Assignment 2 - report - Intro To ML

Mtech - 2nd Sem

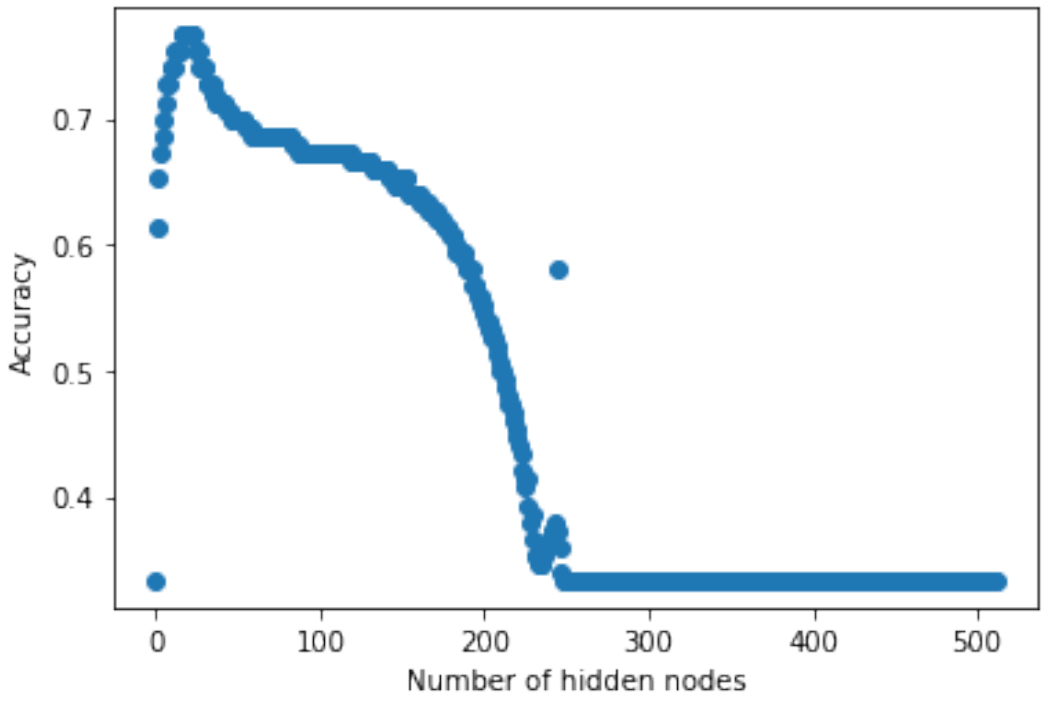
K.Ranjith - MIT2020017

27-03-2021

IIIT Allahabad

problem 1(i) - Iris Dataset

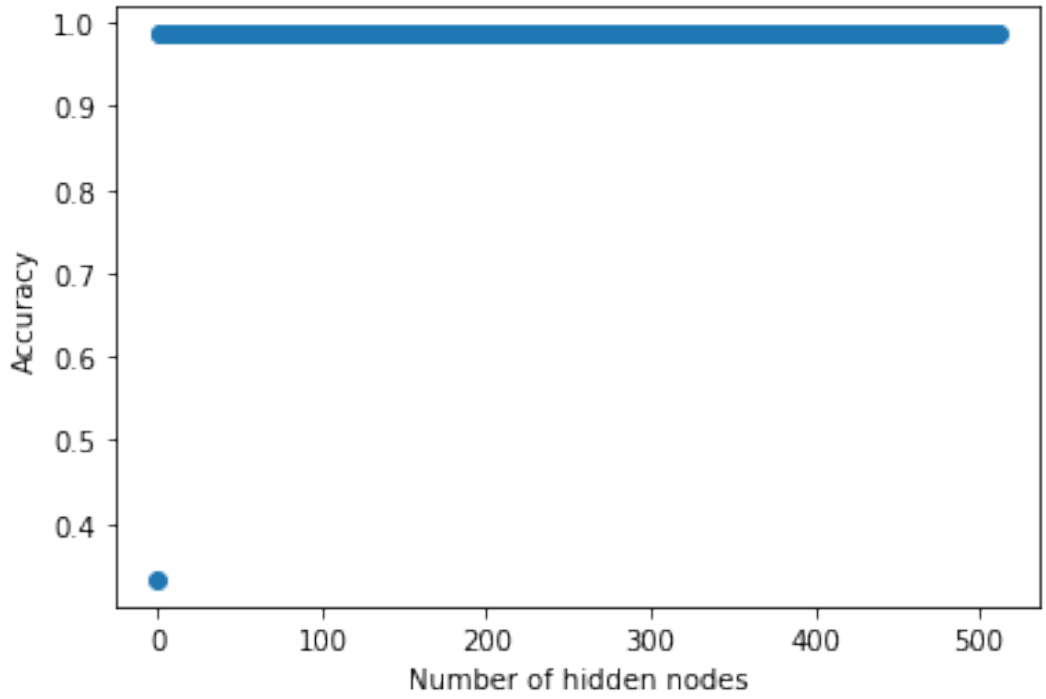
Hyper parameters[data iris, hidden layer nodes = range(512), epoch =1, L-rate= 1, sigmoid activation function , least square loss function]



1. for hidden units = [64, 128, 256, 512] accuracy = [0.68, 0.66, 0.33, 0.33]
2. for hidden unit = 16 we get max accuracy i.e 0.76

problem 1(i) - Iris Dataset

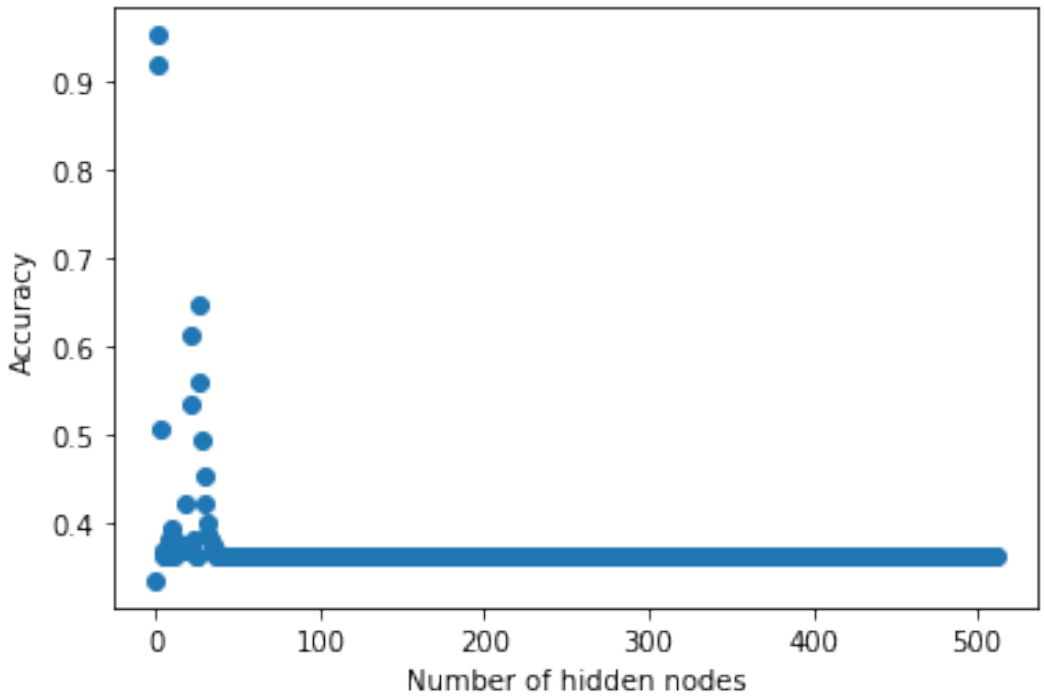
Hyper parameters[data iris, hidden layer nodes = range(512), epoch =1, L-rate= 1, ReLu activation function , least square loss function]



