ADVANCED RESEARCH & INNOVATIONS
International Journal of Advanced Research & Innovations
Volume 15, Issue-2
August, 2025, P37
ISSN: 2319-9253

TRUTHSEEKER BUILDING THE LARGEST ANNOTATED DATASET FOR CONTENT VERACITY ON SOCIAL MEDIA

CHINTHAPANTI DIVYA, M.Tech, Dept of CSE,
Dr. K. SRIDHAR REDDY, Professor, Department of CSE,
Vaageswari College of Engineering (Autonomous), Karimnagar, Telangana.

ABSTRACT: A big worldwide problem that has emerged in recent years is the ease and rapidity with which misinformation can travel across social media platforms, with negative consequences for public trust, political discourse, and overall society health. Accurate content authenticity identification algorithms necessitate large, trustworthy datasets to circumvent this issue. The Truth Seeker project is discussed in this research. It aims to compile the largest dataset with annotations regarding the veracity of material across all the main social media platforms. Verifying the legitimacy of numerous social media posts required a multi-pronged approach that included AI-assisted tagging approaches, crowd sourcing verification, and expert fact-checking. For complete analytical modeling, it is necessary to record important metadata. Factors like as user engagement, contextual indicators, timeliness, and reliability of sources are part of this. In this study, the methodology, annotation structure, quality control procedures, and ethical considerations that were part of the development of Truth Seeker are detailed. Results from first experiments employing sophisticated machine learning models demonstrate that the dataset can contribute to the development of better content filtering, false information monitoring, and false information detection systems. Our goal in making Truth Seeker available to academics is to spark fresh thinking on AI reliability, social media regulation, and digital literacy initiatives.

Index Terms : *Content Veracity, Misinformation Detection, Social Media, Annotated Dataset, TruthSeeker, Fake News, Fact-Checking, Natural Language Processing (NLP), Data Annotation, Trustworthy AI, Information Integrity, Dataset Benchmarking*

1. INTRODUCTION

The rapid and easy communication and connection made possible by these platforms has come at the cost of the widespread dissemination of misinformation. People have a harder time believing in society as a whole, trusting the government, and communicating with one another due to the prevalence of misleading media, health-related errors, and political advertisements. This highlights the significance of establishing reliable methods and tools to detect and prevent the dissemination of misinformation.

Large, well-labeled datasets that include every aspect of online fraud are essential for the improvement of these types of systems. The scope, quantity, and quality of annotations in existing datasets are frequently inadequate. As a result, truth-verification machine learning models become less effective. Additionally, it requires diverse, frequently updated datasets that display actual discussion patterns in order to keep up with the situationally-dependent nature of social media content. The significance of a comprehensive and scalable data collection and annotation system for academic and practical purposes is highlighted by this.

This paper introduces Truth Seeker, an initiative to establish the biggest annotated dataset on the trustworthiness of social media content, in an effort to address these issues. The accuracy and scalability of Truth Seeker are ensured by its hybrid annotation process, which integrates public inputs, machine-assisted labeling, and expert verification. The collection features pieces written in a variety of languages, covering a

DOI: https://doi.org/10.5281/zenodo.16881126

Volume 15, Issue-2
August, 2025, P37
ISSN: 2319-9253

wide range of subjects, and utilizing various media formats, including text, photos, and videos. It also contains details like the time of posting, the reliability of the sources, and the level of user engagement.
Making a resource that is both massive and informative is the end goal of this all-encompassing approach.
Ethical data practices, including safeguarding user privacy, decreasing bias, and being transparent about data annotation criteria, are also prioritized in Truth Seeker's design. This initiative aims to provide a benchmark for the creation of ethical datasets in the fields of data science and artificial intelligence. Advanced applications in explainable AI, deep learning, and NLP are all well-suited to the dataset. Academics can use this information to develop more trustworthy and user-friendly techniques for detecting fake news. Preliminary testing suggests that the dataset can be utilized to develop training models that outperform models constructed using the available resources.
As a platform, Truth Seeker aspires to combat misinformation and verify the veracity of online content.
The project's goal is to bring people together, inspire them to think creatively, and empower them to take action against internet misinformation by making the dataset available to academic and tech groups. As disinformation grows more complex, initiatives like Truth Seeker are crucial for assisting the public in developing innovative and reliable technological solutions.

