

TEAM

- 1) Ansuman Dibyayoti Mohanty (2016A7PS0043H)
- 2) Mridul Bhaskar (2016A7PS0391H)
- 3) Deepak Gupta (2016A7PS0105H)

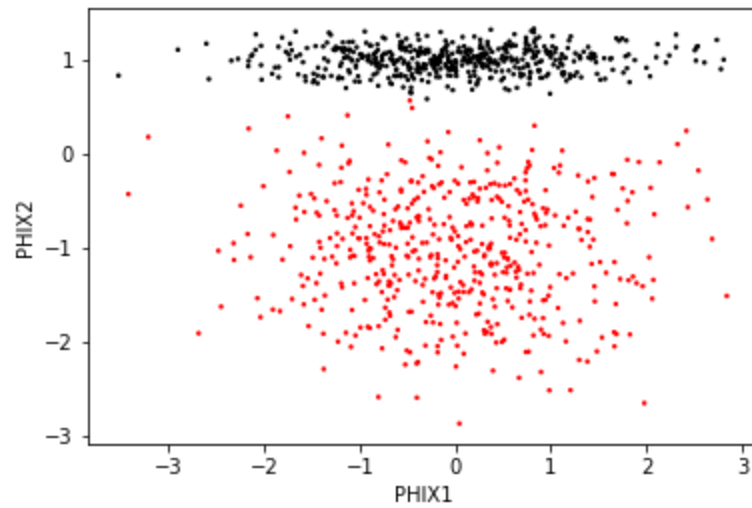
PERCEPTRON

1. The visualisations presented below were done for one random trial in each configuration.
2. Learning rate was kept at 0.01 for visualisation runs.
3. Experimentation with both SGD and BGD was done.
4. Experimentation with learning rate is presented in the table at the end of the section.
5. Dataset 2 doesn't converge since it is not linearly separable.
6. **Effect of random initialization** is checked by plotting boxplots over different runs of Dataset 1 and 3. Dataset 2 was excluded since it doesn't converge.
7. The above was done for SGD and BGD separately.
8. Similarly, the **effect of reshuffling** was checked - however this time the initialization was constant over runs.
9. The GIFs have random initialization and no reshuffling over epochs, with a learning rate of 0.01.

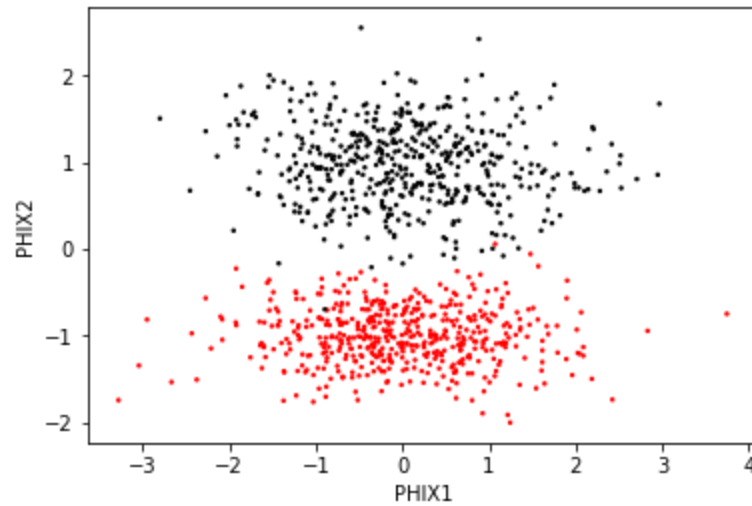
Visualisation:

(Black - Class 1 , Red - Class 0)

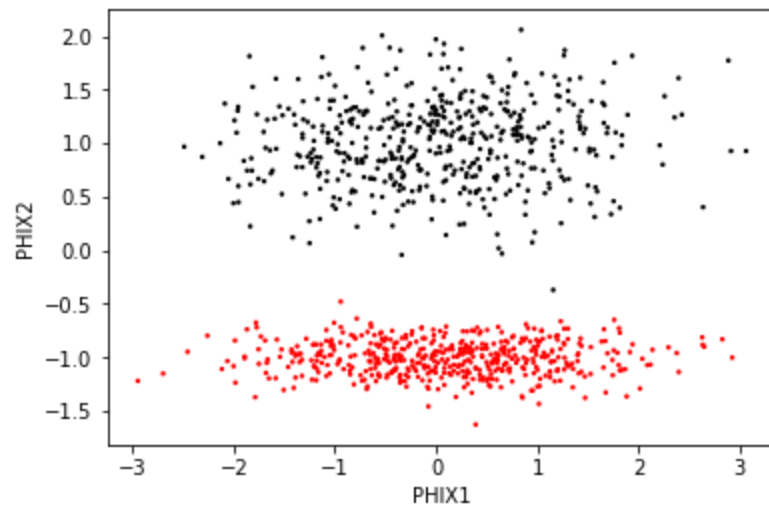
Dataset 1 :



Dataset 2 :



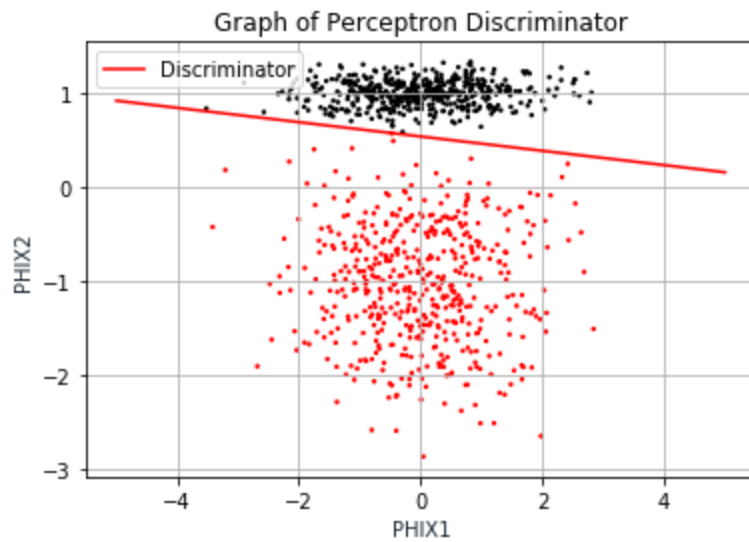
Dataset 3 :



