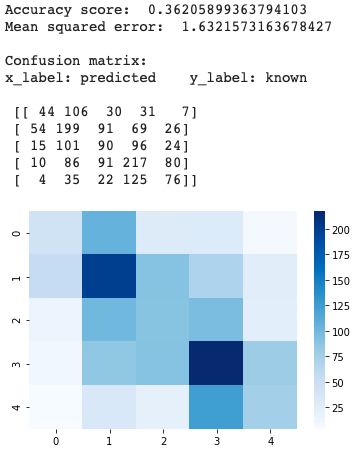
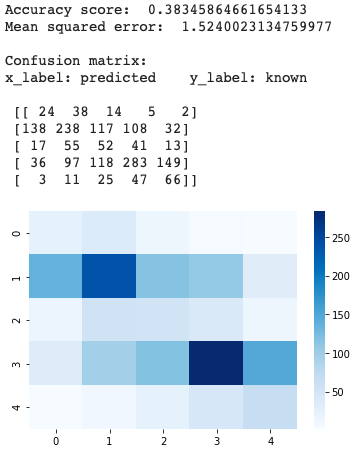
**Deliverable 3**

260860011 Takuya Ishii

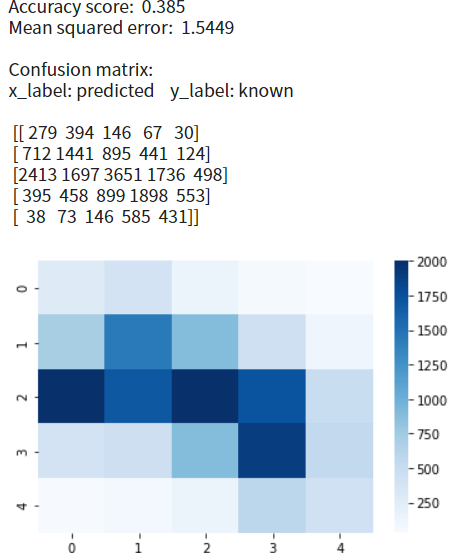
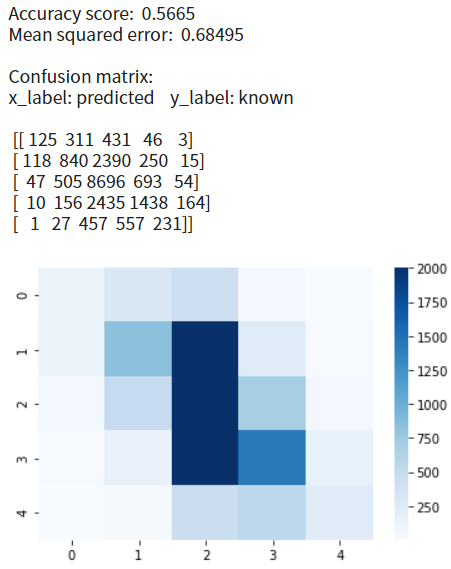
1. **Final Training Result**  
   First, let us recall that the following result was obtained in the preliminary phase (Figure 1).  
     
     
   In the previous analysis, I suspected that excessive reduction of the training data caused these low accuracy scores. So in the final training phase, I first tried to conserve all entries from the original dataset upon preprocessing. This change forced me to reduce the size of the bag of word from 10000 to 3000 because of the memory size limitation of Google Colab (12GB). Despite the effort to increase the accuracy, this modification revealed the imbalance of class-labels with “neutral(2)” being the most frequent, which resulted in the accuracy score lower than expected (*Figure 2*, left). To compensate this imbalance, I initially adopted the complement Naïve Bayes model. However, although it canceled the imbalance in the prediction to a certain degree, the accuracy score didn’t particularly improve (*Figure 2*, right).  
     
     
     
     
     
   As a next approach, I manually removed some “neutral(2)” instances and decreased their number from 79582 to 30000. This improved the accuracy score of the multinomial model from 0.38 to 0.47. Considering this change and the squared matrix pictured, I’d like to use this classifier in the demo (*Figure 3*, left).

**Multinomial Naïve Bayes**

**Logistic Regression**



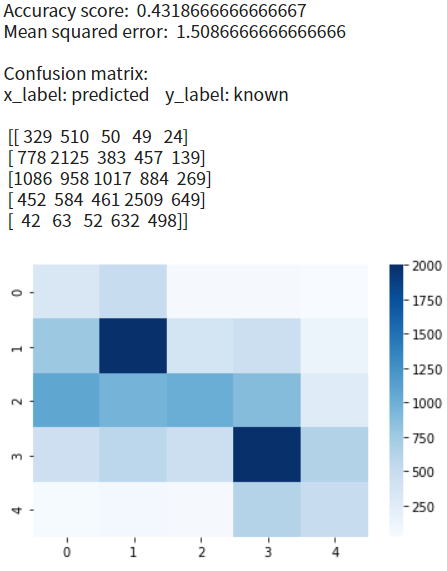
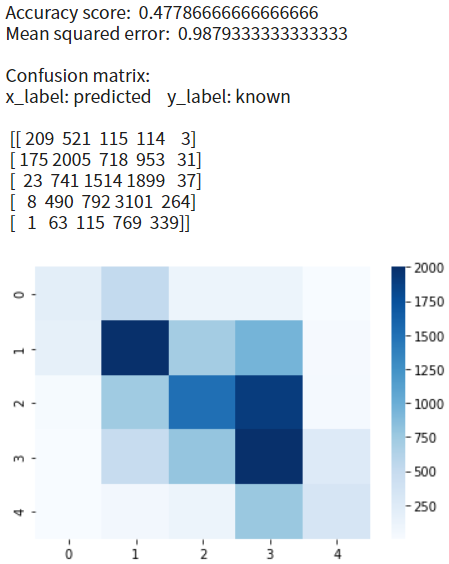
*Figure 1 Preliminary Result*



**Multinomial Naïve Bayes**

**Complement Naïve Bayes**

*Figure 2*



**Multinomial Naïve Bayes**

*Figure 3 Final Result*

**Complement Naïve Bayes**

1. **Final demonstration proposal**My goal is to create a simple webapp where users can type their movie reviews and see their sentiment scores predicted. Since I don’t have any previous experience in web development, I would like to keep it as minimal as possible using Flask. So I will try to replicate the web app explained in the following link and may add some modifications in CSS to improve its user interface, as time permits:  
   <https://medium.com/analytics-vidhya/embedding-sentiment-analysis-model-into-a-web-application-93b76ab6348c>