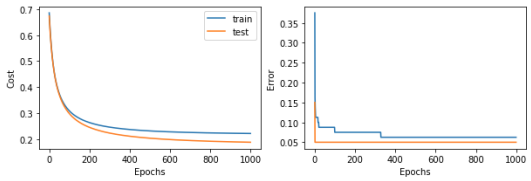
Practical work 04 – **Model Selection**

Banfi Gregory (banfigre@students.zhaw.ch)
Benjamin Kühnis (bkuehnis@hsr.ch)

Exercice 1 Hyper-Parameter Tuning

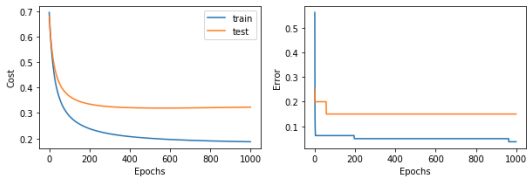
a) See student_data_set.ipynb
b) We implemented 5-fold validation to find out which hyper-parameters best fit the model and give the lowest error. Here below the results. On each validation we tried different polynomial degrees, from 0 to 4.

Polynomial Degree: 0



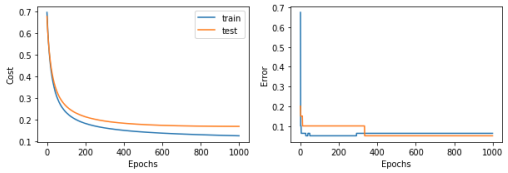
0.22128240625971327 0.1875038541178367 0.0625 0.05
[[-1.29007303 2.55129719 2.91134937]]

Polynomial Degree: 1



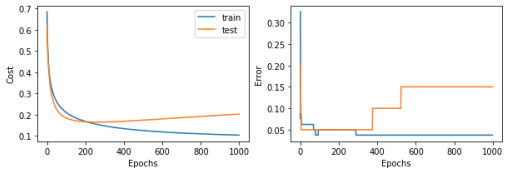
0.1869792530638515 0.32294953652839414 0.0375 0.15
[[-1.22449255 2.77405073 3.14319989]]

Polynomial Degree: 2



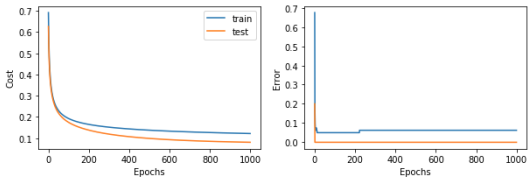
0.12442045877915331 0.16795384041492 23 0.0625 0.05
[[-0.46500588 2.51952398 3.2555084 0.29807519 1.57405383 -1.03662251]]

Polynomial Degree: 3



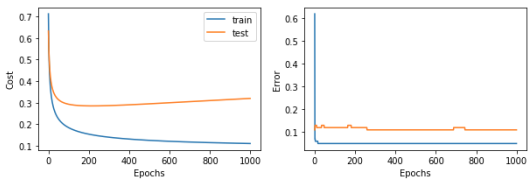
0.10355933914953179 0.20260451103768662 0.0375 0.15
[[0.1879026 1.47553094 2.57844031 -0.89466764 2.09905875 -1.36657599
1.08836372 0.91627728 0.88655496 0.74078773]]

