**Text Classification Using Deep Learning (BERT)**

The analysis of customer reviews is a crucial component of a company’s reputation and as technology evolves, a widening array of tools and methodologies can be used to categorise these reviews. Among these, the application of deep neural network learning models. Businesses can automatically classify customer reviews as positive or negative based on the sentiment expressed in the text.

Data Cleaning

From Data Visualisation, the reviews were classified into 5 different categories but were needed to be organised into ‘helpful and ‘unhelpful’ which are two labels. So ‘3’ as a review out of 5 would be an unbiased rating, which could be unhelpful to the classification process. Data removal of 3 for a biased and identical split using masking and data dropping was implemented.

Data Preprocessing

The first steps regarding data preprocessing were to remove the special characters, HTML tags and punctuation to create readable text for the model. Regarding network selection, BERT the transformer-based model was used to capture the input sequences. Using the NLTK library, stop words were also cleaned out of the text.

Model fitting

As shown in Figure 1, the features were adding into the class to then be called when the model is fit. BERT is an attention-based transformer model, so it inherently includes an embedding layer in its architecture so embedding was explicitly used. Three dense layers were used to customise the model, where 512 units with relu activation by batch normalisation and dropout, then 128 units with relu activation as well as normalisation and dropout and one unit of binary classification with sigmoid classification. Learning rate was also used as an optimizer and early stopping as a call back (Figure 2) to further prevent overfitting in the model.

Overall, the accuracy score came to 0.89 which is captured in Figure 3. Meticulous utilisation of all available features to address overfitting was used to get this to the closest degree which is shown in Figure 4. As all features were exhausted to manage overfitting, there may have been other aspects within data processing that could’ve been further investigated. However, this gave a reliable performance within the real-world scenario of customer reviews.

Appendix

A screen shot of a computer program

Description automatically generated

Figure 1 – class created for BERT function

A screenshot of a computer

Description automatically generated

Figure 2 – model fitting

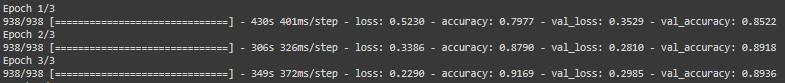


Figure 3 – training outcome

A graph of a graph of a model loss

Description automatically generated with medium confidence

Figure 4 final output of test & train graph