Customer Feedback Sentiment Analysis - NLP Project

1. Project Title

Customer Feedback Sentiment Analysis using NLP and Machine Learning

2. Objective

The objective of this project is to analyze customer feedback data and classify it into sentiments (e.g., Positive/Negative) using Natural Language Processing techniques and Machine Learning models.

3. Dataset Description

The dataset consists of text-based customer feedback. It undergoes preprocessing to remove stopwords, apply stemming, and vectorize using Bag-of-Words via CountVectorizer.

4. Tools & Libraries

- Python
- Pandas, NumPy
- NLTK (stopwords, stemming)
- Scikit-learn
- XGBoost
- WordCloud, Matplotlib, Seaborn

5. Methodology

1. Data Cleaning & Preprocessing
2. Text Normalization (lowercase, remove punctuations)
3. Tokenization & Stopword Removal
4. Stemming using PorterStemmer
5. Vectorization using CountVectorizer
6. Model Training (Random Forest, Decision Tree, XGBoost)
7. Evaluation using Accuracy, Precision, Recall, F1-Score
8. Visualizations like WordCloud and Confusion Matrix

6. Results

All three models were evaluated, with XGBoost providing the best accuracy and performance. The project successfully demonstrates the entire NLP pipeline from preprocessing to evaluation.

7. Conclusion

This project provides a complete workflow for sentiment classification using NLP and ML. It can be improved by using more advanced techniques like TF-IDF, LSTM models, or BERT embeddings.

8. Future Scope

- Add support for multilingual text
- Use deep learning models (e.g., LSTM, BERT)
- Build a real-time sentiment analysis dashboard