

AI-Based Hybrid Fake News Detection System Using Machine Learning and NLP

1st Author

Dhanshree Vaidya

MCA, JSPM University, Pune

Abstract

We all know fake news is everywhere now — thanks to how addictive and fast social media is. News can go viral in seconds, but so can rumors and outright lies. This makes it nearly impossible for anyone to check everything manually. So, we built a fake news detection system that leans on artificial intelligence, machine learning, and natural language processing. Our system takes any piece of news text, decides if it's real, fake, or just uncertain, and gives people a fair shot at figuring out what's trustworthy. First, it cleans up the text — turns everything lowercase, cuts junk characters — then uses TF-IDF to turn words into numbers. These numbers feed into a Logistic Regression model to predict what's real or fake. On top of that, there's a rule-based module checking for classic scammy phrases, plus an uncertainty detector that watches for words like “may” and “expected.” The whole thing runs on Python with FastAPI and Streamlit, so results are instant and user-friendly. Mixing

machine learning, rule checks, and uncertainty handling makes our system more accurate and actually usable in the real world.

Keywords: Fake News Detection, Machine Learning, Natural Language Processing, TF-IDF, Logistic Regression, Text Classification

1. Introduction

The internet changed how everyone finds and shares news. Facebook, Twitter, and endless news sites make it easy to stay updated, but they've also turned into hotbeds for fake news. Fake stories, dressed up to look real, can swing opinions, confuse people, and sometimes even cause major trouble — especially around elections, pandemics, or financial chaos. Fact-checking once meant experts going through news piece by piece, but the sheer volume now makes that impossible. AI and machine learning have stepped up, able to churn through mountains of news and pick

out patterns that set lies apart from facts. But honestly, pure machine learning isn't perfect. It sometimes makes wrong calls with too much confidence, especially with blurry or vague content.

That's why our system mixes things up: we blend machine learning with hard-coded rules and an uncertainty detector. This hybrid approach bumps up accuracy, especially on tricky or ambiguous news pieces.

2. Literature Review

Lots of researchers have taken a shot at catching fake news. Ahmed et al. (2017) used standard models like Naïve Bayes and SVM and showed machines can learn the difference from just text. Wang (2017) put together the LIAR dataset — it's packed with political statements, and nearly everyone uses it for training or testing. Shu and his group (2017) went a step further, combining the content itself with social factors: who's posting, how people engage, etc., and that boosted performance. Conroy et al. (2015) found Logistic Regression worked well for text-based detection. Castillo's team (2011) pointed out how much language patterns matter for credibility. Ruchansky et al. (2017) tested a hybrid model with text, user behavior, and how news spreads, which got even better results. Ma and colleagues (2016) brought in neural networks and time analysis to spot rumors early. Zhou (2020) reviewed the field and showed hybrids really do work best.

For our project, we went with machine learning — specifically TF-IDF and Logistic

Regression — to keep it efficient and fast enough for real-time results.

3. Problem Statement

Most fake news detectors are good, but far from perfect. They make mistakes, especially with news that isn't clearly true or false. Many systems get tripped up by ambiguous statements and can let half-true stories slip by. Deep learning solutions often need heavy computing, which just isn't practical for real-time use. Plus, most still don't have strong validation checks.

We wanted something sharper — a hybrid approach that blends machine learning, rule-based checks, and uncertainty handling. This makes our system more accurate, reliable, and ready for fast, real-world detection.

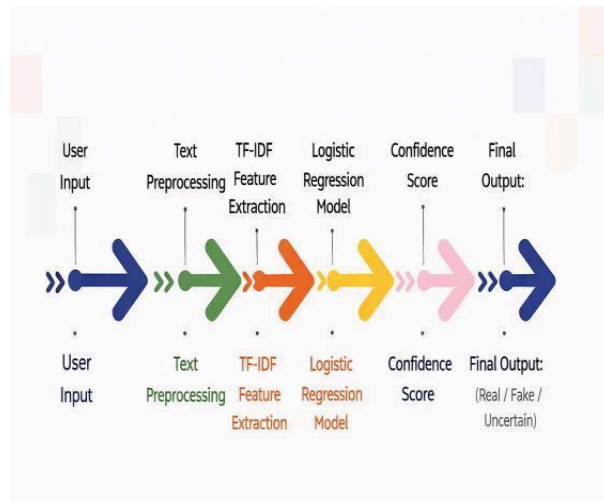
4. Proposed System

Here's how our system works: Users enter a news headline or article, and the text goes through a series of steps. First, it gets cleaned up — everything lowercase, clutter stripped out. TF-IDF turns the cleaned words into numbers. Then, Logistic Regression takes these features to guess if the news is real or fake, and it gives a confidence score.

To catch obvious scams, there's a rule-based module that looks for giveaway phrases like "free money" and "win cash." An uncertainty module hunts for hedging words like "may" and tags those as uncertain. The system combines the results

— machine learning prediction, rule checks, and uncertainty detection — puts the news into one of three buckets (Real, Fake, Uncertain), and shows the answer instantly on a web interface.