BATCH GRADIENT DESCENT

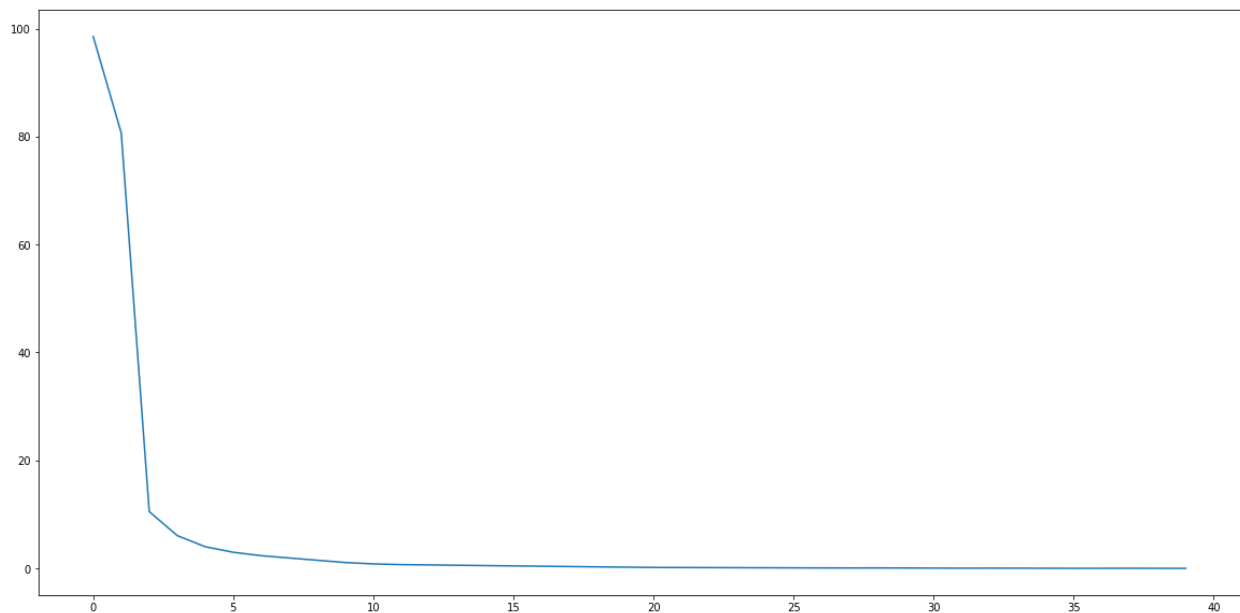
Discriminator after 1000 epochs/convergence whichever was earlier:

Dataset 1 :

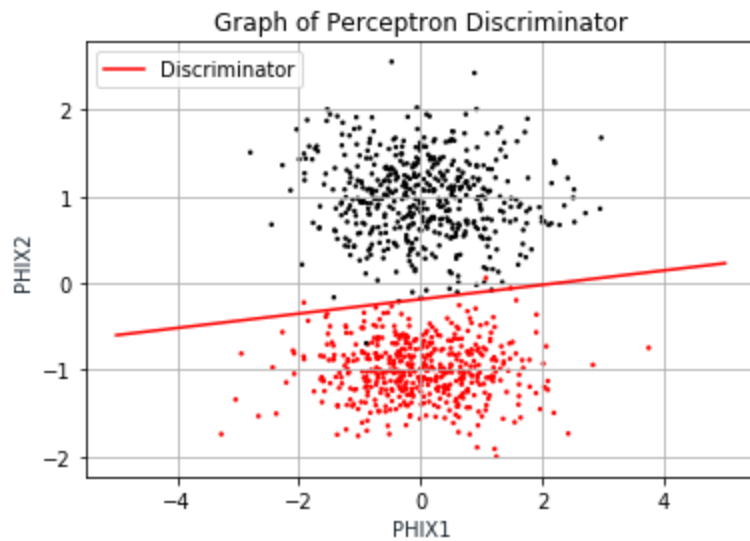


Number of epochs (BGD) = 39

Error vs number of epochs (BGD) :

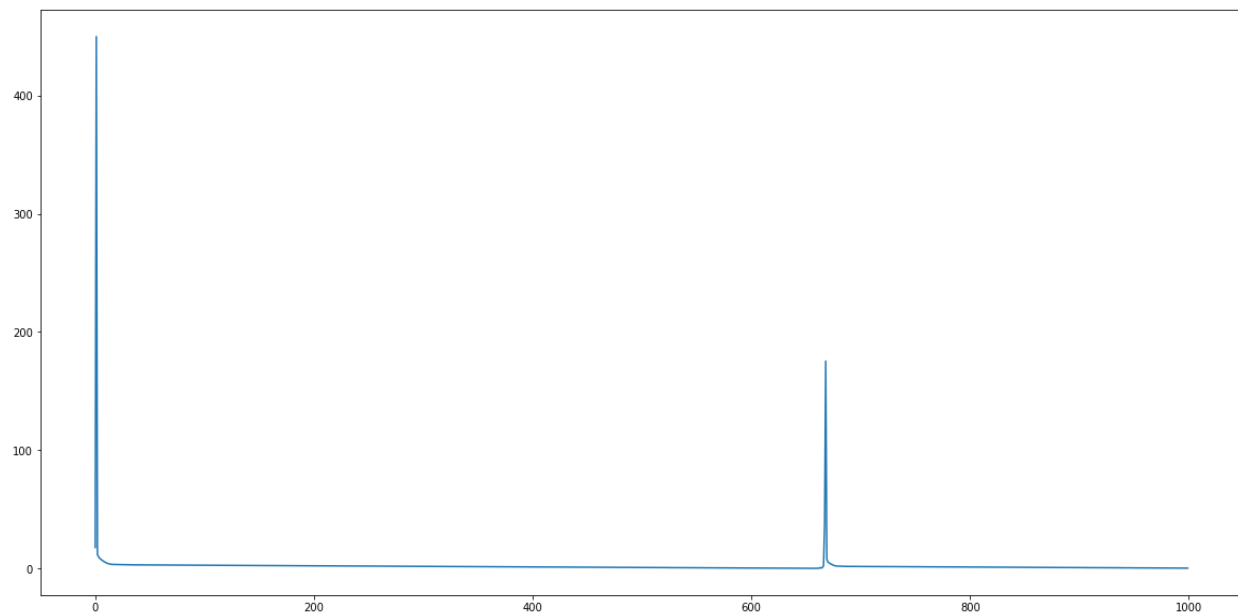


Dataset 2 :

