
Our Approach: A Multi-Dimensional News Trustworthiness Framework

Traditional fake news detectors typically classify news as *real* or *fake* based on text features alone. However, we believe that **misinformation is not just about what's fake—it's also about how it's written and who (or what) created it.**

To address this, we built a system that integrates **three powerful predictors**:

1. **Fake News Detection** – A custom-trained classifier that predicts whether a news article is real or fake.
2. **Sentiment Analysis** – Our model analyzes whether the article tone is **positive**, **negative**, or **moderate/neutral**.
3. **AI-Generation Detection** – A fine-tuned detector that predicts whether the article was written by a human or generated by AI.

Then, we combine all three outputs to assess the **likelihood of misinformation** and provide the user with a **“Safe to Share?” suggestion** — empowering them to make an informed decision.

Why This Matters – Research-Backed Justification

- ♦ **False news spreads faster and farther than truth**, especially on social platforms

Vosoughi et al., MIT, Science (2018) – [Read Study](#)

- ♦ **Negative news grabs more attention and engagement**, even if false

Brady et al., Nature Scientific Reports (2024) – [Read Study](#)

- ♦ **AI-generated content is harder to detect and can amplify misinformation risks**

"Blessing or Curse? A Survey on the Impact of Generative AI on Fake News", arXiv (2024) – [Read Study](#)

By **merging these dimensions**, we're not just identifying fake news — we're estimating **how risky it is to share**.

Misinformation Risk Table (Based on Combined Predictions)

Predicted Real/Fake	Sentiment	AI-Generated	Misinformation Risk	Suggested Action
Real	Positive	Not AI	Low	Likely safe to share
Real	Moderate	Not AI	Low	Likely safe, but verify key claims
Real	Negative	Not AI	Moderate to Low	Verify source before sharing
Real	Positive	AI	Moderate	Check source credibility
Real	Moderate	AI	Moderate	Cross-check with trusted sources
Real	Negative	AI	Moderate to High	Consider fact-checking before sharing
Fake	Positive	Not AI	Moderate	Likely misleading, needs verification
Fake	Moderate	Not AI	Moderate to High	Potential misinformation, verify before sharing
Fake	Negative	Not AI	High	Likely misinformation, avoid sharing
Fake	Positive	AI	High	Likely AI-generated misinformation, avoid sharing
Fake	Moderate	AI	High	High misinformation risk, verify carefully
Fake	Negative	AI	Very High	Highly likely misinformation, do not share