Report for Assignment 4

1. **Introduction**

    This report presents the results of exploring the performance of different embedding layers for a given task. The aim is to determine the impact of embedding layer and vocabulary size on model accuracy.

2. **Methodology**

    By trying different embedding layers and vocabulary sizes, the exploration involves training and testing several models. Ultimately, the models were evaluated based on their accuracy, and the results were recorded.

3. **My Findings and Analysis**

    My findings are illustrated in below table:

Word embeddings	No. of training sample	Test accuracy
Original embedding layer	25000	87%
pretrained layer	100	67%
hypertuned embedding layer 1	1000	63%
hypertuned embedding layer 2	900	66%
hypertuned embedding layer 3	700	53%

The result of my exploration demonstrates that the embedding layer works better than the pretrained word embedding, since the former one achieved a higher accuracy of 87%, while the latter only achieved an accuracy of 67%.

My findings suggest that the choice of embedding layer and its vocabulary size can significantly impact model accuracy. The high accuracy achieved by the large embedding layer indicates that a larger vocabulary may be necessary for certain tasks.

However, hypertuning with too large or too small vocabulary sizes can negatively impact performance, as demonstrated by the significant drop in accuracy observed with a vocabulary size of 700.

4. Limitations and Future Work

One limitation of this exploration is that it only considered a single task. Further research is needed to determine whether these findings hold true for other tasks. Additionally, while the investigation identified the best performing embedding layer and vocabulary size, there may be other factors that could further improve model accuracy.

5. Conclusion

In conclusion, this exploration proves the importance of carefully selecting the appropriate embedding layer and vocabulary size for a given task. The findings suggest that a larger vocabulary size may lead to higher accuracy, but hypertuning with too large or too small vocabulary sizes can negatively impact performance. It is of vital importance to optimize model accuracy with careful consideration and experimentation. Future research should explore other factors that could further improve model performance.