**Prediction of customer ratings from reviews using text classification with Transformer models**

**Abstract**

Online user ratings are important to improve service of a company. They are also instrumental in helping other users to decide which product to buy for their desired price range and help them understand the performance of the product better through the opinion of other users. Normally, a customer writes a review to a product from different aspects and then makes an overall rating. The problem is some customers give rating in contrast with their comments. The other reviewers must read many comments and comprehensive the comments that are different from the rating. We propose the analysis and prediction rating from customer reviews, which is in text using Transformer based neural network architecture [1]. Transformer models are based on self-attention mechanism that we believe to be particularly well suited for language understanding and the problem in hand [2].

We have a novel dataset which is real and targeted to customers in Singapore. The dataset is scraped from [Seedly](https://seedly.sg/), a community platform where customers can post their reviews as per product. Since the data scraped from Seedly is in small size, we shall first train the model using the [Yelp](https://www.yelp.com/dataset/challenge) dataset [3], which is public. We shall compare the performance measures of our model with the already existing models for prediction.

**Methodology**

*Data Acquisition*: We plan to use a scraper-based API to generate the dataset from <https://seedly.sg/>

*Data Exploration*: We shall use visualization tools like matplotlib, ggplot and pandas to visualize a sample of the dataset to make sure that the data is consistent or if skewed, to be aware of when we look at the model accuracy.

*Model Design*:

References

[1] A. Vaswani, “Attention Is All You Need arXiv:1706.03762v5,” no. Nips, 2017.

[2] J. Alammar, “The Illustrated Transformer – Jay Alammar – Visualizing machine learning one concept at a time,” *github*, 2018. [Online]. Available: http://jalammar.github.io/illustrated-transformer/. [Accessed: 07-Oct-2019].

[3] M. Tran, “Predicting Product Ratings From Review Text.” Kaggle.