JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

SOCIAL MEDIA NETWORK REVIEWING OF APPLICATION USING ARTIFICIAL INTELLIGENCE AND BIG DATA

[1]Mr R Vinoth Kumar, [2]Mrs D Maladhy [3]M Praveena,[4]R Rakshana

[1]Head of department, [2]Assistant Professor, [3]Student, [4]Student
[1]Department of information technology,
[1]Rajiv Gandhi college of engineering and technology, Puducherry, India
[2]Department of information technology,
[2]Rajiv Gandhi college of engineering and technology, Puducherry, India

Abstract : Social networks have become the major infrastructure of today's daily activities of people. Peoples can able to interact with each other in these networks and share their interest on resources and give their opinions about these resources or spread their information to world. Trustfulness plays a main role on identifying a suitable product or specific user. The inference mechanism of trustfulness in social networks refers to utilizing available information of a specific user who intends to contact an unknown user. This mostly happens when purchasing a product, deciding to have friendship or other applications which require predicting the reliability of the second party.

IndexTerms - Socialmedia; Machine Learning; Classifiers; K-Nearest Neighbor(KNN); Twitter API

I. INTRODUCTION

Nowadays technology has become a very important part of our lives and most people can't live without it. The Internet provides a platform to share their ideas. Many people are spending a large amount of time on socialmedia. Communicating with people is no exception, as has changed the way people interact more extensively and has given a new dimension to communication. Many people are illegally using these communities. Many youngsters are getting bullied these days. Bullies use various services like Twitter, Facebook,and Email to bully people. Studies show that about 37% of children in India are involved in socialmedia and nearly 14% of bullying occurs regularly.

II. ABOUT THE PROJECT

In recent days, millions of users around the world are connected by means of Online social networks(OSNs), such as Facebook, Twitter, and Weibo. In fact, the number of users in these networks is increasingly growing despite some of them may have decrease in the number of active users on the other hand, users exchange huge amounts of information in social networks every day; based on the level of trust factor is one of the most important issues, when choosing a book to read, we may choose a book that we know its writer, or a book that is suggested by someone who we trust in. Hence, users in social networks share information with other users according to their trust in them.

Literature Survey

In [1], In this paper, we propose a trust evaluation framework based on machine learning to facilitate human decision making by extensively considering multiple trust-related user features and criteria. We first divide user features into four groups according to the empirical analysis, including profile-based features, behavior-based features, feedback-based features, and link-based features. Moreover, four traditional trust evaluation methods are employed to compare with our machine learning based methods. Experiments conducted on a real-world dataset show that the overall performance of our features and methods is superior to the other existing features and traditional approaches.

In [2], The data was labeled using a web service, Amazon's Mechanical Turk. We used the labeled data, in conjunction with machine learning techniques provided by the Weka tool kit, to train a computer to recognize bullying content. Both a C4.5 decision tree learner and an instance-based learner wereable to identify the true positives with 78.5% accuracy.

In [3], This paper, trust evaluation is formalized as a classification problem and a novel approach utilizing machine learning method is presented. Firstly, the trust feature vector is constructed according to the trust related factors. Then by training with

collected sample data which contains trust feature vectors and trust ratings, a trust classifier can be established. Based on this approach, a general trust evaluation framework is proposed for network-level trust evaluation. The experiments in real social networks verify the feasibility of our framework and show that the trained trust classifier has a relatively high predicting accuracy.

In [4], The system is capable of simulating trust between two directly connected individuals on social networks. By integrating the trust values computed through three trust computing components, the system eventually returns a trust score representing the actual trust magnitude from one particular individual to the other.

