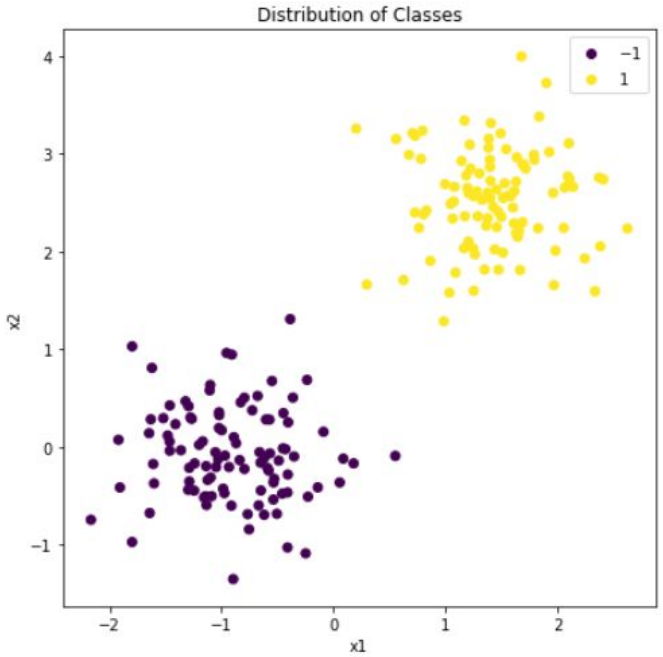
Lab 1a
Artificial Neural Networks and Deep Architectures

25 January 2023

Group 12 - Isabella Rositi and Gustav Thorén

Classification with a single-layer perceptron

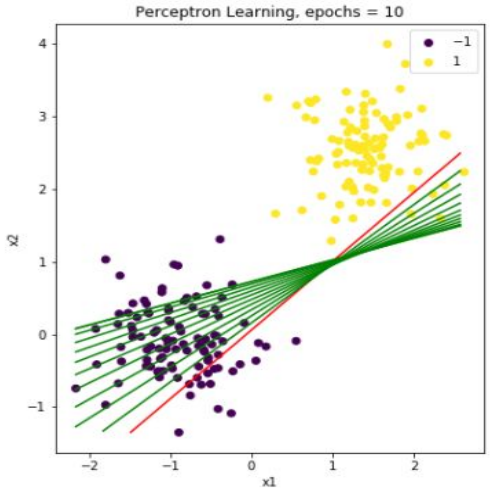


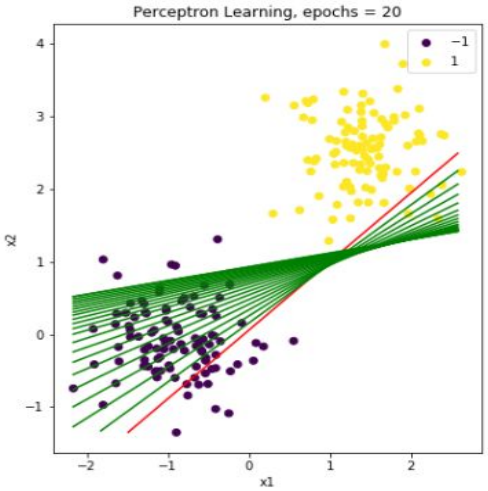
Linearly Separable Data

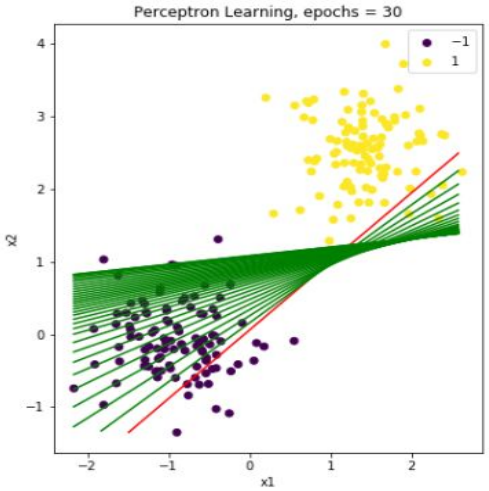
2 classes drawn from two different Gaussian distributions.