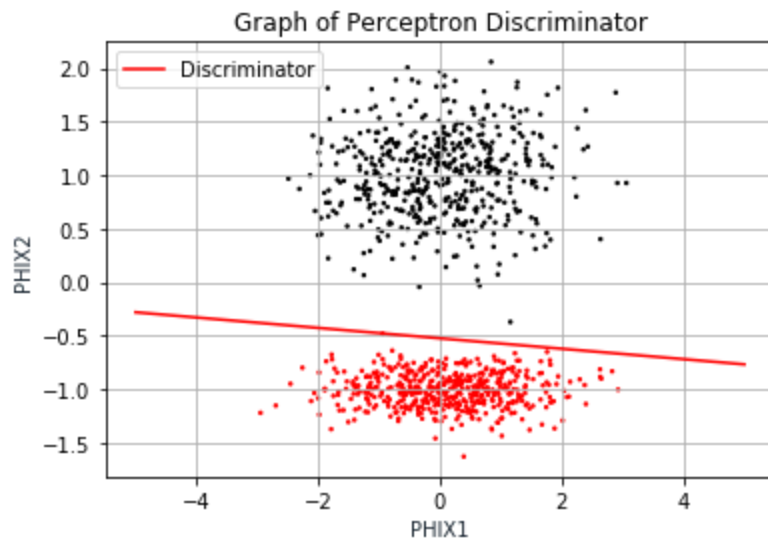


Number of epochs (BGD) = 1000 - Doesn't converge

Error vs number of epochs (BGD):

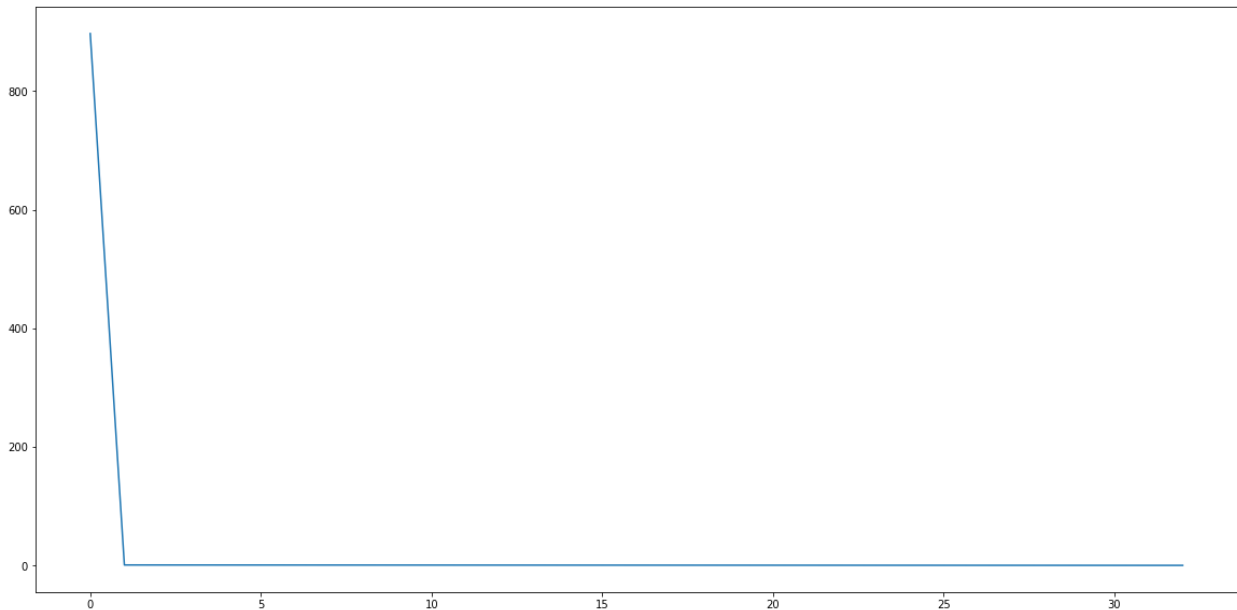


Dataset 3 :



Number of epochs (BGD) = 32

Error vs number of epochs (BGD):

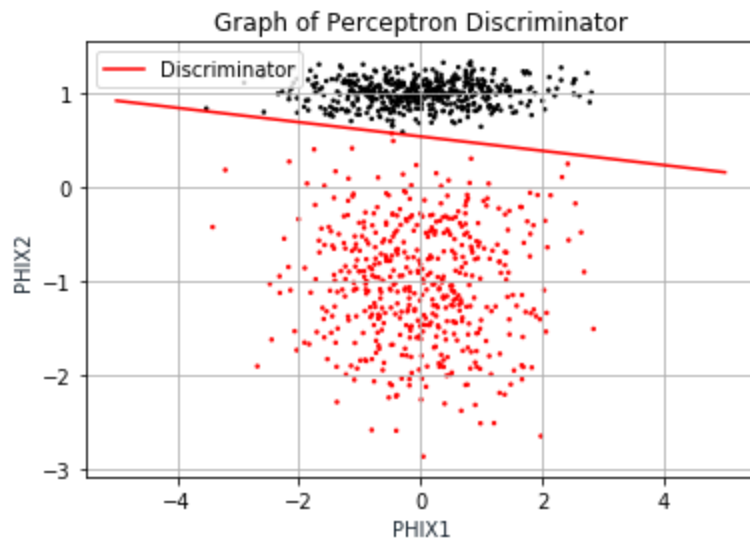


Boxplot comparison of number of epochs on different runs:

