Mythri Partha

Capstone 3: Amazon Review Prediction Project Proposal

Searching for products online is an effective way to gauge how useful a product is to a consumer; they can see reviews and ratings to see what they want from a product and from there they can make a more informed purchase. When a person wants to buy a product online, the first thing they might do is check out the reviews, so that they can decide whether they feel the product they’re looking for would be a good purchase or not. I want to create a model where we analyze at amazon reviews and ratings, create a machine learning method that can see based on the words used as well as the star rating if the review is positive, negative, or neutral, then from there, I want make a prediction model to see if we can predict the rating of a random product.

For this project, I plan to use the supervised and unsupervised modeling methods I learned previously in the course as well as NLP (natural language processing) to observe the reviews and analyze them. My clients are companies who want to predict how their products are being received on amazon. Also, my data and models can be used by consumers who are hesitant to make a purchase before reviews are posted and may want to know if those products would be a good investment for them.

So far, I have found some data of Amazon product reviews on data.world <https://data.world/datafiniti/consumer-reviews-of-amazon-products>, I plan on using NLP to look at the type of the product, the words used in the review, the star ratings, the title of the review, and the number of people who found the review helpful to gauge how the product was received.

My final deliverables will be a completed prediction model, a slide deck presentation, and a report along with images of distributions of data and relationships I find.