MINIMIZING INFLUENCE OF RUMOURS ON SOCIAL NETWORKS USING MACHINE LEARNING ALGORITHMS AND ANALYSIS

T. C. Subash Ponraj\textsuperscript{1}, S. S. Subashka Ramesh\textsuperscript{2}

\textsuperscript{1}Department of Computer Science & Engineering, SRM Institute of Science and Technology, Ramapuram Campus, Chennai, Tamil Nadu, India
\textsuperscript{2}Department of Computer Science & Engineering, SRM Institute of Science and Technology, Ramapuram Campus, Chennai, Tamil Nadu, India

\textsuperscript{1}subashponraj1@gmail.com, \textsuperscript{2}subashka@gmail.com

Corresponding Author: T.C. Subash Ponraj
E-mail: subashponraj1@gmail.com

https://doi.org/10.26782/jmcs.2020.05.00016

Abstract

The advancement of large scale online social networks, online data sharing is turning out to be pervasive consistently. Both positive and negative information is spreading through online social networks. It centres on the negative data issues, for example, online rumours. Blocking of online rumour is one of the major issues in large scale social media networks. Hostile rumours can lead to confusion in the public eye and consequently should be quickly as fast as time permits in the wake of being distinguished. For this we used hybrid SVM, Naive Bayes and KNN algorithm. We will probably limit the impact of the rumour which is the quantity of clients that have acknowledged and sent the rumour by obstructing a specific subset of hubs.

Keywords: Rumour, malicious, Hybrid SVM, Naive Bayes, KNN.

I. Introduction

Network devices that initialize, route and finish the transport of data are called network hubs. Hubs can absorb hosts such as desktops, servers, etc and also network hardware. Two devices are networked together when one device can send and receive information with other device, no matter whether they are connected with one another or not. Mostly, application specific communication protocols are layered over one other.

With the increasing ubiquity of large scale online social media platforms, for example, Twitter, Facebook, and Chinese Sunwido, and so on, a huge number of individuals can become companions and share various forms of data with one another. Social network investigation has surely increased enthusiasm among
scientists. In one case, these online social networks give incredible means for the
dispersion of information which is positive to the social groups, for example, new
innovations and developments. Then again, in another case, they may turn into a
medium for the dispersion of hostile rumours. For instance, some people may write a
post on social media a rumour about a upcoming flood, which may lead to
disturbance among the people and thus may spoil the typical order. So it is essential
to recognize the wellspring of the gossip and furthermore erase the messages
identified with the talk which is adequate to maintain a strategic distance from the
gossip from spreading to increasingly number of individuals.

But in some unfavourable conditions, for instance, fear based online attack, it
might be essential to hinder the related internet based life records to dodge genuine
negative effects. For instance, in 2016, the groups of three out of the forty-nine
unfortunate casualties from the Orlando club shooting scene recorded a case against
Twitter, Facebook and Google for giving "material assistance" to the psychological
warfare relationship of the Islamic State of Iraq and Syria. These associations at that
point took measures to obstruct those records and erased related records and fan
pages on their social organizations to shield the ISIS from spreading noxious
information. Additionally, Facebook have given huge security approaches and
measures to ensure the ability to square records of clients when they are against rules
or in harmful way. Without a doubt, threatening gossip rumours should be ended as
fast as time allows once identified so their negative effect can be restricted. Most of
existing works analyzed the issue of extending the effect of positive information
through online social network. Quick estimation techniques were additionally
proposed to impact boost issue. Interestingly, the negative impact minimization issue
has gained considerably less consideration. There have been many methodologies
being planned for successful blocking of malicious rumours and reducing the cynical
influence.

I. Related Works

Meeyoung Cha, Yajun Wang, Wei Chen, Sejeong Kwon and Kyomin Jung et. al. [I]
proposed that rumour investigation is currently increasing perspective on the grounds
that online web based life empower analysts to look at intently different sorts of data
scattering on the Internet. Right now, investigated social brain science look into
composing on gossipy titbits and endeavoured to perceive the key complexities in the
spread of bits of gossip and non bits of gossip. The outcomes from this examination
can be utilized in improving programmed order of rumours and better appreciate the
talk hypotheses in online social media.

