Name : Nayanathara P.M.C.

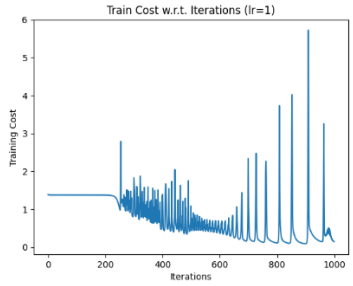
Index No : 210417X

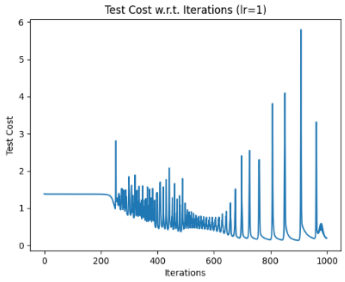
CS3630 : Deep Neural Networks

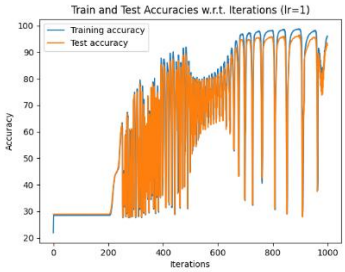
Assignment 01 : Back Propagation

Report : Analysing the influence of the learning rate on the test accuracy of a neural network.

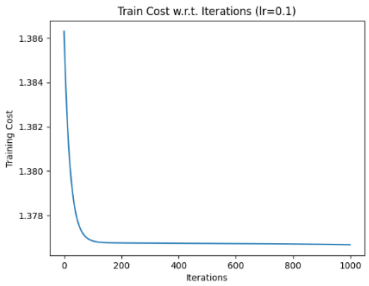
1. Learning rate = 1

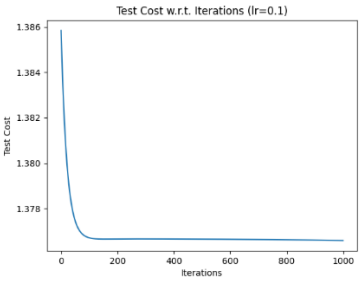


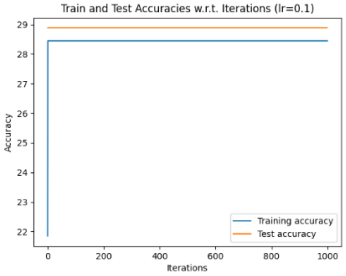




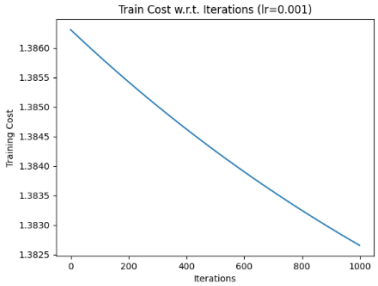
2. Learning rate = 0.1

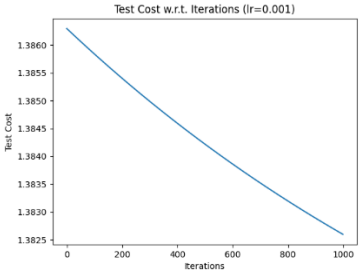


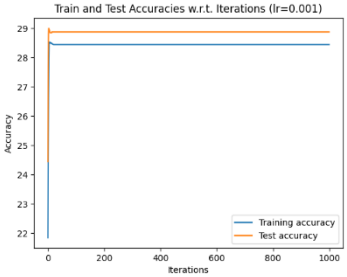




3. Learning rate = 0.001







- **Learning Rate = 1**: The training and testing costs fluctuate wildly showing an erratic behaviour, and the model fails to converge since it takes large steps during gradient descent. Both training and test accuracy are unstable, indicating that this learning rate is too high and causes the model to overshoot the optimal solution. Overall, the high learning rate leads to poor model performance, with low test accuracy and unstable training behaviour. The model struggles to generalize well due to unstable training process.

- **Learning Rate = 0.1**: The model shows more stable and gradual reduction in both training and test costs, converging smoothly. Training and test accuracy steadily increase, making this the optimal learning rate. The gap between training and test costs become smaller, which is a representative of balanced learning. The model balances fast learning with good generalization to the test set.

- **Learning Rate = 0.001**: The gradient descent steps are small, leading to a much slower convergence. The model learns very slowly. Both training and test costs decrease gradually, but the convergence is too slow. Accuracy increases, but not as much as with higher learning rates, indicating the learning process is too slow for practical use. Has a lower chance of overfitting but fails to generalize effectively within the given iterations due to the slow learning process.

Conclusion

- It is important to select an appropriate learning rate to get better accuracy.
- A high learning rate can lead to instability.
- A low learning rate results in slow learning.
- A moderate learning rate provides a good balance ensuring stable convergence and high accuracy.
- Baes on the above graphs, the learning rate of 0.1 seems to be the optimal choice since it offers a better trade-off between convergence speed and generalization.