# Fake News Detection Using NLP

Join me in exploring how natural language processing can be used to identify fake news, from data source to model evaluation.

BY : TEAM NLP


# Data Source

## Social Media

Twitter, Facebook, and other social media platforms are popular sources of news and information.

## News Outlets

Traditional news organizations often have online sites, which provides a wealth of data.

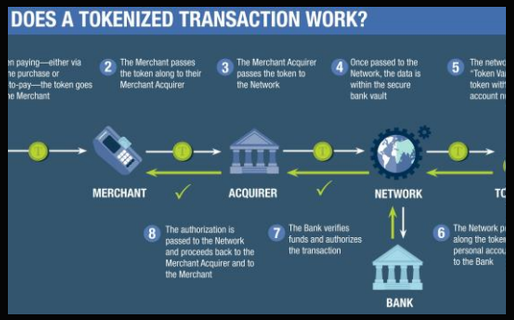
## Web Scraping

Web scraping tools can extract data from any publicly available source online.

# Data Preprocessing

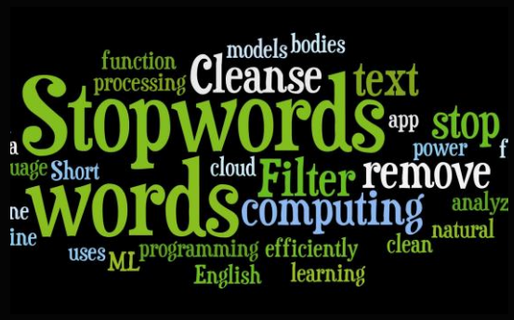

## Data Cleaning

Remove any irrelevant information, such as HTML tags and URLs.



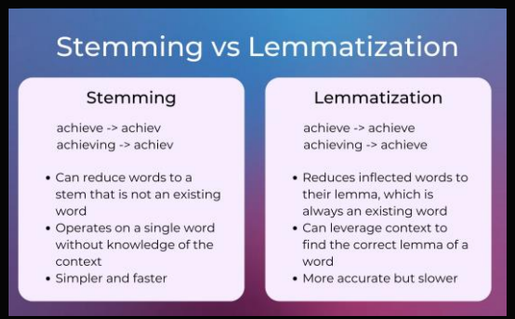

## Data Tokenization

Split up the data into words or phrases so that it can be easily analyzed.




## Stopword Removal

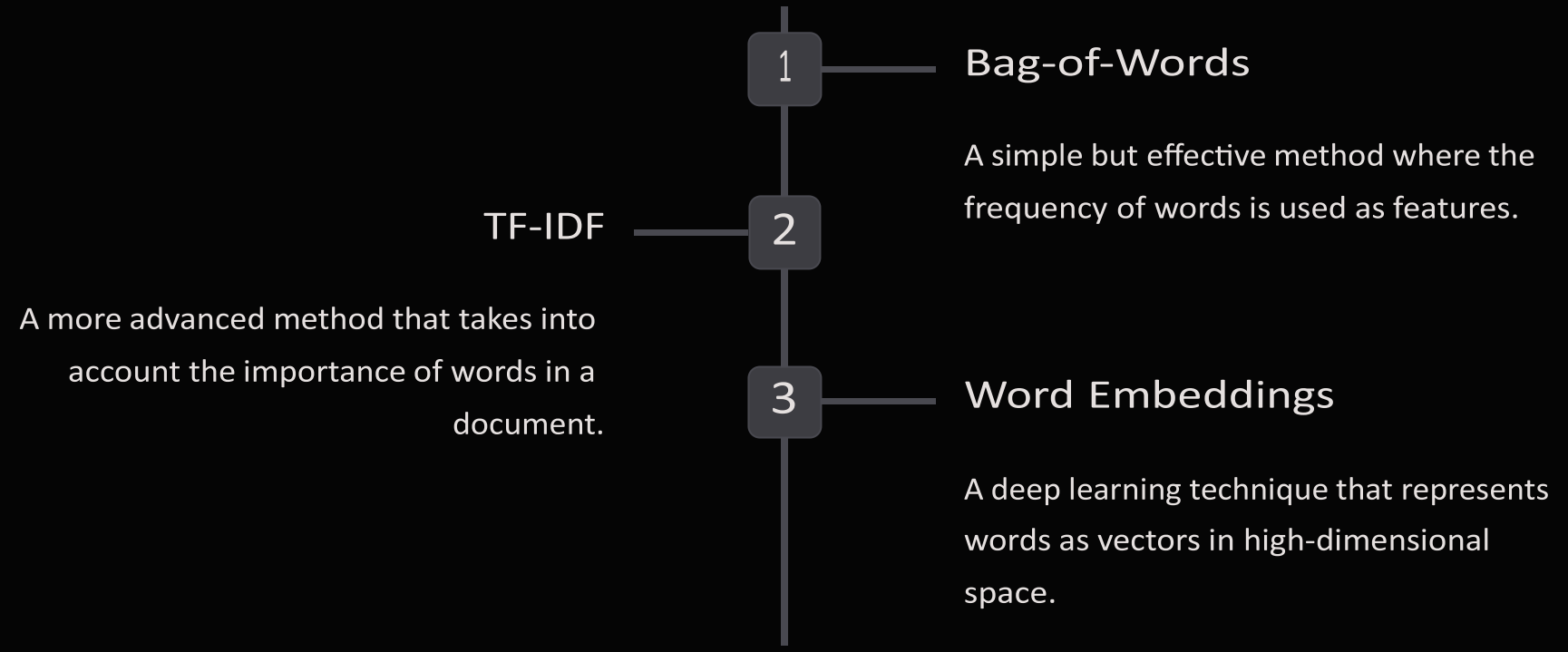
Filter out common words that don't provide useful information.



