## Team Members

**SRI VISHNU. M :** [**srivishnu1103@gmail.com**](mailto:srivishnu1103@gmail.com)  **(Team Leader)**

**SATHIS KUMAR. G :** [**g.sathiskumar22@gmail.com**](mailto:g.sathiskumar22@gmail.com)

**SANJAY KUMAR. R :** [**rsanjaykumar1319@gmail.com**](mailto:rsanjaykumar1319@gmail.com)

**MOHAMMED NOORULLA. S :**

**noorullasnoorul123@gmail.com**

Fake News Detection Using NLP

Detecting fake news using natural language processing (NLP) involves several key steps:

• Feature engineering

• Model training

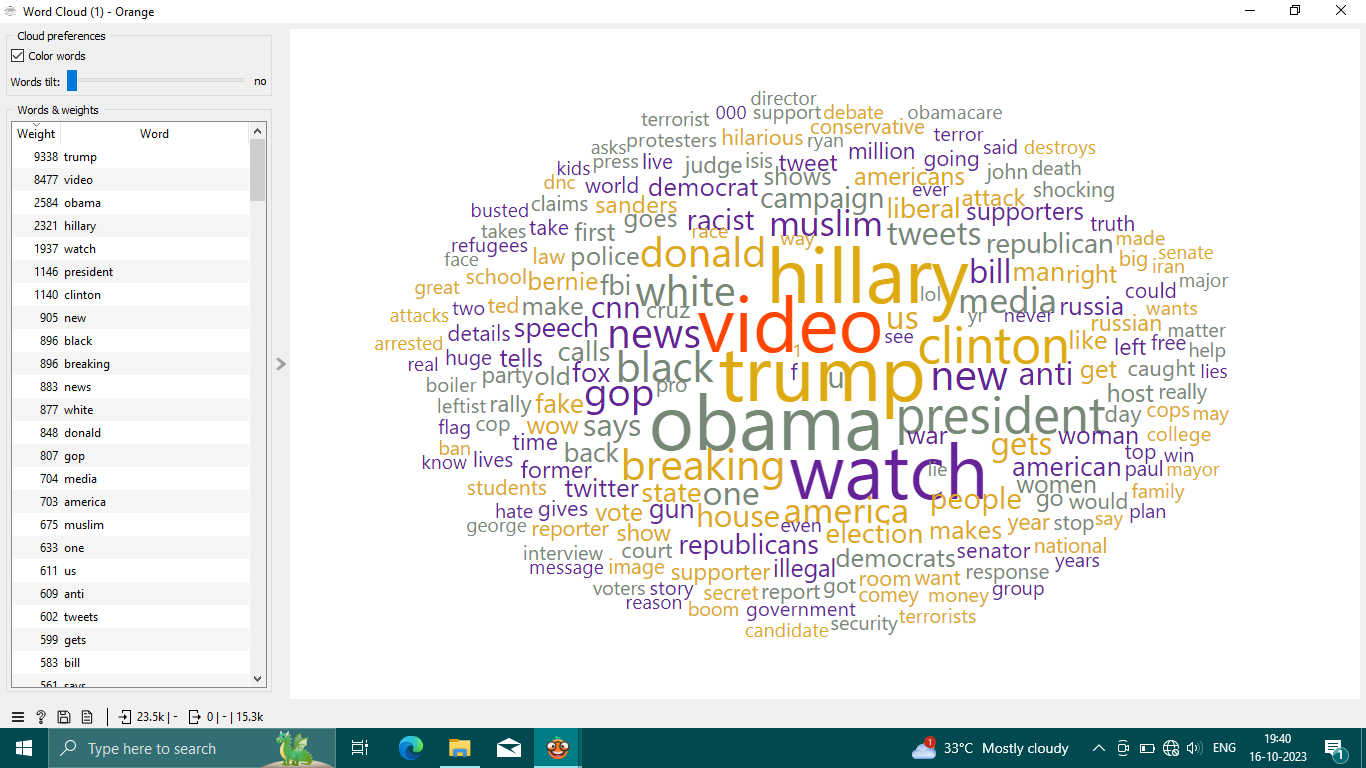
• Evaluation.