System Architecture Flow



5. Methodology

We designed the system as a straightforward but reliable machine learning pipeline.

5.1 Data Preprocessing

News text arrives messy: weird symbols, random casing, too much filler. We clean it, standardize the case, and strip out all unnecessary junk. This way, what goes into the next step is neat and predictable.

5.2 Feature Extraction with TF-IDF

Now, TF-IDF turns the cleaned text into numbers, weighing each word by its importance in the document and across all documents. Rare words that help identify fake news get a higher weight.

5.3 Training with Logistic Regression

With the features ready, we split the dataset for training and testing. Logistic Regression learns the patterns that separate real from fake and is efficient enough to handle lots of news quickly.

5.4 Prediction & Confidence Score

For new news entries, the model repeats the process and predicts the outcome, along with how confident it is in that prediction.

5.5 Rule-Based Validation

The rule module spots classic fake news phrases like “win cash now” or similar clickbait. If it sees these, it marks the news as fake, even if the model thinks differently.

5.6 Uncertainty Detection

When news uses hedging words — “may,” “expected,” etc. — instead of forcing a clear real/fake tag, the system marks it as “uncertain.”

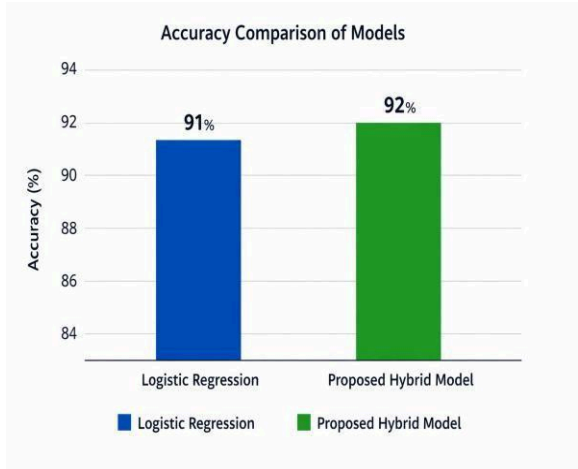
5.7 Final Classification & Output

We pull together outputs from the model, the rule check, and the uncertainty detector. The result goes to the user as Real, Fake, or Uncertain, along with a confidence score, all presented through our web app.

6. Experimental Results & Discussion

We tested everything on a labeled dataset, with part of it for training and part for testing. Logistic Regression on its own hit 91% accuracy. Adding in the rule-based and uncertainty modules nudged this up to 92%.

Model	Accuracy
Logistic Regression	91%
Proposed Hybrid Model	92%



The hybrid setup not only pumps up the accuracy, it's better at catching fuzzy or ambiguous news. Rules nail obvious scams, while uncertainty checks avoid bad guesses on unclear content.

7. Advantages

There's a lot to like about this system. It's fast, light on resources, and skips bulky deep learning. By using a mix of machine learning, rules, and uncertainty checks, it's both accurate and dependable. The uncertainty detection cuts down on flat-out

wrong classifications for tricky news. Plus, the web interface means anyone can use it easily.

8. Limitations

Of course, nothing's perfect. It only handles text (no images or videos, yet). Right now, it's set up for English — so no luck with other languages at the moment. The system's performance depends on the dataset's quality; if that's weak, so are the results. And rule-based detection will only find tricks it already knows — fresh scams can sneak by.

9. Future Work

There's room to grow. We want to make it multilingual, bringing in Hindi, Marathi, and more. Adding image and video capability means folding in computer vision. Direct hooks to social media via APIs would let us monitor breaking misinformation in real time. A mobile app would spread access, and explainable AI could help people actually trust (and understand) the results.

10. Conclusion

This project pulls together machine learning and rule-based checks into one hybrid system for catching fake news. TF-IDF extracts features, Logistic Regression does the classifying, and the extra modules lock down accuracy and reliability. It tackles the big issues with older methods, handles ambiguous news better, and is fast enough to work on modern digital platforms.

11. References

1. Ahmed H., Traore I., Saad S., "Detection of Online Fake News Using Machine Learning Techniques," International Conference on Intelligent Systems, 2017.
2. Wang W. Y., "Liar, Liar Pants on Fire: A New Benchmark Dataset for Fake News Detection," ACL Conference, 2017.
3. Shu K., Sliva A., Wang S., Tang J., Liu H., "Fake News Detection on Social Media," ACM SIGKDD Explorations, 2017.
4. Conroy N. J., Rubin V. L., Chen Y., "Automatic Deception Detection: Methods for Finding Fake News," Proceedings of the Association for Information Science, 2015.
5. Castillo C., Mendoza M., Poblete B., "Information Credibility on Twitter," WWW Conference, 2011.
6. Zhou X., Zafarani R., "A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities," ACM Computing Surveys, 2020.