RECOMMENDATIONS TO AMAZON'S GROCERY & GOURMET FOODS

Introduction

Large wholesale companies like Amazon with millions of unique products in their inventory always demand more sales and better products. In this project we wish to give Amazon an insight to their *Groceries and Gourmet Foods*. The aim is to find products with high ratings and analyse if they have specific features which contribute to their high rating. If so, these features can then be recommended to poorly rated products and thus improve their ratings. As Amazon has a large variety of products, we have narrowed the analysis down to a single category, so we are able to compare the features of the products. We focus the analysis on the category *Candy & Chocolate* since this category has variation in average ratings and hence products that need improvement. The analysis is carried out using various machine learning techniques like topic modelling, regression analysis and machine learning interpretability.

Methods and main findings

In this section we will present the main findings of our analysis. Our analysis is based on a fitted regression model predicting the average rating of the products.

Some of the explanatory features of the regression model are extracted from the product descriptions using the topic modelling method Latent Dirichlet Analysis (LDA). It was found that few topics had maximum likelihood. This indicates that the product descriptions not surprisingly contain similar words since they are all in the category *Candy & Chocolate*. Keeping in mind the result from the grid search, the final number of topics were chosen according to the coherence of the topics and how much the topics affected the predictions. We figured out that it was an iterative method to find topics with high coherency, requiring continuously adding stop words to the text processing step.

We fitted 7 popular regression models and compared their MAE. The best model turned out to be the gradient boosting regressor. The model was interpreted by looking at which features contributed to predicting the average rating both globally and locally. Globally we learned that only one of the LDA topics had some significance in the model, which was a topic that negatively influenced the average rating prediction. However it was only very little significance and we did not find any topics with features that improved the average ratings of a product.

The product improvement analysis turned out to have limited takeaways and thus only providing sparse recommendations to Amazon. Thus we changed our perspective to analyse how the product features including the description affect the sales rank of the product. In this analysis, the description was included by its length and sentiment. Furthermore, we included the length of the title. From the interpretation of the model it was found that products often bought or viewed together with other products are selling more which is not surprising. A more interesting insight was found regarding the product descriptions. It appeared that no or very short product descriptions have a negative impact on the product sales and a longer and elaborate description is preferred. Furthermore, the model interpretation showed that extremely positive product descriptions decreases the amount of sales and products with more neutral descriptions tend to sell better.

Conclusion and recommendations

In this project we wanted to find specific features in high rated *Candy & Chocolate* products and recommend Amazon to use these features on some of their poorly rated products. However through our analysis we were only able to find one topic from the LDA with some impact on the product rating. This topic contained the words *lecithin*, *butter*, *syrup*, *(palm) oil*, *fat* and *emulsifier* and had a negative impact on the product rating. Due to a vague relationship between the topic and the average rating, we can only alert Amazon to avoid these ingredients or using them in the product descriptions. In retrospect, it was a difficult task to extract key product features from the product descriptions using topic modelling. Thus we changed the view of our analysis to focus on how to improve the sales rank of the products. From this analysis we can recommend Amazon to write elaborate product descriptions, since an elaborate description showed to have a positive impact on the amount of sales. Furthermore, our findings showed that the product descriptions are best when neutral and not overselling by using extremely positive words or too many positive words.

Generally it has been difficult if not impossible to validate our recommendations. If we were to validate e.g. removing the negatively impacted words from the product description, we would need the average rating of the product after this removal. However, this data has not been accessible and thus the validation can first be made if Amazon chooses to implement our recommendations. The same holds for the sales rank analysis, where the validation of the impact of the description length and sentiment on the sales rank requires these changes to be applied.