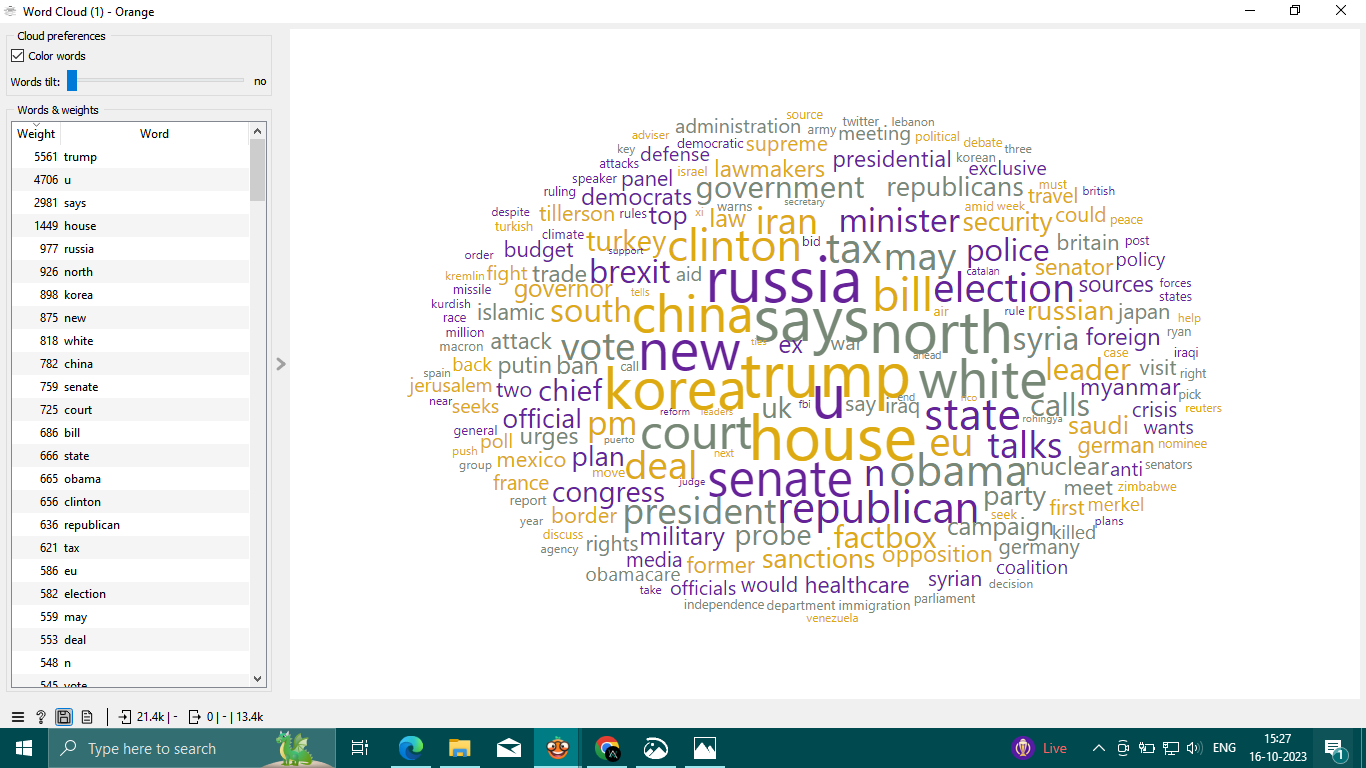
Here's a high-level overview of each step:

Feature Engineering:

Text Preprocessing:

Clean and preprocess the text data, including tasks like tokenization, lowercasing, and removing stop words, punctuation, and HTML tags.



