Project Proposal

Team B

DSCI-644 Fall 2020

Team Members:

Matt Agone
Heather Akers-Healy
Susan George
Jacob Miller
Alex Pataky

Product Purpose:

This product will improve an existing neural network algorithm. The model currently uses sentiment analysis to classify Amazon reviews. The current project will retool this model, potentially making changes to the algorithm with the goal of improving the accuracy of the model, which is currently at a baseline of 60%.

Preliminary Plan:

The performance of the model is not very good, only accurately predicting sentiment 60% of the time. We will improve performance by:

- Revisiting the requirements and formulating the specification
- Architecting and remodeling the system
- Conducting research and experimenting with various models through Azure Machine Learning Studio
- Optimizing preprocessing and features extraction operations
- Verifying the results, evaluating the accuracy and comparing the performance of the new model
- Implementing the model

The project website will be hosted on GitHub. It will host information on collaboration and documentation on requirements, goals, and architecture. Reporting on progress will be done through periodic reports. Requirements, task assignments and deadlines will be tracked in Trello. The final product will be showcased in a presentation. We will also collaborate regularly over Slack.

Motivation:

Customer opinion is vital in drawing useful insights, improving products and services, making informed decisions and taking actionable steps. Sentiment analysis categorizes the unstructured data from customer perceptions driving branding and marketing. An automated solution for categorizing Amazon reviews by sentiment will allow greater flexibility for our customers in utilizing reviews, ultimately making it easier for consumers to glean product information without reading excessive amounts of reviews.

Key Requirements:

(from the viewpoint of users)

- The sentiment analysis model improves upon the accuracy of the existing version
- The new model is trained on the same data inputs as the existing model to allow for direct comparison
- The sentiment analysis model learns from the Amazon application review data
- The sentiment analysis model takes a text review as input and provides the positive/negative class as the output
- A fast, reliable and easy-to-use web-based tool