FAKE NEWS PREDICTION

TEAM MEMBERS

BHADRA V RAJA
RA2111027010077

DANIA B SAM
RA2111027010071

SREELEKSHMY S
RA2111027010118

OVERVIEW

01
Introduction

02
Motivation

03
Problem Statement

04
Literature Review

05
Inference

06
References

INTRODUCTION

- Fake news refers to false or misleading information presented as factual news.

- Machine learning techniques can help analyze large volumes of news articles, social media posts, and other content to assess their credibility and likelihood of being fake.

- Inaccurate news can spread quickly and have serious consequences, influencing public opinion and decision-making.

- Fake news prediction helps in real-world situations like media, journalism, and online platforms.

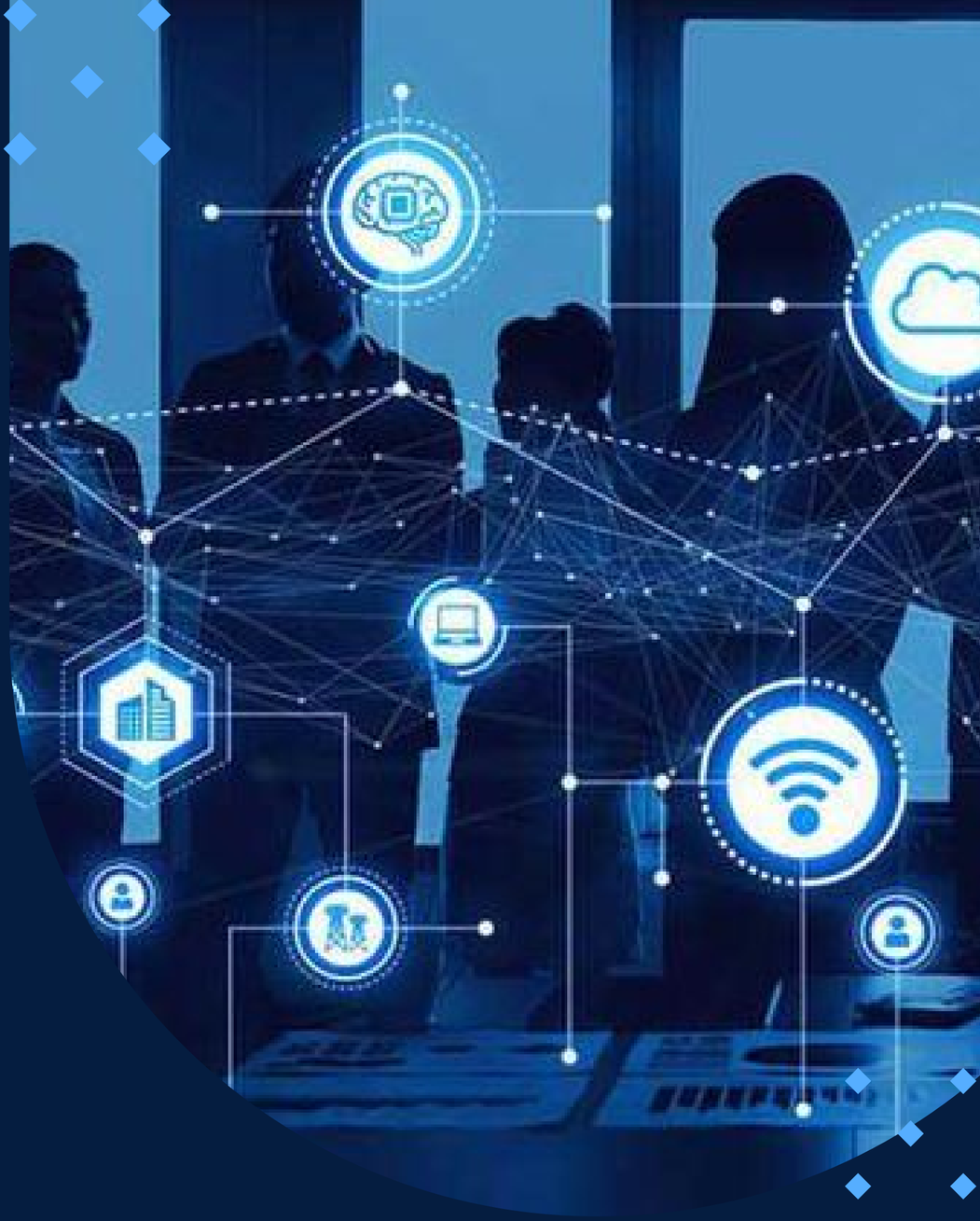
MOTIVATION

64%

Pew Research Center (2019):
from this survey , it was found that 64% of adults said that made-up news has caused "a great deal" of confusion about basic facts of current events.

44%

Reuters Institute Digital News Report (2021):
The report showed that 44% of respondents across 46 countries were concerned about what is real and fake on the internet.



PROBLEM STATEMENT

In today's age of widespread sharing of information online, the spread of fake news is a big issue. It's becoming harder to trust what we read. This project aims to create a machine learning tool using logistic regression and Natural Language Processing (NLP) methods, like stemming. This tool's goal is to predict if news articles are real or fake, helping us make more informed choices about the news we believe.

LITERATURE REVIEW

NAME	TECHNIQUES	ADVANTAGE	DISADVANTAGE
Constructing a User-Centered Fake News Detection Model by Using Classification Algorithms in Machine Learning Techniques	LR,NNET,RF,SVM,CART	solid durability against overfitting (provides regularization to prevent overfitting)	the model fitness inevitably deteriorates due to the increase in noise