Yajun Wang, Kyomin Jung, Meeyoung Cha, Wei Chen and Sejeong Kwon et. al. [II]
recognized features of gossips by looking at the accompanying three parts of
dissemination: worldly, basic, and etymological. For the worldly attributes, they
proposed another occasional time arrangement model that considers every day and
outer stun cycles, where the model exhibits that talk likely have variances after some
time. They similarly recognized key fundamental and etymological differences in the
spread of gossipy titbits and non bits of gossip. Their chose attributes portray bits of
tattles with high exactness and audit in the extension of 87% to 92%.
Vahed Qazvinian Emily Rosengren Dragomir R. Radev Qiaozhu Mei et. al. [III] tended to the issue of rumour discovery in micro blogs and investigated the adequacy of 3 classifications of highlights: content – based, organize – based, and miniaturized scale blog – explicit images for effectively distinguishing gossips. They additionally demonstrated how these highlights are likewise successful in distinguishing incorrectly witnesses that is clients who underwrite talk and furthermore help it to spread. They performed their assessments on in excess of 10,000 physically explained tweets gathered from Twitter and demonstrated how their recovery model accomplishes more than 0.95 in Mean Average Precision (MAP).

Carlos Castillo, Marcelo Mendoza and Barbara Poblete concentrated on programmed techniques for evaluating the validity of a given arrangement of tweets. Especially, they dissected miniaturized scale blog postings identified with trending subjects, and arranged them as believable or not valid, in light of highlights extricated from them. They used features from customers' posting and reposting conduct, from the substance of the posts, and from references to external sources. They surveyed their techniques using innumerable human evaluations about the legitimacy of things on a progressing case of Twitter postings.

Marcelo Mendoza, Barbara and Carlos Castillo dismembered the movement related to the 2010 quake in Chile and investigated Twitter regular after this catastrophe. They in like manner played out an essential examination of certain social wonder, for instance, the spread of false tattles and avowed news. They separate how this information caused through the Twitter sort out, to assess the trustworthiness of Twitter as an information source under uncommon conditions. Their investigation demonstrated that the engendering of tweets that relate to gossips contrasts from tweets that spread news since bits of gossip will in general be addressed more than news by the Twitter people group. Their outcome demonstrated that it is conceivable to identify bits of gossip by utilizing total investigation on tweets.

Social psychologists guessed that rumour mongers have a limited ability to focus, because rumour can prosper just during a brief time limit when there is a requirement for information in the scarcity of correct news from news platforms. Thus, charges and interrogatory proclamations dependent on conditional proof frequently become rumours, as individuals search out data while the realities are covered in uncertainty. The understanding is about the secondary properties of rumour spreading. This examination shows that rumours spread generally in inefficient structures. One investigation concentrated on the different jobs rumour mongers play, for example, a dispatcher who is an individual who carries relevant data to the group and a sceptic who is an individual who doubts over legitimacy of the information. Every one of these is relied upon to influence the structure of rumour network.

The disadvantage of the existing system is that the best known features are not proved to be accurate enough to classify between a rumour and a non – rumour. Phonetic features are most important factors for identifying the rumours but it is found out that existing systems have failed to identify the rumours accurately on the basis of linguistic features than baseline features.

Copyright reserved © J. Mech. Cont.& Math. Sci.
T. C. Subash Ponraj et al
II. Problem Statement

The main aim of this system is to detect and find out the rumours on social media. The spread of rumours has the capability to induce extremely negative impacts on society. Rumour is termed as an explanation whose reality esteem is unsubstantiated or is bogus. False rumours are very menacing as they cause public panic thereby creating a chaos in the society. Rumour detecting websites such as snopes.com, factcheck.org, etc. are some organizations. Because of manual verification, these websites are not understandable in their topical coverage.