1. for hidden units = [64, 128, 256, 512] accuracy = [0.98, 0.98, 0.98, 0.98]
2. We get max accuracy i.e 0.98, for lrate = 1 and number of hidden units = 1

problem 1(i) - Iris Dataset

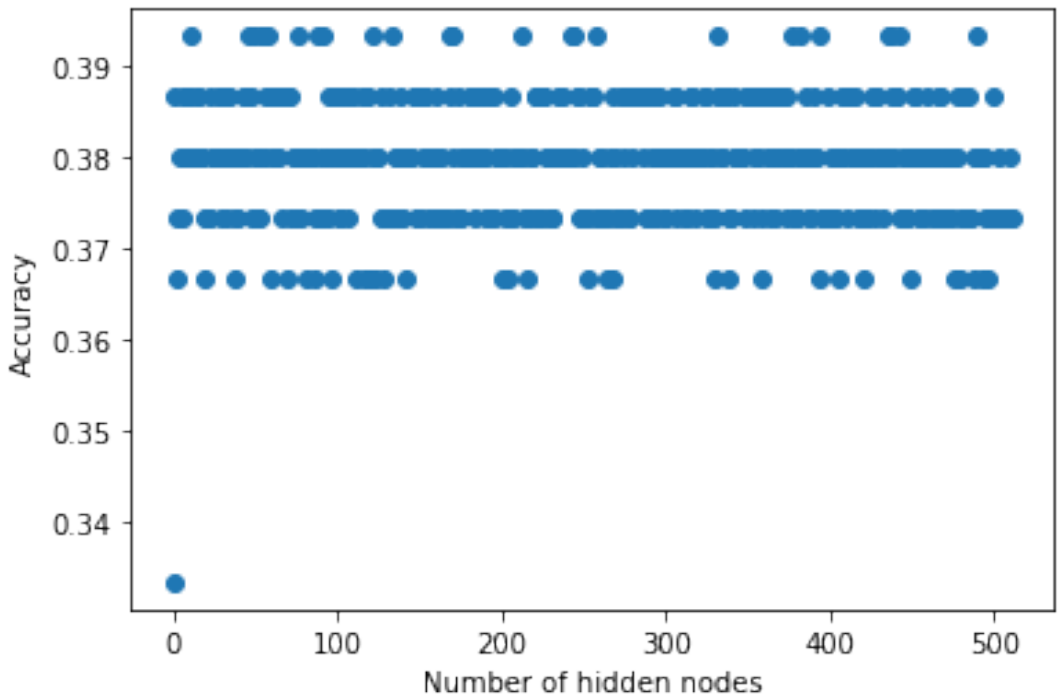
Hyper parameters[data iris, hidden layer nodes = range(512), epoch =1, L-rate= 1, tanh activation function , least square loss function]



1. for hidden units = [64, 128, 256, 512] accuracy = [0.36, 0.36, 0.36, 0.36]
2. We get max accuracy i.e 0.95, for lrate = 1 and number of hidden units = 1

problem 1(i) - Iris Dataset

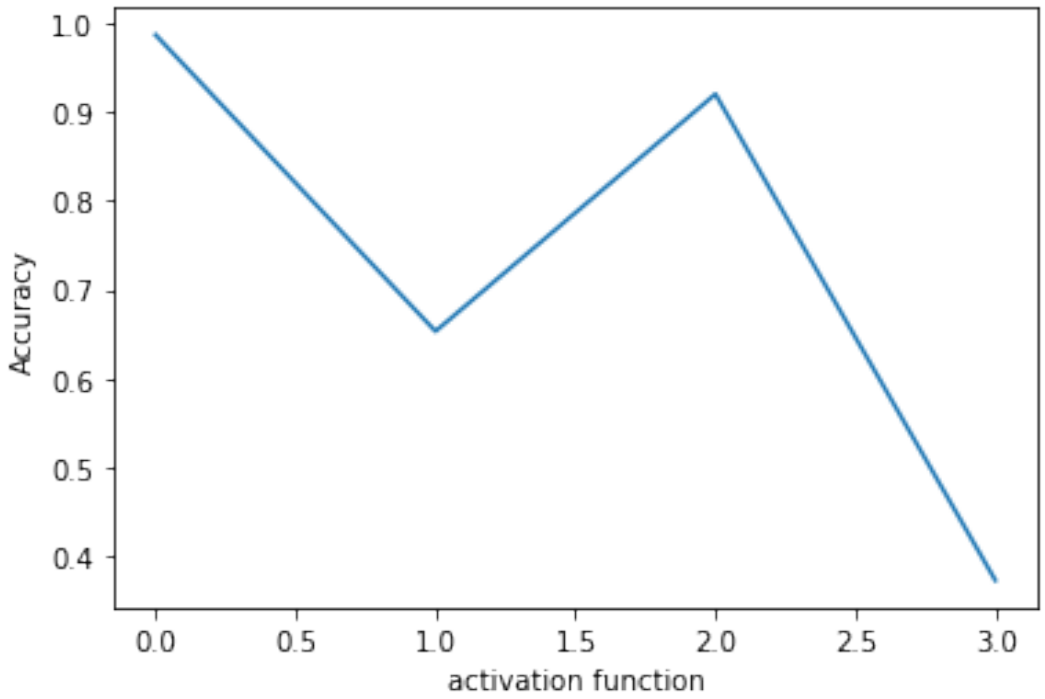
Hyper parameters[data iris, hidden layer nodes = range(512), epoch =1, L-rate= 1, leaky Relu activation function , least square loss function]



1. for hidden units = [64, 128, 256, 512] accuracy = [0.42, 0.36]
2. accuracy when we use Leaky Relu functions fluctuates between o.42 and 0.36 for hidden units range(512)

problem 1(ii) - Iris Dataset

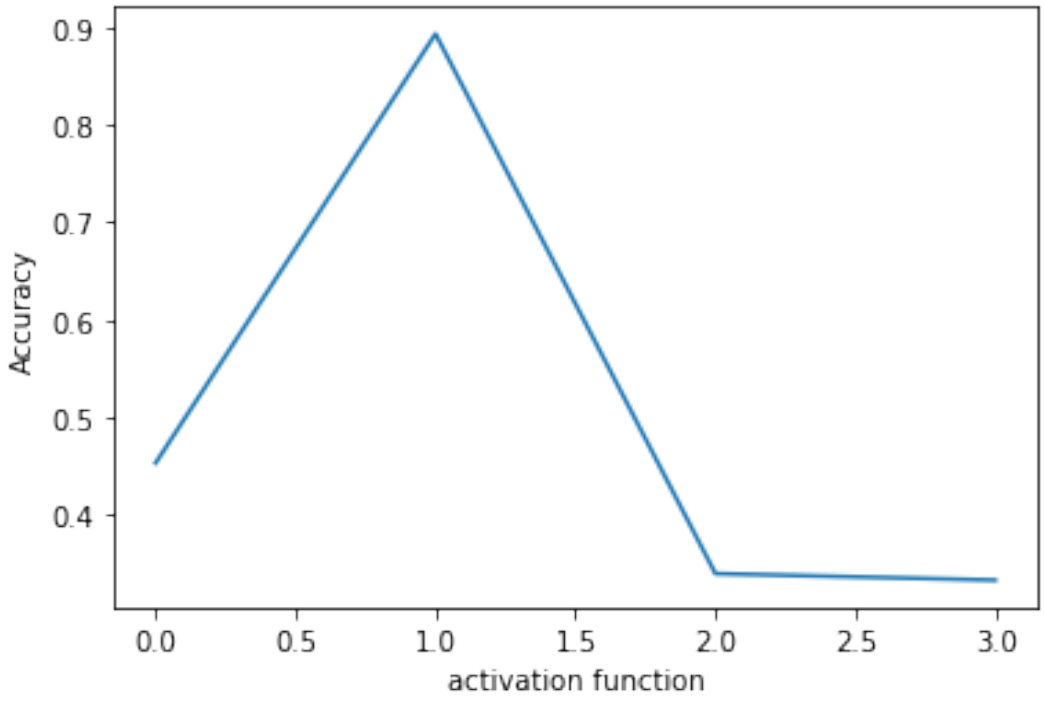
Hyper parameters[data iris, hidden layer nodes = 2, epoch =1, L-rate= 1, activation function = [0=Relu, 1=sigmoid, 2=tanh, 3=LeakyRelu] , least square loss function]



