

ASSIGNMENT 4 Summary

The table below presents the outcomes after refining the parameters of the IMDB class example and conducting training across varying sample sizes ranging from 100 to 200, in increments of 25.

Sample Size	Performance on Test Set	
	Layer Trained from Scratch	Pretrained GloVe Layer
100	49.8	50.9
125	50.2	50.3
150	50	50.8
175	59.8	51.8
200	63	52.6

For smaller sample sizes (100 to 150), the pretrained GloVe layer consistently provides a slight improvement in performance over the layer trained from scratch. This suggests that when the available training data is limited, leveraging a pretrained embedding can harness the broader context and associations learned from larger datasets the GloVe model was originally trained on.

As the sample size increases, the model with the layer trained from scratch starts to outperform the GloVe pretrained layer. Specifically, at a sample size of 175, the model trained from scratch jumps to 59.8% performance, showing a significant improvement compared to the GloVe layer's 51.8%.

At the sample size of 200, the performance of the model trained from scratch continues to improve and reaches 63%, while the GloVe pretrained layer performance improves only slightly to 52.6%. This further improvement for the scratch-trained model could indicate that given more task-specific data, the model is able to learn embeddings that are more relevant and effective for the specific task compared to the more general pretrained GloVe embeddings.

The trend in the data suggests that with an increasing amount of task-specific training data, the benefit of training the embedding layer from scratch grows, possibly because it allows the model to fine-tune the word representations closely to the nuances of the target task. This could be due to the embeddings trained from scratch being more specialized, while pretrained embeddings are more general.