Group Members: Aqdas Zulfiqar 29038
Mahnoor Khurram 29061
Syeda Bisma Urooj Bukhari 30309

Project Proposal

AI-Based Sentiment Analysis System for Product Reviews

1. Introduction

The objective of this project is to develop an AI system capable of analyzing sentiment in product reviews sourced from e-commerce websites. By harnessing cutting-edge AI techniques, the system aims to provide businesses with valuable insights derived from customer opinions, enabling data-driven decision-making processes.

2. System Design and Components

Objective: Design and implement an end-to-end AI system for sentiment analysis of product reviews.

Components:

Data Collection Module: Utilize web scraping techniques to gather product reviews from prominent e-commerce platforms such as Amazon.

Preprocessing Module: Clean and preprocess the raw text data by removing noise, tokenizing text, eliminating stop words, and performing lemmatization.

Feature Extraction: Employ advanced techniques like TF-IDF or Word Embeddings to represent textual data in a meaningful numerical format.

Sentiment Analysis Model: Develop a robust sentiment analysis model using deep learning architectures such as LSTM or BERT to effectively classify the sentiment expressed in the reviews.

User Interface: Design an intuitive web interface that allows users to input product URLs and view sentiment analysis results in an easily interpretable format.

3. Application of AI Techniques and Justification

AI Techniques:

- Deep Learning: LSTM for its ability to model sequential data effectively, capturing contextual information within reviews.
- NLP Libraries: NLTK and spaCy for text preprocessing tasks, ensuring high-quality data for model training.
- Feature Extraction: TF-IDF or Word Embeddings for converting textual data into numerical representations, preserving semantic meaning.
- Framework: TensorFlow or PyTorch for developing and training the sentiment analysis model, providing flexibility and scalability.

Justification:

- **LSTM**: Chosen for its suitability in handling the sequential nature of textual data, enabling the model to capture intricate relationships within reviews.
- **TF-IDF/Word Embeddings**: Selected to transform text data into feature vectors, preserving semantic information and enhancing the model's understanding of the reviews.
- **TensorFlow/PyTorch:** Industry-standard frameworks renowned for their extensive support for deep learning model development, ensuring robust implementation and seamless integration.

4. Validation and Evaluation

Validation Steps:

- Cross-Validation: Split the dataset into training and testing sets, ensuring unbiased evaluation of the model's performance.
- Performance Metrics: Utilize precision, recall, and F1-score to assess the model's accuracy in sentiment classification.

- Misclassification Analysis: Analyze misclassified samples to identify patterns and refine the model accordingly.
- Ground Truth Comparison: Compare predicted sentiments with manually labeled ground truth for a subset of reviews to validate the model's efficacy.

Evaluation Process:

- Train and Test: Train the sentiment analysis model on a diverse dataset of product reviews and evaluate its performance on a separate test set.
- Hyperparameter Tuning: Fine-tune model hyperparameters, including learning rate and dropout rates, to optimize performance.
- Deployment and Feedback: Deploy the system and gather real-time feedback from users to continuously improve the model's accuracy and usability.

5. Project Flow

- Data Collection: Scrape product reviews from e-commerce websites like Amazon.
- Preprocessing: Clean and preprocess the text data to ensure quality and consistency.
- Feature Extraction: Convert textual data into numerical representations using TF-IDF or Word Embeddings.
- Model Training: Train LSTM or BERT model on labeled data to learn sentiment patterns.
- Validation: Validate the model's performance using cross-validation techniques and performance metrics.
- Evaluation: Measure the system's accuracy, precision, recall, and F1-score to assess its effectiveness.
- Deployment: Develop a user-friendly web interface for real-time sentiment analysis of product reviews.
- Feedback Loop: Continuously gather user feedback to refine and enhance the model's performance over time.

6. Conclusion

This AI-based sentiment analysis system offers businesses invaluable insights derived from customer feedback, empowering them to make informed decisions and enhance customer satisfaction. By leveraging state-of-the-art AI techniques and meticulous validation processes, the system ensures reliable and accurate sentiment analysis of product reviews, thereby fulfilling the project's objectives effectively.