledge Technologies and Data-driven Business, 2014, p. 32.

- Majed AlRubaian, Muhammad Al-Qurishi, Sk Md Mizanur Rahman, and A. Alamri, "A Novel Prevention Mechanism for Sybil Attack in Online Social Network," presented at the WSWAN'2015, 2015
- P. T. M. Gayo-Avello, Eni Mustafaraj, Markus Strohmaier, Harald Schoen, D. Peter Gloor, C. Castillo, M. Mendoza, and B. Poblete, "Predicting information credibility in time-sensitive social media,"Internet Research, vol. 23, pp. 560-588, 2013.
- Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," presented at the Proceedings of the 20th international conference on World wide web, Hyderabad, India, 2011.