Hotel Review Sentiment Predictor

This application uses a trained Support Vector Machine (SVM) model to classify hotel reviews into one of three sentiment categories: **Positive, Neutral, or Negative**. It is designed to demonstrate a complete machine learning pipeline, from data preparation to deployment using Streamlit.

Features

- **Sentiment Prediction:** Uses a trained SVM model on text features (TF-IDF) to classify reviews.
- **Improved Neutral Classification:** The model was trained using SMOTE to handle class imbalance, specifically targeting better performance on the Neutral category.
- **Clear Output:** Shows the predicted numeric rating and sentiment with descriptive emojis.
- **Streamlit Interface:** Provides a clean, interactive web interface for user input.
- **Graceful Error Handling:** Includes warnings for empty input or loading failures.

Demo

Hotel Review App (Positive)Hotel Review App (Neutral)Hotel Review App (Negative)

Installation

1. **Clone the repository**:
   git clone https://github.com/Rohan-756/ABSA_for_hotel_reviews.git
   cd hotel-review-sentiment

2. **Create and Activate a Virtual Environment** (optional but recommended):
   python -m venv venv

   Linux/macOS

   source venv/bin/activate

   Windows

   venv\Scripts\activate

3. **Install Dependencies**:
   pip install -r requirements.txt

4. **Download the Model**:
   Ensure your final trained model and vectorizer are saved together (e.g., using joblib.dump((vectorizer, svm), "models/final_svm_model_improved.joblib")) and placed in the newly created models/ directory.

Usage

1. **Run the Streamlit App**:
   streamlit run app.py

2. Enter a hotel review in the text area.

3. Click **Predict Sentiment**.

4. View the predicted rating and sentiment in the output box.

Project Structure

```
hotel-review-sentiment/ │
├─ datasets/
│  └─ tripadvisor_hotel_reviews.csv
│
├─ models/
│  └─ final_svm_model_improved.joblib # Trained SVM model + vectorizer (best version)
│  └─ final_svm_model.joblib
│
├─ pynb files/
│  └─ SVM-skewed-data.ipynb
│  └─ SVM_skewed_data_kaggle.ipynb
│  └─ SVM_skewed_data_kaggle_improved.ipynb # Model training file
│  └─ data-testing.ipynb
│
├─ resources/
│  └─ ML-mini-project-poster.pdf
│  └─ ML-mini-project-report.pdf
│
├─ app.py # Main Streamlit app logic
└─ README.md # Project documentation
```

Requirements

Python 3.8+ and the libraries listed in requirements.txt:

pip install -r requirements.txt

How It Works

1. User enters a hotel review in the text area (app.py).
2. The app loads the pre-trained vectorizer and LinearSVC model from the joblib file.
3. The review text is cleaned (tokenized, stopwords removed, lemmatized).
4. The cleaned review is transformed using the loaded vectorizer (TF-IDF).
5. The SVM model predicts the numeric polarity:
    - 1 → Positive
    - 0 → Neutral
    - -1 → Negative
6. The app displays the result in a visually appealing box.

Contributing

Contributions are welcome! Please follow these steps:

1. Fork the repository.
2. Create a new branch: git checkout -b feature-name.
3. Make your changes and commit: git commit -m "Description".
4. Push to branch: git push origin feature-name.
5. Open a pull request.

License

MIT License ©

Acknowledgements

- Built with Streamlit for the web interface.
- Model trained using scikit-learn.