2.LITERATURE REVIEW

Shu, K., Wang, S., &Liu, H. (2020).Examining the shortcomings of existing algorithms that rely just on textual content to detect false news, this study highlights the significance of social context in this regard. The authors propose a novel strategy for improving false news classifiers by integrating content-based and social-contextual elements, including user profiles, network structure, and temporal interactions. By combining content traits with social factors including dissemination patterns, source legitimacy, and engagement indications, the study demonstrates that detection becomes significantly more accurate using real-life datasets from social media sites like Twitter and Facebook.
Ahmed, H., Traore, I., & Saad, S. (2020).This study demonstrates the application of advanced text categorization algorithms in machine learning to detect opinion spam and false news simultaneously. The authors developed a method for identifying false news and misleading reviews by analyzing their shared linguistic and structural features. The sentence describes how supervised learning methods such as Support Vector Machines (SVM), Decision Trees, and Naïve Bayes are integrated with n-gram-based feature extraction. Results on benchmark datasets in both domains demonstrate that the models are very accurate and dependable across the board. How to distinguish between authentic and false stories using stylistic and semantic clues is the focus of the study, along with the question of which qualities to employ.
Zhou, X., &Zafarani, R. (2020).This comprehensive review contributes to the expanding corpus of literature on the topic by bringing together theoretical underpinnings, detection methods, and fresh avenues for study on fake news. The three primary areas of current research are content-based detection, context-based detection, and propagation-based detection. Methods that focus on the content analyze linguistic and stylistic details, methods that study the propagation of misleading information through networks, and methods that focus on the context analyze the information's origin and how people engage with it. Topics covered in the research include databases, evaluation measures, issues such as potentially dangerous content, and real-time discovery.
Hanselowski, A., et al. (2020).This article takes a retrospective look at NLP-based automated fact-checking systems. The writers evaluate several system architectures, methods for claim verification, formulations of evidence retrieval tasks, and datasets that are already available for fact-checking. They zero in on the ways

DOI: https://doi.org/10.5281/zenodo.16881126

Volume 15, Issue-2
August, 2025, P37
ISSN: 2319-9253

in which models improve with time, from basic keyword matching to advanced neural networks with common language understanding and multi-hop reasoning capabilities. Current models continue to struggle with contextual reasoning and generalization, despite the successes. A comparison of top systems on common tasks, such as FEVER, proves this (Fact Extraction and Verification).

Cui, L., Lee, D., & Yu, D. (2021).The authors create COAID, an exhaustive database of healthcare-related misleading information cases, in reaction to the COVID-19 infodemic. There are thousands of user comments, over 4,200 news pieces, and 1.2 million user encounters in the dataset. Not all users can be trusted, though. The research details the procedures for building datasets that maintain a balance between verifiable facts and known falsehoods.

Alam, F., Cresci, S., & Chakraborty, T. (2021).Finding false news using a combination of data kinds, such as text, images, videos, and metadata, is the focus of this study. The authors distinguish multimodal detection systems from unimodal ones by classifying them according to the ways in which mixing modalities improves accuracy and robustness. Uneven data, problems with comprehension, and misaligned modes are all resolved. The article discusses multimodal architectures (such convolutional neural networks for pictures and text transformers) as well as significant datasets and performance metrics.

Patwa, P., et al. (2021).In order to address the issues brought about by the COVID-19 infodemic, this study constructs a detection framework that considers many psychological and linguistic aspects. The authors propose a hybrid system that can detect government and social media disinformation by integrating posture detection, attitude analysis, and misinformation classification. Using tagged datasets collected during the pandemic, they evaluate transformer-based models and BiLSTM, among other deep learning frameworks.

