**ABSTRACT**

News Tracker is to collect a variety of sources of misinformation on social media and aggregate the stories published so that we could have a clearer sense of the different types of misinformation out there. We wanted to know how frequently it was published and what strategies and narratives were employed to engage audiences on social media. To do this, we built a database of misinformation sites, identified their social media pages, and found any related to that also shared the content. We used the social media API to collect all the stories posted to these pages as well as their engagement data — how many times users liked, shared, and commented on the stories. We continuously monitored this feed to identify patterns in and assess the substance of the posts. Fake news is a false piece of information. Any piece of Fake news can be created to deliberately misinform or deceive readers, promote a biased point of view, particular cause or agenda, and even for the entertainment. Fake news can be promoted by unauthenticated user ID, social media, printing of fake news in newspapers may be due to political pressure and many more. Spreading of fake news can cause discontent among people, riots, and even cause loss of trust between two people and even nations. It also has a huge possibility to exploit general public thinking in a completely different manner. News and media coverage gets hugely distorted due to initialization and spread of fake news.

**INTRODUCTION**

Nowadays, the news media are subject to the social scrutiny because of the lack of credibility, the manipulative media, the misinformation campaigns, and the propagation of fake news point out that the main difference between fake news and true news relies on lack editorial norms and processes that ensure the accuracy and credibility of the information. Thus, to arrange a way that allows guaranteeing these editorial processes (or at least part of them) can suppose a big step in the fight against the above-mentioned issues. To track the evolution of the news reports and the relevant data and information it contains as they change over time, and therefore to trace how the related news evolves, constitute other instruments to face the previous issues. This is not only useful for end readers but for fact-checking agencies and those tools that perform automatic indexing and extraction of relevant information of news. Once the information has been verified and fact-checked, these tools need a way to guarantee that the extracted data have not changed.

**FAKE NEWS**

Fake news refers to information content that is false, misleading or whose source cannot be verified. This content may be generated to intentionally damage reputations, deceive, or to gain attention. The term rose to popularity during the 2016 US Presidential Elections. It was reported that fake news likely influenced the results of the election.

Various types of Fake News include:

* Click bait. Often eye-catching content to capture readers at the expense of being factual.
* Satire/parody. This type of content is considered to be fun and humorous thus considered to be entertaining, yet some readers may interpret the content as fact.
* Propaganda. This is content meant to mislead and influence the reader.
* Biased/partisan/hyper-partisan.
* Oftentimes this is biased political content claiming to be impartial.

**AUTOMATED NEWS TRACKER APPLICATION**

Automated detection systems provide value in terms of automation and scalability. There are various techniques and approaches are implemented and detected in News Tracker Application. And it is worth noting that these approaches often overlap depending on perspective. From my personal perspective, I choose to discuss only two approaches. These two approaches focus on the methods used, as opposed to the content being analyzed. They may also both involve Natural Language Processing (NLP) in their methodology. Natural Language Processing enables computers to understand natural/human language and respond appropriately. Hence, there are two aspects involved:

* Natural Language Understanding
* Natural Language Generation

The two approaches to news tracker application detection are:

* Machine Learning approach
* Deep Learning approach

**PREDICTION**

“Prediction” refers to the output of an algorithm after it has been trained on a historical dataset and applied to new data when forecasting the likelihood of a particular outcome, such as whether or not a customer will churn in 30 days. The algorithm will generate probable values for an unknown variable for each record in the new data, allowing the model builder to identify what that value will most likely be. The word “prediction” can be misleading. In some cases, it really does mean that you are predicting a future outcome, such as when you’re using machine learning to determine the next best action in a marketing campaign. Other times, though, the “prediction” has to do with, for example, whether or not a transaction that already occurred was fraudulent. In that case, the transaction already happened, but you’re making an educated guess about whether or not it was legitimate, allowing you to take the appropriate action.

**PREDICTIONS IMPORTANT**

Machine learning model predictions allow businesses to make highly accurate guesses as to the likely outcomes of a question based on historical data, which can be about all kinds of things – customer churn likelihood, possible fraudulent activity, and more. These provide the business with insight that result in tangible business value. For example, if a model predicts a customer is likely to churn, the business can target them with specific communications and outreach that will prevent the loss of that customer.

**PREDICTION EXPLANATIONS IN MACHINE LEARNING**

Prediction explanations in machine learning explain how or why an AI platform arrived at an outcome. In the past, these models did not explain how a decision was made. This lack of clarity causes the “black box” syndrome, where you have predictions but are unsure of which feature variables affect a model’s outcomes. Machine learning models learn how to make decisions based on rules that it created while analyzing the training dataset. Prediction explanations allow us to understand these rules and how they are applied to new data, as well as what features are the most valuable considerations in determining the output. The prediction explanation assigns each input value a measure of importance, and if you take the sum of all of them, it will add up to 100%.For decision tree models, the prediction explanation follows the prediction path. Still, for others like regression, it is calculated based on aggregating the results of many predictions that use random variations of input data.

**PROPOSED SYSTEM**

The proposed system we are using a hybrid support vector machine and Random forest model for training. We are combining the two machine learning techniques so that we can able to get the accuracy score of above 98% for large datasets also.

**LIMITATIONS**

* Time constrains
* The time under which this work is to be submitted is short.
* This posed as a challenge for the researcher when gathering material for this project.
* Financial constrains
* The researcher had insufficient money to carry out all researches needed for this project.
* The fund for transportation, internet subscription, materials and all needed for this project was inadequate.

**ADVANTAGE**

* It provides higher accuracy through cross validation.
* High stability compare to Existing system.