This assignment compares the performance of two text representation approaches for sentiment classification on the IMDB dataset:

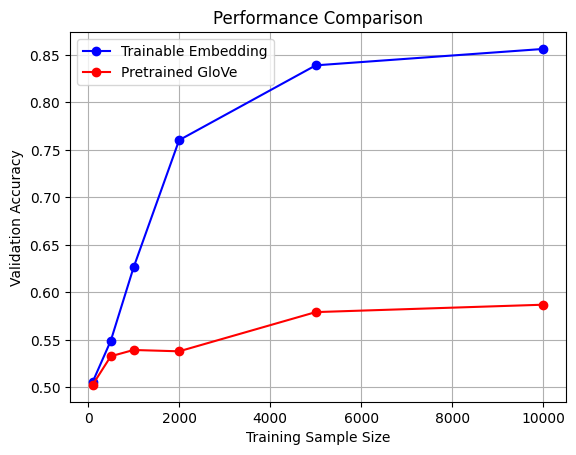
* Trainable Embedding Layer
* Pretrained Glove Embedding

**Experiment Setup**

* Only the top 10,000 most frequent words are used.
* Reviews are truncated after 150 words.
* Training sample sizes vary: 100, 500, 1000, 2000, 5000, and 10000.
* Validation is done on 10,000 unseen samples.
* Both models are trained for 10 epochs with the same architecture (1 dense output layer).

**Result**

|  |  |  |
| --- | --- | --- |
| Training Samples | Trainable Embedding | Pretrained GloVe |
| 100 | 0.505 | 0.502 |
| 500 | 0.549 | 0.533 |
| 1000 | 0.626 | 0.538 |
| 2000 | 0.760 | 0.538 |
| 5000 | 0.839 | 0.579 |
| 10000 | 0.856 | 0.587 |



**Conclusion**

The trainable embeddings have a better performance compared with the pretrained Glove embeddings at every sample size. The performance gap widens as more data is available. This indicates that task-specific embeddings learned during training are more effective for sentiment classification in this context.