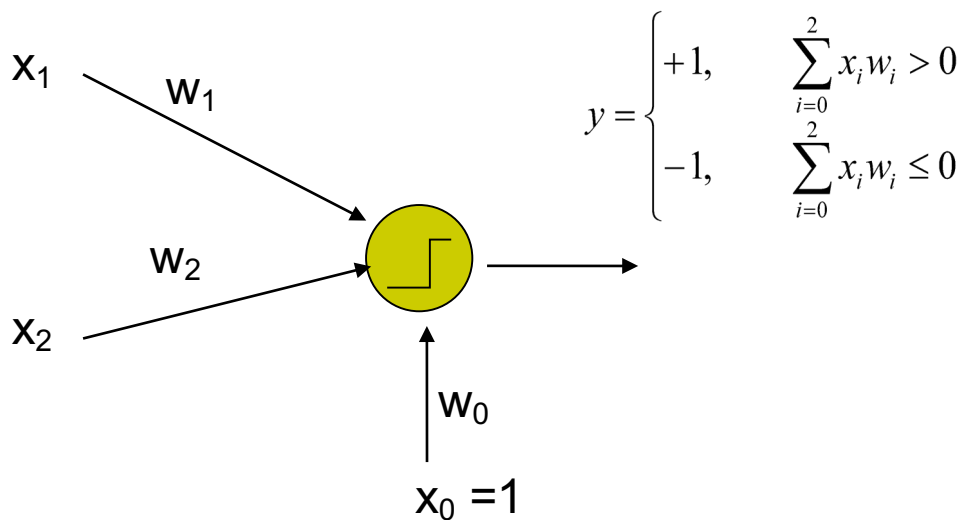


Neural networks and SVM

1) Given a training dataset

x_1	x_2	class
0	0	-1
0	1	+1
1	0	+1
1	1	+1

and a perceptron as follows :



Init: $w_0 = -0.5$, $w_1 = 0.4$, $w_2 = 0.5$ and learning rate $\eta = 0.2$

- Learning a classification model (running by hand) from the training dataset
- Visualizing the training dataset and the classification model in 2D space
- Plotting the ideal SVM model learned from this training dataset

2) Implement the perceptron learning algorithm using **scikit-learn** library in Python. The program requires 4 parameters:

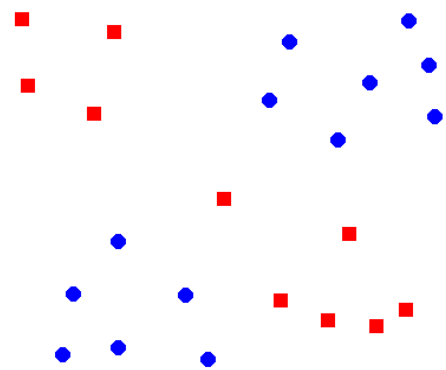
- file name of trainset
- file name of testset
- learning rate η
- maximum number of epochs maxit

The program reports the classification results (accuracy, confusion matrix) with different learning rate η and maximum number of epochs maxit, for 5 datasets:

- Iris (.trn: trainset, .tst: testset)
- Optics (.trn: trainset, .tst: testset)
- Letter (.trn: trainset, .tst: testset)
- Leukemia (.trn: trainset, .tst: testset)
- Fp (.trn: trainset, .tst: testset)

3) Implement the training program using **SVC** in **scikit-learn** library to classify the dataset

X₁	X₂	Class
0.204000	0.834000	0
0.222000	0.730000	0
0.298000	0.822000	0
0.450000	0.842000	0
0.412000	0.732000	0
0.298000	0.640000	0
0.588000	0.298000	0
0.554000	0.398000	0
0.670000	0.466000	0
0.834000	0.426000	0
0.724000	0.368000	0
0.790000	0.262000	0
0.824000	0.338000	0
0.136000	0.260000	1
0.146000	0.374000	1
0.258000	0.422000	1
0.292000	0.282000	1
0.478000	0.568000	1
0.654000	0.776000	1
0.786000	0.758000	1
0.690000	0.628000	1
0.736000	0.786000	1
0.574000	0.742000	1



The program visualises datapoints with the separation boundary of the classifier.

4) Write a program using **SVC** in **scikit-learn** library in Python. The program requires 4 parameters:

- file name of trainset
- file name of testset
- hyper-parameters of SVC (kernel function, kernel parameter, constant C)

The program reports the classification results (accuracy, confusion matrix) with different hyper-parameters, for 5 datasets:

- Iris (.trn: trainset, .tst: testset)
- Optics (.trn: trainset, .tst: testset)
- Letter (.trn: trainset, .tst: testset)
- Leukemia (.trn: trainset, .tst: testset)
- Fp (.trn: trainset, .tst: testset)