



**VELAMMAL**  
INSTITUTE OF TECHNOLOGY

Approved by AICTE - New Delhi  
Affiliated to Anna University - Chennai  
Accredited by NBA & NAAC

## **NAAN MUDHALVAN PROJECT(IBM)**

### **IBM AI 101 ARTIFICIAL INTELLIGENCE-GROUP 1**

**Team name:** Proj\_224823\_Team\_1

**Team members:** SAI VIGNESH S (reg no 113321106079)

THIRUMARAN S (reg no 113321106106)

VIGNESH V (reg no 113321106115)

LINGESH V(reg no 113321106301)

SRINIVASA BHARATH S(reg 113321106094)

**Problem Statement** : Design and develop an NLP-based system that can accurately identify and classify news articles or information as either "fake" or "real" by analyzing the textual content, with the primary goal of mitigating the spread of misinformation and promoting the dissemination of trustworthy information.

## **DATA ACQUISITION:**

Collecting data from a variety of sources, including news websites, social media platforms, and other online media, as part of the data acquisition process for fake news detection using NLP .

## **DATA PREPROCESSING:**

The process of cleaning the data by removing irrelevant information, such as stop words, lowercasing the words punctuations, and special characters.

## **ANALYSIS TECHNIQUE:**

- **Sentiment Analysis**
- **Semantic Analysis**
- **Fact Checking**
- **Named Entity Recognition (NER)**

## **FEATURE EXTRACTION:**

Extracting features from the preprocessed data, such as term frequency-inverse document frequency (TF-IDF) vectors, word embeddings, and bag-of-words representations.

## **MODEL TRAINING:**

Training a machine learning model, such as a decision tree, random forest, or support vector machine (SVM), on the extracted features.

# **MODAL EVALUATION & PREDICTION:**

Evaluating the performance of the trained model using metrics such as accuracy, precision, recall, and F1 score.

Using the trained model to predict the authenticity of new news articles.