

## Labels

Malignant = 1

Benign = -1 for perceptron, FLDA

Benign = 0 for logistic

## Perceptron

---

After 10000 epochs, data gives 90-92% accuracy, 90-94% precision and 85-90% recall when tested using average of 10 different train-test splits.

It is expected that even after further iterations Perceptron will not outperform normalized data.

After normalizing data, the algorithm converges under 10000 epochs giving 94% accuracy, 92% precision and 94% recall.

Rearranging the feature order does not affect the result. The weight vector appears similarly rearranged.

It is found that the data is linearly separable when normalized.

## Fischer's Linear Discriminant Analysis

---

Fischer's LDA does not require training and directly gives 95% accuracy, 95% precision and 91% recall on normalized and imputed data.

Rearranging the feature order does not affect the result. The weight vector appears similarly rearranged.

## Logistic Regression

---

After running for a 1000 epochs it gives maximum of 97% accuracy. It is expected to go even higher with more epochs.

## Batch Gradient Descent

Batch Gradient Descent shows a 74-89% accuracy, 63-89% precision and 76-81% recall with un-normalized data all decreasing with decrease in learning rate.

With normalized data, it gives 88-93% accuracy, 82-92% precision and 89-95% recall.

It is found with normalized data, increasing the learning rate for normalized data decreases the accuracy.

## Stochastic Gradient Descent

Stochastic Gradient Descent shows a 89-91% accuracy, 88-90% precision, 81-88% recall, all except precision decreasing with decrease in learning rate.

With normalized data, it gives 95-97% accuracy, 91-97% precision and 89-96% recall.

## Mini-Batch Gradient Descent

There is no training involved in this model. It is the fastest model we currently have and is quite accurate on both normalized and raw data.

It shows 86-89% accuracy, 84-88% precision, 83-85% recall with raw data.

Shows 95-97% accuracy, 92-97% precision and 94-95% recall with normalized data.

## Comparative Study

After 10000 epochs (except for Fischer) following is the table of observations.

Models	Accuracy	Precision	Recall
Perceptron (PM1)	0.9129032258064516	0.9372268462822859	0.8501716001250819
Perceptron (PM2)	0.9182795698924731	0.9053517145473006	0.8935902155966055
Perceptron (PM3)	0.9521505376344086	0.9406453805430113	0.9412902692543786
Perceptron (PM4)	0.9129032258064516	0.9372268462822859	0.8501716001250819
Fischer (FLDM1)	0.950531914893617	0.952921328366989	0.914998930095255
Fischer (FLDM2)	0.950531914893617	0.952921328366989	0.914998930095255
Batch (LR1)	0.9095744680851063	0.9365079365079365	0.9027777777777778
Stochastic (LR1)	0.9202127659574468	0.9830508474576272	0.8055555555555556
Mini Batch (LR1)	0.9414893617021277	0.9420289855072463	0.9444444444444444
Batch (LR2)	0.9095744680851063	0.8666666666666667	0.9027777777777778
Stochastic (LR2)	0.9680851063829787	1.0	0.9166666666666666
Mini Batch (LR2)	0.973404255319149	1.0	0.9305555555555556

Logistic Regression Metrics have been run for 10000 epochs for only a specific learning rate and threshold and may vary based on others. However, it clearly outperforms both Fischer and Perceptron.

Hence, LR2 is the best performing model.

This could be due to many reasons:

- Logistic regression provides a probabilistic interpretation of the output.
- It is basically perceptron with a sigmoid function.
- Perceptron fails when data is not linearly separable.
- Fischer fails when there are more features.