III Proposed System

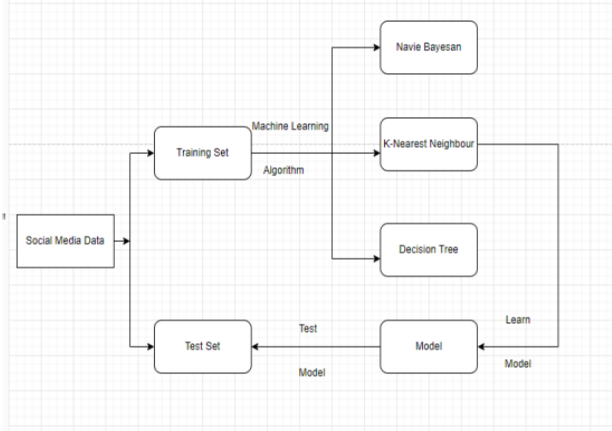

Figure:1

Identifying trustworthy information in the presence of noisy data contributed by numerous unvetted sources from online social media (e.g., Twitter, Facebook, and Instagram) has been a crucial task in the era of big data. This task, referred to as truth discovery, targets at identifying the reliability of the sources and the truthfulness of claims they make without knowing either apriori. In this work, we identified three important challenges that have not been well addressed in the current truth discovery literature. Finally, to classify the trust values, machine learning technique, K-Nearest Neighbor (KNNs) is used instead of traditional weighted sum methods, to express the trust between any two users in the presence of a special pattern. The results show that the accuracy of the proposed method is better than the existing work, and unlike other methods, does not decrease by increasing the number of samples.

3.1 HADOOP

Hadoop is one of the most popular MapReduce implementations. Both input and output pairs of a MapReduce application are managed by an under lying Hadoop Distributed File System (HDFS). At the heart of HDFS is a single Name Node a master server managing the file system namespace and regulates file accesses. The Hadoop runtime system establishes two processes called Job Tracker and Task Tracker. Job-Tracker is responsible for assigning and scheduling tasks; each Task Tracker handles mappers or reducers assigned by Job Tracker.

3.2 MAPREDUCE

MapReduce is a popular data processing paradigm for efficient and fault tolerant workload distribution in large clusters. A MapReduce computation has two phases, namely,the Map phase and the Reduce phase. The Map phase splits an input data into a large number of fragments, which are evenly distributed to Map tasks across a cluster of nodes to process. Each Map task takes in a key-value pair and then generates a set of intermediate key-value pairs. After theMapReduce runtime system groups and sorts all the intermediate values associated with the same intermediate key, the runtime system delivers the intermediate values to Reduce tasks. Each Reduce task takes in all intermediate pairs associated with a particular key and emits a final set of key-value pairs. MapReduce applies the main idea of movingcomputation towards data, scheduling map tasks to the closest nodes where the input data is stored in order to maximize data locality.

3.3 HDFS BALANCER

HDFS data might not always be distributed uniformly across DataNodes. One common reason is addition of new DataNodes to an existing cluster. HDFSprovides a balancer utility that analyzes block placement and balances data across the DataNodes. The balancer moves blocks until the cluster is deemed to be balanced, which means that the utilization of every DataNode differs from the utilization of the cluster by no more than a given threshold percentage. The balancer does not balance between individual volumes on a single DataNode.

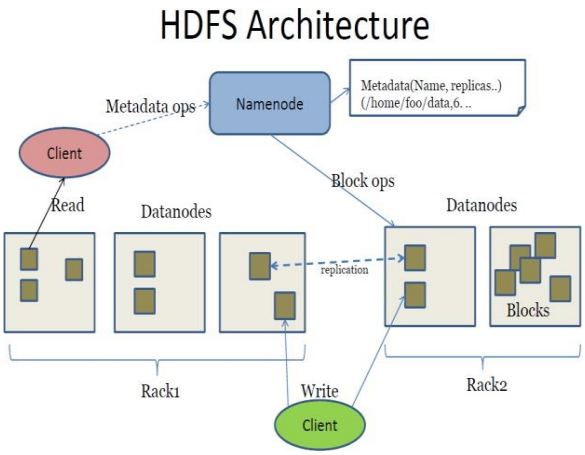

Figure:3.3

3.4 K-NEAREST NEIGHBORS

The k-Nearest Neighbors method (k-NN) is a supervised classifier, like Naive Bayes. The accuracy of this algorithm is highly dependent to the value of k. If k is set toohigh, irrelevant data will be taken into account and thus, the accuracy will decrease. If k is set too low, such as k=1, the one neighbor that will be considered may be noise, and hence, the test data may be labeled incorrectly. In other words, if k is low, the information will be local, and if k is high, the information will be global. In k-NN, to predict the label of a new instance, its Euclidean distance to all of the instances should be calculated. Then, we take the k nearest neighbors and predict its label based on them. Algorithms like k-NN are called instance-based or lazy learners, because: 1) There are no models in these methods to apply toour data. The data should be always available to classify newinstances. Hence, it is called instance based. 2) There is no training and testing in this algorithm and the classification is accomplished in one phase. Thus, it is called lazy.