	Class A (1)	Class B (-1)
Mean	(1.5, 0.5)	(-1, 0)
Standard Deviation	0.5	0.5

Classical Perceptron Learning





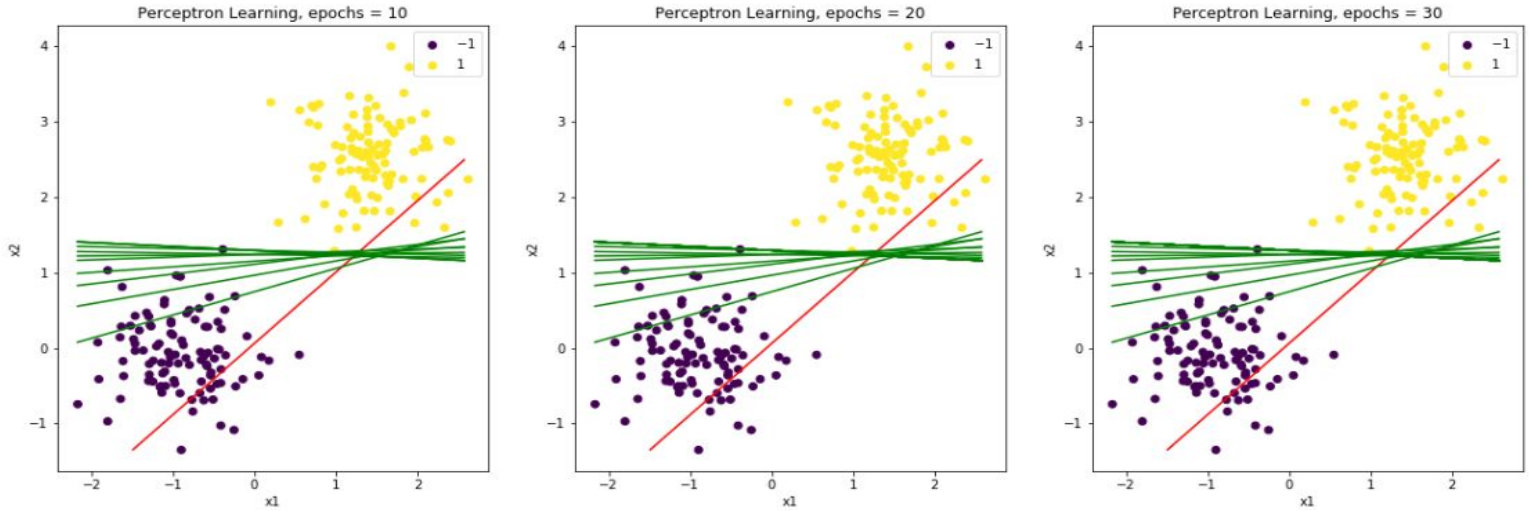


Learning rate = 0.001

Very slow convergence

Epoch	10	20	30
Errors	15	5	2

Classical Perceptron Learning

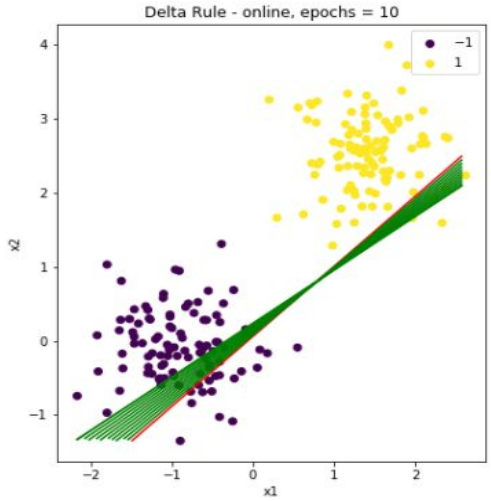


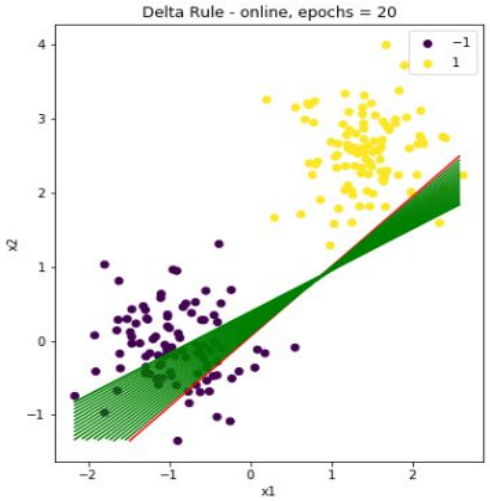
Learning rate = 0.01

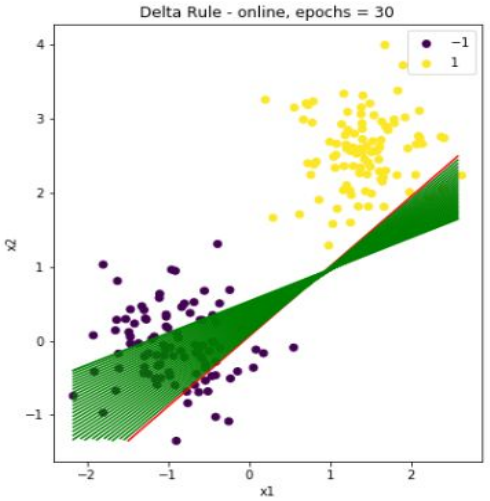
