* **Comparison between pm1 and pm2 :**

Both pm1 and pm2 gives almost the same accuracy as the data set given is random. If the data set is well ordered then it might effect the accuracy in a significant number else the accuracy is nearly same in both cases.

* **Comparison between pm1 and pm3 :**

In pm3 we used normalized data which helps us in bringing all the data into similar scale and helps us analysing the data more accurately so pm3 perform much accurately than pm1.

* **Comparison between pm1 and pm4 :**

In pm4 we just changed the order of features which did not changed the accuracy of the model. i.e both pm1 and pm4 obtained almost equal accuracy.

* Their respective performance is shown in the table at last.
* Based on the value of accuracy we can say that the data is linearly separable. As in PM3 the data is almost linearly separable.
* **Comparison between fldm1 and fldm2 :**

In fldm2 we used these lines

data = pd.read\_csv("fisherdata.csv").dropna()

column\_names = list(data.columns[:-1])

np.random.shuffle(column\_names)

which helps in changing the order of features. Here changing the order of features results in different projection of data which can alter the accuracy little bit. Since we resulted with a different accuracy in fldm2.

* For lr1 the accuracy is low as the data is not normalized
* In lr2 the accuracy of models for particular threshold values as follows

For threshold

0.5 🡪accuracy = 98.94%

0.3 🡪accuracy = 90.43%

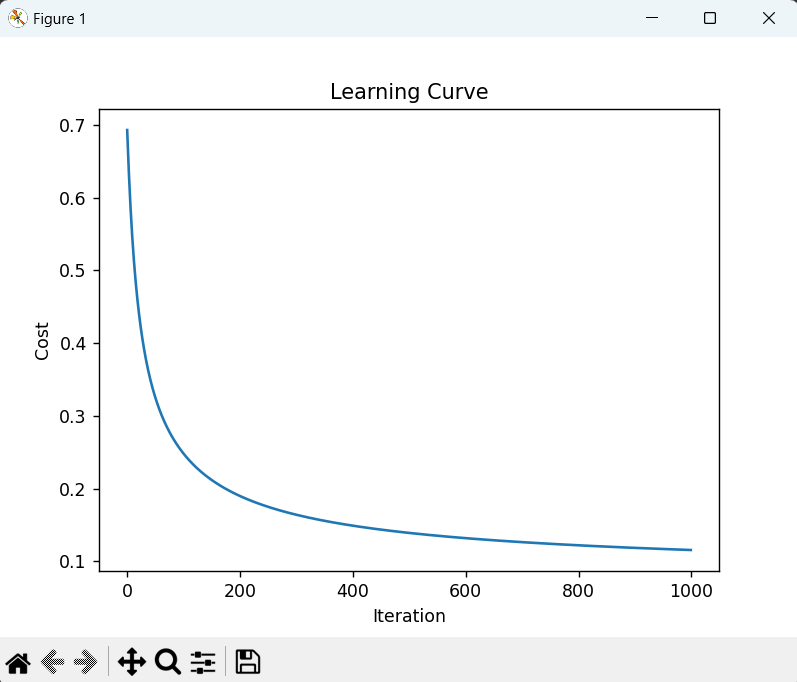
0.4 🡪accuracy = 94.68%

0.6 🡪accuracy = 98.40%

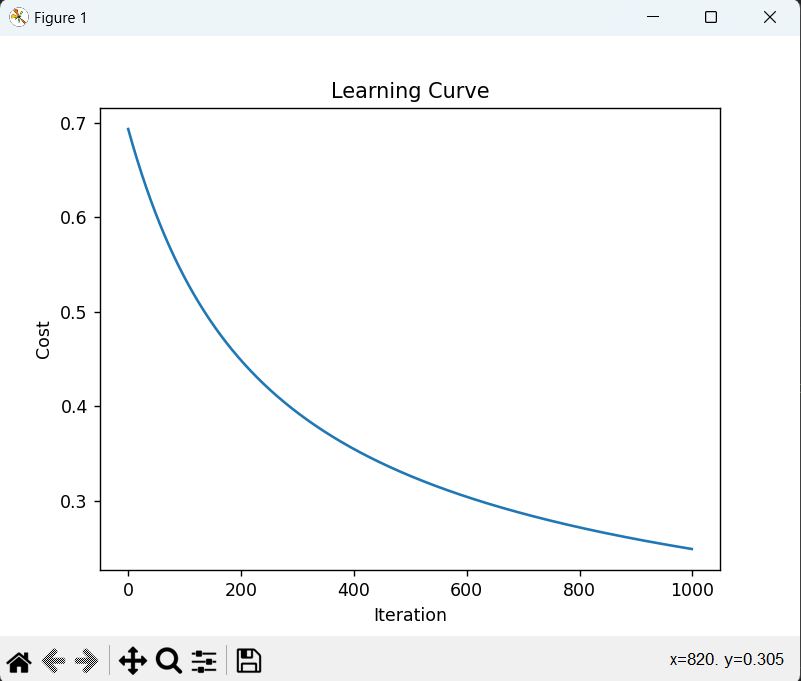
0.7 🡪accuracy =97.87%

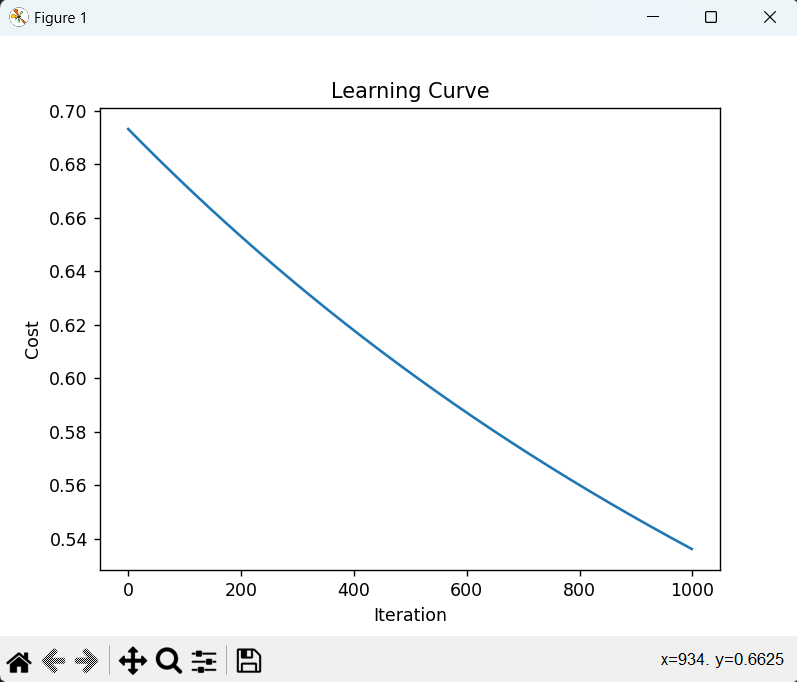
By changing the value of threshold there will be change in precision and recall which will affect the final accuracy. we got highest accuracy when threshold is set to 0.5.

For threshold 0.5 and learning rate=0.01 the plot of cost curve is as follows



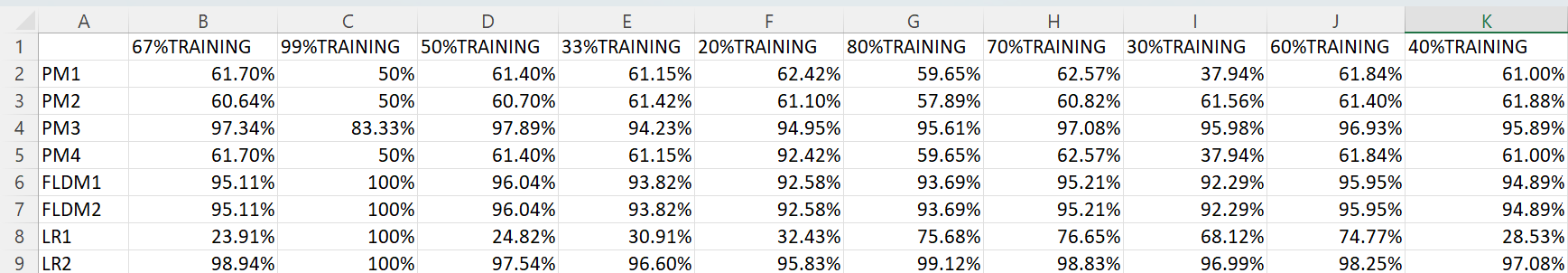
For threshold 0.5 and learning rate=0.001 the plot of cost curve is as follows



For threshold 0.5 and learning rate=0.0001 the plot of cost curve is as follows

* if we decrease the learning rate, the optimization algorithm takes smaller steps, which can result in slower convergence to the optimal set of parameters.
* **COMPARITIVE STUDY**

ACCURACY



* From the above observations we got almost same accuracy in PM3 and LR2 in all cases. For small data sets we can choose PM3 as our model since it can perform the task faster. But for data sets with large number of features we choose LR2 as our model.

DONE BY

1.SRIVARAN KUKATLA(2021A4PS1988H)

2.OMKAR BASANI(2021A7PS0433H)

3.SUTAPALLI LIKHITH BHARGAV(2021A7PS0444H).