CAPSTONE PROJECT

SENTIMENT ANALYSIS OF RESTAURANT REVIEWS

Presented By:

Student Name : Mudgolwad Ravi
College Name & Department : JNTUH COLLEGE OF ENGINEERING SULTANPUR
CSE(AI&ML)

edunet foundation

OUTLINE

- Problem Statement
- Technology used
- Wow factor
- End users
- Result
- Conclusion
- Git-hub Link
- Future scope

PROBLEM STATEMENT

- Analyzing customer reviews of restaurants to determine sentiments as positive or negative.
- This helps businesses identify areas of improvement and strengths based on feedback.
- Efficient sentiment analysis can improve customer satisfaction and loyalty.
- The aim is to automate the process using machine learning techniques

TECHNOLOGY USED

- Programming Language: Python
- IDE: Google Collab
- Libraries:
- - pandas for data handling
- - nltk for text preprocessing
- - scikit-learn for machine learning models

WOW FACTORS

- Comparison of multiple machine learning algorithms: Naive Bayes, Logistic Regression, Random Forest, and SVM.
- Text preprocessing techniques, including stemming and stopword removal.
- Implementation of TF-IDF for vectorizing text data.
- PCA for dimensionality reduction and standardization of features.

END USERS

- Restaurant owners and managers for understanding customer feedback.
- Data analysts for deriving insights from textual data.
- Researchers working on natural language processing.

RESULTS

GaussianNB Accuracy: 0.69

GaussianNB Classification Report:

	precision	recall	f1-score	support
0	0.73	0.56	0.64	96
1	0.67	0.81	0.73	104
accuracy			0.69	200
macro avg	0.70	0.69	0.68	200
weighted avg	0.70	0.69	0.68	200

MultinomialNB Accuracy: 0.745

MultinomialNB Classification Report:

	precision	recall	f1-score	support
0	0.72	0.76	0.74	96
1	0.77	0.73	0.75	104
accuracy			0.74	200
macro avg	0.75	0.75	0.74	200
weighted avg	0.75	0.74	0.75	200

BernoulliNB Accuracy: 0.745

BernoulliNB Classification Report:

	precision	recall	f1-score	support
0	0.73	0.74	0.74	96
1	0.76	0.75	0.75	104
accuracy			0.74	200
macro avg	0.74	0.74	0.74	200
weighted avg	0.75	0.74	0.75	200

Logistic Regression Accuracy: 0.74

Logistic Regression Classification Report:

	precision	recall	f1-score	support
0	0.68	0.85	0.76	96
1	0.82	0.63	0.72	104
accuracy			0.74	200
macro avg	0.75	0.74	0.74	200
weighted avg	0.76	0.74	0.74	200

- Accuracy Comparisons:
- 1. GaussianNB: 69%
- 2. MultinomialNB: 74.5%
- 3. BernoulliNB: 74.5%
- 4. Logistic Regression: 74%
- 5. Random Forest: 70%
- 6. SVM: 74%

- Conclusion: MultinomialNB and BernoulliNB outperform other classifiers in this scenario.

CONCLUSION

- Sentiment analysis effectively identifies positive and negative reviews, aiding restaurants in understanding customer sentiment.
- The best-performing models (MultinomialNB and BernoulliNB) provide a balanced approach to accuracy and generalizability.

GITHUB LINK

- Link: https://github.com/mudgolwadravi/IBM-EduNet-DGT-NLP-AI-.git

FUTURE SCOPE(OPTIONAL)

- Integrating deep learning models such as BERT for better accuracy.
- Expanding the dataset to include multi-language reviews.
- Real-time sentiment analysis for immediate feedback.

THANK YOU

edunet foundation