**

Text Representation:

Convert text data into numerical form. Common techniques include TF-IDF (Term Frequency-Inverse Document Frequency) and word embeddings like Word2Vec or GloVe.

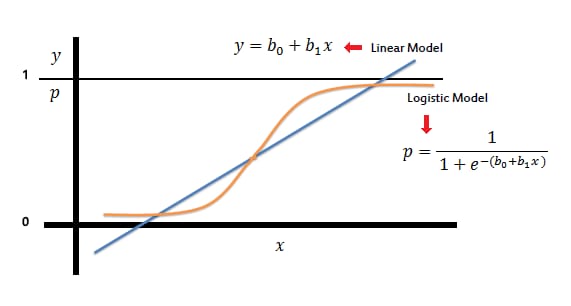
Feature Selection:

Choose relevant features. This can involve using techniques like mutual information or feature importance scores.

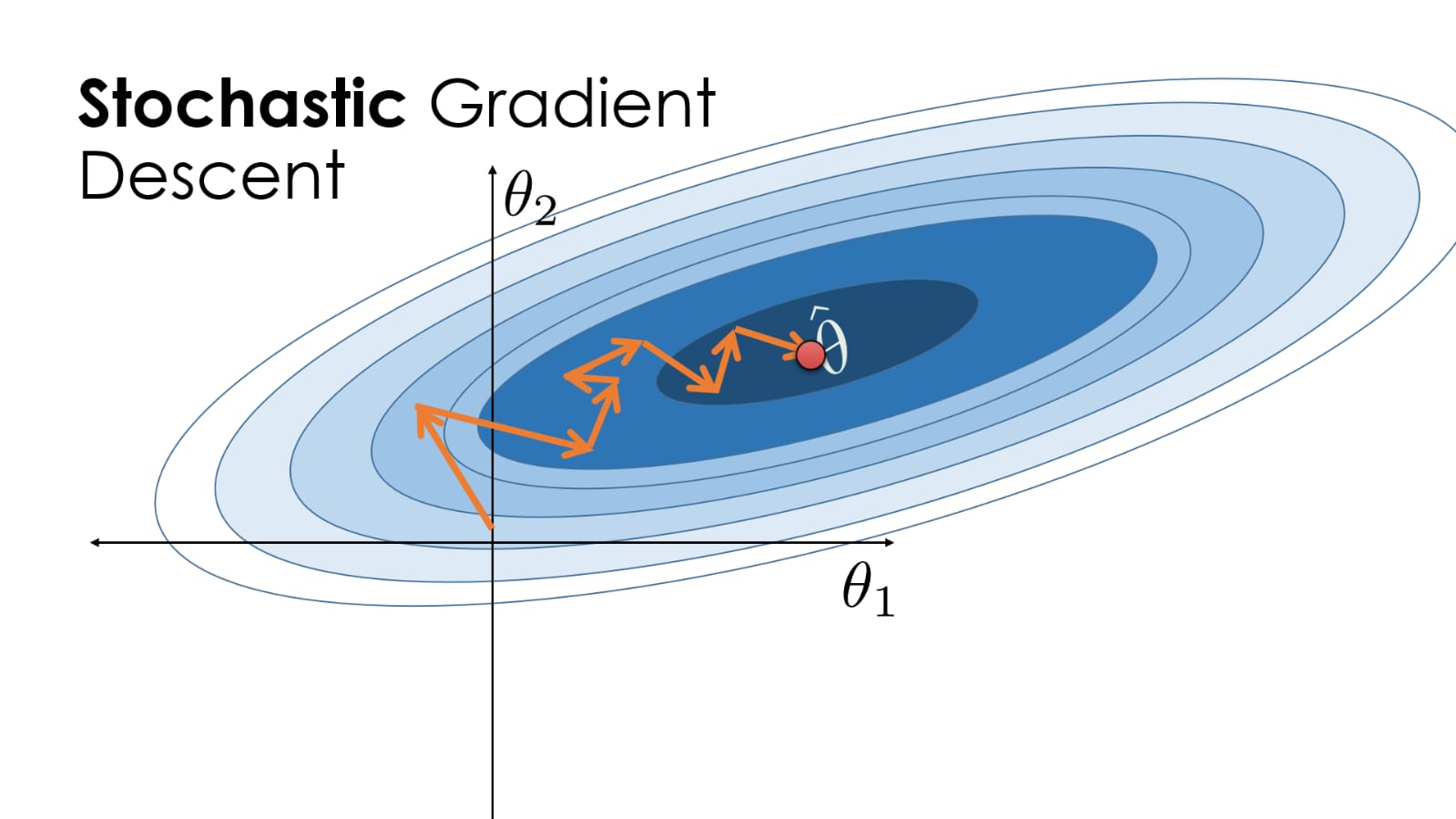
Model Training:

