

STAT6207 Assignment 2

a. Basic

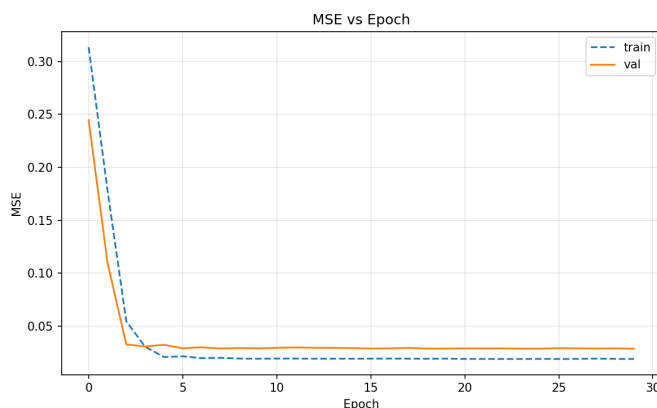


Figure 1: Result of basic.py

The plot from basic.py is a plot of mean-squared error (MSE) vs epochs, where training data and validation data are represented by dashed and solid line respectively.

It is observed that the training MSE decreases during the early epochs and stabilizes afterwards till the end, indicating convergence and consistent learning of the LSTM. The validation curve did not increase at the end. The gap between training and validation maintained is modest, suggesting there is limited overfitting.

b. Advanced

The plot illustrates the performance of a RNN model, with different hyperparameter. In this case, a fixed num of epochs at 15 and two layers are used, while we varied and evaluated the performance of using hidden size 16, 32 and 64.

The plot from advanced.py is also a plot of MSE vs epochs, again training data and validation data are represented by dashed and solid line respectively. The colors of plot represent hidden sizes of 16, 32 and 64 epochs respectively.

Hidden size: 16

Training (dashed): Drops quickly at the start but plateaus sooner at a higher MSE level than larger models, suggesting limited capacity.

Validation (solid): Mirrors the early drop, stabilizes early, and remains above the training curve with a small, steady gap. Overall indicates mild underfitting but very stable generalization.

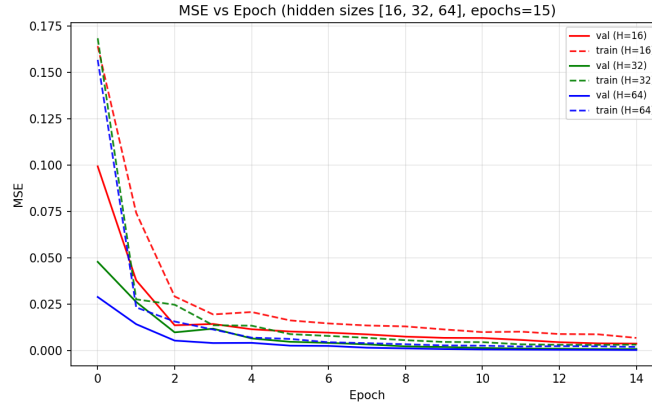


Figure 2: Result of advanced.py

Hidden size: 32

Training (dashed): Rapid initial decline and a lower plateau than H=16, showing the model captures more structure.

Validation (solid): Tracks the training curve closely with a small gap; converges to a noticeably lower MSE than H=16. Good bias–variance trade-off and strong generalization with minimal overfitting.

Hidden size: 64

Training (dashed): Fastest decrease and the lowest training MSE among the three, as expected with higher capacity.

Validation (solid): Continues improving beyond the smaller models and settles at the lowest validation MSE. The train–val gap is slightly larger than H=32, hinting at mild overfitting, but it remains controlled.

Overall summary

All three settings show a steep MSE drop in the first few epochs followed by a gentle plateau by ~epoch 10–15.

Increasing hidden size consistently lowers the final validation MSE, with diminishing returns from 32 to 64.

H=32 offers a strong accuracy/efficiency balance; H=64 typically achieves the best final validation MSE with a small increase in overfitting risk.