AI Dataset Deconstruction

Two CSV files one having only fake news and one only having True news

Both CSV file consists of four Categories

Title (Data - String)

Text (Data – String)

Subject (Data – String)

Date (Data – String & Integer)

Application of Dataset

Goal Objective:

Are there Recognizable patterns between True and Fake news Articles

Potential Machine Learning Models:

SUPERVISED LEARNING

Logistic Regression

Support Vecetor Machines(SVM)

* Good for text Categorization

K-Nearest Neighbours

* Pattern Recognition
* Sensitive to choice of K

Naïve Bayes

* Text classification
* Well performing

UNSUPERVISED LEARNING

Linear Discriminant Analysis(LDA)

* Pattern Recognition