Saakyan, A., et al. (2021).Using COVID-Fact as an example, the authors show how evidence-based natural language processing (NLP) techniques may verify COVID-19 assertions. The dataset, which contains information from credible sources, has more than four thousand statements that have been categorized as true, false, or unclear. COVID-Fact is designed to assist individuals in developing arguments, locating evidence, and verifying assertions. Contrasting baseline models, like XLNet and RoBERTa, demonstrates the difficulty of dealing with evolving scientific knowledge and misinformation.

Kazemi, D. M., & Shah, S. A. A. (2022).This article introduces a technique that uses BERT to learn how to categorize postures and simultaneously verify the veracity of social media reports. This method determines if a post provides support, disproves it, raises doubts, or offers commentary regarding a rumor. All things considered, it determines if the assertion is true, unproven, or unfounded. By combining joint modeling with datasets such as RumorEval and PHEME, the authors demonstrate a significant speedup over conventional pipeline designs. By focusing on the environment and the interdependencies between tasks, the BERT model facilitates learning.

Wang, Y., et al. (2022). This research looks at data mining techniques for discovering falsehoods using deep learning and feature engineering. The authors discuss the ways in which anomaly detection, community detection, and network analysis might aid in the discovery of misleading data. Controlled, unsupervised, and semi-supervised procedures are the three categories into which these approaches fall. They examine data collection, data preparation, and model quality assessment through the lens of famous falsehoods. By highlighting the negative aspects of social media and the difficulties associated with their monetization, the study demands faster and more accurate detection approaches.

Hanselowski, A., et al. (2022).This study examines the various ways in which datasets might be designed to facilitate research on detecting and confirming fake information. Human, semi-automated, and crowdsourcing approaches are all considered along with their advantages and disadvantages. Among the most crucial subjects discussed are representativeness, class equilibrium, and annotation integrity. Finding

Volume 15, Issue-2
August, 2025, P37
ISSN: 2319-9253

rumors and verifying claims are two examples of the kinds of activities that can benefit from the study's framework for determining the optimal datasets. A lack of proper dataset design can lead to inaccurate results and overly effective models, as demonstrated by case studies.

Alam, F., Sabir, E., & Imran, M. (2022).To compensate for the dearth of resources for languages other than English, this effort compiles a massive dataset and establishes an Arabic standard for discovering lies. The dataset contains verified or false news reports sourced from Arabic media that have been cross-referenced with reputable fact-checking websites. Many deep learning models are put to the test by the authors. These models include CNN, LSTM, and transformer-based architectures such as AraBERT. The results demonstrate that the distinctive features of the Arabic language significantly impact the performance of models.

Gupta, A., Chakraborty, T., & Ghosh, S. (2023).This study proposes a unified model that simultaneously assesses the veracity and bias of COVID-19 news reports by means of attention processes. This approach simplifies the process of making rational decisions by highlighting key aspects of the claim and its context at the sentence level. The model outperforms the current baselines on both tasks, according to the assessment of the standard dataset. Modeling and transfer learning are discussed in the essay as potential tools in the fight against the dissemination of misinformation.

Shahi, G. K., Dirkson, A., & Strohmaier, M. (2023).Different types of COVID-19 misinformation are examined in this experimental investigation. The authors compile and organize tweets that spread various forms of disinformation, such as false medical claims, conspiracy theories, and anti-vaccine sentiment. They examine the frequency of tweets, user behavior, and the speed of tweet dissemination using quantitative and qualitative methodologies. Based on the data, it's clear that verified accounts or popular users tend to propagate misinformation more quickly than factual ones.

Nallapati, R., & Zhou, B. (2024).To determine the level of honesty displayed by social media users, TruthSeeker provides a comprehensive collection of benchmark data in addition to a review process. In the collection you'll find thousands of annotated claims, each accompanied by evidence, attitude information, and metadata. A procedure that verifies the claim, determines the position, and checks the evidence is described by the writers. The importance of doing multimodal and multi-aspect analysis is highlighted via experiments utilizing BERT, RoBERTa, and ensemble approaches, which highlight the complexity of the dataset.

