Objective News: Locating and Delivering Accurate News with NLP and NNs

Harry Yin

February 9, 2025

Contents

1. Abstract

In an era where misinformation and bias increasingly shape public opinion, it is increasingly important that reliable information can be found. While there are existing approaches in AI that focus on fake news detection and bias classification, there are no tools that actually improve objectivity and delivery of news content.

Approach: This proposal is for a project that implements a comprehensive news processing system that integrates textual clustering with unique methodology; fake news detection using CNN-LSTM models; innovative objectification/bias removal using rule-based NLP analysing and/or fine-tuned summarization models; and opinion extraction with fallacy predictors and bias scoring, all joined together with an interactive and transparent user interface.

Impact: The comprehensive approach taken to mitigate misinformation and promote accurate, non-bias information with a variety of perspectives serves as a foundation for advancing trustworthy, automated news evaluation.

2. Introduction

Misinformation often clouds the truth, harming both the individual and society. The individual may face mental distress due to online information or even physical distress after receiving incorrect information about a life-threatening topic while society is negatively affected by a build-up of polarity and poor choices. This has been seen recently, whether it be during the Covid-19 pandemics of the 2024 presidential election.

In the status quo, there has been a lot of development in locating fake news and bias, for instance, Sastrawan et al. (2022) and Powers et al. (2024). However, there has been little to no development in developing methods to deliver the news in an objective, non-bias manner while also depicting a variety of different perspectives and opinions.

Here, I am proposing a new methodology to implement a comprehensive news processing system that integrates textual clustering; fake news detection using CNN-LSTM models; objectification/bias removal using rule-based NLP analyzing and/or fine-tuned summarization models; opinion extraction with fallacy predictors and bias scorings; and an interactive and transparent user interface. By not only having detection but also delivery, this project works to improve both trust in media and the methods used for generating non-bias text.

3. Methodology

As a preface, a Beta of this project has been released on Github by me! (Yin, 2024). Anything that has been implemented before will be marked as so along with the current methodology, results, and limitations. Below is a visual depicting how the process of news analysis would generally work.

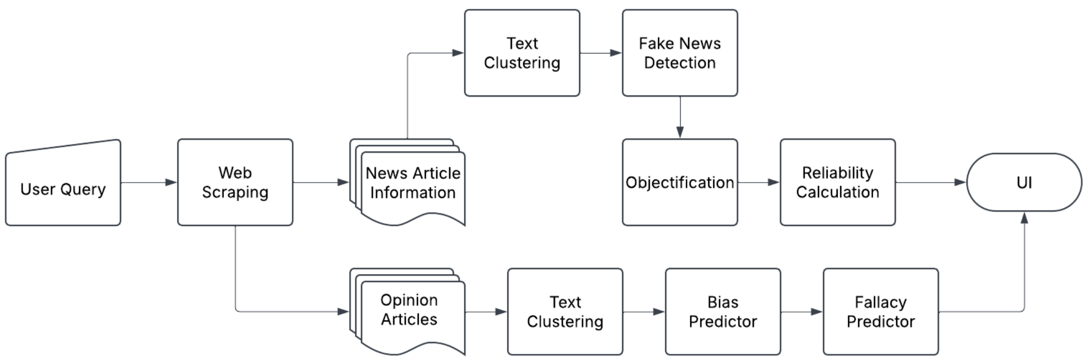

Figure 1: Diagram of the news analysis process

First, from a user query, web scraping is performed to fetch either news or opinion articles. News articles undergo textual clustering, fake news detection, objectification, and a reliability calculation. Opinion articles go through opinion extraction which includes textual clustering. Finally, everything is displayed with a UI.

3.1. Information Retrieval

Objective: The objective of this process is to gather information from the internet to analyze. This includes primarily finding and fetching online articles and extracting their text, date or publish, and other important information.

***Current* Method:** Links are retrieved from Google queries done with requests and bs4. Then, Trafilatura is used to retrieve the information from said links. In some cases, newspaper3k is used as a fallback in case Trafilatura fails.