Fast convergence
Premature stopping

Epoch	10	20	30
Errors	0	0	0

Delta Rule - Online mode





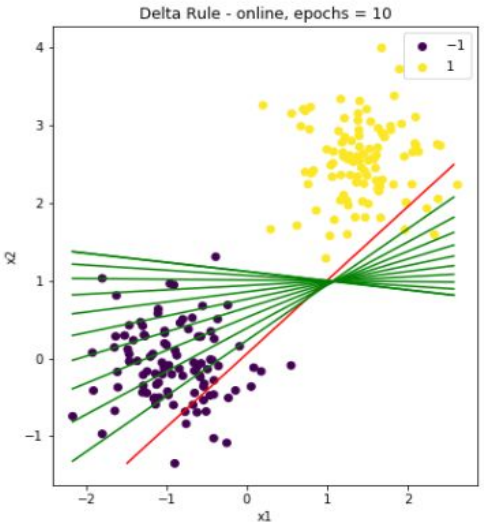


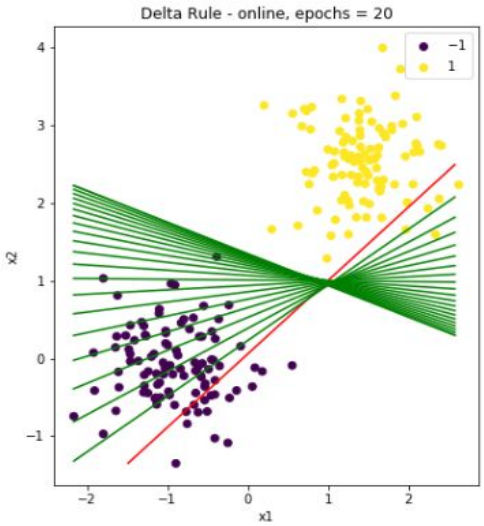
Learning rate = 0.0001

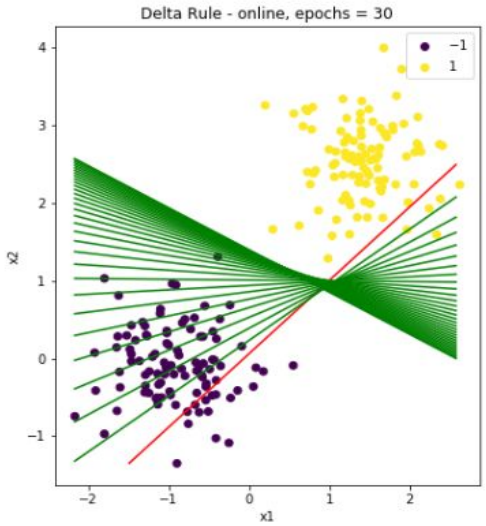
Very slow convergence

Epoch	10	20	30
MSE	1.23	0.78	0.49

Delta Rule - Online mode





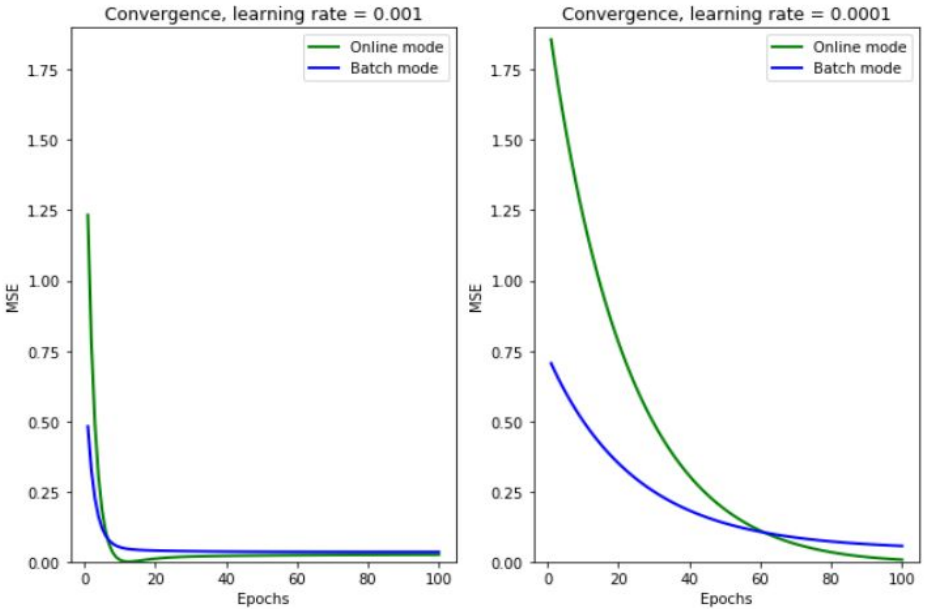


Learning rate = 0.001

