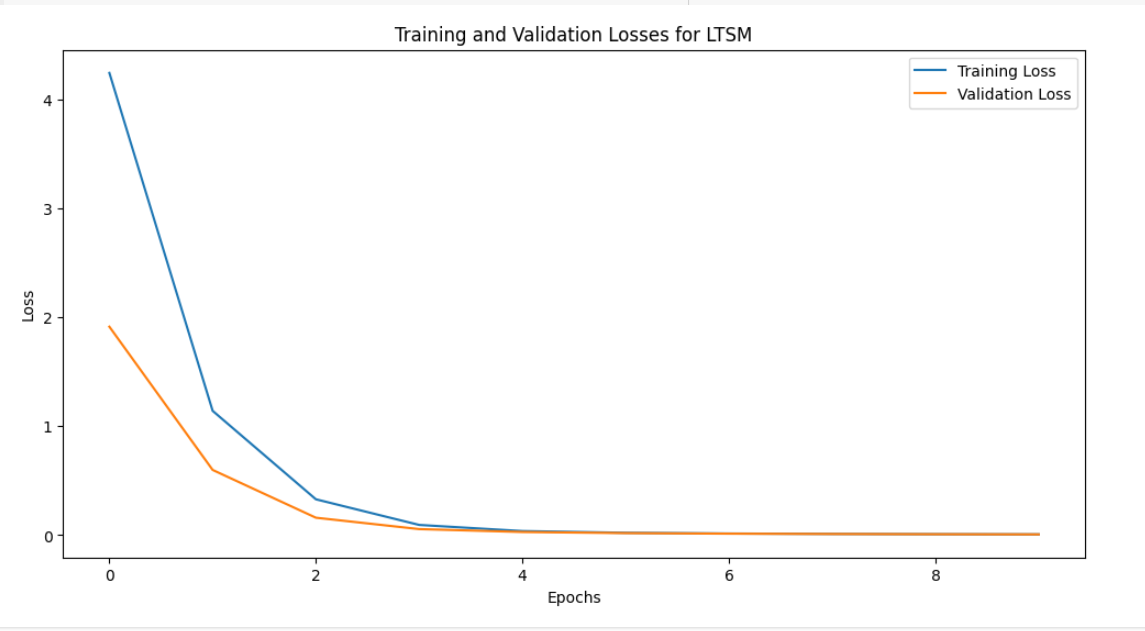
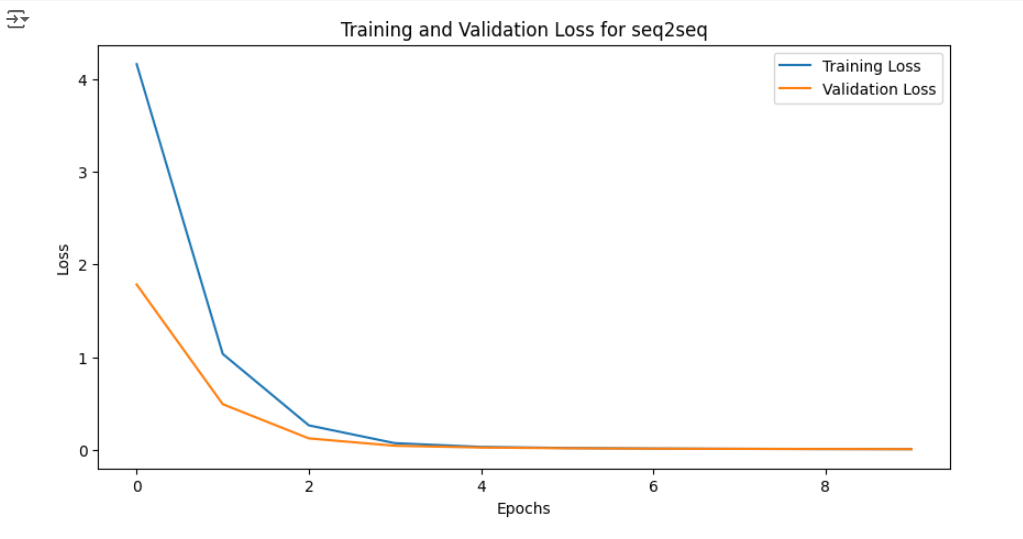
**LTSM Training Curve Plot**



**Seq2seqTraining Curve Plot**



**Observation and Conclusion**

In comparing the performance of the LSTM-based and Seq2Seq models over 10 epochs, there are subtle but notable differences in training and validation losses that highlight the strengths of each model. During the initial epoch, both models started with similar training and validation losses, with the LSTM model showing a Training Loss of 4.25 and a Validation Loss of 1.92, while the Seq2Seq model began with a Training Loss of 4.16 and a Validation Loss of 1.78. These values suggest that both models are reasonably well-initialized for the dataset; however, the Seq2Seq model’s marginally lower losses imply it may be slightly more adapted to the data from the outset.

As training progressed into the middle epochs (Epochs 2-5), both models showed a rapid decrease in training and validation losses, which is a sign that they were effectively learning and generalizing to the validation data. By Epoch 5, the LSTM model achieved a Training Loss of 0.04 and a Validation Loss of 0.03, while the Seq2Seq model recorded a Training Loss of 0.03 and a Validation Loss of 0.02. These middle-epoch results further reflect the Seq2Seq model’s slight edge in learning efficiency and generalization, as it consistently maintained lower losses than the LSTM model. This trend in favor of the Seq2Seq model continued, albeit by a small margin.

In the later epochs (Epochs 6-10), both models continued to converge well, achieving notably low losses. By the final epoch, the LSTM model’s Training Loss reached 0.008 with a Validation Loss of 0.008, while the Seq2Seq model recorded a Training Loss of 0.007 and a Validation Loss of 0.0067. This slight advantage in the Seq2Seq model’s validation loss indicates that it may generalize marginally better on unseen data, suggesting a small but meaningful improvement over the LSTM model in terms of generalization.

While both the LSTM-based and Seq2Seq models demonstrate effective learning and convergence, the Seq2Seq model consistently outperforms the LSTM model across training and validation losses. This advantage, though small, suggests that for this specific dataset and task, the Seq2Seq architecture is slightly more effective in capturing the underlying data patterns and is likely to offer better generalization on unseen data. In summary, while the LSTM model performs well, the Seq2Seq model’s lower losses and efficient learning indicate it is the stronger performer for this translation task.