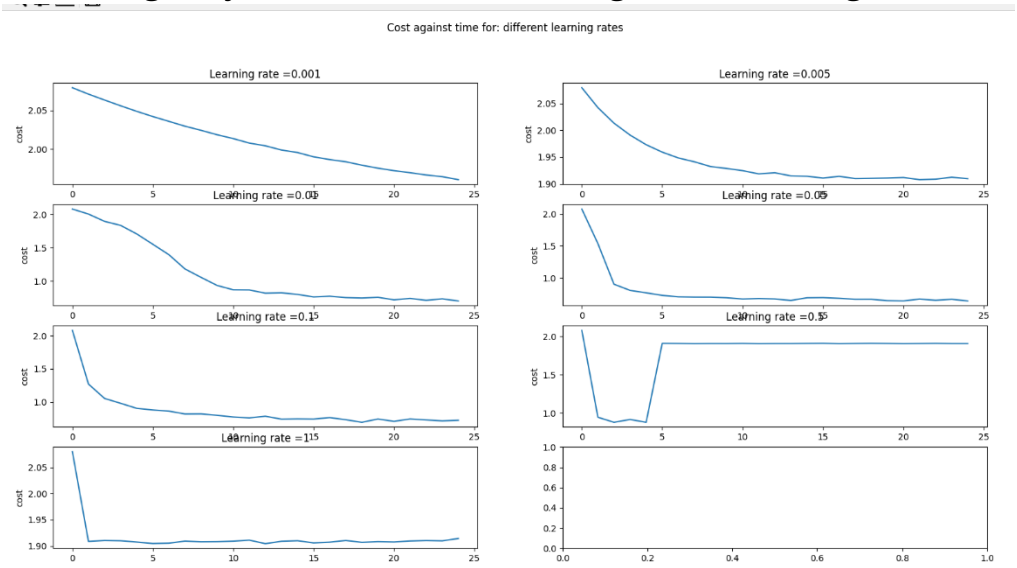
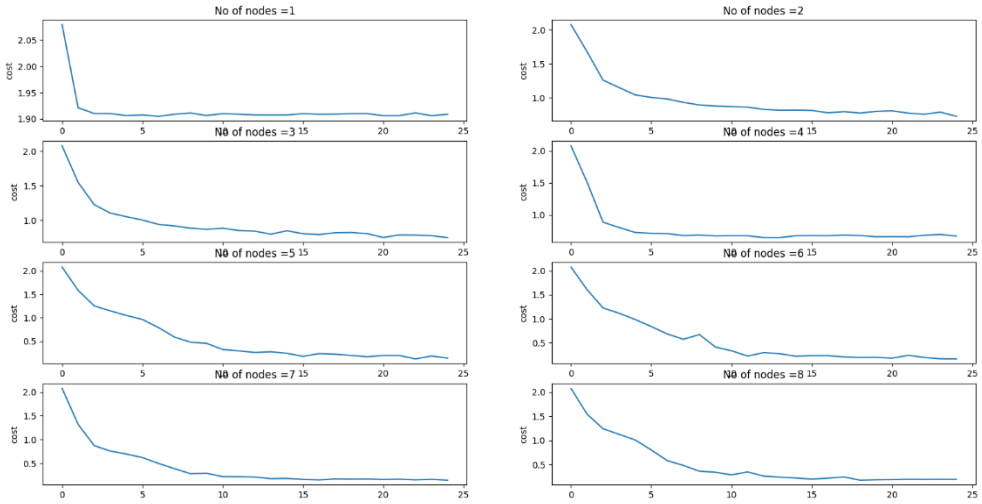
Assignment #5 Neural Networks

Name : Mohammed Almotasem Sec:2 Bn:19

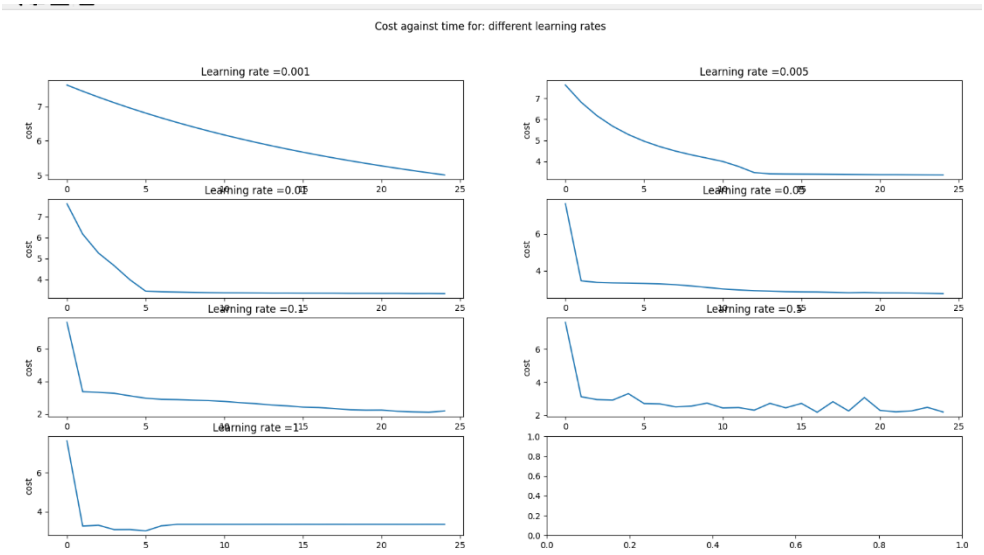
- Results:
  1. By using different learning rates and a fixed no of nodes in the single layer on the Iris dataset we get the following