Keeps training even after all data points are correctly classified

Epoch	10	20	30
MSE	0.018	0.011	0.006

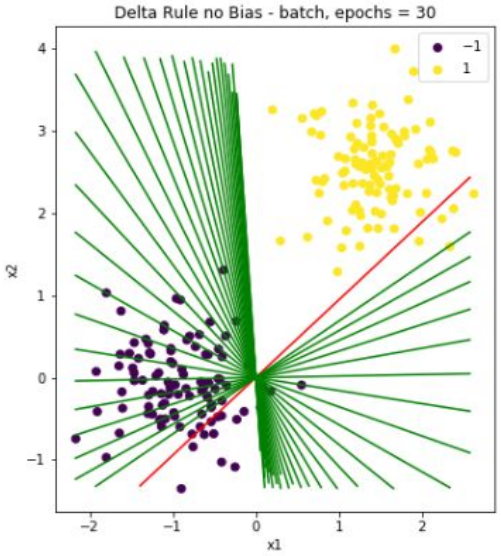
Delta Rule - Online mode vs Batch mode

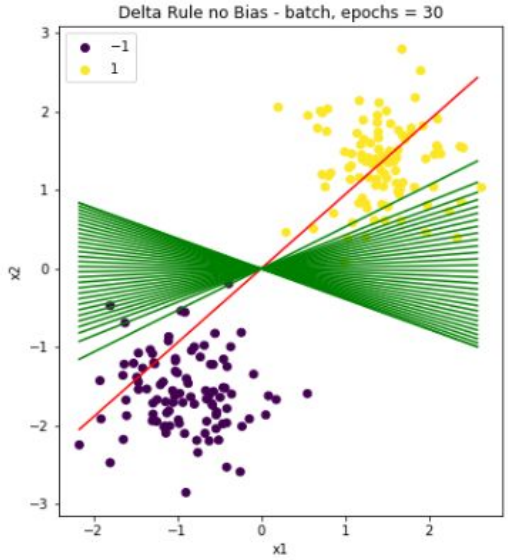


Convergence

- Online mode is faster to converge
- Batch mode produces lower errors at the beginning
- Both methods are susceptible to random initialization of weights

