



---

## **Credibility Analysis of Social Contents With 3 Way Deep Learning – Literature Review**

*Dhivya R<sup>1</sup>, Dr. Gogineni Rajesh Chandra<sup>2</sup>*

<sup>1</sup>Ph.D. Research Scholar, Computer Science Engineering Department, Shri Jagdishprasad Jhabarmal Tibrewala University, Rajasthan.

Research Guide, Computer Science Engineering Department, Shri Jagdishprasad Jhabarmal Tibrewala University, Rajasthan

---

### **ABSTRACT :**

After reviewing existing literature, it is evident that rumours and fake content tend to spread faster than genuine and credible information. Filtering out credible content from broad spectrum of posted information is significant to maintain the uprightness of microblogging websites and minimizing losses due to spread of fake content. There's need for an instrument or model that reduce gossip & also predict authenticity in shared information. Numerous researchers proposed many techniques to identify unauthenticated content, yet there is a room for growth and novel in present methods for determining the originality of content given in social media

**Keywords :** social media, Tweets, Credibility, Sentiment Analysis.

---

### **INTRODUCTION :**

Emergence of social media namely twitter, Sina Weibo, & Facebook has greatly facilitated exchange and updating of information. Twitter has played a key role in spreading information far and wide across the globe. With smartphones, people can easily post tweets and send messages instantly from anywhere, at any time. At the time of crisis information accessibility through social media made it as a valuable network of communication resource for ordinary people. However, it is important to authenticate the content that is shared online. False information's were widespread on online social media. While some of the content posted online seems harmless and easily ignored, there are instance where it can lead to significant chaos and irreversible consequences. Hence, it is essential to create model that can assess credibility of content posted.

---

### **LITERATURE REVIEW :**

Many researchers have been working for finding the credibility of content posted on social media, which had made it an essential and inquisitive research topic over the past decade. The dissemination of fake content is harmful for the well-being of both the user and micro blogging services in a long run.

#### ***Filter process of Fake Content to enhance credibility :***

Scholars and practitioners have explored a myriad of techniques for filtering fake content, ranging from manual fact-checking to automated detection algorithms. Manual fact-checking involves human verification of information through rigorous investigative processes, such as cross-referencing multiple sources, analyzing evidence, and consulting subject matter experts. While manual fact-checking is labour-intensive and time-consuming, it remains a gold standard for verifying the credibility of content, particularly in cases where automated methods are insufficient or unreliable

Automated detection algorithms leverage machine learning, natural language processing, and data analytics to identify patterns, anomalies, and indicators of fake content. These algorithms analyze various features of content, such as linguistic cues, metadata, and network dynamics, to distinguish between genuine and deceptive information (Pennycook & Rand, 2019). Supervised machine learning techniques train algorithms on labelled datasets of fake and genuine content, enabling them to classify new content based on learned patterns. Unsupervised techniques, such as anomaly detection and clustering, identify outliers or anomalies in content distribution and behavior, signalling potential instances of fake content

Social network analysis plays a crucial role in filtering fake content by examining the propagation dynamics and network structures associated with misinformation and disinformation. Network-based approaches analyze the diffusion patterns, user interactions, and community structures within online platforms to identify sources of fake content and influential actors involved in its dissemination

**Assessment of Credibility Using User Characteristics :**

The CredRank algorithm represents a significant advancement in the field of misinformation detection and credibility assessment on social media platforms. Developed by researchers at the Qatar Computing Research Institute (QCRI), CredRank is a machine learning-based approach that aims to identify and prioritize credible information sources on Twitter

The algorithm leverages various features, including user behavior, content characteristics, and network structure, to assign credibility scores to Twitter accounts, helping users distinguish between trustworthy and unreliable sources of information. CredRank addresses the challenges of misinformation by providing a scalable and automated solution for assessing credibility in real-time, contributing to efforts to combat fake news and promote informed public discourse

**Analyzing credibility through content characteristics :**

Content characteristics refer to the attributes, properties, and qualities of information that can be analyzed to assess credibility. These characteristics encompass various dimensions, including textual features, visual elements, source attributes, and contextual factors, that provide valuable cues about the credibility, reliability, and trustworthiness of information. Textual features such as language style, readability, coherence, and source attribution offer insights into the quality and authenticity of content

The data collection process continued till 3.5 years from the date of the launch of Twitter, which includes around 54 million users' profile and 1.7 billion tweets with a 1.9 billion follow link between them. For investigating the rumored and non-rumored content the websites like "times.com", "snopes.com", "pcmang.com" and "urbanlegends.about.com" were taken into consideration. Furthermore, tweets were outsourced to four different human annotators for validation purpose and intra-class correlation was measured between the agreement score of the annotators. In this paper, three main features were taken into consideration for detection rumors from tweets which were having structural, temporal, and linguistic properties. The machine learning classifiers have been built using SVM, Decision tree and Random Forest for evaluating the rumored and non-rumored tweets based on above-mentioned features set and achieved a high level of precision and recall with 87 to 92 percent of accuracy.