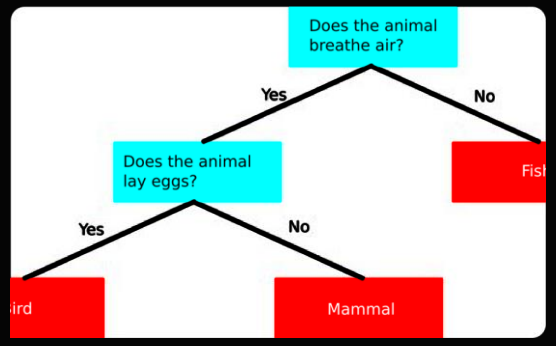

## Stemming & Lemmatizing

Reduce words to their root stem or base form so that they can be analyzed more accurately.

# Feature Extraction

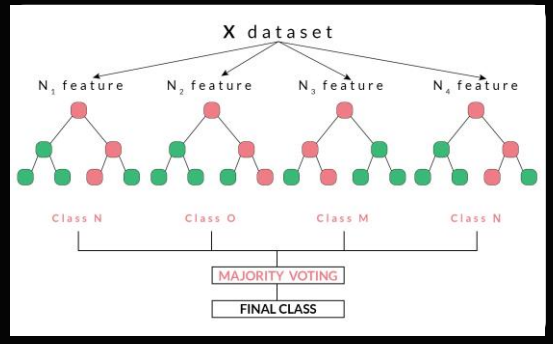


# Model Selection



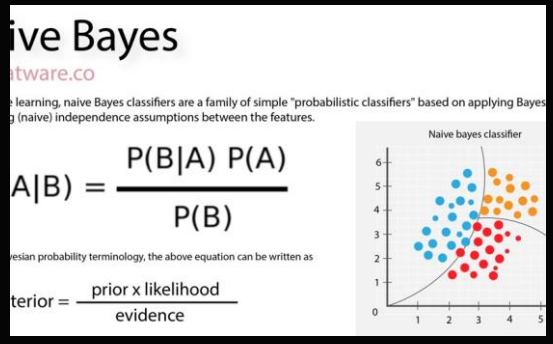

## Decision Tree

Creates a tree-like model of decisions and their possible consequences.




## Random Forest

Creates multiple decision trees to improve accuracy and reduce overfitting.




## Naive Bayes

A probabilistic algorithm that makes predictions based on a set of probabilities.

# Model Training

## 1 Split Data

Divide the dataset into a training set and a testing set. The training set teaches the model and the testing set evaluates its accuracy.

## 2 Train the Model

Feed the training data into the selected algorithm and adjust the parameters to optimize performance.

## 3 Repeat

Train and test the model multiple times, adjusting hyperparameters each time until optimal accuracy is achieved.

# Model Evaluation

## Accuracy

How well the model predicted the correct class.

## Precision

How often the model predicted the correct class when it actually was that class.

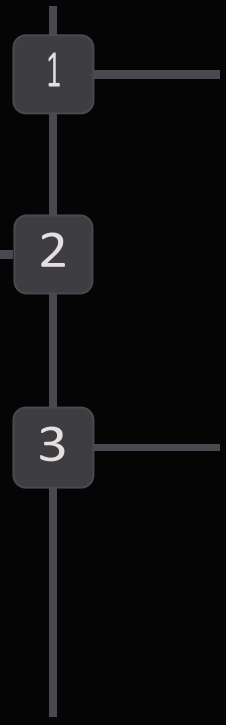
## Recall

How often the model correctly identified a class when it was actually that class.

## F1 Score

A combination of recall and precision that takes into account false positive and false negative rates.

# Application of NLP for Fake News Detection




## Sentiment Analysis

Detecting and analyzing the emotions behind a piece of writing can give insight into its credibility.

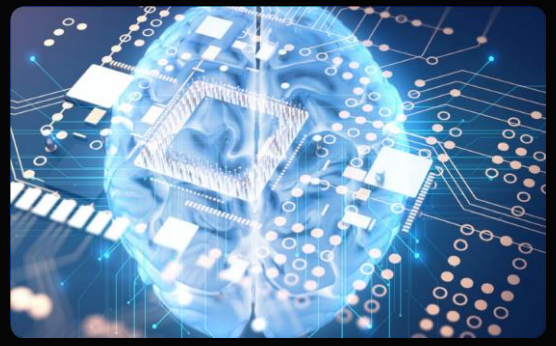
## Named Entity Recognition

Identifying people, places, and organizations mentioned in an article can help verify its authenticity.

## Topic Modeling

Analyzing the words and phrases used in an article can help identify its topic and any possible biases.

# The Future of Fake News Detection

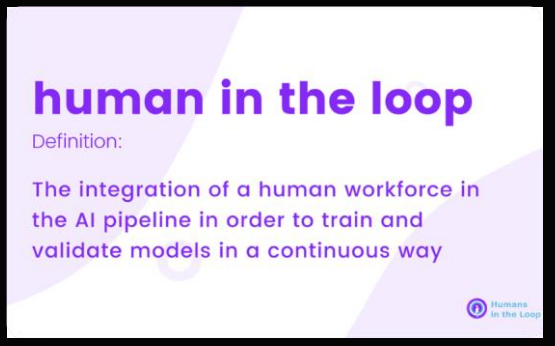


## Deep Learning

Using neural networks with many layers to detect complex patterns and relationships in data.


## Explainable AI

Developing models that provide transparent explanations for their decisions.




## Human-in-the-Loop

Combining the strengths of AI and human judgement to create more accurate models.