ASSESSMENT PROPOSAL

Domain of interest

The domain of interest is Sentiment Analysis which focuses on reviews from customers for Amazon Fine Foods.

Problem Definition

To understand and critically analyze the sentiments expressed by customers in their reviews of Amazon Fine Foods. This involves classifying review as positive, negative and neutral to obtain understanding of client satisfaction, opinions and possible potential areas for development.

How AI will fit for solving the problem

Specifically with the use of Machine Learning and Natural Language Processing (NLP) techniques to employ sentiment analysis. These techniques will be used to analyse textual data from customers' reviews. By training a model on labelled dataset of reviews, the AI system can learn to classify new reviews in to the right sentiment categories.

Postulate hypotheses for aiming to solve the domain problem

1. The positive reviews has a correlation with customer satisfaction, higher product ratings and a repeat in purchases.
2. Negative sentiment in reviews may highlight areas for product improvement or customer enhancements.
3. Identifying key topics within reviews can provide insights into customer preferences and concerns.
4. The language use in negative reviews are associated with dissatisfaction and issues.

Think of an agent design which may solve the problem using AI techniques

An NLP model trained on a labeled dataset of customer reviews will make up the agent. The agent will use NLP libraries such as NLTK for sentiment analysis and two techniques. VADER which uses a bag of words approach, and a pretrained model, RoBERTa. Reviews should be able to be categorized by the model as neutral, negative, or positive. Furthermore, an application called Streamlit will be created to serve as an interface for users to submit new reviews and obtain sentiment analysis outcomes.

Aim and Objectives

Aim: Develop a Sentiment Analysis system for customer reviews on Amazon Fine Foods.

Objectives;

1. Collect and process customer review data.
2. Train an NLP model on a labeled dataset of Amazon Fine Foods reviews.

3. Develop a Streamlit application for user input and sentiment analysis.
4. Deploy the application for real-time sentiment analysis

Dataset description (include URL of data source)

This dataset is an Amazon Fine Food Reviews dataset which was originally published on Kaggle by a user named Laowingkin. It consists of reviews of fine foods from Amazon between October 1999 and October 2012. This dataset is commonly used for sentiment analysis, natural language processing tasks and building recommendation systems. It provides a rich resource for exploring sentiment trends over a period of time and understanding customer behaviour in online product reviews.

https://www.kaggle.com/code/laowingkin/amazon-fine-food-review-sentiment-analysis/input