STOCHASTIC GRADIENT DESCENT

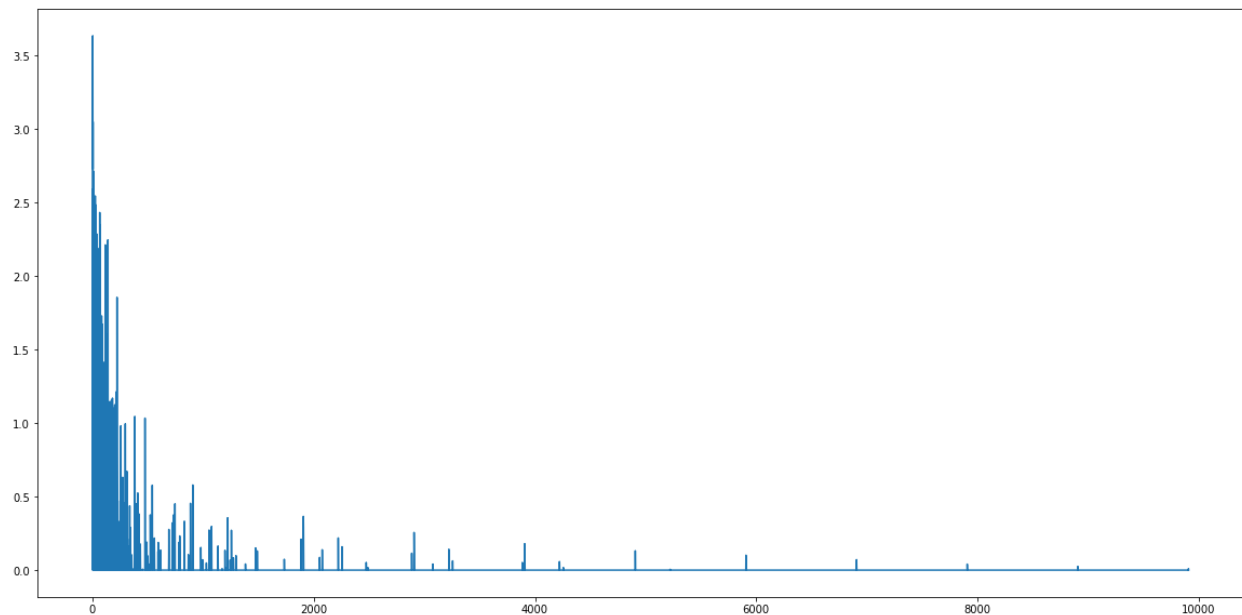
Discriminator after 10e5 epochs/convergence whichever was earlier:

Dataset 1 :

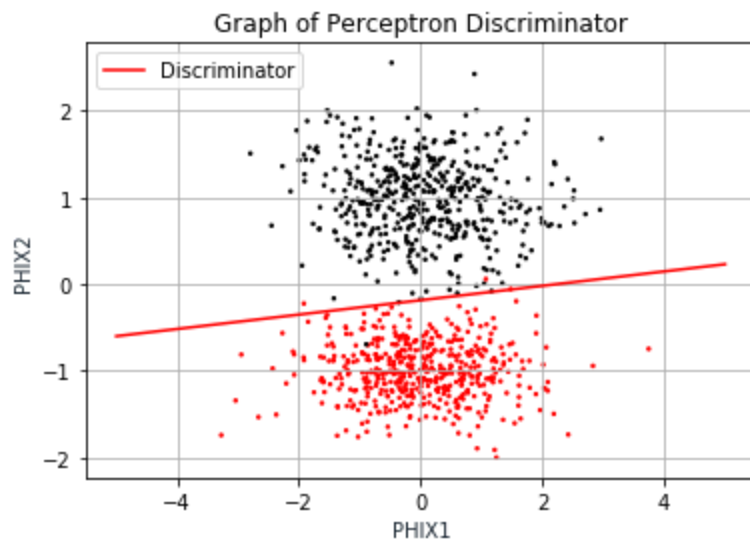


Number of epochs (SGD) = 9908

Error vs number of epochs (SGD) :

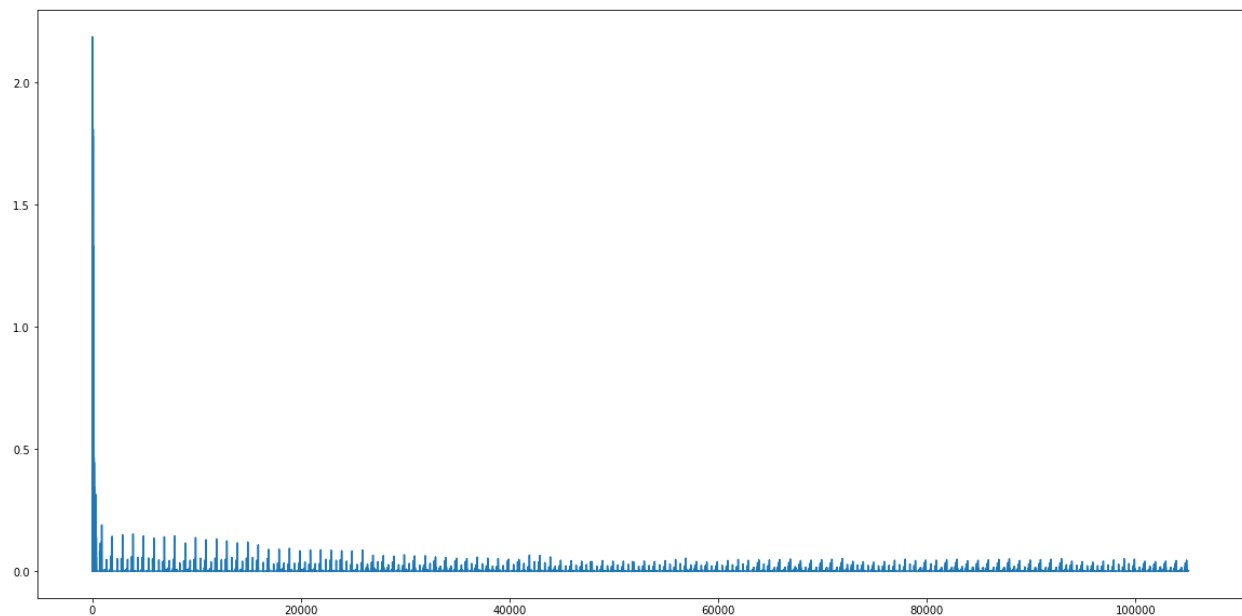


Dataset 2 :