Evidence-Aware Multilingual Fake News Detection	LSTM	this detection system applicable to various languages and regions	using the evidences and working with multiple languages can increase the complexity of the detection system
Fake Online Reviews: A Unified Detection Model Using Deception Theories	logistic regression (LR), Naïve Bayes (NB), decision tree (DT), and random forest (RF)	By considering both verbal and non-verbal features, the model's accuracy in distinguishing between fake and genuine reviews is improved.	using neural networks could have given better predictions.

LITERATURE REVIEW

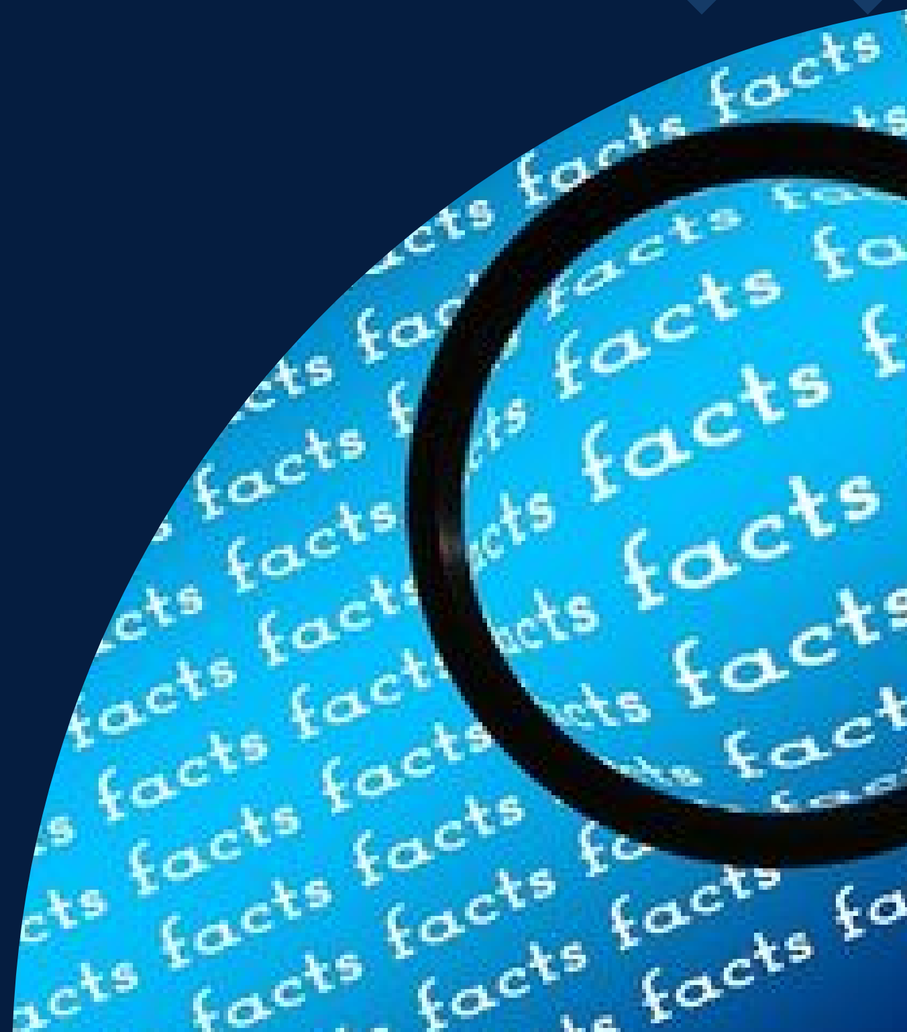
NAME	TECHNIQUES	ADVANTAGE	DISADVANTAGE
Fake News Detection and Prediction Using Machine Learning Algorithms	k-nearest neighbor	By using multiple search engines, the accuracy of the system in detecting fake or real news is significantly enhanced, reaching up to 90%.	Using many search engines can take up a lot of computer power and time because it requires a lot of work and energy to gather information from all of them.
Fake News Detection on Social Media: A Data Mining Perspective	Latent Dirichlet Allocation (LDA)	can handle a large amount of social media data, making it suitable for analyzing the vast volume of contents	fake news and real news can talk about similar things, the method might mistakenly think real news is fake
Fake News Detection: A Deep Learning Approach	TF-IDF and BoW	The model works well for headlines and articles with stances like 'agree', 'discuss', and 'unrelated'	indicates a limitation in effectively identifying content where the headline and article contradict each other.

