Task: Sentiment Analysis of Product Reviews

Model:

Distributor: Hugging Face

https://huggingface.co/fabriceyhc/bert-base-uncased-amazon\_polarity

Chosen model: **fabriceyhc-bert-base-uncased-amazon-polarity**

I chose this fabriceyhc-bert-base-uncased-amazon-polaritybecause it was finetuned on Amazon reviews, meaning it should be capable of performing sentiment analysis on product reviews with little to no additional training. It makes use of BERT, which is good for sentiment analysis, and it achieved high accuracy and low loss based on its evaluations. While it is only capable of binary classification of positive and negative, it is good enough for my needs.

<https://nlp.stanford.edu/sentiment/index.html>

Managing the model:

Descriptive tags were added to the model.

Solution Development:

Because the model accepts input in a {“input” : [list\_of\_text]} format, I only needed to extract the text form the input csv file and put them into a list. Azure handled the tokenization of the input. After I tested the model with the processed data, I copied the test results, which were in JSON format, onto a new json file.

Accuracy score: 0.9795918367346939

Future improvements:

Finding a dataset with more descriptive review to use as test input