Introduction

The goal of this report is to compare the two models for sentiment analysis on the "playstation" dataset. The first model used a Naive Baye classifier while the second model applied the CNN neural network to determine a review of its products: Playstation 4 and Xbox, part of sentence classification problem. The training of highly performant deep learning model CNN has been enabled due to the availability of computation power and large datasets.

It is worth mentioning the setup when making train-split in both models are the same as random_state = 42 and test_size = 0.2 and thus the comparison is not biased. It is also noted that the data for Xbox is not large enough (only 29 negative reviews and 83 positive reviews) to be trained using CNN and therefore the deep learning method only trained the Playstation 4 data.

Model Comparison

1. Performance Comparison

The model using Naive Baye classifier have test accuracy of 0.806, F1 Score of 0.887, and ROC AUC score of 0.590 while those figures in CNN model are 0.888, 0.926, 0.909 respectively . From those measure metrics, we can see that the model built from CNN algorithm performs better than Naive Baye classification model.

This is because the CNN architecture used pre-trained word embeddings, increasing the accuracy of the text classification model. Using pre-trained word embeddings is advantageous because we can customize word embedding to train on millions of words. Another reason is the pooling operation to create pooling layer greatly reduces the number of weights in the network and, thus, prevents overfitting of the network. CNN as one of the deep learning models learn hundreds of thousands of parameters (weights and biases). Knowing the total numbers of learnable parameters helps to determine the required sample size (number of training data points) to avoid overfitting.

2. Computation Power requirement

With the current setup and the same laptop configuration , the CPU and Wall times to train with CNN method are required longer than Naive Baye classification model (almost 21 minutes CPU times and 11 minutes Wall times).

Though there are benefits of using CNN method in sentence classification task, obtaining such a huge dataset would be a challenge since training a large dataset re quire a lot of computation power.

3. Applicability in real-world

For the particular matter of sentiment analysis , Naïve Bayes classifier have a very serious assumption, which is all the features used are mutually independent so that each feature contributes independently to the probability of the class. In the real world, it is almost not true, especially for sentiment classification unless the features are handcrafted manually.

Model Suggestion

Based on the model evaluation and comparison between both models it is suggested that the company should use one of the deep learning methods: RNN or in the scope of this assignment CNN.
In the next move, instead of choosing classifier, it makes more sense to choose how to generate features, some models that consider word order like RNN or bag-of-words to take word counted into consideration could come out better results.