LITERATURE REVIEW

NAME	TECHNIQUES	ADVANTAGE	DISADVANTAGE
Efficient Fake News Detection Mechanism Using Enhanced Deep Learning Model	Convolutional Neural Network (CNN), Recurrent Neural Network (RNN)	High accuracy in identifying complex patterns and contextual clues in fake news	Requires substantial computational resources and potentially longer training times.
A Comparative Study of Machine Learning and Deep Learning Techniques for Fake News Detection	Random Forest, Support Vector Machine, LSTM (Long Short-Term Memory)	Provides insights into the performance trade-offs between traditional ML and deep learning	May not fully exploit the potential of the latest deep learning innovations
Machine Learning-Based Identification of COVID-19 Fake News Using Biomedical Information Extraction	Logistic Regression, Naive Bayes	Easy to interpret, handles linear relations,Simple, computationally efficient	May not achieve the same level of accuracy as complex DL models.

Inference

- **The application of machine learning techniques, specifically logistic regression, in conjunction with fundamental Natural Language Processing (NLP) principles and stemming, provides an effective solution for predicting and identifying fake news articles**
- Logistic regression, a powerful classification algorithm, allows us to build a predictive model that can distinguish between real and fake news based on textual content.
- The integration of NLP techniques further enhances the model's predictive capabilities. NLP involves the analysis and processing of human language, allowing us to extract meaningful insights from text
- The integration of NLP techniques further enhances the model's predictive capabilities. NLP involves the analysis and processing of human language, allowing us to extract meaningful insights from text



REFERENCES

https://www.analyticsvidhya.com/blog/2023/02/tackling-fake-news-with-machine-learning/#:~:text=Here%2C%20we%20first%20create%20a,a%20corresponding%20label%20vector%2C%20y.

https://arxiv.org/ftp/arxiv/papers/2102/2102.04458.pdf

https://sist.sathyabama.ac.in/sist_naac/documents/1.3.4/b.e-cse-batchno-214.pdf

https://iopscience.iop.org/article/10.1088/1757-899X/1099/1/012040/pdff

https://data-flair.training/blogs/advanced-python-project-detecting-fake-news/

https://www.simplilearn.com/tutorials/machine-learning-tutorial/how-to-create-a-fake-news-detection-system