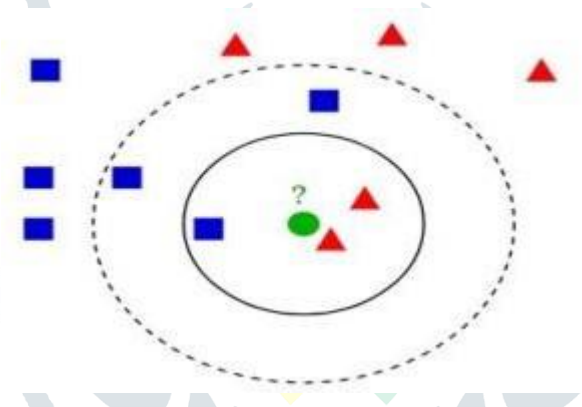

Figure:3.4

IV MODULES

- ✓ 1. KNN
- ✓ 2. Map Module
- ✓ 3. Reduce Module

V FUNCTIONAL REQUIREMENTS

HARDWARE	SOFTWARE
I3PROCESSOR 300 GB. 1.44 Mb. 15 VGA Colour. Logitech. 4 gb.	-Windows XP/Ubuntu java for Maper and Reducer Php,Javascript (Intelligent Graph) Hadoop Cluster Virtual Box Oracle tool

VI SYSTEM ARCHITECTURE

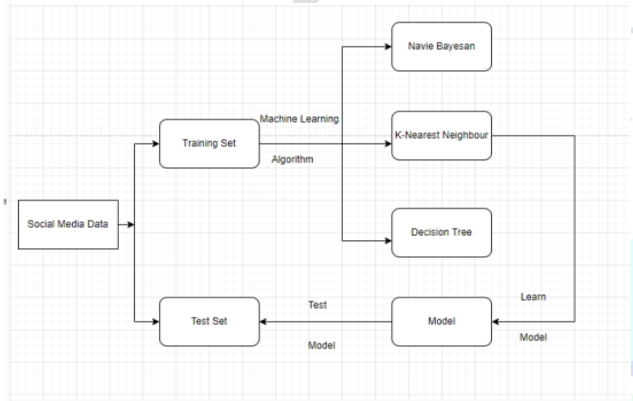

Fig 6.1: Architectural diagram of 4 modules

CONCLUSIONS

The literature depicts that future industries will be fueled by AI through the support of smart machines, Big Data, IIoT, robots, high-speed communication architectures, blockchain, and the broader transition of the economy. We believe such a rigorous analysis of the domain will provide a baseline for future research.

REFERENCES

[1] Xu Chen, Yuyu Yuan, Lilei Lu, and Jincui Yang, "A Multidimensional Trust Evaluation Framework for Online Social Networks Based on Machine Learning", IEEE Translation and Content Mining, IEEE Access, ISSN: 2169- 3536 , Vol. 7, Jul 2019, pp.175499-175513.

[2] I.Hemalatha, Dr.G.P.Saradhi Varma and Dr. A.Govardhan, "Social network analysis and mining using machine learning techniques", Elsevier, Vol. 5, Jun 2019, pp. 603-607.

[3] Z. Kang and P. Li, "A machine learning based trust evaluation framework for online social networks," in IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications, Beijing, 24-26 Sep. 2018.

[4] S. Adali, R. Escriva, M. K. Goldberg, M. Hayvanovych, M. Magdon-Ismail, B. K. Szymanski, W. A. Wallace, and G. Williams, "Measuring behavioral trust in social networks," in IEEE International Conference on Intelligence and Security Informatics (ISI'10), Vancouver, BC, Canada, pp. 150–152, 23-26 May 2017.

[5] S. NEPAL, W. SHERCHAN, AND C. PARIS, "STRUST: A TRUST MODEL FOR SOCIAL NETWORKS," IN 10TH IEEE

[6] S. Trifunovic, F. Legendre, and C. Anastasiades, "Social trust in opportunistic networks," in INFOCOM IEEE Conference on Computer Communications Workshops, San Diego, CA, USA, 15-19 Mar. 2017.

[7] J. Zhan and X. Fang, "A novel trust computing system for social," in IEEE International Conference on Privacy, Boston, MA, USA, 9-11 Oct. 2017.

[8] F. Alam and A. Paul, "A computational model for trust and reputation relationship in social network," in International Conference on Recent Trends in Information Technology (ICRTIT), Chennai, India, 8-9 Apr. 2016.

[9] G. Yin, F. Jiang, S. Cheng, X. Li, and X. He, "AUTrust: a practical trust measurement for adjacent users in social network," in IEEE Second International Conference on Cloud and Green Computing, Xiangtan, China, pp. 360–367, 1-3 Nov. 2016.