Alex Damian

Kevin Morales

SE785

Professor Skevoulis

Truthteller (Tentative Name)

Fake News Detector Web App Utilizing ML Models

1.Proposal

The Problem

With the rise and advancement of generative AI, news media can be very difficult to filter through. Being able to read every news story with 100% certainty that the information in the article is factual, is going to become more difficult. While the term “fake news” has a certain stigma in today’s culture, the overall goal of the application will be to filter out new articles that are either reliable or unreliable using a trainable ML model. This software will be web based and allow any user to drop a file of an article into the web app and the AI model, using data we have trained it on will determine whether the article is real or fake. Personally, we have always looked to find the correct sources when looking through news media. Having a tool like this would not only make it easier to understand the overall media landscape but allows others to accurately understand what is true and what is not while also using AI and Machine Learning in an ethical and resourceful way.

2. Introduction

We have always had misinformation, but the new advent of AI-generated content makes it easier, faster and more believable than ever before. Many (traditional fact-checking) organizations and human review processes are much too slow relative to the volume and velocity of content.

Our approach is to utilize machine learning to provide fast, scalable and available verification. Unlike Web pages which are disputed by human experts as a fact-checking Web page that uses them, *Truthteller* incrementally adapts using the answer set of large training data and learning to get better in detecting lies. The application’s mission is twofold:

1. Give the average news consumer the power to make informed decisions about news.
2. Help to combat misinformation in general by providing a transparent and ethical AI driven solution.

3. Software

Truthteller will be engineered into a lightweight web-based platform that runs on both desktop and mobile. The key components include:

* Frontend (UI): It is a minimal and user-friendly design where users can paste the text, upload the files, or enter the URLs of articles.
* Backend (ML Model + API): A machine learning pipeline model trained on news and misinformation datasets of labeled articles. It'll read content used by NLP (Sentiment analysis, semantic similarity, linguistic capabilities).
* Training and Evaluation: Iterative learning process for the model to improve its performance taking in feedback from output, you know, just keep adding data to make it better.
* Output: Easy-to-understand results categorizing content (*e.g., likely reliable*, *unreliable*, or *requires further review*) along with confidence scores.

4. Prospective Users

* Everyday Readers: People who want a quick way to check accuracy of things they read online.
* Students and Scholars: People who value reliable information in their academic studies.
* Educators and Institutions: Schools, colleges, and universities that offer classes on media literacy.
* Journalists and newsrooms: The source verification tool for media pros everywhere.
* Policy and Advocacy Groups: Groups working to fight misinformation and promote fact-based discussion.

5. Alternatives

Although current fact-checkers (e.g., Snopes, PolitiFact) have done a good job in tackling misinformation, they are review-based and cannot scale with the continuous generation of new rumors. There are browser extensions and misinformation detection tools, but many do not provide information on the methods used, have small datasets or cannot be accessed by non-expert users. Filling such a void, *Truthteller* proposes to provide:

\*Automated, scalable and timely detection source=output.sink 9.15.1 Google provides an automated, scalable system for detecting phishing pages on the web -- FP-Detect--to index that are both automated detectors(covering half of the crawling set) or man ual %Table V: Details of Google's implementation in CIT commons (remainder of the crawling set).

* It is an easy to use and transparent system.
* A wider reaching availability to those beyond academia or professional journalisAm.