III. Proposed Work

We proposed a framework which when given a tweet predicts whether it is related to a claimed rumour and if it is a rumour then, whether the user believes it or not. By comparing the tweets message with a news article which is extracted from a reliable news website we classified two methods for detecting rumoured data before it starts to spread to a large community of people. The feature-based approach gives quantitative results which are not based on the language or tone of tweet message. The linguistic approach returns qualitative results after processing the language and the content of the message of the tweet and categorize the words in positive and negative models based on whether the words or gestures such as hash tags, emoticons expression, and URL content are denying the rumour or accepting it. We grouped all the tweets of a single domain as an input and collectively measured the score of sentiment and linguistic categories. We developed a model for three way classification task for classifying sentiments into positive, negative and neutral. Our detection approach uses a dataset. Dataset is a repository of tweets which are related to different rumour topics. We have not cleaned the data during pre-processing because punctuation marks tell if the rumour is questioned or if the user is surprised to read such a rumour. Capitalization is necessary as it shows the impact of the rumour on the user.

We proposed the hybrid SVM, Naive Bayes and KNN algorithms and Stop words and Stemming Techniques based on different nodes.

Fig. 1: Flow Diagram
IV. Implementation

Proposed system has following modules:

i. Admin
   - Login
   - Upload Dataset: Admin will upload the dataset which contains the news and tweets.
   - View Dataset: Admin can view all the data in the dataset uploaded by him.
   - View rumour and non-rumour tweets: View all rumour and non-rumour tweets.
   - Block or Activate users: Block the users who post the rumour tweet. Admin can block and activate the users.
   - Generate graph: Admin can generate the graph based on list of rumour and non-rumour tweets that has been posted.

ii. Users
   - Register: If the user is a new user, then he has to register himself first with their details and twitter API keys (consumer secret key, access token and token secret key). Users need to register with correct value credentials.
   - Login: User can login to their account: User can view their followers’ counts, following counts and favourite counts also.
   - View all Users: View all registered users.
   - View all tweets and post new tweets: User can view all tweets posted by other users and also user can post the new tweet to their account. It can be posted to real twitter account. At the time of tweet posting, automatically we predict the rumour or non-rumour tweets by using machine learning algorithms.
   - Search News: Users can search the news and can view the news uploaded by the registered users.
   - Performance Analysis: Users can predict the accuracy of algorithm (precision, recall, f-measure and performance values).

V. Algorithms

Hybrid SVM Algorithm

Hybrid Support Vector Machine calculation is utilized to discover a hyper plane in a Nth-dimensional space where Nth is the quantity of features that uniquely orders the data points. In order to seclude the two classes of the data centers there are various possible hyper planes which can be picked. Bolster vectors are information focuses which are organized closer to the hyper plane and effects the position and course of the hyper plane. With the assistance of these support vectors, we can...
expand the edge of the classifier. The goal is to locate a plane that has the most extreme edge, that is, the maximum separation between data points of the two classes. Increasing the distance from the edge gives some support so future information focuses can be ordered with progressive level of certainty. Hyper planes are choice cut-off points which help in organizing the information focuses. Focuses falling on either side of the hyper plane can be doled out to different classes and the component of the hyper plane depends on the amount of highlights. If the number of features are 2, at that point the hyper plane it looks simply like a line and if the number of features are 3, by then the hyper plane looks like a two dimensional plane.

**Fig. 2:** Hybrid SVM Classification Approach

In hybrid SVM, we take the yield of the straight capacity and if the yield is more than 1, we identify it with one class and if the yield is -1, we distinguish it with another class. Since the threshold values are changed to 1 and – 1. In this manner we get the support scope of qualities (-1, 1) which goes about as margin. This algorithm is utilized to distinguish the negative data issues, for example, the online rumours by blocking a specific subset of nodes. We consider the requirement of client experience utility. In our system, we consider the impact of blocking time to client involvement with genuine interpersonal organizations. In this way, we propose a blocking time requirement into the conventional talk impact minimization target work.

Hybrid SVM is a non-linear classifier which is regularly revealed as delivering better arrangement results as compared to other techniques. The thought behind the technique is to non-linearly map the information to some high dimensional space, where the information can be linearly isolated, accordingly giving extraordinary characterization or relapse. One of the limitations of the algorithm is the enormous number of support vectors utilized from the preparation set to perform order undertakings.

**Navies Bayes Algorithm**

Naval forces Bayes classifier is a probabilistic AI figuring that is used for performing grouping.
\[
P(A|B) = \frac{P(B|A)P(A)}{P(B)}
\] (1)

A and B are events and \(P(B) \neq 0\).

