- **Yassine Sfaihi**
- www.linkedin.com/in/yassinesfaihi
- www.github.com/yassinesfaihi
- www.kaggle.com/yassinesfaih

[Author] : Yassine Sfaihi

Daily News-Driven Stock Market Prediction through Textual Data Analysis and Machine Learning

Tools Used

- pandas
- seaborn
- matplotlib
- nltk
- sklearn
- gensim
- numpy

Text Preprocessing Techniques

- Removing stop words
- Stemming
- Vectorizing the text data
- Tokenizing the words
- Lemmatizing the words
- Removing punctuation and special characters

Feature Generation Techniques

- Term frequency-inverse document frequency (TF-IDF)
- Word count
- Character count
- Punctuation count

Machine Learning Models Used

- Logistic Regression
- Support Vector Classifier (SVC)
- Random Forest
- Gradient Boosting
- K-Nearest Neighbors (KNN)
- Decision Tree

- **Yassine Sfaihi**
- www.linkedin.com/in/yassinesfaihi
- www.github.com/yassinesfaihi
- www.kaggle.com/yassinesfaih

Model Evaluation Metrics

- Accuracy
- Precision
- Recall
- F1 Score

Hyperparameter Tuning

- Grid Search on SVC

Key Insights

- The impact of text preprocessing and feature generation techniques on the performance of the models
- The importance of hyperparameter tuning in improving the performance of machine learning models