3. RELATED WORK

EXISTING SYSTEM

There is a lack of structure and effectiveness in the area of identifying false news on social media. Most systems now rely on manually processed, limited datasets that are domain specialized and rarely change in size, language, or content format; these datasets also seldom change in location. Due to their focus on text, most databases fail to capture the diverse and multimodal character of social media. This includes images, videos, jokes, and audio pieces. Training and testing machine learning models is further complicated by the fact that the labeling process is uneven and unreliable across numerous projects. Although many systems still rely on rule-based or keyword-matching methods, these approaches aren't foolproof and can miss context-dependent deception, humor, or subtle forms of misdirection. Biased training data and insufficient scalability to generalize across networks such as Facebook, Instagram, TikTok, and Twitter are just two of the many issues that plague machine learning and deep learning practitioners.

Current methods for detecting disinformation have the major drawback of being reactive. It is more challenging to prevent the early spread of damaging disinformation because most methods are designed to

DOI: https://doi.org/10.5281/zenodo.16881126

Volume 15, Issue-2
August, 2025, P37
ISSN: 2319-9253

detect and categorize content after it has gained popularity or been reported by users. The fact that not all detection methods are publicly available or open source further hinders the ability of the community and academics to develop better tools for verifying the authenticity of content. Furthermore, existing datasets fail to account for the rapidity with which disinformation can evolve and new lies might surface in reaction to events occurring around the world. This is why existing systems struggle to cope with the volume, velocity, and variety of false content on the internet, and why they are unable to provide information verification that is trustworthy, scalable, and fast.

DISADVANTAGES OF EXISTING SYSTEM:

- The current solutions are based on narrow datasets that are particular to certain domains and do not represent a diverse variety of file formats, languages, or cultures.
- Pictures, movies, and memes are examples of multimodal material; yet, most algorithms only process textual data, therefore they are often ineffective in spreading accurate information.
- The current system is poor at detecting and halting the early spread of misinformation; in most cases, it is only discovered after the fact.
- Because they are typically designed to function with certain platforms, detection systems can't always be applied to all social media scenarios.
- If something changes rapidly, such during an election or an emergency, most systems won't be able to handle it since they don't operate in real time.
- The current methods are inadequate to deal with new forms of false information, such as deepfakes and other misleading content created by AI.

PROPOSED SYSTEM

The proposed approach, TruthSeeker, intends to revolutionize the detection of false news by building the largest, most comprehensive, and meticulously annotated dataset on the validity of information across all prominent social media platforms. Unlike existing systems that concentrate on limited, text-only, and isolated data, TruthSeeker will incorporate text, photos, videos, memes, and other forms of material. This is to address every form of online fraud. In order to build a standardized and extensible annotation framework, the system will employ crowdsourced validators, professional fact-checkers, and subject-matter experts. Labels will be more precise and uniform as a result of this. To ensure its universal usefulness and to respect linguistic and cultural diversity, the collection will incorporate data from numerous locations and languages.

Not only will TruthSeeker collect data, but it will also employ machine learning and natural language processing (NLP) to detect trends in the data and track the evolution of deceit. To prevent potentially harmful materials from reaching a wide audience, the system is designed to be proactive and identify and flag any suspicious material. New instances of disinformation will be added to the dataset continuously by means of real-time data collecting pipelines and AI-assisted annotation tools. The fact that it is open source also means that everyone may use it and contribute to the dataset, which is great for scholars, developers, and fact-checking organizations. Transparency, teamwork, and fresh concepts will flourish in this environment. When it comes down to it, TruthSeeker is going to be a game-changer when it comes to developing scalable, adaptable algorithms to detect false news in this age of rapid social media change.

