E-COMMERCE PRODUCT REVIEW SENTIMENT ANALYZER

AN INTERNSHIP PROJECT

10Alytics

Introduction

The project aims to develop a sentiment analyzer for e-commerce product reviews sourced from Amazon, focusing on the Electronics category. The analysis will involves Data Science covering data acquisition, preprocessing, model development, and deployment.
This project is for 4 weeks.

10Alytics

About Company

Amazon is a globally renowned e-commerce platform that connects consumers with millions of products at competitive prices.
Amazon offers a diverse range of goods, spanning electronics, fashion, home essentials, and more, catering to the needs of customers worldwide. With a focus on affordability, convenience, and reliability, Amazon provides a seamless shopping experience, empowering individuals and businesses to discover, purchase, and sell quality products from trusted sellers across the globe.

10Alytics

Problem Statement

In today's digital age, e-commerce platforms face the challenge of analyzing vast amounts of customer feedback to understand product sentiment accurately. Understanding customer sentiment is crucial for businesses to make informed decisions about product improvements, marketing strategies, and customer satisfaction. However, manually analyzing thousands of product reviews is time-consuming and inefficient. Therefore, there is a need for an automated sentiment analysis solution to process and interpret these reviews efficiently.

10Alytics

10Alytics

Project Objectives

The company aims to develop a robust sentiment analysis system for e-commerce product reviews, specifically targeting the Electronics category on Amazon. By leveraging Data Science techniques, our objective is to build an end-to-end solution that can extract, preprocess, analyze, and visualize customer sentiment from textual reviews. This solution will provide valuable insights to our business stakeholders, empowering them to make data-driven decisions and enhance the overall customer experience.

Data Scientists

10Alytics

1. **Data Acquisition**: Obtain text reviews and star ratings from database, comprising more than 10,000 rows.
2. **Data Preprocessing**:
   - Text Cleaning: Remove noise, special characters, and irrelevant information.
   - Tokenization: Split text into tokens.
3. **Feature Engineering**:
   - Bag of Words Vectors: Convert text data into numerical feature vectors.
   - TF-IDF: Transform text data into TF-IDF matrices.
4. **Model Development**:
   - Utilize pre-trained models (e.g., Vader) or train custom models (e.g., Naive Bayes) for sentiment analysis.
5. **Model Evaluation**:
   - Assess model performance using accuracy and F1 score metrics.
   - Optionally, conduct hyperparameter tuning for improved performance.
6. **Model Deployment**:
   - Set up pipelines for text cleaning, feature engineering, and inference.
   - Deploy the trained model:
     - Locally using a Flask application.
     - Via a cloud service (e.g., Streamlit), with assistance if needed.

Deliverables

- **Fully Functional Sentiment Analyzer**: Develop a reliable sentiment analysis model capable of accurately classifying the sentiment of e-commerce product reviews.
- **Presentation Materials**: Prepare professional presentation slides summarizing the project objectives, methodology, results, and future recommendations for stakeholders and internal teams.
- **Quality Assurance and Testing**: Conduct rigorous testing and validation of the sentiment analysis model and data pipelines to ensure accuracy, reliability, and scalability.
- **Deployment and Integration**: Deploy the sentiment analysis system on both local and cloud-based environments, integrating it with existing company infrastructure and tools for seamless operation.
- **Performance Metrics and Reporting**: Define key performance metrics (e.g., accuracy, F1 score) for evaluating the sentiment analysis model's performance and generate regular reports to track system performance and effectiveness.

You Are Expected To Present A Comprehensive Report Of This Task To The Stakeholders