Product Description and Review Analysis using NLP

Overview

This project focuses on analyzing and optimizing e-commerce product descriptions and customer reviews using Natural Language Processing (NLP) techniques. The aim is to enhance the understanding and presentation of product features by summarizing product descriptions, extracting relevant keywords, and classifying customer reviews for sentiment analysis. This can help businesses improve product discovery and customer experience in e-commerce platforms.

Features

- **Product Description Analysis**: Automatically processes product descriptions to highlight key features.
- **Customer Review Analysis**: Analyzes customer reviews to identify sentiment (positive/negative), trends, and potential areas for improvement.
- **Text Summarization**: Extracts concise and relevant key product features from lengthy descriptions.
- **Feature Extraction**: Uses techniques like TF-IDF and POS tagging to identify the most important features in product descriptions.
- **Sentiment Classification**: Applies machine learning models to classify customer reviews and predict their sentiment.
- **Visualization**: Creates word clouds and other visualizations to represent key terms and insights.

Dataset

The dataset used for this analysis is the **Amazon Product Dataset**, which includes the following columns:

- **Product Name**: The name of the product.
- **Product URL**: A link to the product page on Amazon.
- **Product Image**: The URL of the product image.
- **Total Number of Reviews**: The total number of customer reviews.
- **Product Description**: A detailed description of the product.
- **Features**: Specific features or specifications of the product.
- **Customer Reviews**: Text of customer feedback or reviews.

Dependencies

- Python 3
- pandas
- numpy
- scikit-learn
- matplotlib
- wordcloud
- textblob

Setup

1. Clone this repository or download the notebook file.

Install the necessary dependencies:
Copy code
pip install pandas numpy scikit-learn matplotlib wordcloud textblob

2. Run the Jupyter notebook
   (`PRODUCT_DESCRIPTION_AND_REVIEW_ANALYSIS.ipynb`) to start the analysis
   process.

Usage

The notebook includes the following steps:

1. **Data Preprocessing**: Cleaning and preparing the data for analysis (tokenization, lemmatization, stopword removal).
2. **Text Feature Extraction**: Applying TF-IDF and POS tagging to extract key features.
3. **Sentiment Analysis**: Using machine learning models to classify customer reviews as positive or negative.
4. **Visualization**: Generating word clouds and plotting key terms.

Results

The project outputs visualizations, such as word clouds, and statistical reports, including classification metrics (accuracy, precision, recall, and F1-score) to evaluate the model's performance.

Future Work

- **Voice-Activated Search Optimization**: Adapt the product descriptions for better compatibility with voice search.
- **Multimodal Data Integration**: Combine textual data with images and videos for a richer analysis.
- **Cross-Language Support**: Expand the model to support multiple languages for global e-commerce platforms.

Google colab link:- ∞ PRODUCT DESCRIPTION AND REVIEW ANALYSIS