1. Accuracy is great when we use relu activation function with least square loss function

problem 1(ii) - Iris Dataset

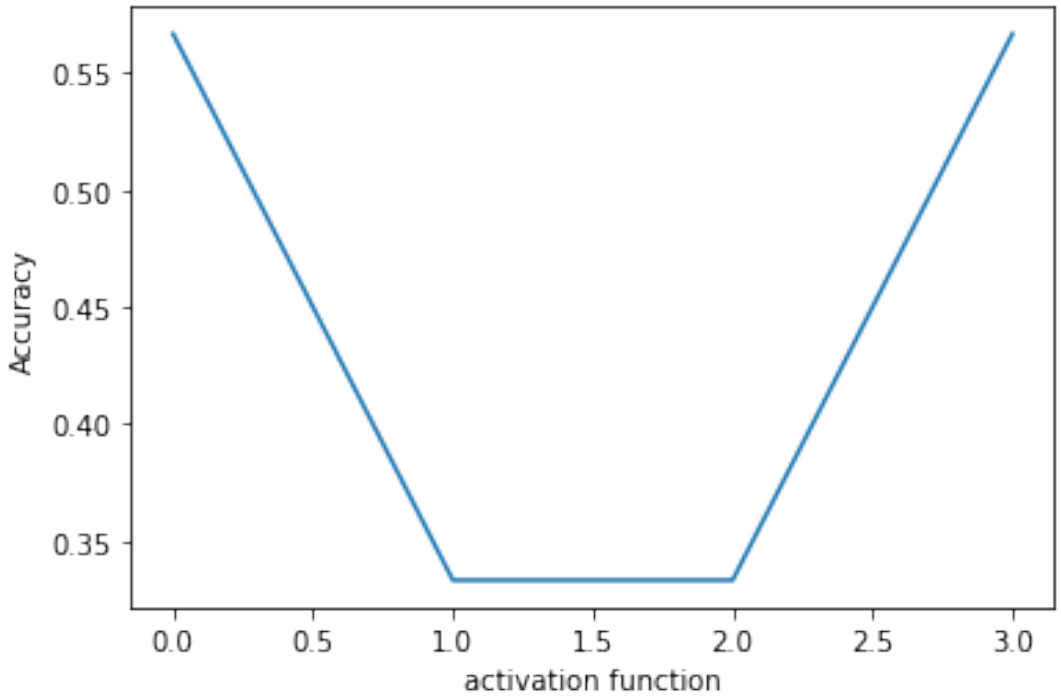
Hyper parameters[data iris, hidden layer nodes = 2, epoch =1, L-rate= 1, activation function = [0=Relu, 1=sigmoid, 2=tanh, 3=LeakyRelu] , Multi-Class Cross-Entropy Loss]



1. Accuracy is great when we use sigmoid activation function with Multi-Class Cross-Entropy Loss function

problem 1(ii) - Iris Dataset

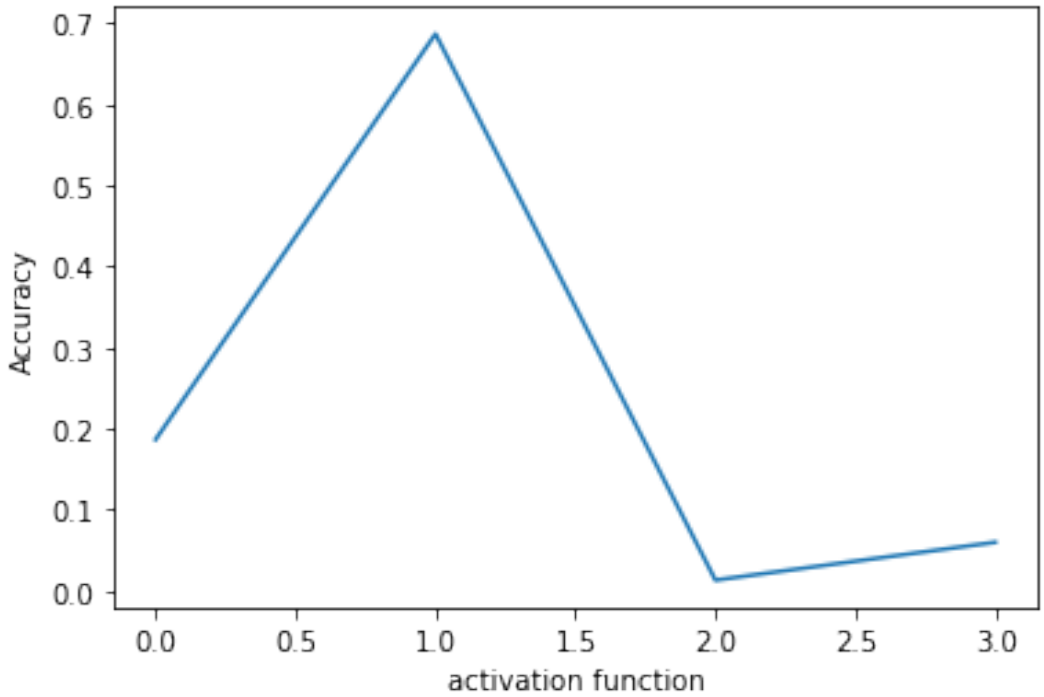
Hyper parameters[data iris, hidden layer nodes = 2, epoch =1, L-rate= 1, activation function = [0=Relu, 1=sigmoid, 2=tanh, 3=LeakyRelu] , Sparse Multi-Class Cross-Entropy Loss]



1. Accuracy is great when we use Relu and Leaky Relu activation function with Sparse Multi-Class Cross-Entropy Loss function

problem 1(ii) - Iris Dataset

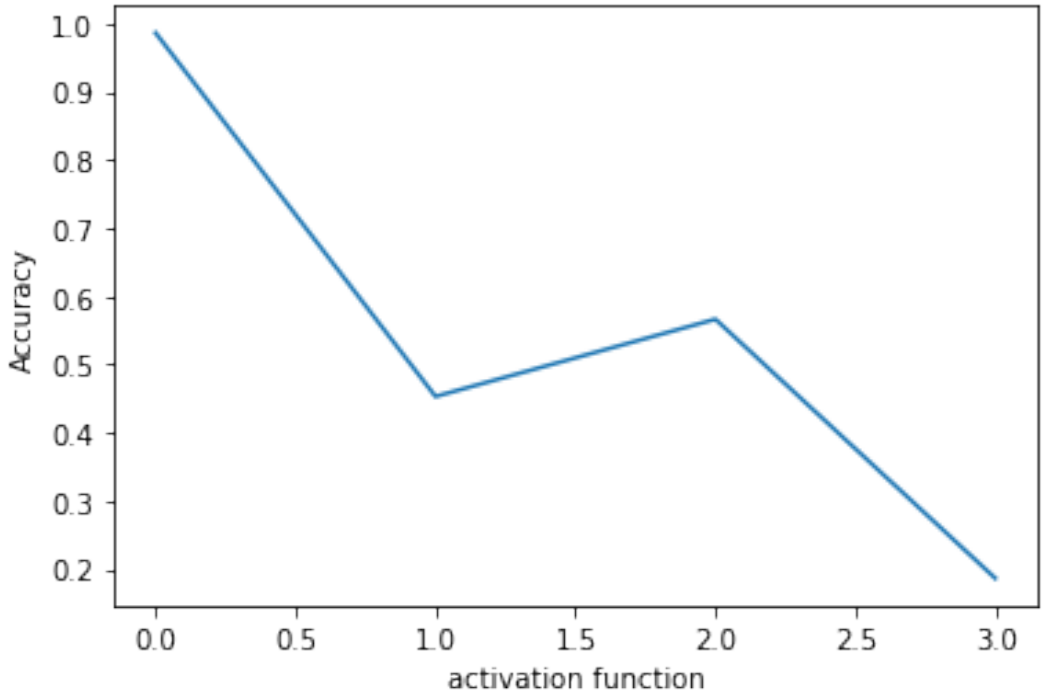
Hyper parameters[data iris, hidden layer nodes = 2, epoch =1, L-rate= 1, activation function = [0=Relu, 1=sigmoid, 2=tanh, 3=LeakyRelu] , Kullback Leibler Divergence Loss]