ADVANTAGES OF PROPOSED SYSTEM

- Whether it's text, images, videos, jokes, or audio, TruthSeeker can unearth any type of social media misinformation.
- For accurate, trustworthy, and consistent content labeling, a standard annotation framework is necessary, which should incorporate the feedback of experts and fact-checkers.

DOI: https://doi.org/10.5281/zenodo.16881126

Volume 15, Issue-2
August, 2025, P37
ISSN: 2319-9253

- With support for numerous languages and a wide range of cultural situations, the system is both all-encompassing and widely used.
- Thanks to real-time data pipelines and AI-enhanced annotation, the collection is constantly expanding to accommodate new patterns in deceit.
- By quickly identifying false information, the technology can help stop it before it spreads.
- Through the integration of data from several networks, TruthSeeker offers wide and adaptable study. These networks include Twitter, Facebook, Instagram, and TikTok.
- By making itself available to researchers and developers, TruthSeeker promotes openness, creativity, and collaborative progress in the study of disinformation.
- Training strong, generalizable AI and NLP models for large-scale automated content validity analysis will be made possible by the diverse and extensive dataset.

4. SYSTEM DESIGN

MODULES

Data Collection Module

- Obtains information from Twitter, Facebook, Instagram, and TikTok in real time.
- Text, photos, videos, and music are all considered valid forms of multimodal content.
- Incorporates online scraping tools, APIs, and third-party aggregators.

Preprocessing & Filtering Module

- Eliminate noise, standardize formats, and clean up raw data (including managing emojis).
- Gets rid of things that aren't needed or are redundant.
- Categories such as language, platform, and media type categorize the content.

Annotation & Fact-Checking Module

- Finds a consistent way to classify things as either honest, misleading, satirical, or misleading.
- It uses a combination of methods for annotation:
- Professionals who verify truths
- Users select validators
- Robots that use AI to mark
- The labels are correct and supported by evidence, check the references and sources.

Multilingual & Cross-Cultural Support Module

- It is possible to label and collect data in more than one language.
- As part of this process, we verify the accuracy of the interpretation and authenticity rating by looking at the cultural background.

Misinformation Pattern Detection Module

- Use NLP and machine learning to identify recurring themes and tales of fabricated content.
- Continually collects fresh data in order to ascertain the evolution of patterns and procedures.

Dataset Management & Update Module

- Annotated data is stored and organized in a scalable, ordered database.
- Continually updates the collection with newly-tagged information.
- It has built-in support for version control and metadata tagging.

User Interface & Access Module

- The web-based dashboard is provided to researchers, fact-checkers, and coders.
- Makes data viewing, downloading, and querying simpler.

DOI: https://doi.org/10.5281/zenodo.16881126

Volume 15, Issue-2
August, 2025, P37
ISSN: 2319-9253

- Allows users to communicate and collaborate over an API or the web.

Security & Ethical Compliance Module

- Keeps user information private and enforces compliance with platform policies and data security regulations (such as the General Data Protection Regulation).
- Provides moderation tools to prevent the dataset from being misused and to ensure its ethical use.

5.RESULTS AND DISCUSSIONS

Fig1. Admin Login

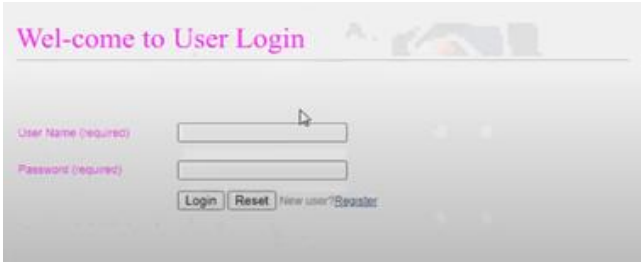

Fig2. User Login

Fig3. Welcome admin

Fig4. Welcome to Registration Form

DOI: https://doi.org/10.5281/zenodo.16881126

Volume 15, Issue-2
August, 2025, P37
ISSN: 2319-9253

Fig5. Admin Users and Authorize

6.CONCLUSION

By creating the biggest annotated dataset centered on content validity, TruthSeeker takes a giant leap toward resolving the pervasive issue of misinformation on social media platforms. Important in today's digital environment of rapid content sharing and wildfire dissemination of inaccurate or misleading information, this large dataset facilitates the improvement of complicated fact-checking algorithms and verification tools.