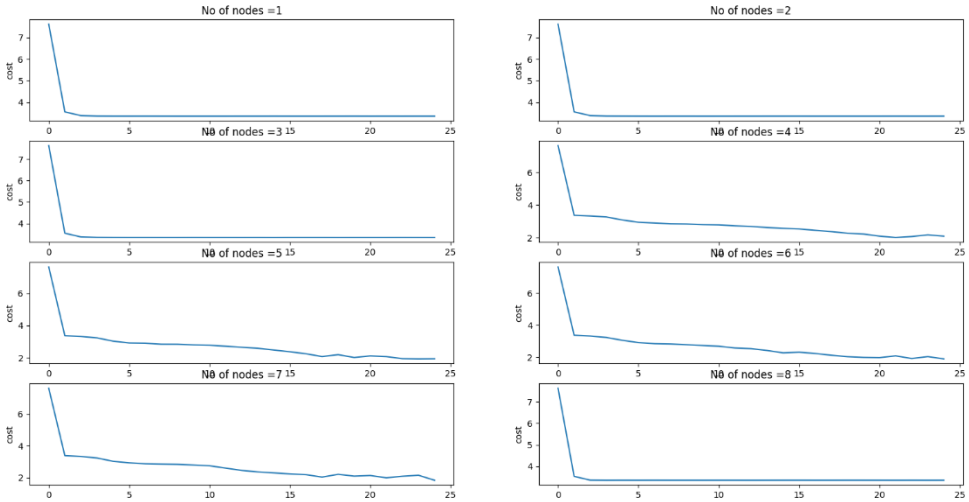
  2. By using the best learning rate found and different no of nodes we get the following



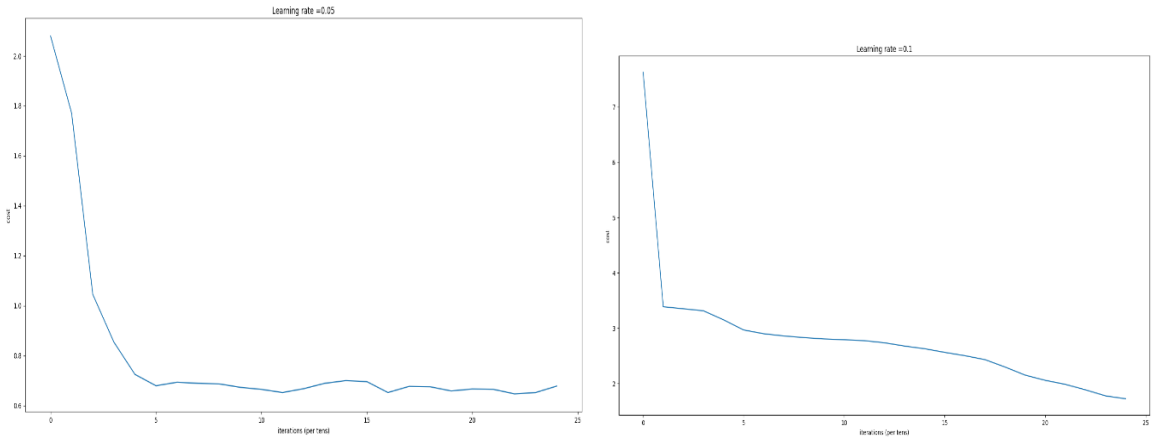
  3. By using different learning rates, 6-nodes for the first layer and 4-nodes in the second layer on the Vowels dataset we get the following



4. By using the best learning rate found and different no of nodes for both layers we get the following



5. By repeating the previous experiments with applying the momentum term we get the following



- Discuss:

1. From results 1,3 we deduce that using different learning rates will affect the performance of the model where Smaller learning rates require more training epochs given the smaller changes made to the weights each update, whereas larger learning rates result in rapid changes and require fewer training epochs. Also, a learning rate that is too large can cause the model to converge too quickly to a suboptimal solution, whereas a learning rate that is too small can cause the process to get stuck.
2. From results 2,4 we deduce that using different no of nodes in the hidden layers of the network will affect the performance of the model where Using too few neurons in the hidden layers will result in underfitting. While using too many neurons in the hidden layers can result in several problems. First, too many neurons in the hidden layers may result in overfitting. Second too many neurons will cause the model to consume more time training
3. By comparing the best no of nodes with that calculated from the heuristic value we get that:
    a. Iris dataset:
        i. Best no of nodes found:5
        ii. Nodes calculated = 150/(10*(4+3)) = 2
    b. Vowels dataset:
        i. Best no of nodes found:7
        ii. Nodes calculated = 990/(10*(12+11)) = 5
4. From result 5 we deduce that using the momentum term of optimization affects the model's performance where it allows the model to converge more faster with a safety factor that it won't suddenly diverge from the global maximum point which will save more time in training