

# FAKE NEWS DETECTION

MD ASIF RAZA (12500118075)

MD AAMIR SHABBIR (12500118077)

MD FIROZ ALAM (12500118125)

# CONTENTS

- ▶ Introduction
- ▶ Major problem
- ▶ Purpose
- ▶ Architecture
- ▶ Methodology
- ▶ Techniques
- ▶ Results
- ▶ Conclusion
- ▶ References

# INTRODUCTION



- ▶ Fake news exist way before from social media but it multifold when social media was introduced.
- ▶ Fake news is a news designed to deliberately spread hoaxes,propaganda and disinformation.
- ▶ Fake News stories usually spread through social media sites like Facebook,Twitter etc.

# BACKGROUND

- Social media is used for news reading.
- Source of the news
  - Professions used to distribute the news in the past
  - Nowadays, everybody want to be a journalist.
- People are profiting by clickbait's and publishing fake news on online
- More clicks contribute to more money for content publishers.

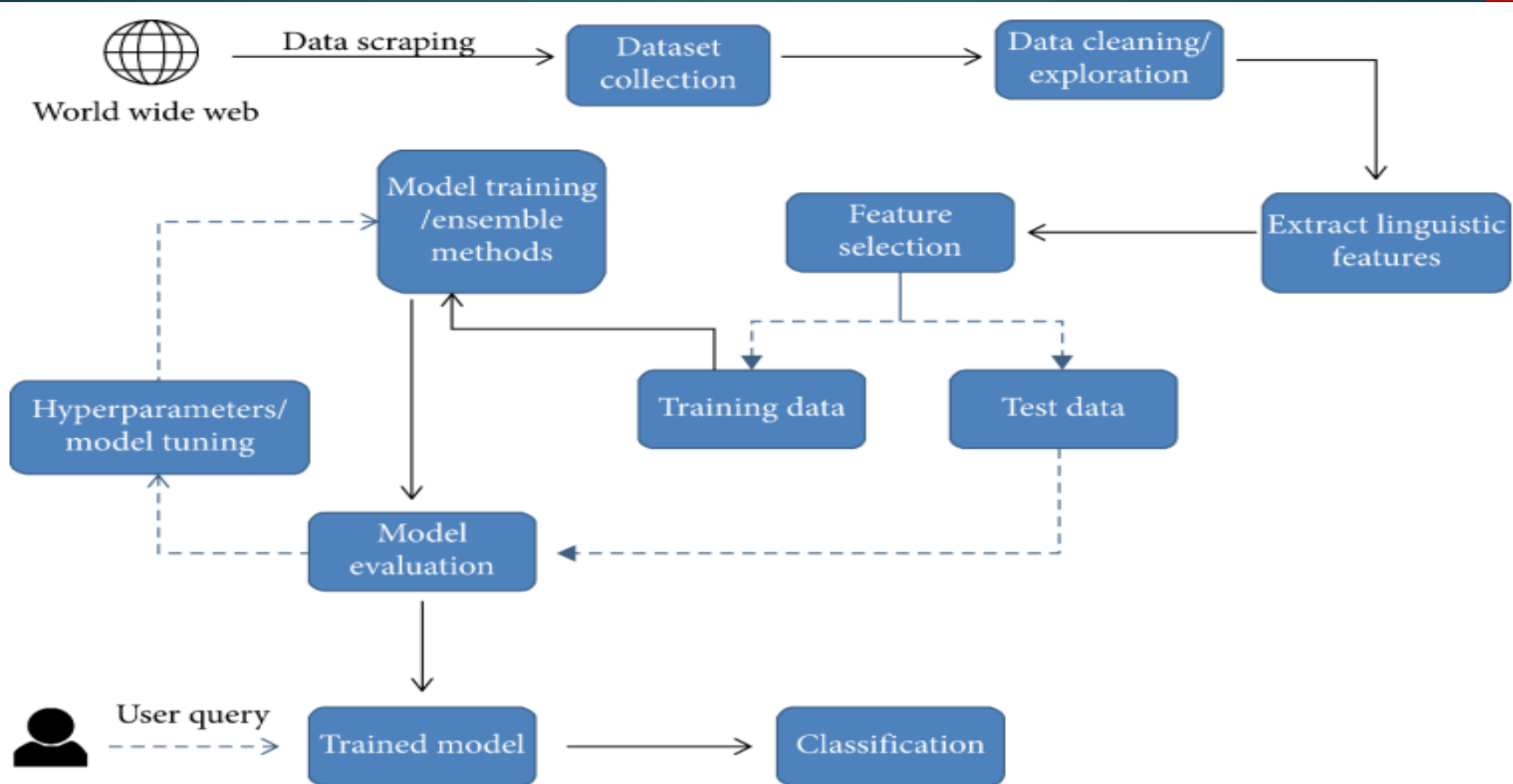
# MAJOR PROBLEM

- ▶ By clicking on a clickbait, users are led to a page that contains false information.
- ▶ Fake news influences people's perceptions.
- ▶ The rise of Fake news has become a global problem that even major tech companies like Facebook and google are struggling to solve. It can be difficult to determine whether a text is factual without additional context and human judgement.

# PURPOSE

- ▶ This project aims to develop a method for detecting and classifying Fake news stories using Natural Language Processing.
- ▶ The main goal is to identify fake news, which is a classic text classification issue.
- ▶ We gathered our data, preprocessed the text, and translated our articles into supervised model features.
- ▶ Our goal is to develop a model that classifies a given news articles as either fake or true.

# ARCHITECTURE



**Figure 1:** Workflow for training algorithms and classification of news articles.

# DELIMITATIONS

- ▶ Our system does not guarantee 100 % accuracy
- ▶ The system is unable to test data that is unrelated to the training database.