***Current* Result:** This process fetches information with sufficient quality. It retrieves all links necessary and most of the information along with all necessary critical information like date, author, and text.

***Current* Limitation:** There is a limit on retrieval speed due to possible blocking and there is no option to diversify sources of information, often resulting in the same sources/articles being used. Furthermore, there is no way to bypass paywalled sources, resulting in occasional failures.

New Proposed Method: Use multiple search engines to gather links (NewsAPI, Google News API, Bing News API, Duck Duck Go API), implement RSS searches (mainly using Google RSS scraping), and use source filtering to ensure no source overlaps and mainly "reliable" sources are committed to from (locate the sources of each link and given a limit of `n` sources, sort the sources by reliability, remove overlaps, and return `n` sources). Also, possibly incorporate archive.ph to bypass paywalls. The expected result is a more diversified link pool with information coming from more reliable sources.

3.2. Textual Clustering

Objective: The objective of this process is to cluster texts based on general argument or piece of information. For instance, when reporting on a broad topic such as natural disasters, information about the damages will be grouped separately compared to information about the cleanup.

***Current* Method:** Preprocessing: Preprocessed the text through lemmatization and stopword removal.

Embedding: Used `sentence_transformer` models (`all-MiniLM-L6-v2` and `all-mpnet-base-v2`) to embed individual sentences and applied untrained self-attention [1].

Context: Applied context by extracting embeddings of surrounding sentences and averaging them with individual sentence embeddings with weights.

Clustering: Clustered using either KMeans or Agglomerative Clustering, running through `n_clusters` [2, max_clusters] to find the best cluster on a metric weighed heavily by silhouette scoring and marginally by Davies-Bouldin and Calinski-Harabasz scores.

***Current* Result:** Below is a chart with the scores calculated across multiple clustering tests.

Method	Preproc	Attention	Context	Avg Silhouette	Avg % Improvement
KMeans (Baseline)	False	False	False	0.064	0.000%
KMeans	True	False	False	0.055	-12.440%
KMeans	False	True	False	0.234	159.900%
KMeans	False	False	True	0.082	15.037%
KMeans	True	True	False	0.233	157.867%
KMeans	**True**	**True**	**True**	0.261	**189.620%**
KMeans	False	True	True	0.268	194.033%
KMeans	True	False	True	0.075	2.833%
Agglomerative	False	False	False	0.074	8.537%
Agglomerative	True	False	False	0.063	-9.436%
Agglomerative	False	True	False	0.239	162.530%
Agglomerative	False	False	True	0.096	22.037%
Agglomerative	True	True	False	0.235	161.020%
Agglomerative	**True**	**True**	**True**	0.273	**197.700%**
Agglomerative	False	True	True	0.263	189.833%
Agglomerative	True	False	True	0.085	10.093%

While the silhouette scores are the only ones in displayed on the chart, average improcement is calculated through all three metrics of silhouette, Davies-Bouldin, and Calinski-Harabasz scores.

Some interesting occurrences to note are that: 1. Applying attention, even without any fine-tuning, and context played a significant role in improving the clustering and 2. Applying preprocessing seemed to occasionally make the clustering worse.

***Current* Limitation:** More research needs to be done into the merits and demerits of preprocessing and the benefits of untrained self-attention to textual clustering. Furthermore, outlier information is classified into existing clusters, meaning there is some dissonance with the clustering results and subsequent summarization of the information.

[1] Applying untrained self-attention is implemented to test if it improves clustering even if untrained, independent of trained attention algorithms (see Singh (2022) for applying trained self-attention to text clustering).

Improved Proposed Method: Apply UMAP post-embedding to lower dimensionality, increase efficiency, reduce noise (Barla, 2023). Change the usage of KMeans/Agglomerative Clustering to HDBSCAN. HDBSCAN has the advantage of using a hierarchical clustering method, which given Agglomerative clustering is ~8% better than other clustering methods, and the outlier identification of DBSCAN (Borrelli, 2025). The expected result is improved cluster solidarity as outliers are not counted in existing clusters.

