potterblack-ke / **Sentiment-Analysis-Twitter-NLP**

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

0 stars 0 forks 0 watching Branches Activity
Tags

Public repository

1 Branch 0 Tags Go to file t Go to file Add file Code

Mistari254-crypto Update README.md dacb563 · 7 hours ago

data	Add files via upload	8 hours ago
images	Add files via upload	8 hours ago
README.md	Update README.md	7 hours ago
Sentiment-Analysis-NLP-v5.ip...	Add files via upload	8 hours ago
Twitter-Sentiment-Analysis-u...	Add files via upload	8 hours ago

README

Sentiment-Analysis-Twitter-NLP

Sentiment Classification of Tweets Using NLP Team Project by: James Wachira, Tim Musungu, Vivian Kwamboka, Calvin Mutua, Hashim Ibrahim

Project Overview

This project aims to classify tweets about products into Positive, Neutral, or Negative sentiment using Natural Language Processing (NLP) and machine learning techniques.

Businesses can use these insights to:

Understand customer opinions and brand perception.

Improve products and services based on feedback.

Monitor reputation in real-time on social media.

Data

Source: Tweets dataset (tweets.csv) containing:

Raw tweet text

Labeled sentiment (Positive / Neutral / Negative)

Methodology

Text Preprocessing

Tokenization with TweetTokenizer

Conversion to lowercase

Removal of special characters

Stopword removal

Lemmatization with WordNetLemmatizer

Feature Engineering

Vectorization:

CountVectorizer (word counts)

TF-IDF Vectorizer (term importance across corpus)

Optional dimensionality reduction with PCA

Modeling

Models Used:

Multinomial Naïve Bayes

Decision Tree Classifier

Random Forest Classifier

Additional Techniques:

SMOTE for handling imbalanced sentiment classes

Pipelines for clean preprocessing & modeling

GridSearchCV for hyperparameter tuning

Results & Insights

Best performing models:

Random Forest

Naïve Bayes

SMOTE significantly improved results for minority classes.

Preprocessing boosted model performance considerably.

Random Forest provided good balance of accuracy and interpretability.

Business Recommendations

Amplify messaging around topics with positive sentiment.

Proactively engage with negative sentiment tweets to protect brand reputation.

Deploy the model as part of a real-time sentiment monitoring dashboard for continuous insights.

Installation & Requirements

Python Libraries:

bash Copy Edit pip install pandas numpy scikit-learn nltk imbalanced-learn matplotlib seaborn

Repository Structure

kotlin Copy Edit ├── data/ │ └── tweets.csv ├── notebooks/ │ └── twitter-NLP-project.ipynb ├── Sentiment Classification of Tweets Using NLP.pptx └── README.md

Acknowledgments

Thanks to the Moringa School instructors and TMs for their guidance in this NLP project.

License

Releases

No releases published

Create a new release

Packages

No packages published
Publish your first package

Contributors 2


potterblack-ke potterblack


Mistari254-crypto calvin-mutua

Languages

- **Jupyter Notebook** 100.0%