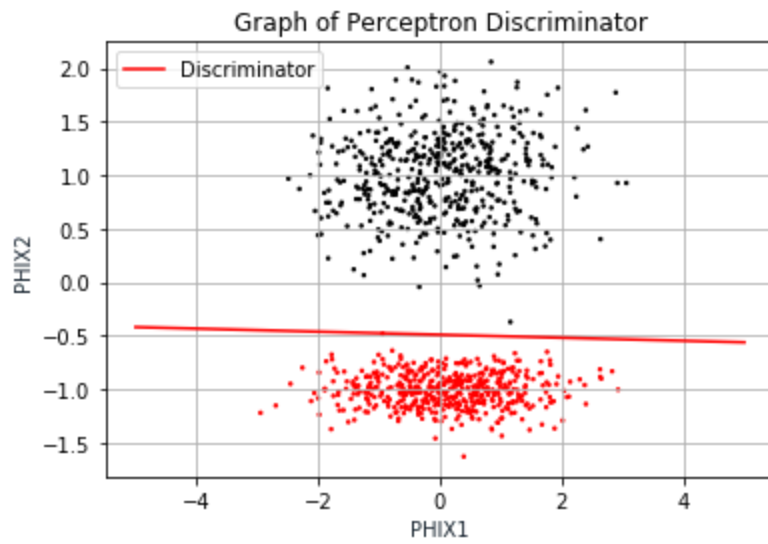


Number of epochs (SGD) = 10^5 - Doesn't converge

Error vs number of epochs (SGD):

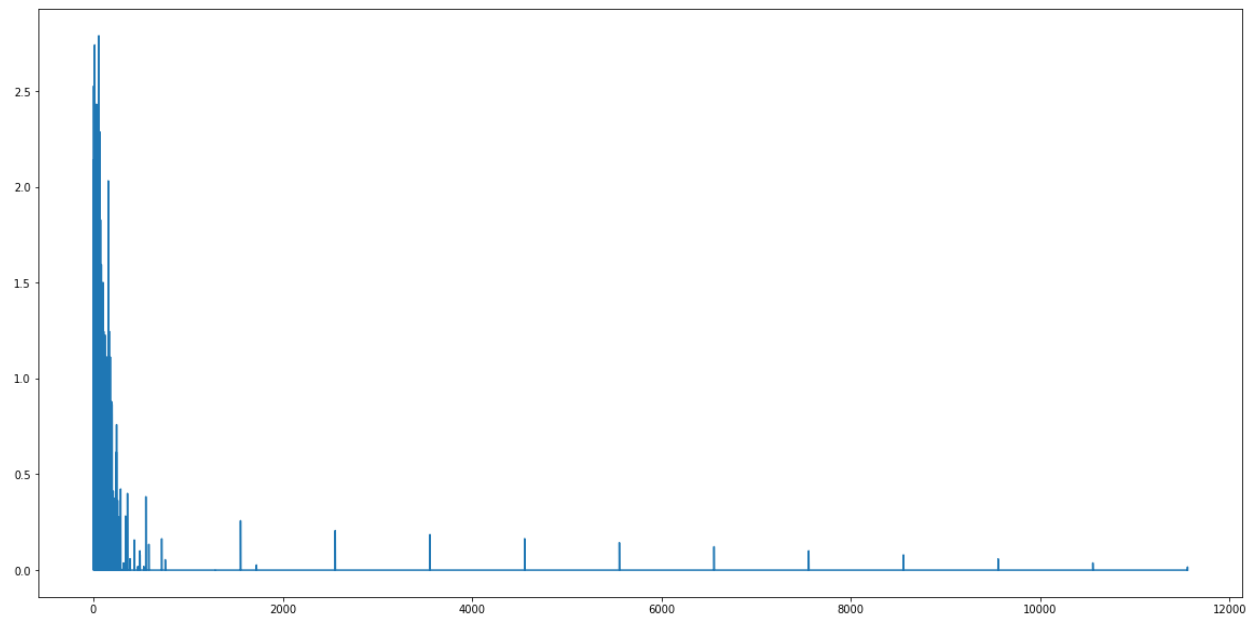


Dataset 3 :



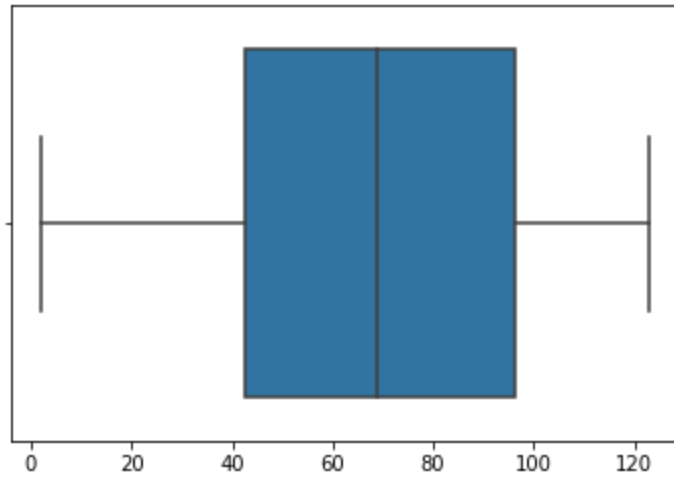
Number of epochs (SGD) = 11556

Error vs number of epochs (SGD):