Selecting Models:

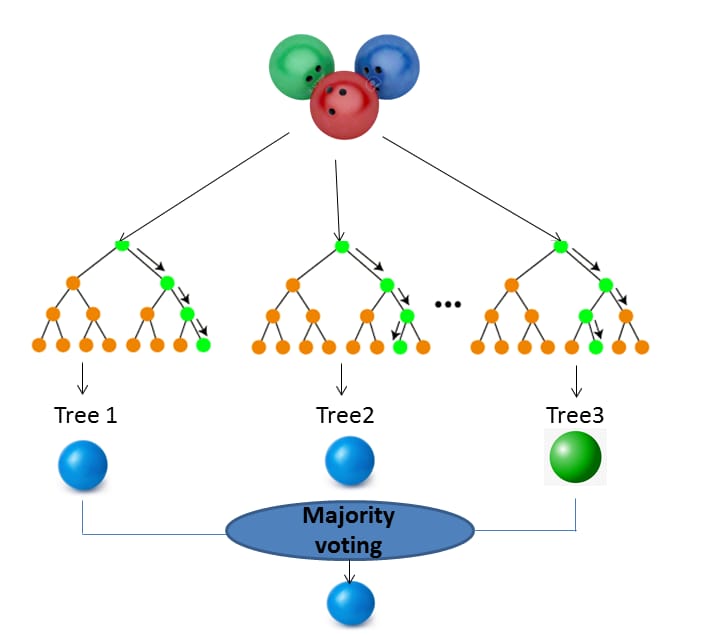
Choose appropriate machine learning or deep learning models for your task. Common choices include logistic regression model, stochastic gradient descent, random forest, support vector machine, naïve bayes.