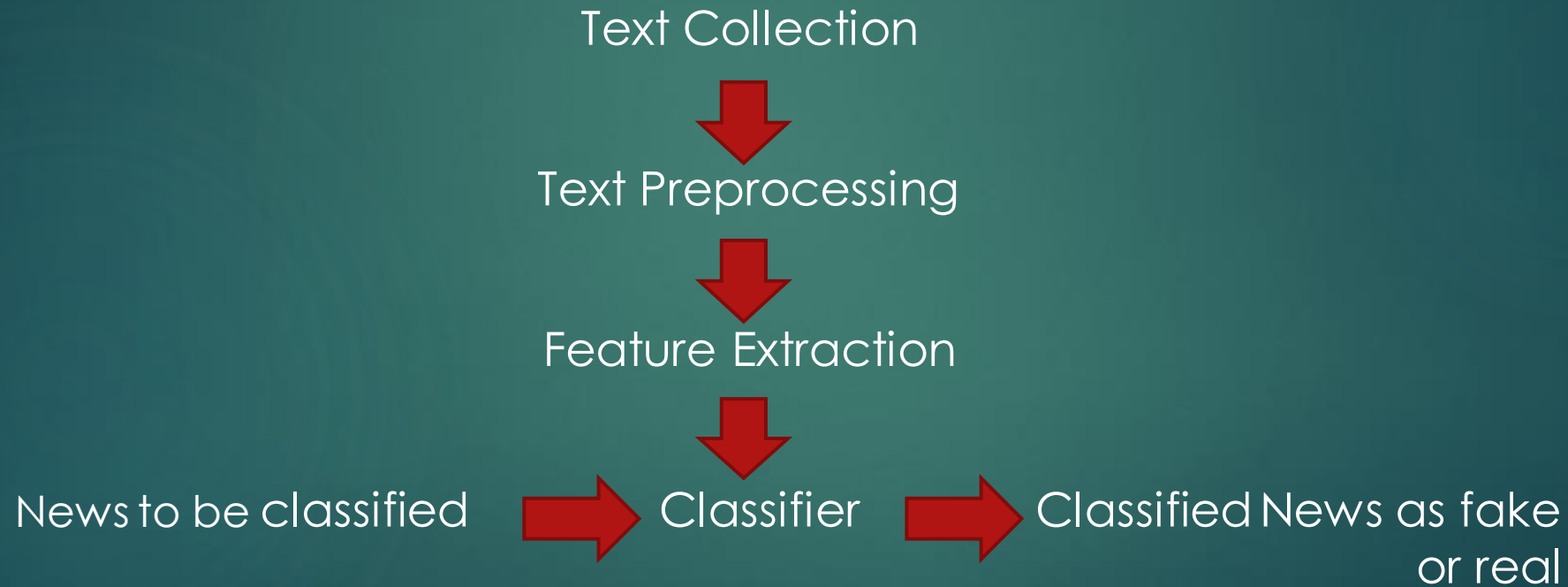


# TYPES OF FAKE NEWS

- ▶ Visual based type: Visual based are mainly photoshopped images and videos which are posted in social medias.
- ▶ Linguistic based type: Linguistic based are mainly the manipulation of text and string content. This issues is with blogs, news, or emails.

# METHADODOLOGY

The fake news model detection is built using steps like text collection, Text preprocessing ,Feature Extraction and then finally classification using different classifier.



# TECHNIQUES

- ▶ We used Ensemble Learner, RF for feature extraction.
- ▶ We trained our data by using SVM, KNN and LR algorithms.
- ▶ We tested the efficiency of the classifier using accuracy.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} .$$

# SVM(support vector machine)

- ▶ Support vector machine is a supervised machine learning algorithm used for both classification, regression and outliers detection..
- ▶ The objective of an SVM model is to estimate a hyperplane on the basis of feature set to classify data points.
- ▶ The dimension of hyperplane varies according to the number of features.

# K-NN ALGORITHM

- ▶ The K-NN supervised machine learning algorithm denotes to K Nearest Neighbors algorithm.
- ▶ KNN works by calculating the distances between a query and all the example in the data
- ▶ KNN can be used for classification and regression problems.
- ▶ KNN significantly slows when the amount of data increases

# RESULT

We can evaluate machine learning algorithm using various metrics like:

1. Accuracy
2. Precision
3. Recall
4. F1-Score

Hence we evaluate and analyse the

Result based on these metrics for different datasets, classifier and different methods of feature extraction methodology.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} .$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} .$$


$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} .$$

$$\text{F1 - score} = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} .$$