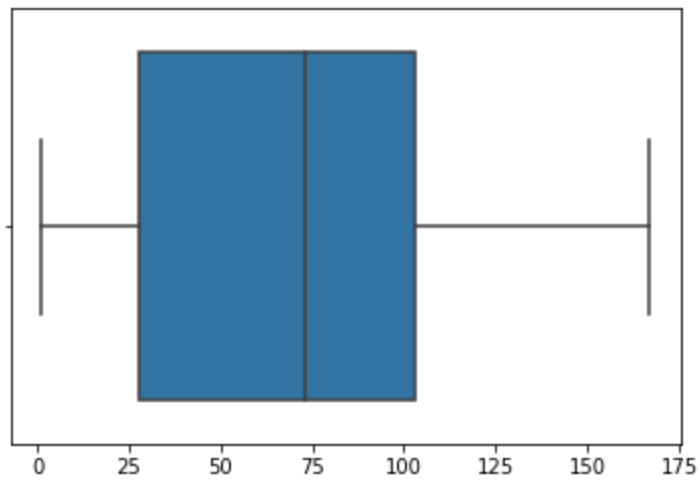


Boxplot comparison of number of epochs on different runs:
Random Initialization

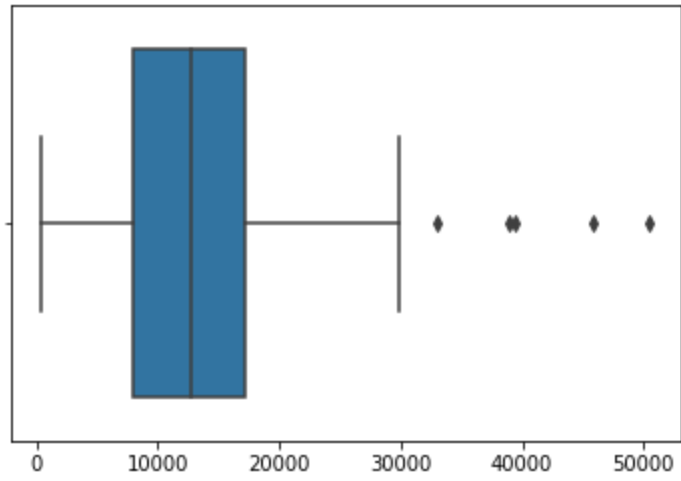
BGD- Dataset 1



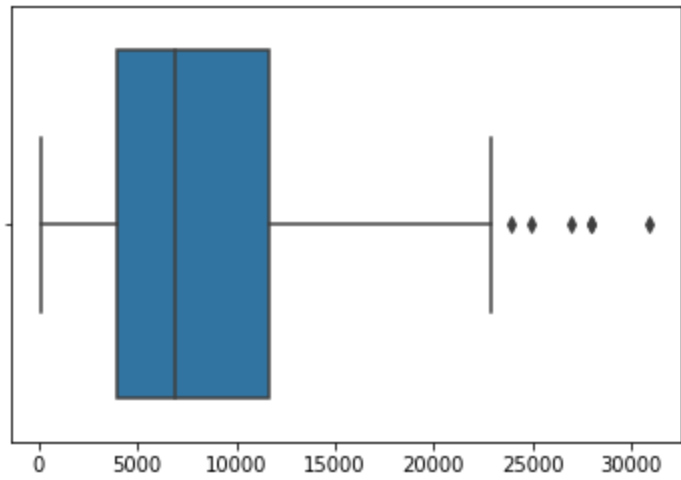
BGD - Dataset 3



SGD - Dataset 1

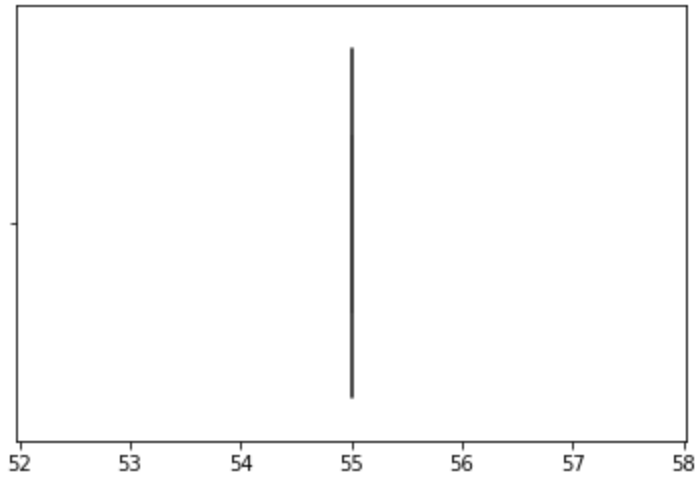


SGD - Dataset 3

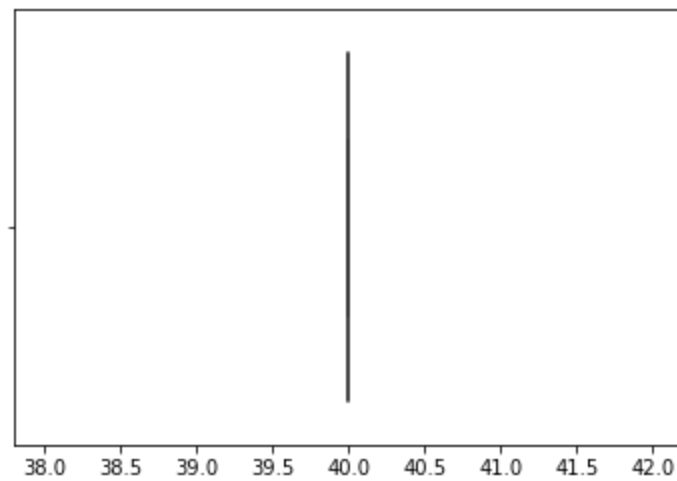


Reshuffling on every run

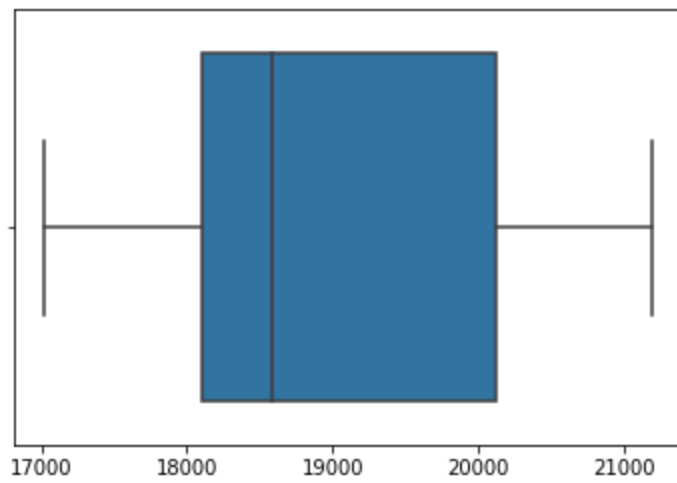
BGD - Dataset 1



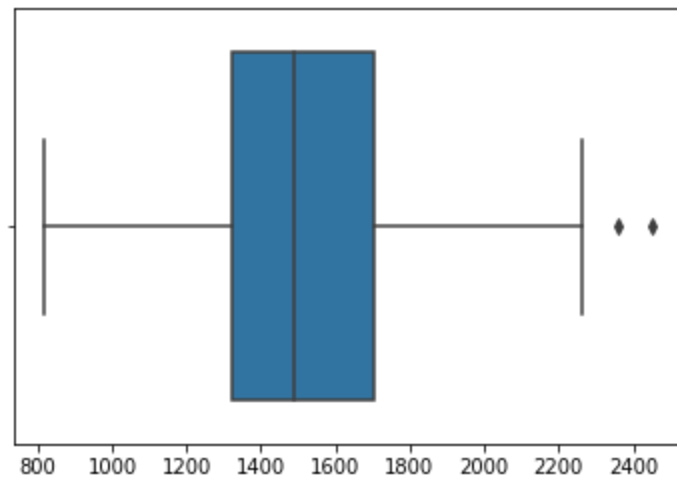
BGD - Dataset 3



SGD - Dataset 1



SGD - Dataset 3



Dataset	Batch Gradient Descent				
	Convergence	Difference observed on reshuffling data on each pass over dataset	Number of epochs to convergence in learning rates (av of 10 random trials)		
			0.01	0.001	0.0001
