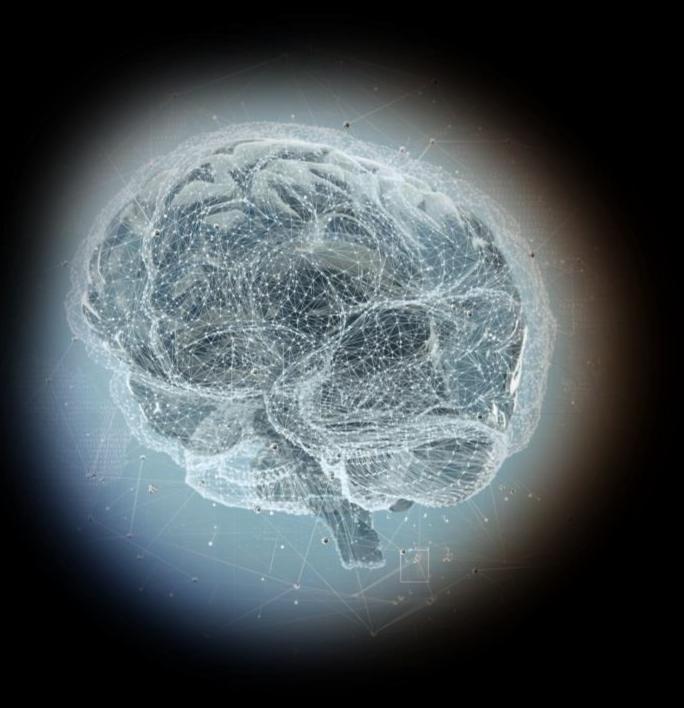# ARTIFICIAL INTELLIGENCE

PROJECT
TITLE:FAKE NEWS
DETECTION
USING NLP

MENTOR:

DR. M. RUBAN GLADWIN, M.E, PH.D, (AP/ECE)

- OUR TEAM MEMBERS:

- AKILA.J                             -951321106001

- DAYANA.S                      -951321106007

- DIVYA DHARSHINI.R      -951321106008

- JERLIN PRABHV. J          -951321106018

- RATHINA MALA. B         -951321106039

# INTRODUCTION

- Intelligence:"The capacity to learn and solve problems'. Artificial intelligence is the simulation of humans intelligence by machines.

# PROBLEM DEFINITION:

- The fake news dataset is one of the classictext analytics datasets available on Kaggle. Itconsists of genuine and fake articles' titles and textfrom different authors. Our job is to create a modelwhich predicts whether a given news is real or fake.

- The problem is to develop a fake news detectionmodel using a Kaggle dataset. The goal is to distinguisbetween genuine and fake news articles based on theirtitles and text. This project involves using naturallanguage processing (NLP) techniques to preprocessthe text data, building a machine learning model forclassification, and evaluating the model's performance.Input: News items, social contexts and associated sideinformation

Output: One of two labels: 'fake' or 'real'.

# DESIGN THINKING:

- 1. Data Source: Choose the fake news datasetavailable on Kaggle, containing articles titles andtext, along with their labels (genuine or fake).

- 2. Data Preprocessing: Clean and preprocess thetextual data to prepare it for analysis.

- 3. Feature Extraction: Utilize techniques like TF-IDF(Term Frequency-Inverse Document Frequency) orword embeddings to convert text into numericalfeatures.

- 4. Model Selection: Select a suitable classificationalgorithm (e.g., Logistic Regression, Random Forest,oor Neural Networks) for the fake news detectiontask.

- 5. Model Training: Train the selected model using thepreprocessed data.

- 6. Evaluation: Evaluate the model's performance usingmetrics like accuracy, precision, recall, F1-score

# INNOVATION TECHNIQUES:

- It is vital to recognize and differentiate between false and accurate news. One method is to have anexpert decide, and fact checks every piece of information, but this takes time and needs expertise thatcannot be shared. Secondly, we can use machine learning and artificial intelligence tools to automatethe identification of fake news.Online news information includes various unstructured format data (such as documents, videos, adiaudio), but we will concentrate on text format news here. With the progress of machine learning andNatural language processing, we can now recognize the misleading and false character of an article or statement.