1. Accuracy is great when we use sigmoid activation function with Kullback Leibler Divergence Loss function

problem 1(iii) - Iris Dataset

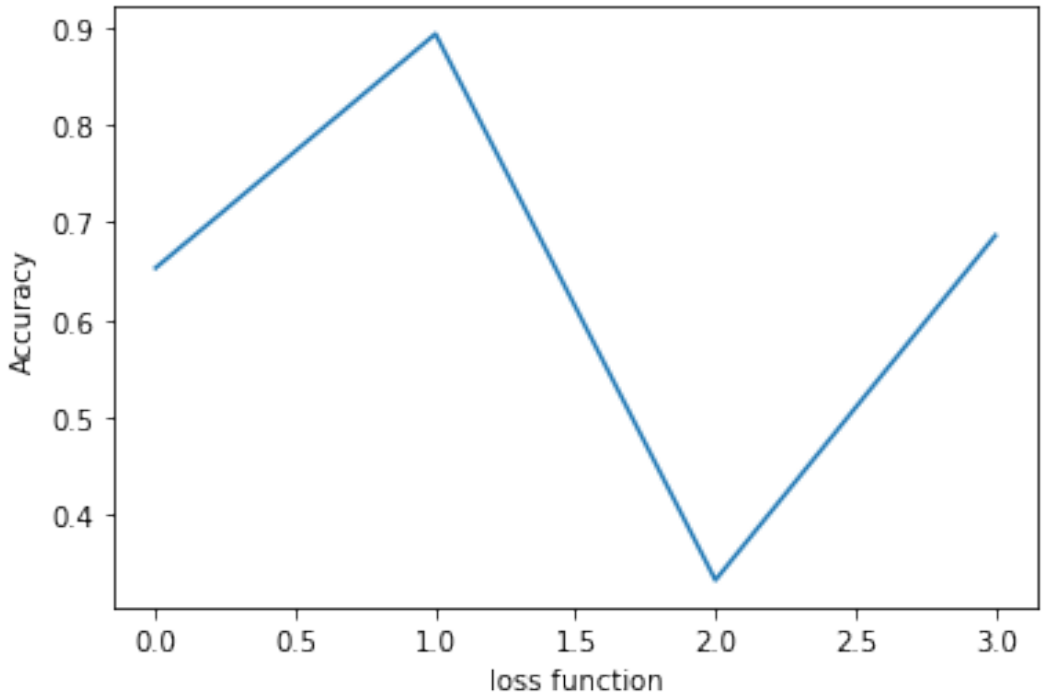
Hyper parameters[data iris, hidden layer nodes = 2, epoch =1, L-rate= 1, activation function =Relu , loss function [0=MSE, 1=Multi-Class Cross-Entropy Loss, 2=Sparse Multiclass Cross-Entropy Loss, 3=Kullback Leibler Divergence Loss]



1. Accuracy is great when we use Relu activation function with MSE loss function

problem 1(iii) - Iris Dataset

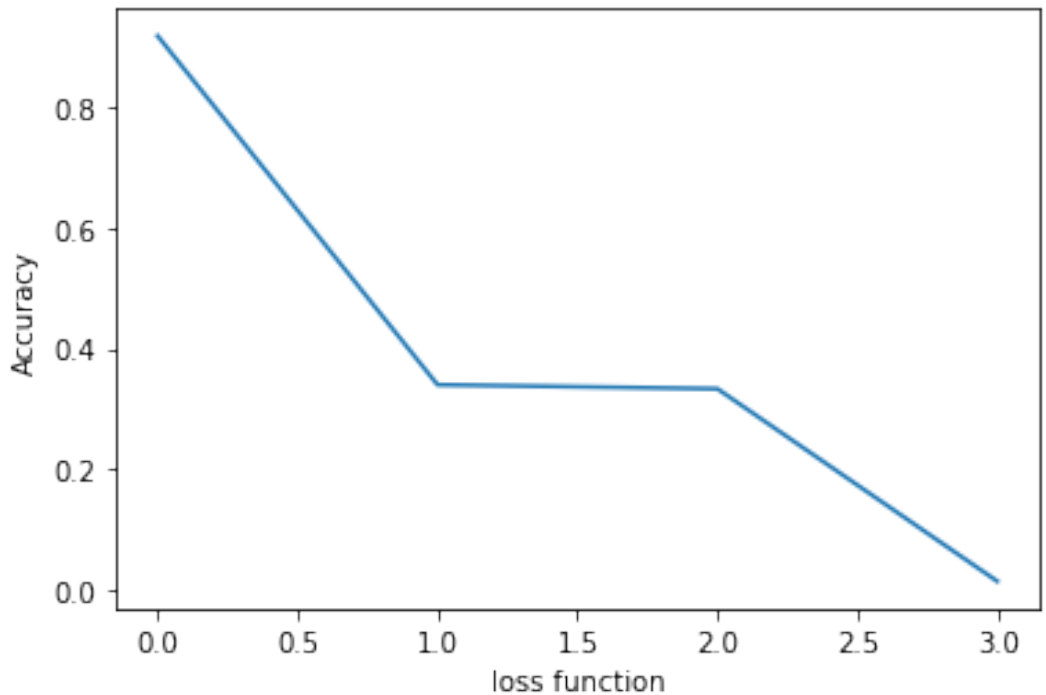
Hyper parameters[data iris, hidden layer nodes = 2, epoch =1, L-rate= 1, activation function =sigmoid , loss function [0=MSE, 1=Multi-Class Cross-Entropy Loss, 2=Sparse Multiclass Cross-Entropy Loss, 3=Kullback Leibler Divergence Loss]



1. Accuracy is great when we use sigmoid activation function with Multi-Class Cross-Entropy Loss

problem 1(iii) - Iris Dataset

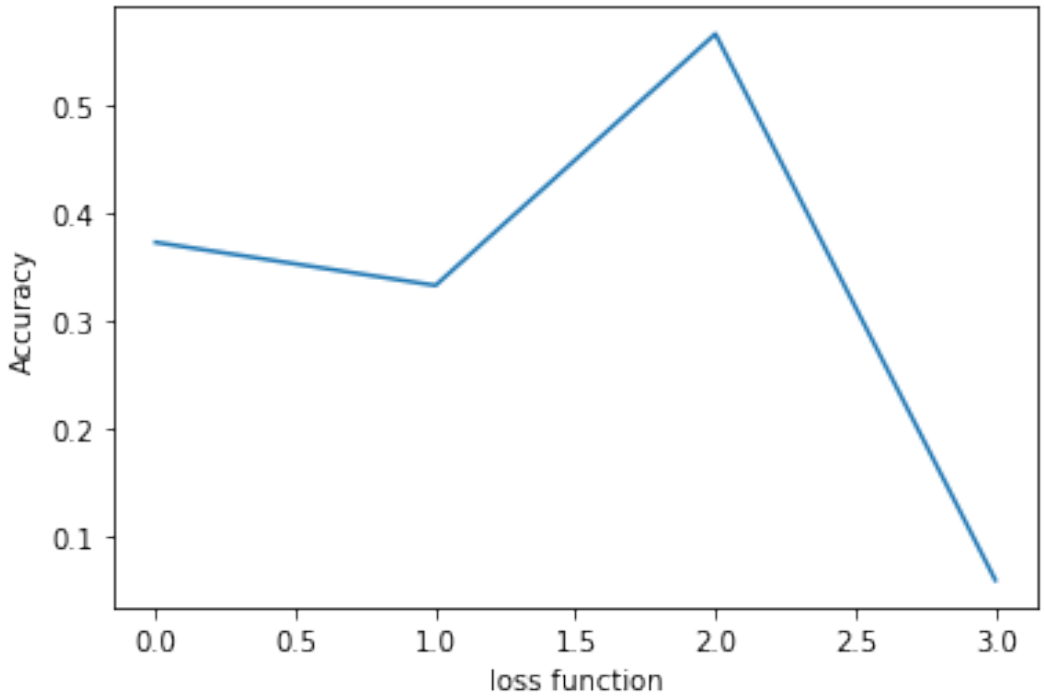
Hyper parameters[data iris, hidden layer nodes = 2, epoch =1, L-rate= 1, activation function =tanh , loss function [0=MSE, 1=Multi-Class Cross-Entropy Loss, 2=Sparse Multiclass Cross-Entropy Loss, 3=Kullback Leibler Divergence Loss]



