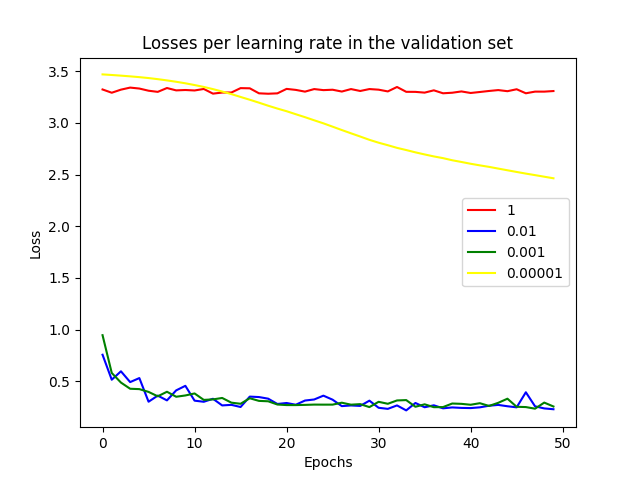
**IML Project 4 Report - Nir Ellor**

** 6.1.2.1:**

The best learning rate in the experiment is **0.01**, as it

Convergence fast with a low, stable loss.

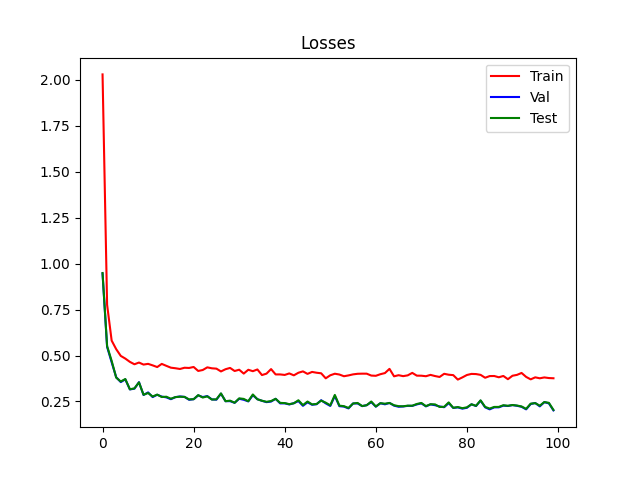
A learning rate of 1 remains consistently high (~3.5)

without proper convergence. A learning rate of 0.001

performs well but converges more slowly compared to

0.01, while 0.00001 is too low, results in very slow learning

and relatively high loss after 50 epochs. A too high learning

**** rate leads to overshooting the optimal solution, causing instability and potential divergence. Conversely, a too low learning rate slows down convergence, making training inefficient and increasing the risk of not converging.

**6.1.2.2:**

The train loss stabilizes around 0.4 by ~60 epochs. While

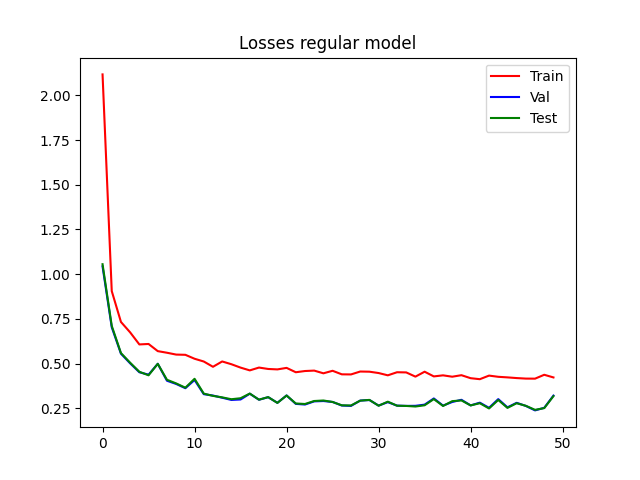
validation and test losses stabilize around 0.25 by ~20 epochs,

indicating slight underfitting. With too many epochs, we might

not find any improvement in the loss. Moreover, the model risks

overfitting to the training dataset, results in low training loss

but high validation /test loss. With too few epochs, the model suffers from lack of expressiveness, failing to learn important patterns, leading to high losses on both train, validation and test sets.

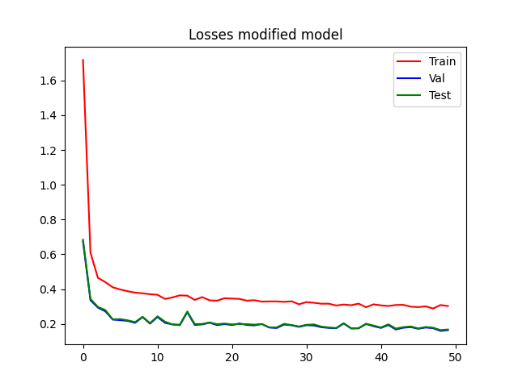


**6.1.2.3:**

Batch normalization helps by normalizing the activations of each hidden

layer, which stabilizes the learning process and accelerates convergence.

In the modified model, the train loss decreases more rapidly and stabilizes

at a lower value (~0.3–0.4) compared to the regular model (~0.5). The validation

and test losses in both models are similar (~0.2–0.25), but the modified

model shows less fluctuation, indicating improved stability and

generalization.

**6.1.2.4:**

**(i):**

**(ii):**

**(iii):**

**6.2.1.1:**

**6.2.1.2:**

**6.2.1.3:**

**6.2.1.4:**

**6.2.1.5:**

**6.2.1.6:**

**6.2.1.7:**

**7.6.1:**

**7.6.2**: