A

SYNOPSIS

of

MINOR PROJECT

on

Sentiment Analysis For Product Rating

Submitted by

Nitya Vijayvargiya(21EGICA024)

Project Guide
Ms.Charu Kavadia

Head of Department
Dr. Mayank Patel

Geetanjali Institute of Technical Studies, Dabok , Udaipur (Raj.)
Department of Computer Science and Engineering
July,2024

Problem Statement:

Sentiment Analysis for Product Rating

Brief Description:

Online reviews offer a wealth of customer opinions, but star ratings alone provides limited detail. Sentiment analysis can automatically categorise reviews as positive, negative, or neutral, offering valuable insights into customer sentiment. Constructing a Python machine learning model algorithms for sentiment analysis of product reviews using Random Forest approach and determine its effectiveness. The metrics evaluated using metrics like accuracy, precision, recall, and F1-score assess the model's ability to correctly classify reviews.

Objective and Scope:

This project aims to develop a Python-based machine learning model capable of accurately predicting product sentiment. This project is build to analyze existing reviews.

NLP techniques like text cleaning, tokenization, and stemming will prepare the review data. A Random Forest algorithm, trained on labeled reviews (positive, negative, neutral), will learn to classify reviews based on sentiment.

Methodology:

Data Acquisition: Utilize existing customer reviews (text data).

Data Preprocessing: Apply NLP techniques:

- Clean text (remove irrelevant characters).
- Tokenize reviews (split into words).
- Apply stemming (reduce words to root form).

Model Training: Train a Random Forest algorithm on labeled reviews (positive, negative, neutral).

Model Evaluation: Assess performance using accuracy, precision, recall, and F1-score

Hardware and Software Requirements:

Software:

- Platform: Web Browsers(Chrome) to access Google Colab
- Pre-existing customer review dataset (text format).

Hardware:

- Modern multi-core CPU
- Minimum 2GB RAM
- ROM required : 32GB ROM
- SSD for faster data processing and storage

Technologies:

Programming Language: Python (powerful and versatile for machine learning)

Machine Learning: Random Forest Algorithm (robust for sentiment classification)

Natural Language Processing (NLP): Techniques for data cleaning, tokenization, and stemming.

Evaluation Metrics: Accuracy, Precision, Recall, and F1-score (measure model performance).

Testing Techniques:

- **Data Source:** Utilize Amazon product reviews as the dataset for sentiment analysis.
- **Data Collection:** Use Amazon's Product Advertising API or web scraping techniques to gather review data.
- **Data Cleaning:** Implement NLP preprocessing techniques like tokenization, stop-word removal, and lemmatization.
- **Model Testing:** Split data into training and testing sets
- **Evaluation Metrics:** Measure model performance using accuracy, precision, recall, and F1-score.

- **Scalability Testing:** Test model performance on larger datasets to ensure scalability.
- **Deployment Testing:** Verify model functionality in a real-world environment, ensuring it can handle live data and provide accurate predictions.

Project Contribution:

Taking Data from Amazon product reviews.Project Contribution involves actively participating tasks, collaborating with Mentor and delivering assigned responsibilities effectively to achieve project goals, thereby advancing the project's overall success and impact.