Delta Rule - No Bias



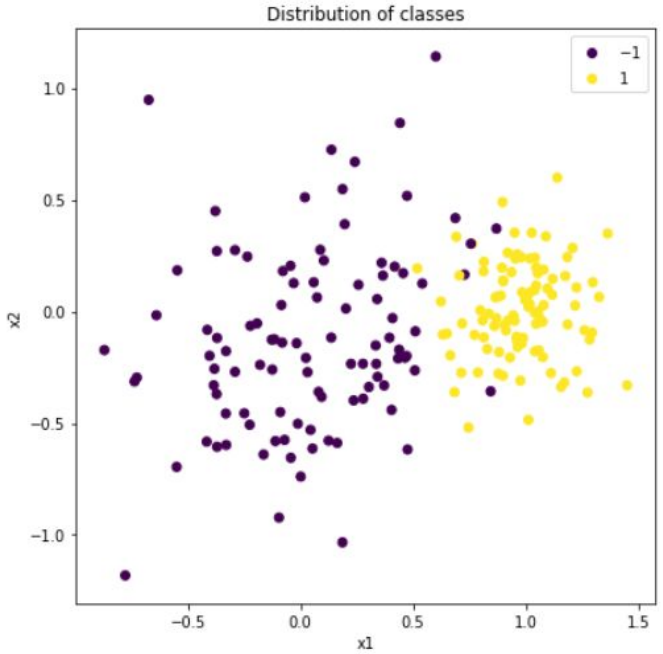


Learning rate = 0.001

If data not spaced correctly in the 2D plane, no optimal solution

Data set	Original	Adapted
MSE	0.091	0.034

Classification of non-linearly separable data

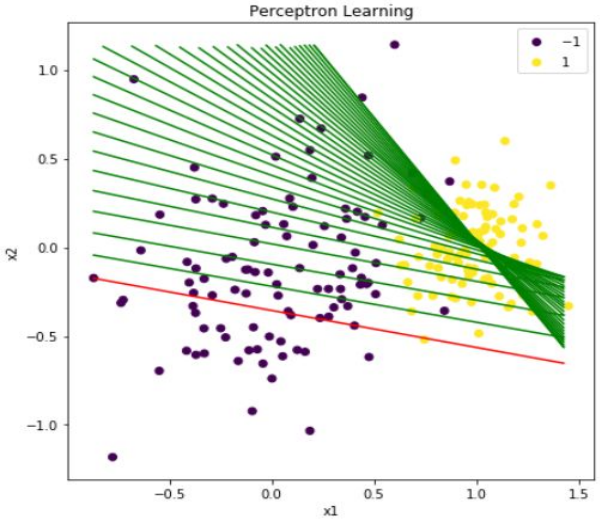


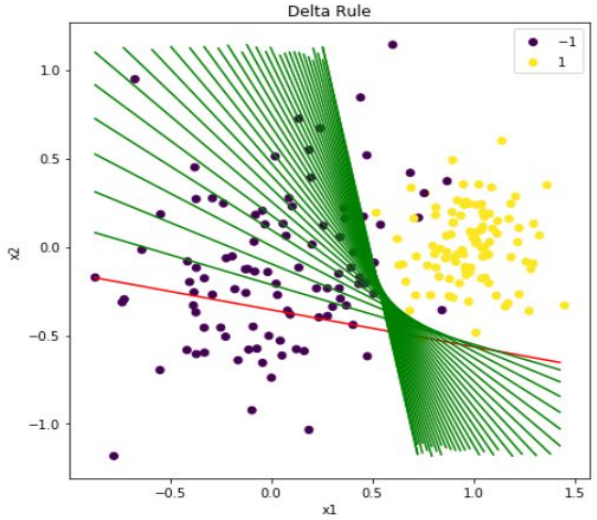
Non-Linearly Separable Data

2 classes drawn from two different Gaussian distributions.