# SOFTWARE TOOL:

- 1)pandas(gateway into the world of data science)

- 2)numby(fundamental Python package for scientific computations)

- 3)sklearn(the most useful and robust library for machine learning in Python and classification of the function)

- 4)TfidfVectorizer

## Programming language:

- 1)Python

# SOURCE CODE:

- # This Python 3 environment comes with many helpful analytics libraries installed

- # It is defined by the kaggle/python Docker image: https://github.com/kaggle/docker-python

- # For example, here's several helpful packages to load

- import warnings

- warnings.filterwarnings('ignore')

- import numpy as np # linear algebra

- import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

```python
import nltk

from nltk.sentiment import SentimentIntensityAnalyzer

import warnings

warnings.filterwarnings("ignore")

from sklearn.model_selection import train_test_split

from sklearn.feature_extraction.text import TfidfVectorizer

from sklearn.svm import SVC

from sklearn.metrics import accuracy_score, classification_report
```

```
# Input data files are available in the read-only "../input/" directory

# For example, running this (by clicking run or pressing Shift+Enter) will list all files under the input
directory

import os

for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames: print(os.path.join(dirname, filename))# You can write up to 20GB to the
current directory (/kaggle/working/) that gets preserved as output when you create a version using
"Save & Run All"

# You can also write temporary files to /kaggle/temp/, but they won't be saved outside of the
current session
```

/kaggle/input/fake-and-real-news-dataset/True.csv

/kaggle/input/fake-and-real-news-dataset/Fake.csv

Source code:

Loading Data:

true = pd.read_csv('/kaggle/input/fake-and-real-news-dataset/True.csv')

fake = pd.read_csv('/kaggle/input/fake-and-real-news-dataset/Fake.csv')In [3]:

# linkcode
# fake['Category'] = 'fake'

# Fake['Category'] = 'fake'

output:

| s.n | title | text | subject | date | category |
|-----|-------|------|---------|------|----------|
| 0 | Donald Trump Sends Out Embarrassing New 'Year'... | Donald Trump just couldn t wish all Americans ... | News | December 31, 2017 | fake |
| 1 | Drunk Bragging Trump Staffer Started Russian ... | House Intelligence Committee Chairman Devin Nu... | News | December 31, 2017 | fake |
| 2 | Sheriff David Clarke Becomes An Internet Joke... | On Friday, it was revealed that former Milwauk... | News | December 30, 2017 | fake |
| 3 | Trump Is So Obsessed He Even Has Obama's Name... | On Christmas day, Donald Trump announced that ... | News | December 29, 2017 | fake |
| 4 | Pope Francis Just Called Out Donald Trump Dur... | Pope Francis used his annual Christmas Day mes... | News | December 25, 2017 | fake |

# true['Category'] = 'true'

| s.no | Title | text | subject | date | Category |
|---|---|---|---|---|---|
| 0 | As U.S. budget fight looms, Republicans flip t... | WASHINGTON (Reuters) - The head of a conservat... | politicsNews | December 31, 2017 | true |
| 1 | U.S. military to accept transgender recruits o... | WASHINGTON (Reuters) - Transgender people will... | politicsNews | December 29, 2017 | true |
| 2 | Senior U.S. Republican senator: 'Let Mr. Muell... | WASHINGTON (Reuters) - The special counsel inv... | politicsNews | December 31, 2017 | true |
| 3 | FBI Russia probe helped by Australian diplomat... | WASHINGTON (Reuters) - Trump campaign adviser ... | politicsNews | December 30, 2017 | true |
| 4 | Trump wants Postal Service to charge 'much mor... | SEATTLE/WASHINGTON (Reuters) - President Donal... | politicsNews | December 29, 2017 | true |

# REFERENCE:

1)Artificial Intelligence for Humans by Jeff Heaton https://youtu.be/f7ckMaEJvh4?si=0EYjBfRCUjoKahgf

2)https://youtu.be/2o-_bMKGi_o?si=y_L8U6EVfIU_tc93

3)https://youtu.be/DdgymahD7Zg?si=MTWrP0W1s7QvSQid

4)Combating Fake News with Computational Intelligence Techniques

books.googl

Mohamed Lahby, Al-Sakib Khan Pathan, Yassine Maleh · 2021

5)https://books.google.co.in/books?id=GfFUEAAAQBAJ&printsec=frontcover&dq=fake+news+detection+usi

ng+python+and+nlp+reference&hl=en&newbks=1&newbks_redir=1&sa=X&ved=2ahUKEwj7hOiU3uOB
AxWsbWwGHdD7BFYQ6AF6BAgFEAI

Thank you