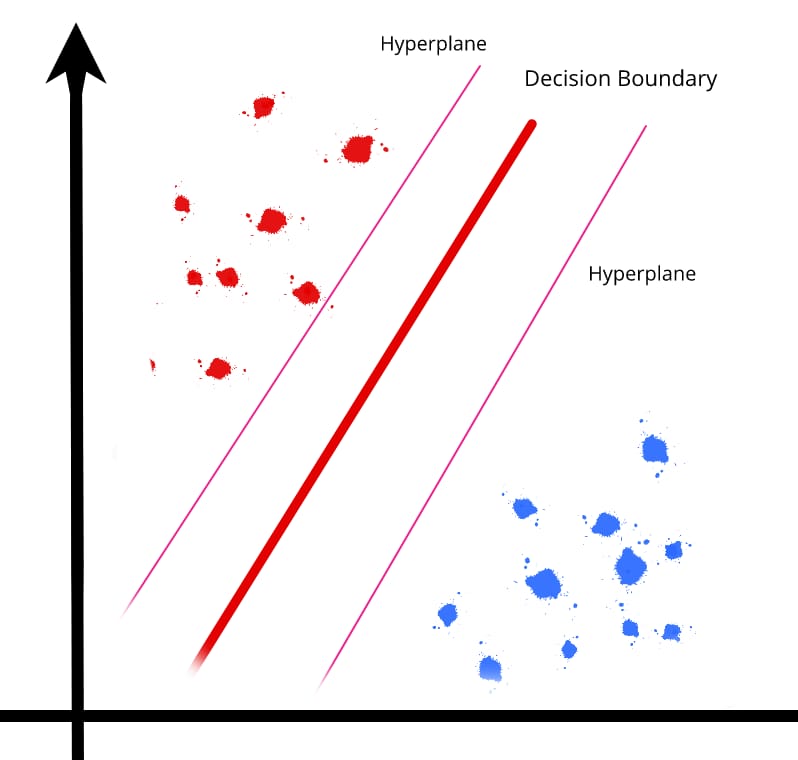
Logistic regression model

**

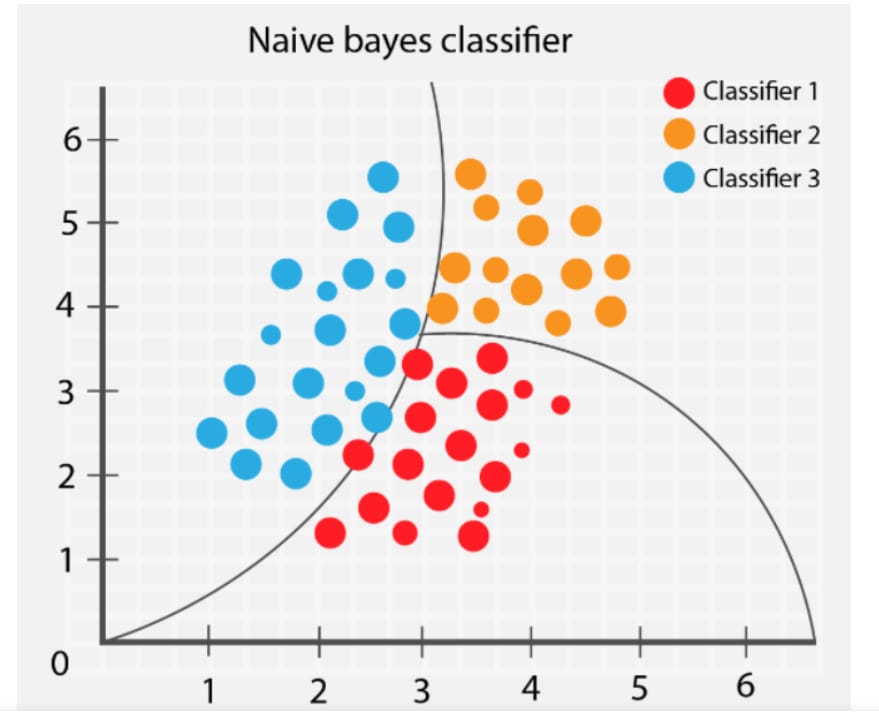
Stochastic gradient descent

**

Random forest

**

Support vector machine

**

Naïve bayes classifier

Data Splitting:

Divide your dataset into training, validation, and test sets to train and evaluate the model.

Model Training:

Train your selected model on the training data using the chosen features and representations.

Hyperparameter Tuning:

Optimize hyperparameters to improve the model's performance using techniques like grid search or random search.

Evaluation:

Metrics:

Choose appropriate evaluation metrics. Common metrics for binary classification problems like fake news detection include accuracy, precision, recall, F1-score, and area under the ROC curve (AUC).

Confusion Matrix:

Analyze the model's performance using a confusion matrix to understand true positives, true negatives, false positives, and false negatives.

Cross-Validation:

Perform k-fold cross-validation to assess the model's stability and generalization.

Bias and Fairness:

Check for bias and fairness issues in your model's predictions to ensure that it's not exhibiting discriminatory behavior.

Ensemble Methods:

Consider using ensemble techniques like bagging or boosting to improve model performance further.

Conclusion:

It is essential to keep in mind the dynamic nature of fake news, and your model may require continuous monitoring and updating as new misinformation techniques evolve. Moreover, the quality and size of your dataset play a crucial role in the success of your fake news detection