3.3. Fake News Detection

Objective: The objective of this process is to predict fake news using NNs. Furthermore, another goal is to test different architectures to determine the best one for the task given similar amounts of training.

Proposed Method: Train these models architectures: CNN-LSTM; CNN-BiLSTM; Transformer-LSTM, (Rai et al., 2022); Transformer-BiLSTM on these datasets: LIAR, LIAR2, ISOT, and various other datasets from Kaggle/Huggingface to lower the chances of bias. Scoring will be on a confidence metric due to the variations between each dataset in terms of nuance.

Expected Result: Locating the best model for predicting fake news with at least ~97.5% (Sastrawan et al., 2022) accuracy on benchmark tests in fake news prediction. The Onion can be used as reference for fake news and other sources for real news.

3.4. Objectification / Bias Removal

Objective: The objective of this process is to remove subjectivity and bias from a given text without altering its base meaning. Although there has been a focus on bias mitigation in text generation with efforts like *MBIAS* by Raza et al. (2024), little attention has been paid to the mitigation of bias of the specifically existing text. For instance, instead of having a very emotive and skewed text, post-objectification the text will be presented with the necessary information conveyed in a professional tone.

***Current* Method:** Using NLP to locate emotive/strong adjectives and adverbs. If the word serves an important structural purpose (ex: “They were *happy*.”), replace it with a less emotive dictionary synonym or use an encoder-only model to replace the word. If it does not serve an important structural purpose (ex: “The *happy* people. . . ”), remove it and remove any unnecessary conjugations, punctuation, etc. that serve no purpose with the removal of the word.

***Current* Result:** Using Textblob subjectivity as an evaluation metric, objectivity post rule-base objectification improved by an average of 29.80% across artificially generated 300 reviews. Below is an example of a few of outputs.

Original Text	New Text	Original Objectivity	New Objectivity
This reckless president made a disastrous decision that harmed the country.	This president made a decision that harmed the country.	0.199	1.0
This unreliable product was horrible! I hate this stupid product!	This product was horrible! I hate this product!	0.0333	0.050

***Current* Limitation:** Occasionally, rule-based objectification would worsen objectivity due to the removal of a critical adjective and there are cases where grammatical/punctuation mistakes would occur due to missing/incomplete rule-checks.

Improved Proposed Method: There are two different proposals, one improvement on the current method and one new method.

The first proposal involves comprehensive objectivity/bias checks per removal/replacement of a word. The current method involves measuring a single word for subjectivity, however, such a method often fails to grasp the objectivity comprehensively. Thus, by checking the entire sentence after removing/replacing a word with a both an objectivity and bias score and comparing it to the sentence without the alterations, a deeper grasp of what should be replaced/removed can be achieved.

The second proposal involves fine-tuning an existing summarization model, preferably `sshleifer/distilbart-cnn-12-6` (or other medium-sized BART/T5 models), using bias and objectivity scores to train it to output more objective and less bias scores. This has the advantage over rule-based objectification as it can capture and retain the nuance. Possible complications include: computational restraints, under-objectification (no significant impact in objectification / bias removal), or over-objectification (objectification / bias removal to a point where the text loses its meaning).

The expected result is, for the first proposal, a more consistent objectifier that can objectify with fewer mistakes in removing critical adjectives. The expected result for the second proposal is a more comprehensive and innovative but more uncertain and resource intensive objectification.

3.5. Reliability Calculation

Objective: Uses a mix of information to comprehensively determine the reliability of a certain piece of information. Information may include source, temporal, and bias scores to determine overall trustworthy-ness of a piece of information.

***Current* Method:** Uses a CSV data with pre-assigned source reliability (Media, 2024), the objectivity of the text, and Gaussian-like weighing of individual sources to calculate general reliability of a single piece of information.

$$w_i = exp(\frac{2o^2}{-(r_i + r_{min})^2})$$

is used to calculate the weighting of individual sources with an emphasis on the smallest score, where r_{min} is the min score, o is standard deviation, and r_i is the individual score.

$$R_g = \frac{\sum_{i=1}^{N} w_i r_i}{\sum_{i=1}^{N} w_i}$$

is used to calculate general reliability by averaging the reliability with weights.

$$R_f = R_g * (1 + a * \frac{N_p}{N})$$

is used to calculate final reliability while accounting for outliers, where a is the outlier penalty, N_p is the number of scores that exceed $2r_{min}$, and N is the total number of scores.