	Class A (1)	Class B (-1)
Mean	(1, 0)	(0, -0.1)
Standard Deviation	0.2	0.4

Classification of non-linearly separable data



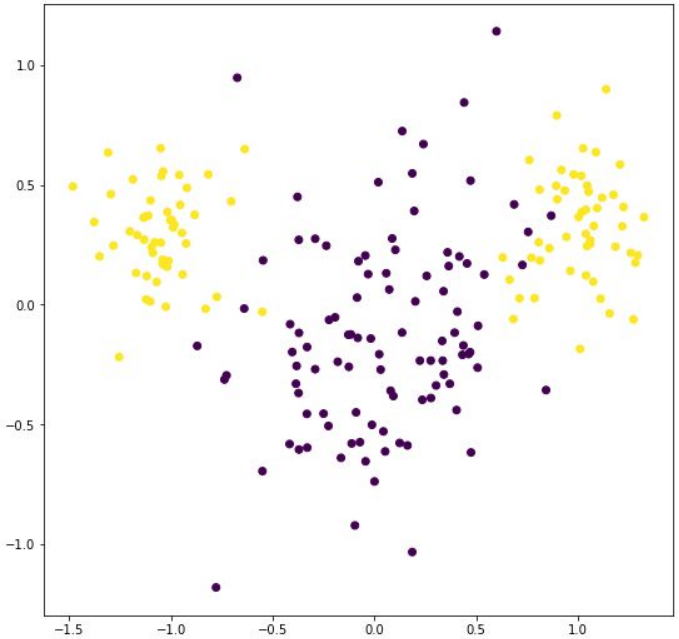


Learning rate = 0.001
Epochs = 30

Limitation: no optimal solution

Rule	Perceptron Learning	Delta Rule
Error	28 Errors	0.033

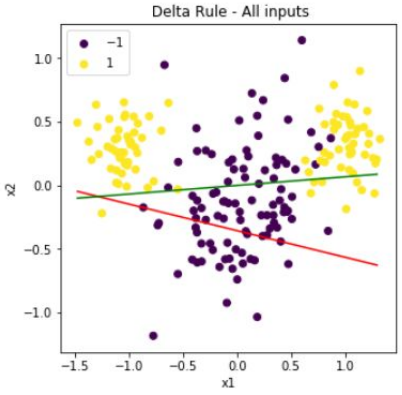
Classification of non-linearly separable data

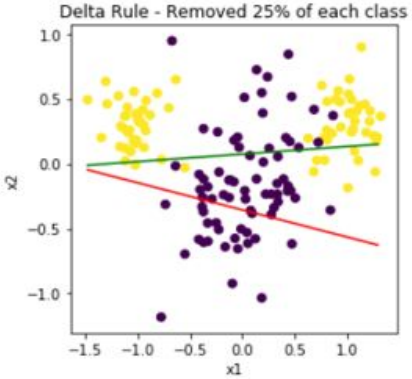


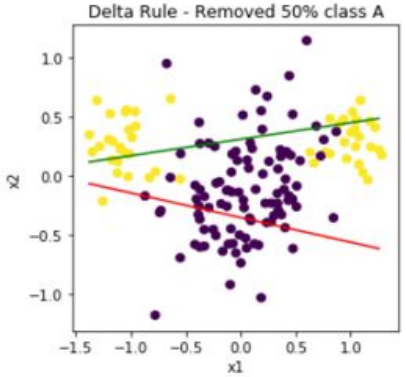
Non-Linearly Separable Data

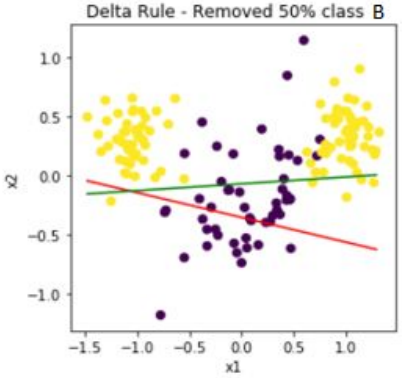
2 classes drawn from three different Gaussian distributions.

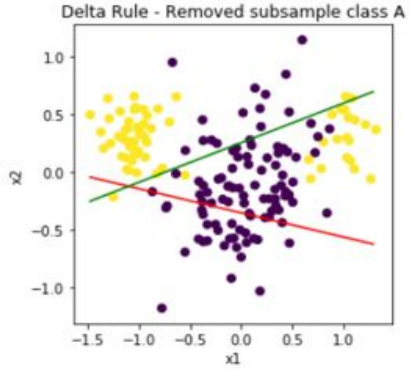
	Class A (1)	Class B (-1)
Mean	50% (1, 0.3), 50% (-1, 0.3)	(0, -0.1)
Standard Deviation	0.2	0.4











Non-Linearly Separable Data

Removed data	Sensitivity	Specificity
No data removed	0.39	0.39
25% class A,B	0.55	0.55
50% class A	0.38	0.57
50% class B	0.6	0.56
20% class A x1<0, 80% of class A x1>0	0.46	0.43

The decision boundary depends heavily on the input data

Final remarks

- Intuitive model
- Limitations of single-layer perceptron

Thank you