ADVANCED MACHINE LEARNING

ASSIGNMENT-03

INTRODUCTION:

Recurrent Neural Networks (RNNs) possess the capability to retain an internal state or memory, enabling them to sequentially process input sequences by considering the preceding elements. Consequently, RNNs prove to be a valuable tool in handling sequential data, including text and speech. Furthermore, recurrent neural networks (RNNs) have effectively tackled a diverse set of natural language processing (NLP) tasks, including language modelling, machine translation, sentiment analysis, and named entity recognition.

Table: Performance Results

Sample size	**One hot encoded sequence**		**Embedded Layer**		**Embedded Masked**		**Pre-Trained**	
	Test Loss	**Test Accuracy**	**Test Loss**	**Test Accuracy**	**Test Loss**	**Test Accuracy**	**Test Loss**	**Test Accuracy**
Initial	0.3096	0.880	0.3973	0.858	0.3016	0.878	0.2982	0.873
100	0.4422	0.804	0.4779	0.775	0.4476	0.797	0.4799	0.767
500	0.3558	0.847	0.3782	0.833	0.3680	0.840	0.3730	0.833
1000	0.3893	0.836	0.3705	0.833	0.3733	0.839	0.3765	0.839
2000	0.3587	0.846	0.3806	0.838	0.3836	0.831	0.3672	0.837
4000	0.4063	0.833	0.3966	0.832	0.3760	0.840	0.3708	0.833
8000	0.3528	0.846	0.3751	0.835	0.3726	0.837	0.3688	0.839

Summary of the results:

- The table presents a comparative analysis of the performance of a neural network model on a certain task, considering different sample sizes and embedding methods. The evaluation encompassed four distinct forms of embeddings, namely: one hot encoded sequence, embedded layer, embedded masked, and pre-trained embeddings.
- The first results indicate that the model employed a sequence embedding technique using single hot encoding, resulting in a test accuracy of 0.880. Nevertheless, when the sample size was increased to 100, all embeddings exhibited a decline in test accuracy and an increase in test loss, except for the embedded layer, embedded masked, and pre-trained embeddings, which demonstrated similar outcomes.
- The test loss demonstrated a decrease across all four embedding ways as the sample size was augmented to 500. Among these approaches, the embedded layer exhibited the lowest test loss value of 0.3782, accompanied by the greatest test accuracy rate of 0.833. The results were consistent when examining the sample. The size was augmented to 1000, wherever the inserted layer exhibited the most minimal test loss of 0.3705 and the highest test accuracy of 0.833.

- The embedded layer and embedded masked demonstrated superior performance compared to the other embeddings as the sample size was increased to 2000 and 4000, respectively. When the sample size was increased to 8000, all embeddings exhibited similar test loss and accuracy. Specifically, the embedded layer demonstrated test losses of 0.3751 and test accuracies of 0.835.

Conclusion:

Based on the data presented in the table, several inferences can be inferred.

The preliminary findings indicate that the one hot encoded sequence embedding achieved the lowest test loss of 0.3096 and the maximum test accuracy of 0.880.

As the sample size was progressively raised from 100 to 500 and 1000, it was constantly observed that the embedded layer exhibited superior performance, characterised by the lowest test loss and maximum test accuracy.

In the case of bigger sample numbers, specifically 2000, 4000, and 8000, it was seen that the embedded layer and embedded masked had superior performance in comparison to the remaining embeddings.

In general, the embedded layer exhibited constant performance across various sample sizes, but the other embeddings displayed varied outcomes.

When the sample size reached a sufficient magnitude, the performance of all embeddings exhibited convergence.

Hence, the utilisation of the embedded layer is the preferred approach for this recurrent neural network (RNN) model, as it continuously shown superior performance across various sample sizes.

In order to enhance the performance of the model, it is advisable to utilise a larger dataset for training purposes. Furthermore, it is crucial to select the cutoff duration and number of training samples with careful consideration of the specific requirements and limits of the job or application under consideration.

In order to determine the optimal model parameters, it may be necessary to conduct further testing and engage in a comprehensive evaluation process. In general, it is feasible to train models with a reduced amount of data. However, it is crucial to acknowledge potential constraints and implement additional strategies, such as data augmentation, transfer learning, or utilisation of pre-trained models, in order to enhance the model's efficacy.