Also uses temporal reliability, calculated based on coverage (how long-standing is a piece of information relative to other information) and recency (how recent is the information) as a weight.

$$S_C = \sqrt{\frac{C_i - C_{min}}{C_r}}$$

is used to calculate coverage scores, where C_i is the individual coverage length, C_{min} is the smallest coverage length, and C_r is the coverage

$$S_L = \min\left(1, \max\left(0, \frac{|D_i - D_{median}|}{\frac{(D_{max} - D_{min})}{2}}\right)\right)$$

is used to calculate recency scores, where D_i is the individual date, D_{median} is the median last date, D_{min} is the most recent date, and D_{max} is the oldest date.

***Current* Result:** The process outputs a numerical value denoting how "reliable" a piece of information. It is more comprehensive compared to most methods of binary methods of determining reliability.

***Current* Limitation:** The scoring lacks interpretability; other a smaller score representing higher reliability, there is no clear metric as to how reliable a score is. Furthermore, there is no score breakdown, costing transparency.

Improved Proposed Method: Instead of calculating one singular score, divide the score into multiple sections to allow for better understanding of the reasoning behind the scoring. Sections will include source, bias/objectivity, temporal reliability, and reportability (how reported information is) reliability. In addition, define explicit reliability categories depending on numerical values. The expected result is a more clear and distinct scoring method with greater transparency.

3.6. Opinion Extraction

Objective: Locate opinions and arguments on a topic to allow for greater insight into different perspectives on a topic.

Proposed Method: Scrape the web for specifically "opinion" articles on the topic and extract information. Classify the information into news (newly reported information), opinion (personal belief), or fact (established information) using zero-shot classification with `facebook/bart-large-mnli`. Then, group the opinions with textual clustering. Determine the general bias of the opinions and look for logical fallacies within the arguments using models like `q3fer/distilbert-base-fallacy-classification` (Jin et al., 2022).

Expected Result: A series of opinions can be retrieved, grouped, and presented with the information including their political bias and any fallacies in the argument to provide a clearer understanding of different perspectives and issues within arguments.

3.7. User Interface

Objective: The objective of this aspect is to create:

1. An easy-to-use interface that is both fast and reliable
2. A transparent process that increases trust

***Current* Method:** The current application is hosted on Streamlit with the UI also being built on Streamlit. The usage is query based, meaning a user inputs a text and the web is searched based on the text.

***Current* Result:** Streamlit provides a decently accessible user interface. Furthermore, it is extremely easy to host on.

***Current* Limitation:** There are massive computational constraints, meaning the processes are both slow and unreliable. There is also little transparency and little customizability in the Streamlit process.

Improved Proposed Method: Changing hosting platform and build a UI using a different method, possibly React-based frontend and AWS for hosting. Focus on a 3-section UI including the main points of information and their reliability, located opinions, and transparency about the process which details the thought process of the program.

4. Logistics

4.1. Timeline

This project will likely take upwards of 5½ months to complete. The following is an estimate of the time allocation. Month 1: Information Retrieval and Textual Clustering Month 2: Fake News Detection Month 3: Objectification / Bias Removal Month 4: Opinion Extraction Month 5: Reliability Calculations and UI Final 2 weeks: Buffer period

4.2. Resources

The main necessary resources will be APIs and computational resources. Necessary APIs will likely only be for information retrieval (ex: Google News API, Bing News API, and NewsAPI).

However, computational resource may entail cloud services like AWS, Google Collab Pro, or other similar services to allow for training/fine-tuning of some of the larger transformer models.

