Misinformation Detector

Team Blue

1. The Problem **2**
1.1 Proposed Solution 2
2. Product Description **2**
2.1 Features and Capabilities 3
2.2 Major Components 3
3. Who is the intended consumer? **3**
What is the structure of the prototype? **3**
Changes and developments **4**
4. Product Prototype Description **4**
4.1 Prototype Architecture 5
4.2 Prototype Features and Capabilities 5
4.3 Prototype Development Challenges 7
Glossary **7**
References **7**

1. The Problem

The problem is that, especially in recent years and with the spread of information being increasingly feasible due to things such as social media, misinformation continues to be a huge problem around the world and is quite noticeable in countries such as the USA. Misinformation tends to target a variety of topics, including, but not limited to, healthcare, politics, business, and general everyday life, and can often be done with malicious intent, which is disinformation or without it, which counts as misinformation. As a result, it can possibly cause confusion, spread harm, and erode public trust in which the latter can be seen as trust in mass media has recently decreased. This is partly because the average reader lacks the necessary tools to help fact-check information on a variety of topics, and the significant amount of information that is currently on the internet would make manual checks extremely difficult to do regularly. For additional context, according to the article "Americans' Trust in Media Near Record Low", only 34% of Americans trust mass media to report in an unbiased, factual manner. Also, according to the article "How Misinformation on Social Media Has Changed News.", misinformation spreads up to ten times faster than credible reporting.

1.1 Proposed Solution

Our proposed solution to this issue was to create a fact-checking site that did things differently. Most fact-checking sites take the articles and show you related articles on the topic that either agree or disagree with the topic presented by the user, or the user inputs a keyword into a search and finds topics about said subject, but not many sites provide a credibility score on the related article or topic given. The solution we present is a web server capable of using Machine learning and algorithms to analyze the content of a presented article and give it a numeric score to its credibility. This will be a way to present a combative force against misinformation and disinformation. The product shall be called the Misinformation Detector.

2. Product Description

The solution starts with a user reading an article, the content is then questioned by the user, "Is this article truthful? Can I trust the author? Is the content based on fact or on opinion?". The user takes the link to the article and submits it to our website. The website then uses algorithms to determine content, context, and look for other articles either related to the topic or written by the author to create a "profile" of their previous works and the legitimacy of their claims. From there, it can then compare the article to other articles that are trusted and proven factual, and give a numerical score between 1 and 100, providing links to related and trusted articles on the given topic for further information. The scoring would be listed as 1 being the lowest score, the content is not trustworthy and contains false information, or the author is an untrustworthy source. 100 would be the highest score, the content is factual, or the author is a trustworthy

source. Every score in between is an indicator of varying degrees of misinformation or disinformation. The user can then form a more educated opinion on the topic given.

2.1 Features and Capabilities

The credibility analysis web application is designed to automatically evaluate the trustworthiness of online articles. Users can submit links from news sites, social media posts, or reference pages such as Wikipedia. The system scans and analyzes the content by comparing it with verified sources, assessing factors such as author reliability, factual accuracy, and tone. Once processed, the tool generates a credibility score along with a brief explanation to show how that score was determined. The system updates regularly as new information becomes available, ensuring that results stay current and accurate.

2.2 Major Components

The application is built around several main components that work together to evaluate content. The Content Ingestion component gathers article data from submitted URLs and prepares it for analysis. The Automated Misinformation Detection component runs the text through models and algorithms that check for inconsistencies, misleading language, and factual claims. The Source Credibility and Analysis component maintains a database of trusted and untrusted sources, assigning each one a credibility score based on its history and reputation. Finally, the User Interaction component manages user input and displays the credibility results in a clear, easy-to-understand format.

3. Who is the intended consumer?

The main target of this solution is users who wish to research a given topic and find more information to conclude if the article they have read contains valid information or not. This solution is for anyone and everyone who holds any question or doubt about the information they read, and wants to know more about the topic discussed. This could be utilized by literally anyone and everyone on the open internet.

What is the structure of the prototype?

The prototype structure starts with the frontend page, where a user can paste a link to an article. That link gets sent to our backend, which handles content ingestion, preprocessing, and running the text through the smaller algorithms. We use a small database with a limited set of trusted sources so the system can compare the article and come up with a score.

Changes and developments

Some of the changes, developments, and challenges that may occur over the construction of the prototype may include developing a proper database which may change over time to include or remove pieces of data, utilizing and implementing machine learning algorithms to learn to detect misinformation and issues to be worked on involving whether it can do well with any new data and any problems with changes to pre-established knowledge, and having the credibility score work in tandem with the aforementioned algorithms and give the proper score.

4. Product Prototype Description

The prototype is a simplified, functional version of the Misinformation Detector. Its goal is to demonstrate the core capabilities of the final product while reducing or simulating certain features that would require larger datasets, more advanced algorithms, or real-time performance.

