

Amazon Pet Supplies Product Reviews Sentiment Analysis using BERT and RoBERTa models

Project Proposal

Amanda (Man Kuei) Chen
MS in Business Analytics

Introduction

As a business person with a marketing background, I plan to use NLP techniques to address common marketing problems. I will be using CRISP-DM framework to perform this project.

Business problem understanding

Businesses appreciate receiving customer feedback to enhance their operations and boost profits. However, when confronted with a substantial volume of reviews, how can businesses efficiently identify negative feedback and promptly resolve customer problems? Using sentiment analysis in customer reviews automatically categorizes and prioritizes customer needs.

Data understanding

Amazon product reviews by pet supplies category.

Data source: <https://cseweb.ucsd.edu/~jmcauley/pdfs/emnlp19a.pdf>

Data preparation

NLP packages: NLTK for pre-processing; spaCy for task like NER to extract product names; Transformers library from Hugging Face for BERT and RoBERTa models; Scikit-learn for metrics.

NLP tasks: EDA analysis like term frequency analysis; Named entity recognition to identify product's name; Sentiment analysis for analyzing the customer's reviews and classify them into positive/negative/neutral opinions.

Modeling

BERT and RoBERTa models for comparison

Evaluation

Using Accuracy and F1 score to see how well the model perform

Deployment

Generate a final report and presentation.

Tentative schedule

Week 1: Data collection, cleaning, preprocessing and EDA analysis

Week 2: Train and fine-tune the models for sentiment analysis

Week 3: Evaluate the models and select the best performance

Week 4: Finalize report and presentation.