Dataset 1	Yes. Linearly separable.	No	67.8	121.6	606.2
Dataset 2	No. Not linearly separable.	No	-	-	-
Dataset 3	Yes. Linearly separable.	No	110.2	86.2	442.6

****For effects of random initialization and reordering, please refer to the boxplots.**

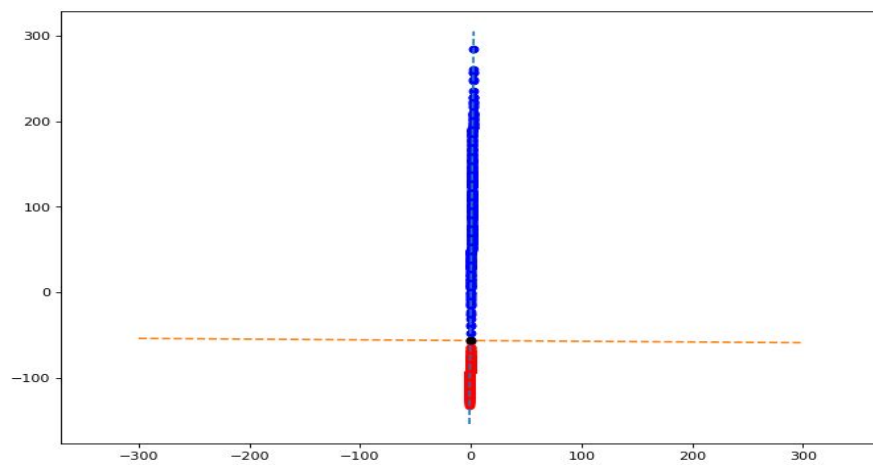
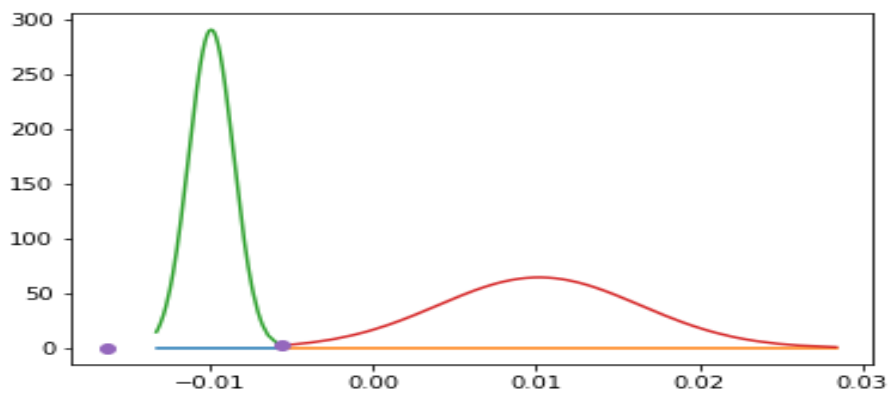
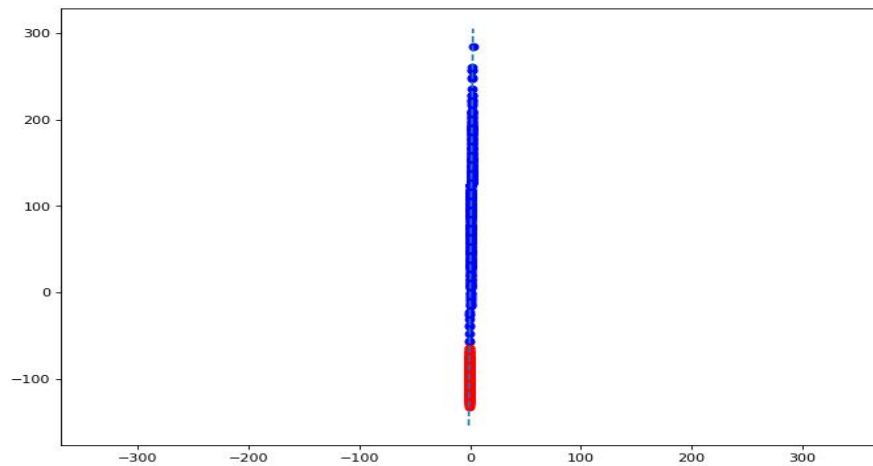
Dataset	Stochastic Gradient Descent				
	Convergence	Difference observed on reshuffling data on each pass over dataset	Number of epochs to convergence in learning rates (average of 10 trials)		
			0.1	0.01	0.001
Dataset 1	Yes. Linearly separable.	Yes	13690.8	19028.8	191334.5
Dataset 2	No. Not linearly separable.	Yes	-	-	-
Dataset 3	Yes. Linearly separable.	Yes	227.8	1678.5	16161.4