1. Accuracy is great when we use tanh activation function with MSE loss function

problem 1(iii) - Iris Dataset

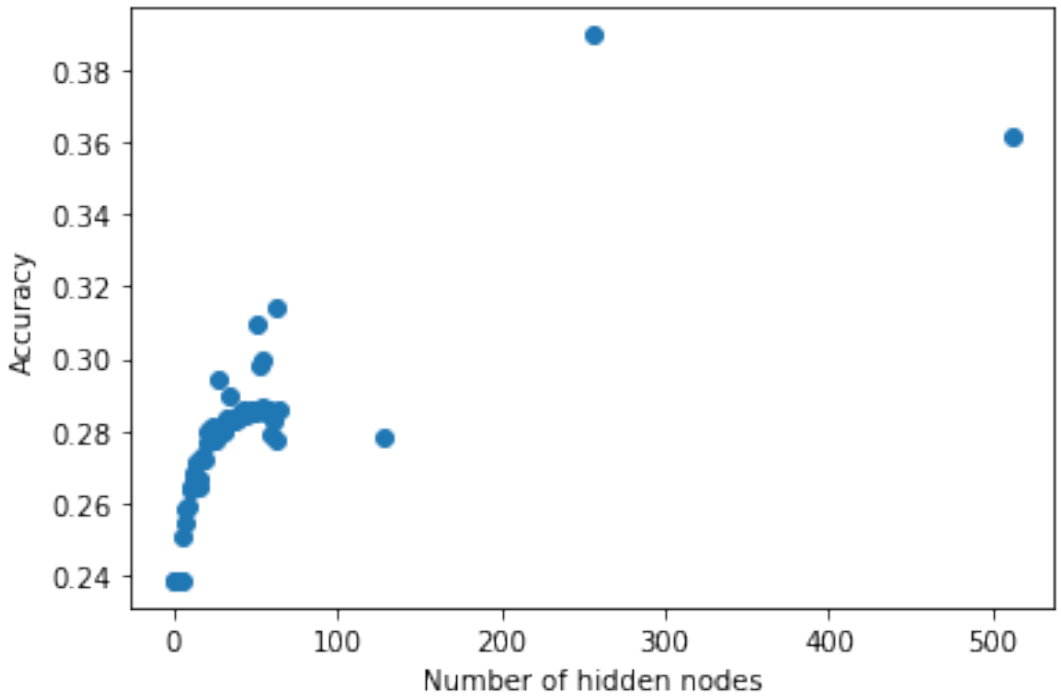
Hyper parameters[data iris, hidden layer nodes = 2, epoch =1, L-rate= 1, activation function =Leaky Relu , loss function [0=MSE, 1=Multi-Class Cross-Entropy Loss, 2=Sparse Multiclass Cross-Entropy Loss, 3=Kullback Leibler Divergence Loss]



1. Accuracy is great when we use Leaky Relu activation function with Sparse Multiclass Cross-Entropy Loss

problem 1(i) - landsat

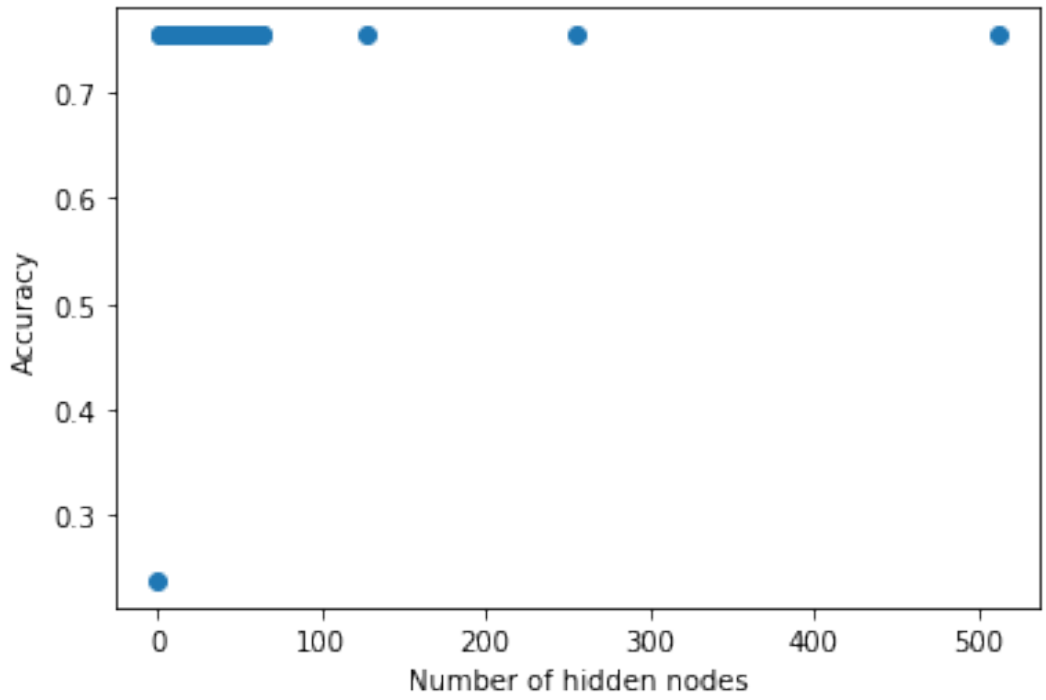
Hyper parameters[data landsat, hidden layer nodes = range(512), epoch =1, L-rate= 1, sigmoid activation function , least square loss function]



1. for hidden units = [64, 128, 256, 512] accuracy = [0.28, 0.27, 0.38, 0.36]
2. for hidden unit = 256 we get max accuracy i.e 0.38

problem 1(i) - Landsat

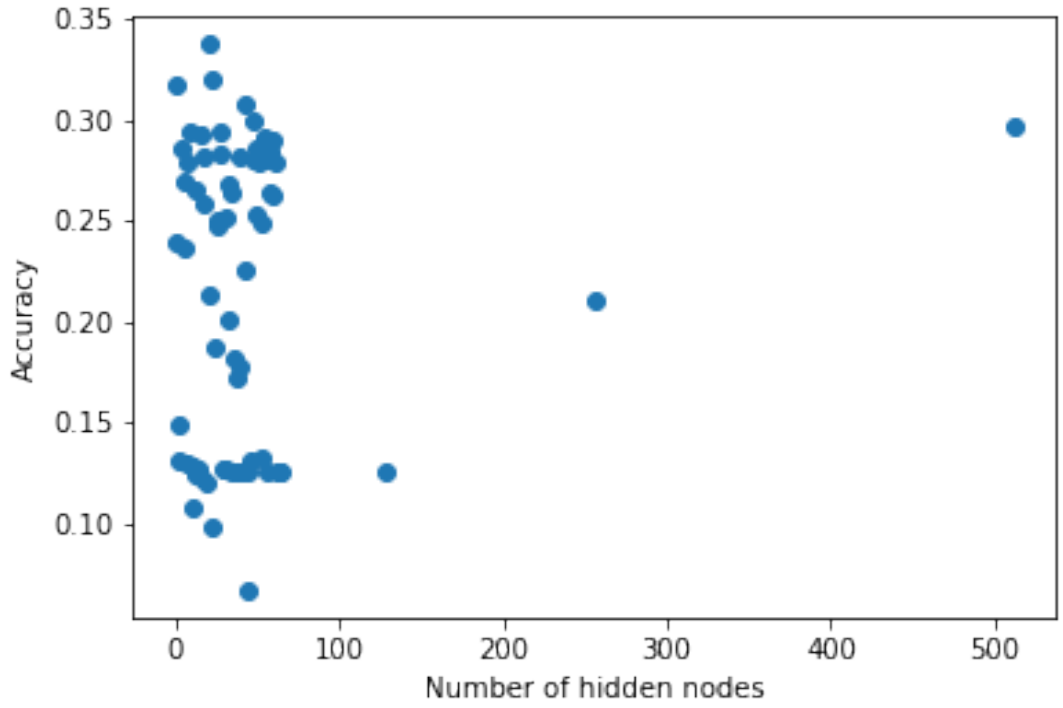
Hyper parameters[data landsat, hidden layer nodes = range(512), epoch =1, L-rate= 1, ReLu activation function , least square loss function]