A multitude of models exist for models that I will not train:

1. `sshleifer/distilbart-cnn-12-6` for summarization and fine-tuning
2. `facebook/bart-large-mnli` for zero-shot classification
3. `q3fer/distilbert-base-fallacy-classification` for bias classification
4. `all-mpnet-base-v2` for sentence embedding

4.3. Risks

The three main risks of this project are Fake News Detection, Objectification / Bias Removal, and Opinion Extraction.

For fake news prediction, if a working model is not produced within 3 weeks of the given month, the project will fall back to using pretrained fake news prediction models.

For Objectification / Bias Removal, if the fine-tuning of the model shows no significant results within 2 weeks of the given month, the fallback will be focusing the rule-based objectification.

For Opinion Extraction, if a working methodology is unable to be produced within 2 weeks of the given month, this aspect of the project will be dropped in favor of fine-tuning and improving the other parts of the project.

5. Conclusion

Summary: The comprehensive solution proposed with ObjectiveNews is the first attempt to deliver objective and accurate news instead of just determining their validity. New textual clustering techniques, objectification methods, and reliability calculations all serve as tools to all serve as basic tools to transform how news is processed, evaluated, and delivered.

Future Direction: In the future, I hope to implement four new features: multilingual support, multimedia content, news feeds, and Google extensions. Multilingual support could expand the scope of the project and offer foreign perspectives on an issue, multimedia content could provide more insights and perspectives, news feeds could increase ease of use, and a Google Extension could allow for easy objectification for news articles.

Closing Statement: I hope that this project can serve as a tool to deliver more reliable information and allow people to beyond polarization, yellow journalism, and misinformation. In a world where these issues, along with AI, are growing in magnitude, may this project serve as a set in the right direction, away from hate and towards understanding and empathy.

References

Nilesh Barla. Dimensionality Reduction for Machine Learning. *Neptune*, August 2023. URL https://neptune.ai/blog/dimensionality-reduction.

David Borrelli. Clustering sentence embeddings to identify intents in short text | Towards Data Science. *Towards Data Science*, January 2025. URL https://towardsdatascience.com/clustering-sentence-embeddings-to-identify-intents-in-short-text-48d22d3bf02e.

Zhijing Jin, Abhinav Lalwani, Tejas Vaidhya, Xiaoyu Shen, Yiwen Ding, Zhiheng Lyu, Mrinmaya Sachan, Rada Mihalcea, and Bernhard Schölkopf. Logical Fallacy Detection. *arXiv*, February 2022. doi: 10.48550/arXiv.2202.13758.

Ad Fontes Media. Media bias chart. *Ad Fontes Media Reports*, 2024. URL https://app.adfontesmedia.com/chart/interactive.

Maximus Powers, Umang Mavani, Harshitha Reddy Jonala, Ansh Tiwari, and Hua Wei. GUS-Net: Social Bias Classification in Text with Generalizations, Unfairness, and Stereotypes. *arXiv*, October 2024. doi: 10.48550/arXiv.2410.08388.

Nishant Rai, Deepika Kumar, Naman Kaushik, Chandan Raj, and Ahad Ali. Fake News Classification using transformer based enhanced LSTM and BERT. *International Journal of Cognitive Computing in Engineering*, 3:98–105, June 2022. ISSN 2666-3074. doi: 10.1016/j.ijcce.2022.03.003.

Shaina Raza, Ananya Raval, and Veronica Chatrath. MBIAS: Mitigating Bias in Large Language Models While Retaining Context. *arXiv*, May 2024. doi: 10.48550/arXiv.2405.11290.

I. Kadek Sastrawan, I. P. A. Bayupati, and Dewa Made Sri Arsa. Detection of fake news using deep learning cnn–rnn based methods. *ICT Express*, 8(3):396–408, September 2022. ISSN 2405-9595. doi: 10.1016/j.icte.2021.10.003.

Lovedeep Singh. Clustering Text Using Attention. *arXiv*, January 2022. doi: 10.1109/ICCCNT51525.2021.9579951.

Harry Yin. Github - xild076/objectivenews: Tools made to find and create the most objective news possible. *GitHub Repository*, 2024. URL https://github.com/Xild076/ObjectiveNews.