****For effects of random initialization and reordering, please refer to the boxplots.**

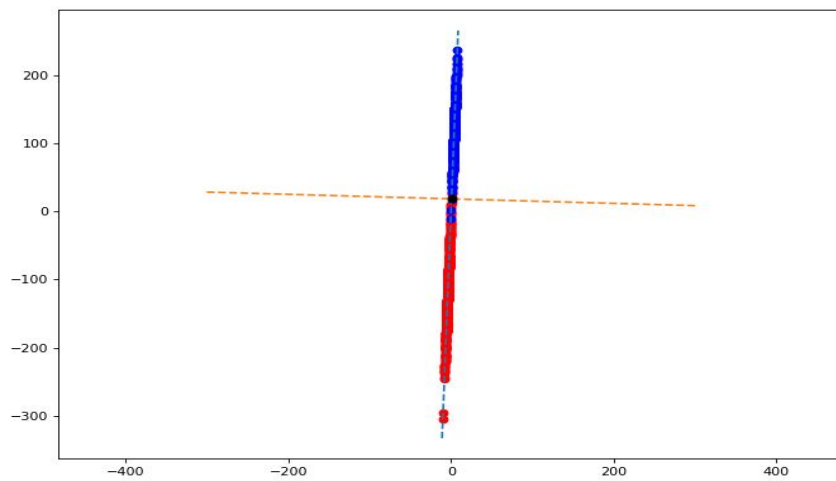
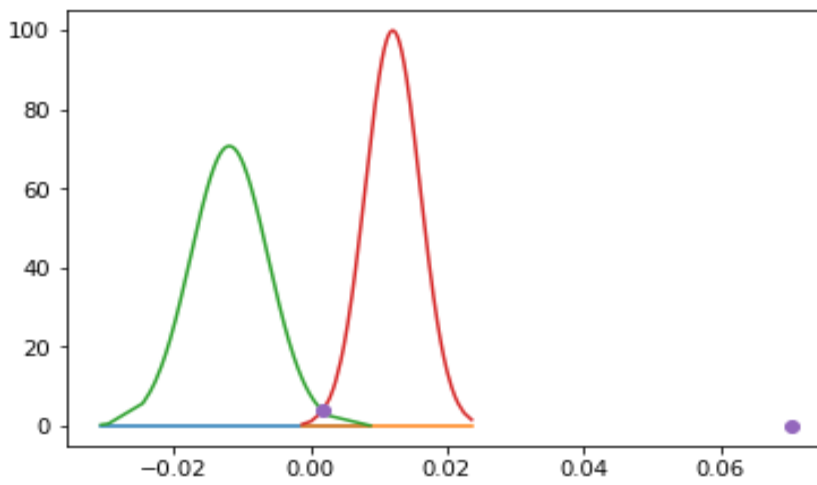
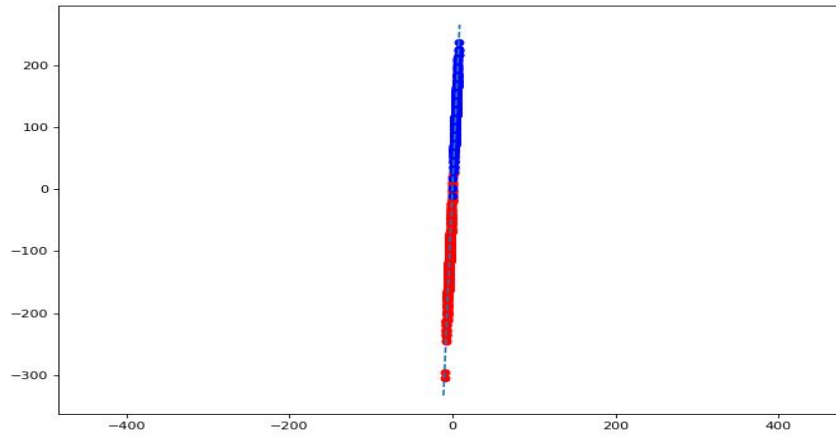
Fisher Linear Discriminant Analysis

Datasets	Direction Vector	Slope of Direction Vector in 2-D	Intersection Point(1-D)	Intersection Point(2-D)
Dataset-1	$-8.280661 \times 10^{-5} i + -0.0098807604 j$	119.3233284305982 2	-0.0056264518786 54835	$-4.7151335602990036 \times 10^{-5} i + -0.005626254304096943 j$
Dataset-2	$-0.0003978597311258453 i + -0.012032862744630085 j$	30.24398249749991	0.00183084803417 78857	$-4.7151335602990036 \times 10^{-5} i + -0.005626254304096943 j$
Dataset-3	$0.0003966623171311423 i + -0.02252931703518684 j$	-56.797220361465094	0.00885568080703 3544	$0.00015589334283101506 i + -0.008854308545658619 j$

Dataset-1



Dataset-2



Dataset-3

