# Recurrent Neural Network Analysis
# Assignment Report

The experiment conducts fair, apples-to-apples comparison between two models with identical 100-dimensional architectures. The only difference is the source of their embedding weights.

Model A (Scratch): An Embedding layer (100 dimensions) with 1,034,113 trainable parameters.

Model B (GloVe): An Embedding layer (100 dimensions) with its 1,000,000 embedding parameters frozen. It has only 34,113 trainable parameters.

Approach that works better with 100 training samples:

In this experiment, the "Embedding model" performed better than the pre-trained GloVe model.

Test Accuracy (Scratch, 100 samples): 0.6245
Test Accuracy (GloVe, 100 samples): 0.5764

This result is counter-intuitive but provides a perfect illustration of a core machine learning concept: high variance and small dataset induced overfitting.

Theory suggests the GloVe model should win because it has pre-trained knowledge. However, the "Scratch" model won because it has more trainable parameters and it used this massive capacity to memorize and overfit to the tiny 100-sample training set.

The log provides clear evidence of this. The "Scratch" model's training accuracy (in Epoch 20) hit 1.0000, meaning it achieved 100% perfection on the 100 training samples. Though its validation accuracy was wildly unstable, 0.6311 being the best achieved.

In this specific run, this "lucky" overfit model just happened to generalize better to the test set, achieving 62.45%. The GloVe model, with 97% fewer trainable parameters, was far less flexible and could not overfit in this way, resulting in a more stable but lower score of 57.64%.

Point at which the embedding layer give better performance:

The script found the crossover point 250 samples.

The script's logic (if val_acc_scratch > val_acc_glove: break) executed perfectly. Because the "Scratch" model had a higher validation accuracy, the script correctly identified this as the crossover point and stopped the experiment. This is not a meaningful crossover but a statistical one. It confirms that at 250 samples, the "Scratch" model's performance is unreliable, high-variance, and in this specific instance, it "lucked" into a higher score.

Training with small sample sizes can lead to overfitting and thus provide irregular results. Although the embedding dimension was changed to equalize the Scratch model with the Glove model, the sheer number advantage in training parameters made sure that the Scratch outperform the Glove even though the test accuracy was low for both.