# CONCLUSION

- ▶ The task of classifying news manually requires in-depth knowledge of the domain and expertise to identify anomalies in the text. In this research, we discussed the problem of classifying fake news articles using machine learning models and ensemble techniques. The data we used in our work is collected from the World Wide Web and contains news articles from various domains to cover most of the news rather than specifically classifying political news. The primary aim of the research is to identify patterns in text that differentiate fake articles from true news. We extracted different textual features from the articles using an LIWC tool and used the feature set as an input to the models. The learning models were trained and parameter-tuned to obtain optimal accuracy. Some models have achieved comparatively higher accuracy than others. We used multiple performance metrics to compare the results for each algorithm. The ensemble learners have shown an overall better score on all performance metrics as compared to the individual learners.



- 
- ▶ Fake news detection has many open issues that require attention of researchers. For instance, in order to reduce the spread of fake news, identifying key elements involved in the spread of news is an important step. Graph theory and machine learning techniques can be employed to identify the key sources involved in spread of fake news. Likewise, real time fake news identification in videos can be another possible future direction



# REFERENCES

- ▶ A. Douglas, “News consumption and the new electronic media,” *The International Journal of Press/Politics*, vol. 11, no. 1, pp. 29–52, 2006. View at: [Publisher Site](#) | [Google Scholar](#)
- ▶ J. Wong, “Almost all the traffic to fake news sites is from facebook, new data show,” 2016. View at: [Google Scholar](#)
- ▶ A. D. Holan, 2016 Lie of the Year: Fake News, Politifact, Washington, DC, USA, 2016
- ▶ . H. Ahmed, I. Traore, and S. Saad, “Detection of online fake news using n-gram analysis and machine learning techniques,” in *Proceedings of the International Conference on Intelligent,*



THANK YOU