1. for hidden units = [64, 128, 256, 512] accuracy = [0.74, 0.74, 0.74, 0.74]
2. We get max accuracy i.e 0.74, for lrate = 1 and number of hidden units = 1

problem 1(i) - Land Sat

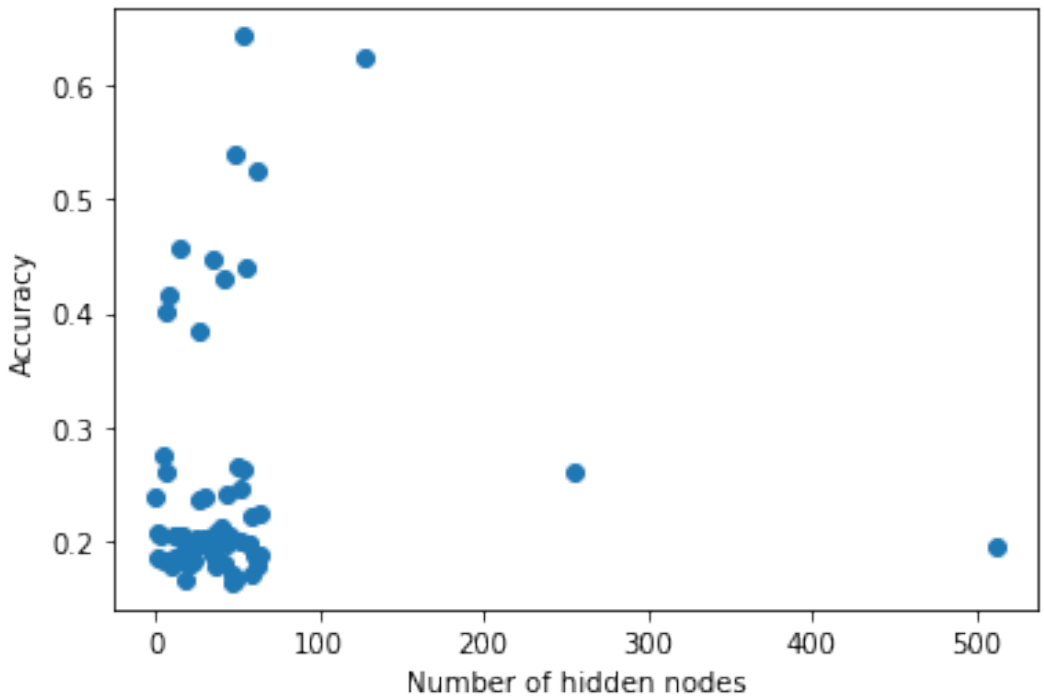
Hyper parameters[data Landsat, hidden layer nodes = range(512), epoch =1, L-rate= 1, tanh activation function , least square loss function]



1. for hidden units = [64, 128, 256, 512] accuracy = [0.13, 0.13, 0.22, 0.30]
2. We get max accuracy i.e 0.35, there are lot of fluctuations in accuracy is observed

problem 1(i) - landsat

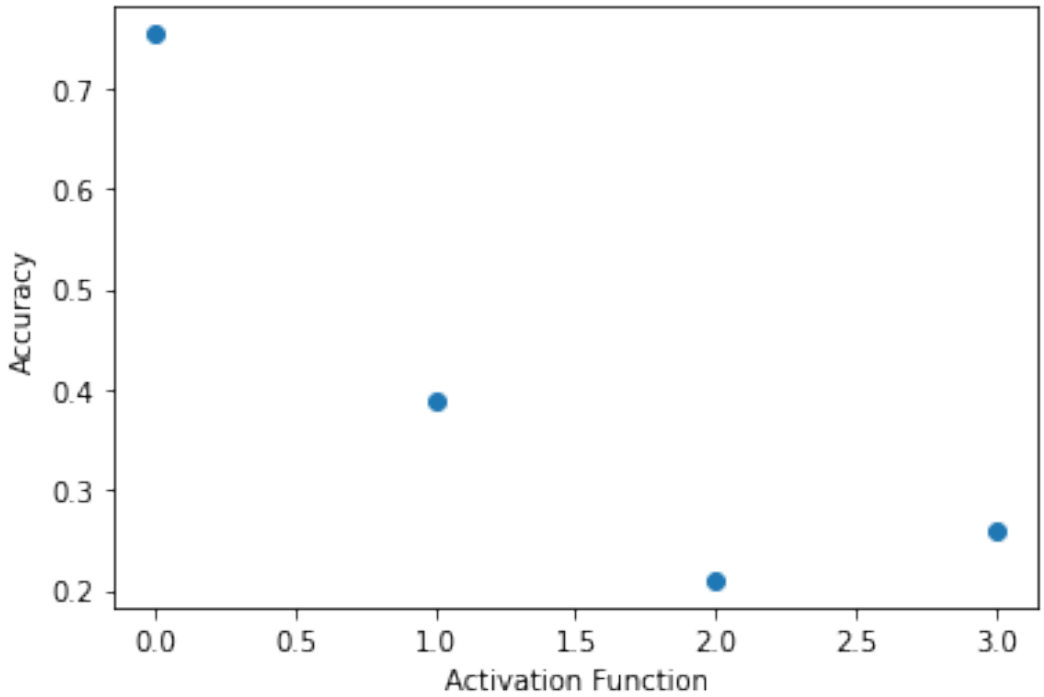
Hyper parameters[data landsat, hidden layer nodes = range(512), epoch =1, L-rate= 1, leaky Relu activation function , least square loss function]



1. for hidden units = [64, 128, 256, 512] accuracy = [0.18, 0.62, 0.25, 0.19]
2. accuracy when we use Leaky Relu functions fluctuates between 0.19 and 0.65 for hidden units range(512)

problem 1(ii) - landsat

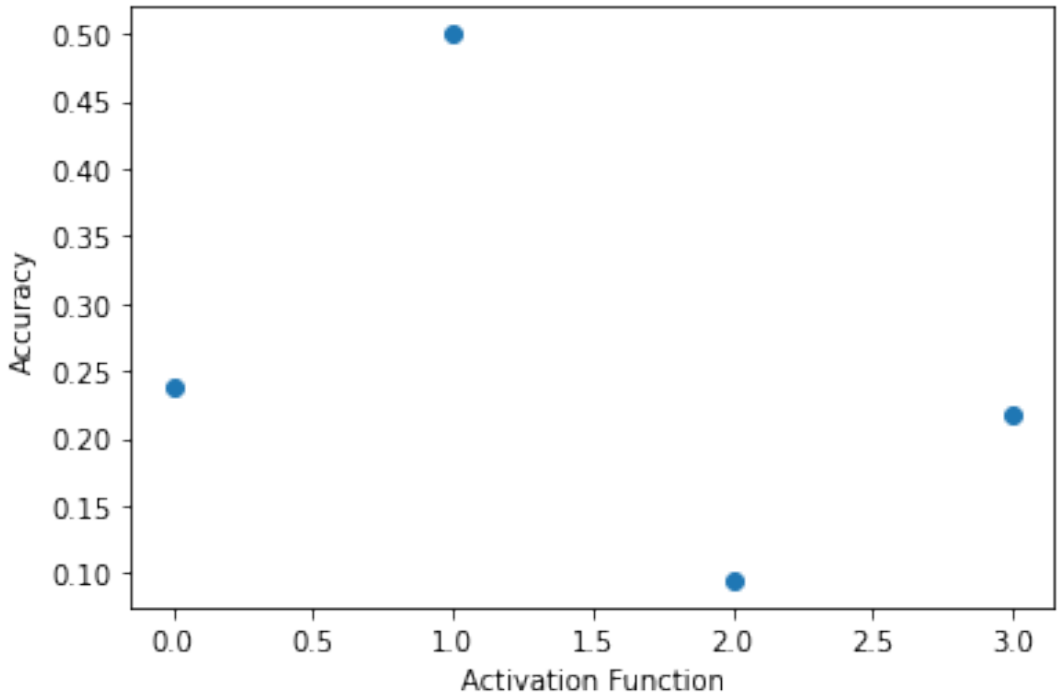
Hyper parameters[data landsat, hidden layer nodes = 256, epoch =1, L-rate= 1, activation function = [0=Relu, 1=sigmoid, 2=tanh, 3=LeakyRelu] , least square loss function]