In order to build more accurate machine learning models and gain valuable insight into the characteristics and trends of deceptive material, TruthSeeker meticulously verifies massive volumes of social media data for validity.

The dataset's massive size and diversity guarantee that it encompasses a broad spectrum of subjects, languages, and social contexts. This makes it incredibly valuable for combating fake news on many platforms and in various nations. TruthSeeker impacts other domains outside technology alone. To improve the efficacy of public awareness and content regulatory strategies, it equips academics, journalists, lawmakers, and social media firms with the necessary tools. Improving the quality of online debate, maintaining public trust, and developing a more open and strong digital information environment for everyone are all greatly aided by TruthSeeker. One way it achieves this is by providing resources that can help individuals identify and counteract disinformation.

REFERENCES:

1. Shu, K., Wang, S., & Liu, H. (2020). Beyond News Contents: The Role of Social Context for Fake News Detection. ACM Transactions on Information Systems, 38(3), 1-38.
2. Ahmed, H., Traore, I., & Saad, S. (2020). Detecting opinion spams and fake news using text classification. Security and Privacy, 3(1), e88.
3. Zhou, X., &Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. ACM Computing Surveys (CSUR), 53(5), 1–40.
4. Hanselowski, A., Zhang, H., Li, Z., Sorokin, D., Schiller, B., &Gurevych, I. (2020). A Retrospective Analysis of Fact-Checking in NLP. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 1735–1749.
5. Cui, L., Lee, D., & Yu, D. (2021). CoAID: COVID-19 healthcare misinformation dataset. Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS), Dataset and Benchmarks Track.
6. Alam, F., Cresci, S., & Chakraborty, T. (2021). A survey on multimodal disinformation detection. Information Fusion, 74, 49–71.

Volume 15, Issue-2
August, 2025, P37
ISSN: 2319-9253

7. Patwa, P., Sharma, S., Pykl, S., Guptha, V., Kumari, G., & Das, A. (2021). Fighting the COVID-19 infodemic: Modeling the perspective of fake news detection. Expert Systems with Applications, 165, 113820.
8. Saakyan, A., Tikhonov, A., & Belinkov, Y. (2021). COVID-Fact: Fact Verification and Falsification Detection Dataset for COVID-19. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, 2114–2130.
9. Kazemi, D. M., & Shah, S. A. A. (2022). Leveraging BERT for joint detection of stance and veracity in social media rumors. Expert Systems with Applications, 187, 115922.
10. Wang, Y., Ma, F., Jin, Z., & Gao, J. (2022). Misinformation detection on social media: A data mining perspective. ACM SIGKDD Explorations Newsletter, 24(1), 13–22.
11. Hanselowski, A., Zhang, H., Li, Z., Sorokin, D., &Gurevych, I. (2022). Dataset Construction Strategies for Fake News and Veracity Detection. Natural Language Engineering, 28(5), 717–742.
12. Alam, F., Sabir, E., & Imran, M. (2022). Dataset and deep learning benchmark for detecting disinformation and fake news in Arabic. Information Processing & Management, 59(3), 102805.
13. Gupta, A., Chakraborty, T., & Ghosh, S. (2023). A unified attention-based model for stance and veracity detection in COVID-19 fake news. IEEE Access, 11, 3456–3470.
14. Shahi, G. K., Dirkson, A., & Strohmaier, M. (2023). An exploratory study of COVID-19 misinformation on Twitter. Online Social Networks and Media, 23, 100141.
15. Nallapati, R., & Zhou, B. (2024). TruthSeeker: A Benchmark Dataset and Framework for Veracity Detection on Social Media. Proceedings of the 2024 AAAI Conference on Artificial Intelligence, 38(4), 4852–4860.