Capabilities that are reduced or simulated in the prototype include: handling large-scale article ingestion, processing multiple simultaneous requests, integrating with external platforms, and using fully trained NLP models. Instead, the prototype uses a limited, static database of trusted sources, smaller-scale algorithms, and a simplified scoring system to illustrate the process of assessing credibility.

Feature	Prototype	Real-World Product
Database of Sources	Limited, static set of trusted sources	Global, continuously updated
Machine Learning	Simplified ML algorithms and keyword checks	Full NLP models analyzing context & content
Real-Time Analysis	Single article at a time	Multiple articles simultaneously
User Interface	Basic page for link submissions	Fully interactive and scalable web portal
Credibilty scoring	Simplified scoring using limited data	Automated, fully weighted, adjustable
Integration with External Platforms	Not implemented/simulated	APIs for news and social media
Reporting and Recommendations	Basic related article suggestions	Detailed suggestions & related articles

4.1 Prototype Architecture

The prototype is structured to demonstrate the essential features of the finished product. The architecture includes three main components:

Frontend Web Interface:

The frontend web interface allows users to submit article URLs and view the resulting credibility score. It also displays links to related verified articles, helping users explore trustworthy information on the same topic. The interface is built using simple web technologies such as HTML, CSS, and JavaScript to ensure accessibility and ease of use.

Backend Analysis Server:

The backend analysis server is responsible for handling content ingestion and preprocessing, including tokenization and basic natural language processing. It runs simplified algorithms to evaluate the article's reliability and generate a credibility score based on comparisons with data from trusted sources.

Prototype MFCD Description:

The system begins with the User Interface, where a user submits a URL to an article they wish to verify. That input is sent to the Backend Server, which handles content ingestion — retrieving and preparing the article text for analysis. The Credibility Analysis Engine then evaluates the content by comparing it to entries in the Trusted Source Database. This engine applies simplified algorithms to assess factual consistency, author reliability, and the presence of biased or misleading language.

After analysis, the system generates a Credibility Score (1 to 100) along with brief explanations and links to related verified articles. These results are then returned to the User Interface for analysis by the user to determine if they wish to continue with the selected article or not. The results are sent to the database to process by the machine learning model. When a source is fed into the machine learning model, it will be used for future inquiries.

4.2 Prototype Features and Capabilities

The prototype demonstrates the following:

- **Article Submission & Scoring:** Users submit an article URL and receive a numeric credibility score (1–100).

- **Source Verification:** Checks the article's author and publication against a small database of trusted/untrusted sources.
- **Content Comparison:** Compares submitted articles with verified content in the database.
- **Related Articles:** Provides links to trusted articles on the same topic.

Significance:

- Demonstrates how users can evaluate misinformation and make more informed decisions.
- Provides a proof-of-concept for automated credibility scoring.
- Even with reduced capabilities, the prototype shows that the system can identify untrustworthy content and highlight reliable sources.

Risk Mitigation:

- Using a limited dataset and simplified algorithms reduces the risk of misclassification.
- Demonstrates functionality without requiring full-scale real-time processing, mitigating technical complexity risks.

Functional Goals and Objectives:

- Enable article submission and scoring.
- Provide insight into the credibility of the article and author.
- Suggest verified related articles to support further research.

4.3 Prototype Development Challenges

Expected challenges include:

- **Knowledge Gaps:** Limited experience with advanced NLP and ML algorithms may reduce accuracy.
- **Data Constraints:** The small, static source database may not cover all topics or authors.
- **Algorithm Limitations:** Simplified algorithms may misinterpret nuanced or opinion-heavy content.
- **Technology Constraints:** Real-time processing, multiple user requests, and platform integration are not feasible at the prototype stage.
- **Score Calibration:** Ensuring that the numeric score meaningfully reflects credibility is challenging with limited data.

Glossary

Algorithm - A process or set of rules to be followed in calculations or other problem solving operations
Disinformation - false or inaccurate information, often shared with intent of causing harm
Misinformation - false or inaccurate information, often shared without the intent of causing harm
Machine Learning (ML) - A subsection of AI that is focused on how al
Natural Language Processing (NLP) -

References

1.) Gallup. “Americans’ Trust in Media Near Record Low” Gallup, 2022, https://news.gallup.com/poll/403166/americans-trust-media-remains-near-record-low.aspx
2.) PIRG. “How Misinformation on Social Media Has Changed News.” PIRG, 2025, https://pirg.org/edfund/articles/misinformation-on-social-media/
3.) Bergmann, Dave. “What Is Machine Learning.” *Www.ibm.com*, https://www.ibm.com/think/topics/machine-learning#:~:text=Author,fundamental%20goal

%20of%20machine%20learning. Accessed 13 Nov. 2025.