1. Accuracy is great when we use relu activation function with least square loss function

problem 1(ii) - landsat

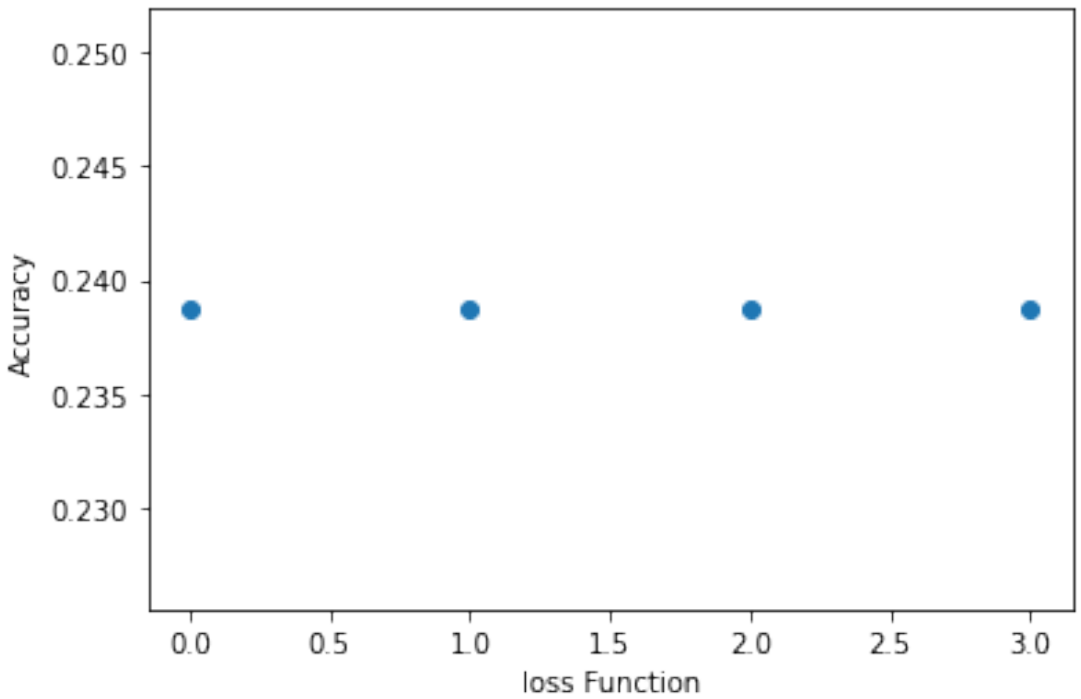
Hyper parameters[data landsat, hidden layer nodes = 256, epoch =1, L-rate= 1, activation function = [0=Relu, 1=sigmoid, 2=tanh, 3=LeakyRelu] , Multi-Class Cross-Entropy Loss]



1. Accuracy is great when we use sigmoid activation function with Multi-Class Cross-Entropy Loss function

problem 1(ii) - land sat

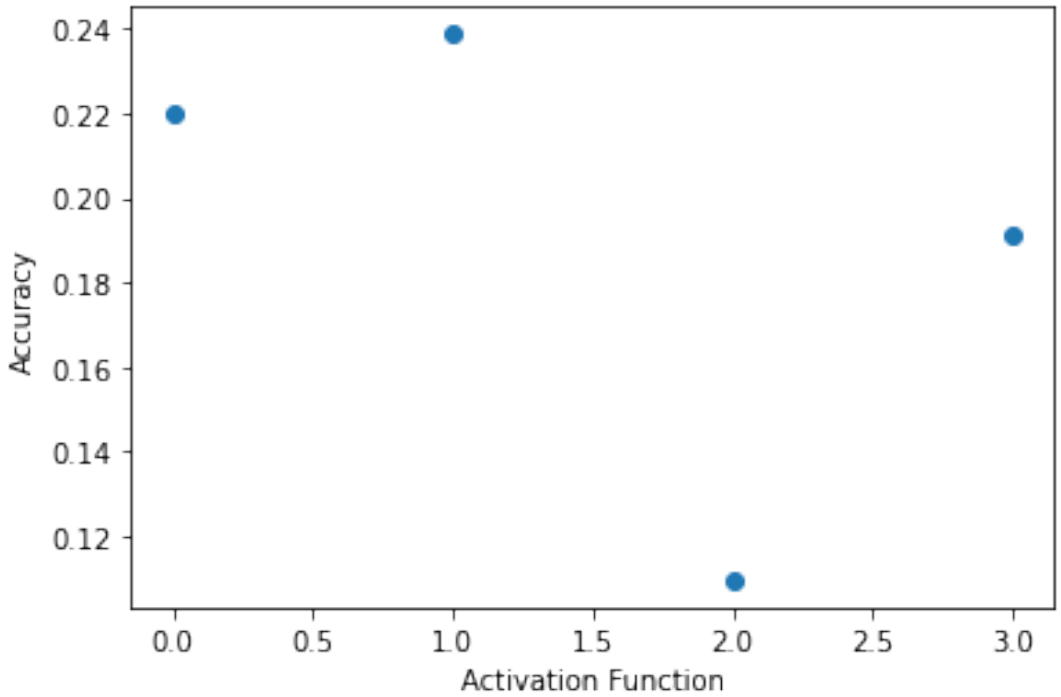
Hyper parameters[data landsat, hidden layer nodes = 256, epoch =1, L-rate= 1, activation function = [0=Relu, 1=sigmoid, 2=tanh, 3=LeakyRelu] , Sparse Multi-Class Cross-Entropy Loss]



1. The choice of activation function has no effect on accuracy when we use Sparse cross entropy loss function

problem 1(ii) - landsat

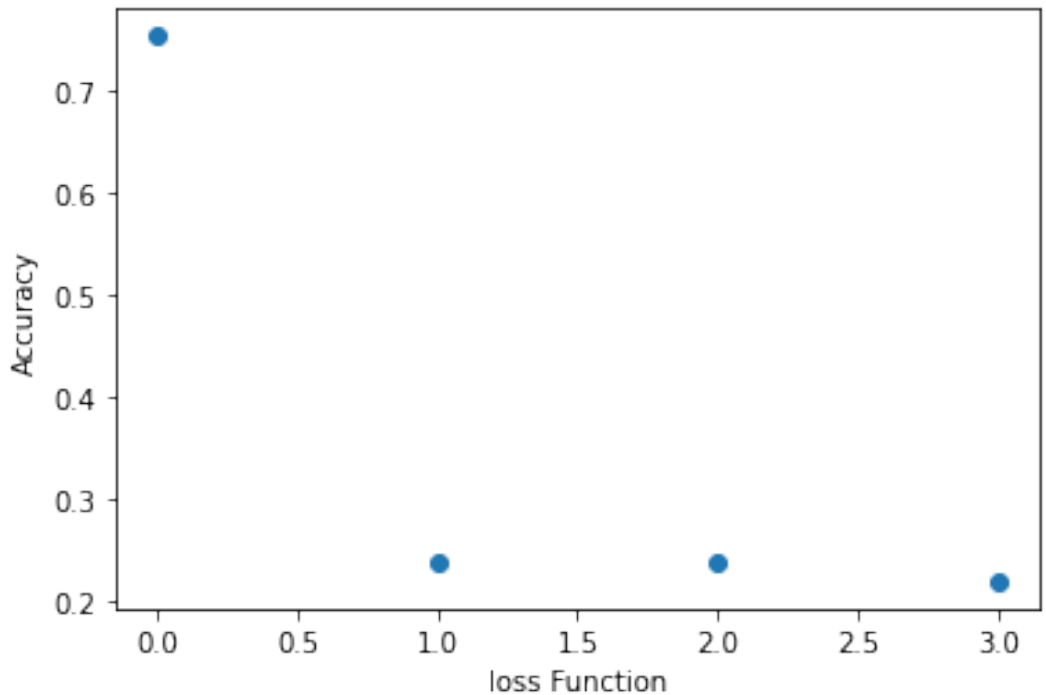
Hyper parameters[data landsat, hidden layer nodes = 256, epoch =1, L-rate= 1, activation function = [0=Relu, 1=sigmoid, 2=tanh, 3=LeakyRelu] , Kullback Leibler Divergence Loss]