**Text Preprocessing :**

Text preprocessing also involves handling stopwords, which are common words that often occur frequently but carry little semantic meaning, such as "the," "and," and "is." Removing stopwords from text data helps reduce noise and improve the efficiency of NLP tasks by focusing on the most relevant and informative words. However, the selection of stopwords may vary depending on the specific task and domain, requiring careful consideration and customization. Additionally, text preprocessing techniques such as n-gram extraction, which involves extracting contiguous sequences of n tokens from text data, can capture contextual information and relationships between words. By applying n-gram extraction and other text preprocessing techniques, researchers can enrich text representation and capture subtle linguistic patterns that contribute to improved NLP performance.

**CONCLUSION :**

Online Social Networks such as Twitter provides a platform for communicating, interacting, and collaborating with the people from all over the world. It provides a similar way of connecting with people, exchanging thoughts and learn from their experiences. Nowadays, people turn towards social media platform more easily and frequently than tradition news media or TV channels, easily after any high impact events or occasions. The purpose behind making these platforms is to accelerate the communication process and to the provide real time information to the masses. However, fake or rumored content often disseminates with the real one and it was found that rumors tend to spread faster than the real content as fake content tends to question more.

**REFERENCES :**

1. Houshmand Hirani-Mehr "Applications of Deep Learning to Sentiment Analysis of Movie Reviews" Published 2015, Corpus ID: 1724810, Computer Science, Stanford University.
2. Chaitanya Bhagat, Dr. Deepak Mane "Survey On Text Categorization Using Sentiment Analysis" INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 8, ISSUE 08, AUGUST 2019 ISSN 2277-8616.
3. Kasper Welbersa, Wouter Van Atteveldt, and Kenneth Benoit "Text Analysis in R" COMMUNICATION METHODS AND MEASURES 2017, VOL. 11, NO. 4, 245–265 doi.org/10.1080/19312458.2017.1387238.
4. Alexander Pak & Paroubek, Patrick "Twitter as a Corpus for Sentiment Analysis and Opinion Mining" Proceedings of LREC, European Language Resources Association (ELRA), 10 May 2010.
5. Aliza Sarlan, Chayanit Nadam, Shuib Basri "Twitter Sentiment Analysis" 2014 International Conference on Information Technology and Multimedia (ICIMU), November 18 – 20, 2014, Putrajaya, Malaysia.
6. Carter T. Butts "Social Network Analysis with sna" Journal of Statistical Software, February 2008, Volume 24, Issue 6.
7. Marlo Souza and Renata Vieira "Sentiment Analysis on Twitter Data for Portuguese Language" Springer-Verlag Berlin Heidelberg 2012, PROPOR 2012, LNAI 7243, pp. 241–247, 2012.

8. Balakrishnan Gokulakrishnan, Pavalanathan Priyanthan, Thiruchittampalam Ragavan, Nadarajah Prasath, AShehanPerera “Opinion Mining and Sentiment Analysis on a Twitter Data Stream” The International Conference on Advances in ICT for Emerging Regions - ICTer 2012 : 182-188.
9. Sandeep Sricharan Mukkuand Radhika Mamidi “ACTSA: Annotated Corpus for Telugu Sentiment Analysis” Proceedings of the First Workshop on Building Linguistically Generalizable NLP Systems, pages 54–58 Copenhagen, Denmark, September 8, 2017.
10. SreekavithaParupalli, Vijjini Anvesh Rao and Radhika Mamidi “BCSAT : A Benchmark Corpus for Sentiment Analysis in Telugu Using Word-level Annotations” Proceedings of ACL 2018, Student Research Workshop, pages 99–104 Melbourne, Australia, July 15 - 20, 2018.
11. King, J. Li and K. T. Chan, “A Brief Survey of Computational Approaches in Social Computing”, in Proc. of Int. Joint Conf. on Neural Network, 2009, pp. 2699-2706.
12. S. R. Barahate and V. M. Shelake, “A Survey and Future Vision of Data mining in Educational Field”, in Proc. 2nd Int. Conf. on Advanced Computing & Communication Technology, 2012, pp. 96-100.