Jagriti

Savoring coherence through awareness

X

SOCIAL SERVICE COUNCIL

IIT (BHU) VARANASI

Problem statement

In today's digital age, misinformation has become a major concern, particularly with the rapid spread of fake news on platforms like social media, websites, and other online sources. Misinformation can influence public opinion, distort reality, compromise social stability, and impact important decision-making processes in areas like health, politics, and economics. The challenge is exacerbated by the ease with which false or misleading information can be shared and amplified across digital networks.

The aim is to classify news articles/posts into real and fake categories based on textual content, providing a framework to combat misinformation and foster trust in digital information.

To address this growing issue, an automated system is required to detect and differentiate between real and fake news articles. The goal is to build a model capable of classifying news articles.

Dataset overview

Title:

- The headline of the article is often crafted to grab attention. Titles can convey the overall intent or tone (e.g., neutral, biased, sensational).
- Example: "Ex-CIA head says Trump remarks on Russia interfere with democracy."
- Analysis of title features like sentiment, keyword frequency, and length (short, impactful vs. long, descriptive) may help in classification.

Text:

- This is the main content or body of the article, providing detailed information. Articles vary in length and depth of detail, with some being factual while others include opinions, exaggerations, or false claims.
- Example: "Former CIA director John Brennan criticized Trump's remarks, stating they undermine democratic principles."
- Techniques such as tokenization, sentiment analysis, or topic modeling can be applied here to extract meaningful patterns for classification.

Subject:

- Acategorical attribute representing the general theme or topic of the article, such as politics, news, or government affairs.
- Example categories include: politicsNews, Government News, left-news.
- Subject labels may indicate biases in coverage, as certain themes are more prone to polarizing or fabricated content.

Date:

- Represents when the article was published. Temporal patterns, like surges in fake news during elections or major events, could be useful for understanding the context of misinformation.

Annotations

***Labels**:*

- Binary labels (0 or 1) indicate whether the article is fake or real news.
- Example: An article labeled 1 is considered real, while one labeled 0 is fake.
- These labels form the ground truth for training and testing the classification model.

Indicators

Title Length:

- Analyzing the number of characters or words in a title can provide insights.
- Sensational or fake news articles often have longer, exaggerated headlines compared to concise, factual ones.

Text Length:

- The body length in terms of word count or sentences can be another useful indicator. Fake news articles might include verbose descriptions to obscure a lack of factual content.

Keyword Density:

- Examining the frequency of specific words (e.g., "shocking," "breaking," "exclusive") can reveal patterns of sensationalism often associated with fake news.

Sentiment Analysis:

- Measuring the sentiment (positive, negative, neutral) of the title and text.
- Fake news may use highly emotional language to provoke readers.

Linguistic Features:

- Attributes like grammar quality, readability, and use of punctuation (e.g., excessive use of exclamation marks) can differentiate between real and fake news.
- These indicators, when combined with text processing techniques, can provide a robust set of features for training machine learning models to detect fake news.

Link to dataset:

https://drive.google.com/drive/folders/1pfShmHPgUwxW3UMyFP-poABHfavh5B42?usp=sharing

Objective:

- Build a classification model based on textual content to predict the likelihood of different indicators.
- The model should employ evaluation metrics such as accuracy, precision, recall, F1 score, and area under the receiver operating characteristic curve (AUC-ROC).

Evaluation and Benchmarks:

- Participants are encouraged to use standard evaluation metrics:

Accuracy:

- The proportion of correctly classified instances.

Precision:

- The ability of the model to correctly identify positive instances among all instances predicted as positive.

Recall (Sensitivity):

- The ability of the model to correctly identify positive instances among all actual positive instances.
- F1 Score: The harmonic mean of precision and recall, providing a balanced measure.

AUC-ROC:

- The area under the receiver operating characteristic curve, measuring the model's ability to distinguish between positive and negative instances.
- The models will be assessed based on the above metrics. After the submissions, we will also evaluate the model on our dataset..

EXAMPLE FORMAT

The following format is used for the dataset:

Title	Text	Subject	Date	Label
"Trump's new policy reviewed"	"Details on the new policy..."	politicsNews	November 5, 2023	1
"Shocking claims about health!"	"Evidence suggests otherwise..."	health	July 22, 2021	0

This standardized format ensures consistency in data representation, facilitating the training and evaluation of the classification model.

Output Requirements

- For evaluation, the model should output results in a structured format. For instance, if given test data enclosed in ‘{}’, the results should be saved in a result.txt file with corresponding classifications also enclosed in ‘{}’.

Example Output:
{
}
["Trump's new policy reviewed", 1],
["Shocking claims about health!", 0]

Additional Testing:

- Participants are welcome to conduct additional testing further using their own or relevant external datasets to validate their models' robustness and generalization capabilities. Results from these other tests should be presented along with the primary evaluation metrics.