We are attempting to discover likelihood of event A when the event B is valid. Event B is additionally named as proof. \(P(A)\) is the earlier likelihood of A. The proof is a property estimation of an obscure case. \(P(A|B)\) is a post likelihood of B.

The dataset is segregated into two segments which are include grid and the reaction vector. Highlight network contains all the lines of dataset in which every segment includes the estimation of ward highlights. Reaction vector contains the estimation of class variable or yield for each column of highlight lattice.

**K-Nearest Neighbour**

K-Nearest neighbour cluster calculation algorithm is the simple and most notable unsupervised learning algorithm. Unsupervised algorithms generate results from dataset utilizing just input vectors without alluding to the training data. A cluster is a collection of data points which are grouped together in light of specific similarities between them. K alludes to the number of centroids that we need in the dataset. A centroid is an imaginary or a real area which represents the central point of the cluster group. Each data point in the dataset is allocated to each of the groups by diminishing the intra cluster total of squares. K – Nearest neighbour algorithm detects k number of centroids and assigns every data point to the nearest cluster by fixing the centroids very small.

K-Nearest neighbor calculation starts with an underlying arrangement of whimsically chose centroids which are utilized as the underlying information focuses for each bunch and afterward iteratively figures to enhance the places of the centroids. It ends the bunch creation and improvement if the centroids are balanced out or the quantity of emphases that have been characterized has accomplished. K-means algorithm follows following steps:

- Specify the number of clusters and denote it as K.
- The centroids are introduced by rearranging the dataset. Select K information focuses arbitrarily for the centroids.
- Iterate continuously till the centroids doesn’t change.
- Calculate the sum of the square of distance between the data points and the centroids.
- Allocate data points to the centroid.
- Take the normal of all information focuses that have a place with each group and ascertain the centroids for the bunches by taking the normal of the considerable number of information focuses.
VI. Result

In the above graph we have compared all the three algorithms on the basis of accuracy. As we can see that hybrid SVM algorithm has maximum accuracy of 0.8 followed by Naive Bayes which is having accuracy of 0.76 and then KNN algorithm which is having accuracy of 0.72.
Fig. 5: Count of Rumour and Non-Rumour tweets

The graph above shows that according to the given number of tweets, for instance if we give any 5 tweets then out of them, we have found out that there are 3% of tweets which are rumoured and 2% of tweets are non-rumoured.

VII. Conclusion

We investigated the blocking of online rumour issue in social networks. We proposed a framework which when posted a tweet by someone predicts whether it is related to a claimed rumour and if it is a rumour then, whether the user believes it or not. A unique rumour dissemination model which consolidates both worldwide rumour ubiquity and individual tendency is likewise displayed dependent on the model. We used 3 machine learning algorithms which are hybrid SVM algorithm, Naive Bayes algorithm and KNN algorithm for the classification purpose. As compared to the existing systems, our system has able to accurately identify the rumours from other types of information when combined with a set of linguistic features. Our genuine qualities of the tweets assume a noticeable job in arrangement of gossip as recognized to the most popular attributes that have broke down previously.

The future work of our system will be that we are planning to create a higher end rumour blocking framework which will more accurately predict the rumours. We will also try to assess how to prevent the propagation of rumours at later stages of dispersion.
References

I. Castillo, C., Mendoza, M., & Poblete, B. (2011), March. Information credibility on twitter. In Proceedings of the 20th international conference on World wide web (pp. 675-684).

II. Kwon S Cha, M. Jung, K., Chen, W. & Wang, Y. (2013), November. Aspects of rumor spreading on a microblog network. In: International Conference on Social Informatics. Springer, Cham.

III. Kwon S Cha, M., Jung, K., Chen, W., & Wang, Y. (2013), December. Prominent features of rumor propagation in online social media. In: 2013 IEEE 13th International Conference on Data IEEE.

IV. Mendoza, M., Poblete, B., & Castillo, C. (2010), July. Twitter under crisis: Can we trust what we RT?. In: Proceedings of the first workshop on social media analytics (pp. 71-79).

V. Qazvinian, V., Rosengren, E., Radev, D. R., & Mei, Q. (2011), July. Rumor has it: Identifying misinformation in microblogs. In Proceedings of the conference on empirical methods in natural language processing (NPL). Association for Computational Linguistics.