1. Accuracy is great when we use sigmoid activation function with Kullback Leibler Divergence Loss function

problem 1(iii) - land sat

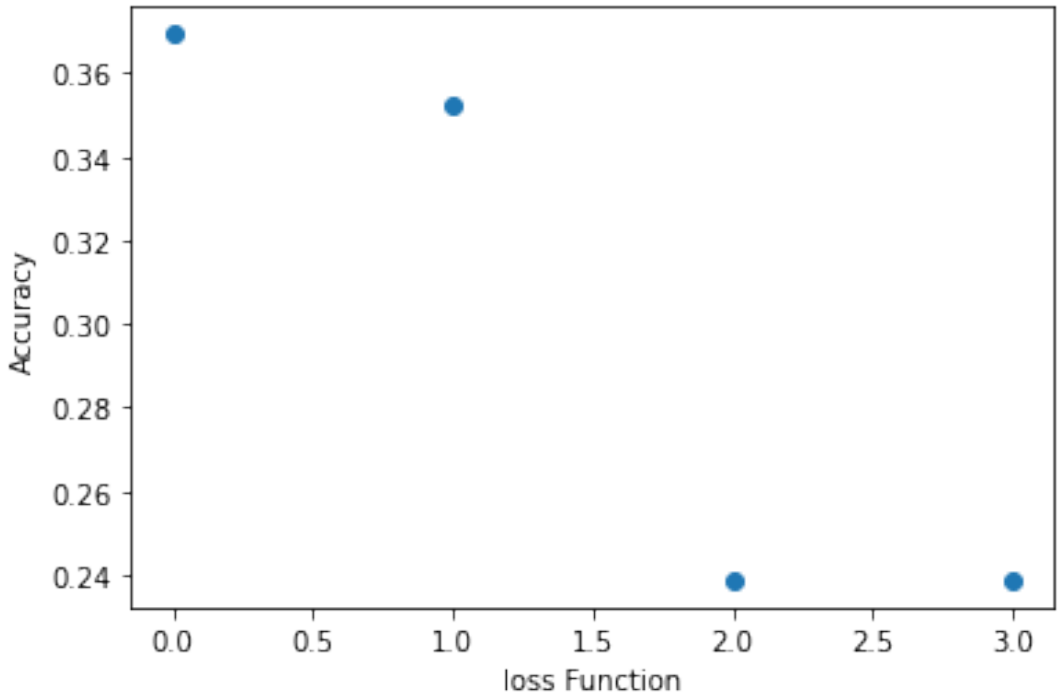
Hyper parameters[data landsat, hidden layer nodes = 256, epoch =1, L-rate= 1, activation function =Relu , loss function [0=MSE, 1=Multi-Class Cross-Entropy Loss, 2=Sparse Multiclass Cross-Entropy Loss, 3=Kullback Leibler Divergence Loss]



1. Accuracy is great when we use Relu activation function with MSE loss function

problem 1(iii) - landsat

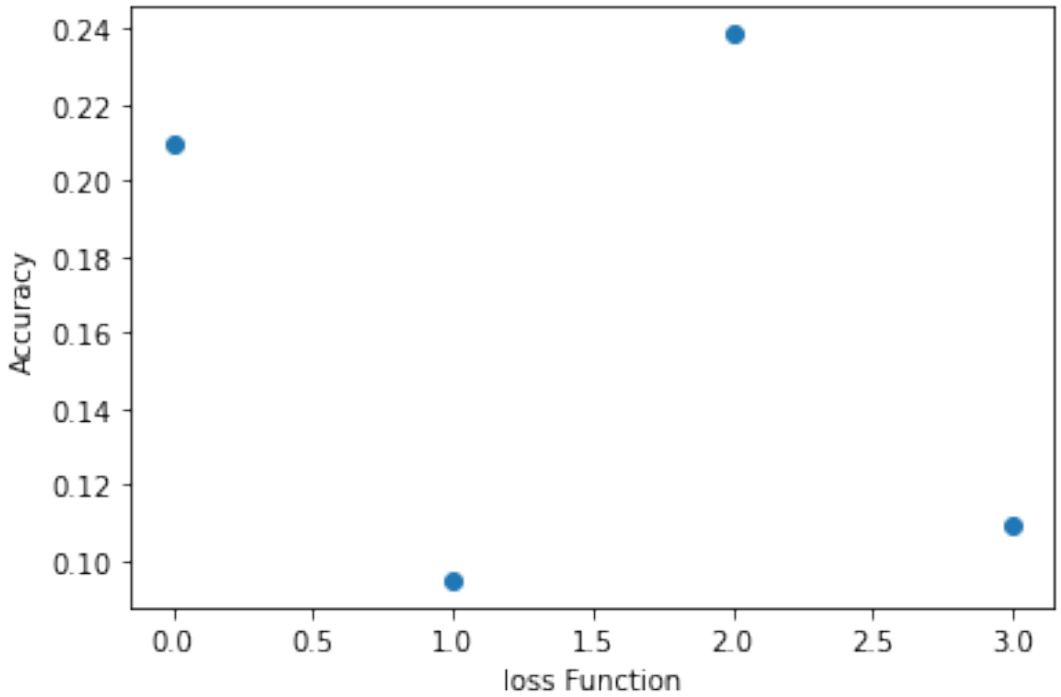
Hyper parameters[data landsat, hidden layer nodes = 256, epoch =1, L-rate= 1, activation function =sigmoid , loss function [0=MSE, 1=Multi-Class Cross-Entropy Loss, 2=Sparse Multiclass Cross-Entropy Loss, 3=Kullback Leibler Divergence Loss]



1. Accuracy is better when we use sigmoid activation function with Multi-Class Cross-Entropy Loss and MSE

problem 1(iii) - landsat

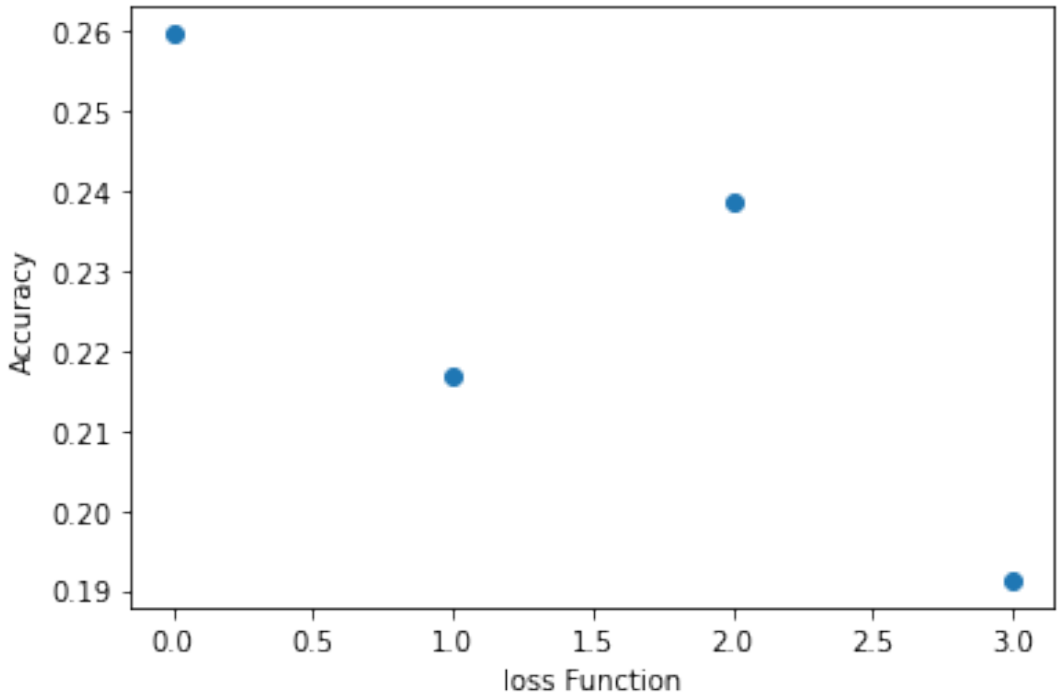
Hyper parameters[data landsat, hidden layer nodes = 256, epoch =1, L-rate= 1, activation function =tanh , loss function [0=MSE, 1=Multi-Class Cross-Entropy Loss, 2=Sparse Multiclass Cross-Entropy Loss, 3=Kullback Leibler Divergence Loss]



1. Accuracy is better when we use tanh activation function with MSE loss function and Spare Multi class loss functions

problem 1(iii) - landsat

Hyper parameters[data landsat, hidden layer nodes = 2, epoch =1, L-rate= 1, activation function =Leaky Relu , loss function [0=MSE, 1=Multi-Class Cross-Entropy Loss, 2=Sparse Multiclass Cross-Entropy Loss, 3=Kullback Leibler Divergence Loss]



1. Accuracy is great when we use Leaky Relu activation function with Sparse Multiclass Cross-Entropy Loss and MSE

PROBLEM 1 - OBSERVATIONS

1. For Iris data we get maximum accuracy ratio i.e 0.98 when we use relu activation function with MSE loss.
2. For landsat data we get maximum accuracy ratio i.e 0.74 when we use relu activation function with MSE loss.
3. These (activation function, loss function) pairs are producing better results (relu, MSE), (sigmoid, Cross entropy), (tanh, MSE), (Leaky Relu, Sparse Cross Entropy).
4. Accuracy is insensitive with number of hidden units when we use Relu and Leaku Relu activation functions.
5. Accuracy is low with large number of hidden units when we use sigmoid and Tanh activation functions.