Polynomial Degree: 4

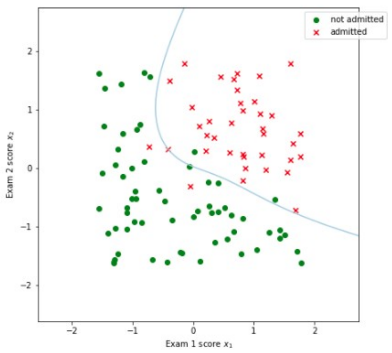


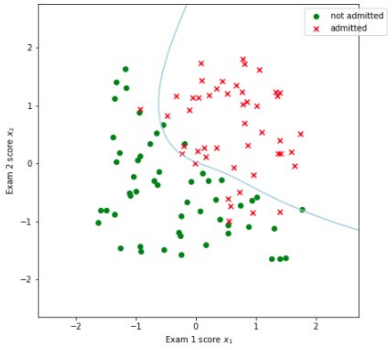
0.1220275423889674 0.08082937093336799 0.0625 0.0
[[-0.25779451 0.61356609 2.40240572 -0.52990882 0.91601862 -0.37662823
1.6038414 0.55284175 0.63381847 1.27391304 0.1890501 0.46192417
-0.73309565 1.08587773 -0.39225376]]

Lowest error 0.0
With degree 4



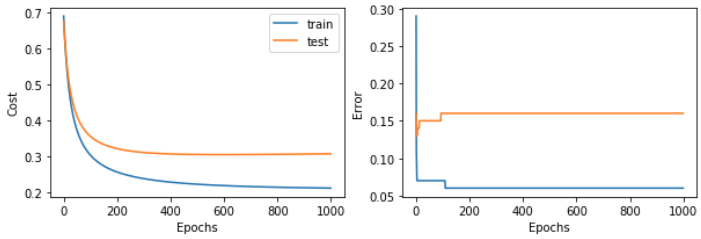
0.11088503442767504 0.32007591013233266 0.05 0.11
[[-0.06598332 1.01504668 2.46368407 -0.30982128 0.9220483 -0.43528822
1.56417326 0.58766596 0.55350278 1.36571142 0.06239889 0.50959773
-0.70528562 1.05319363 -0.47925387]]





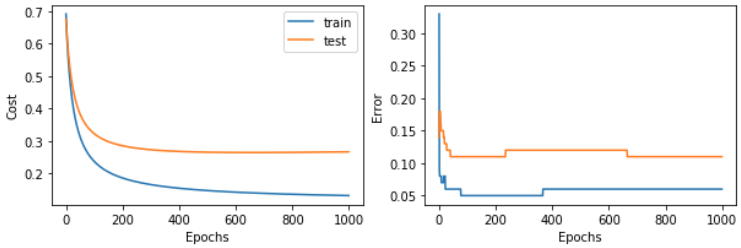
c) In the images we see different plots with different parameter degrees. The best result we got with a polynomial of degree2. Analyzing the plots we can see that when we increase the degree above 3 a larger error and cost with respect to the test set. On the other side the error for the training set still decreases because the model fit the data very well. The conclusion is that we are going into overfitting, the model is to depending on the training data and does not perform so good on the test data.

Degree1



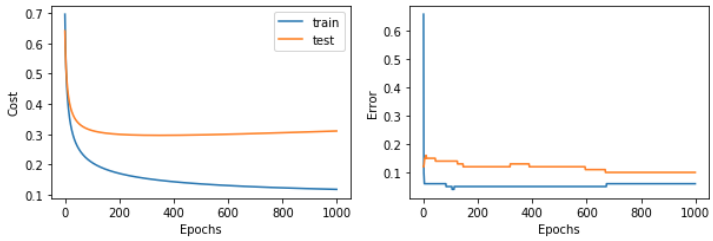
0.21309812918484897 0.308135926751895 0.06 0.16
[[-1.39511096 2.64552943 2.90021686]]

Degree2



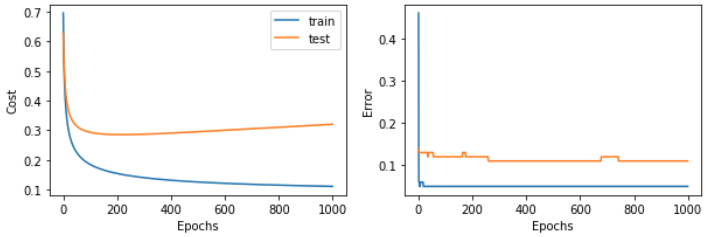
0.13189786119906438 0.26650221025769716 0.06 0.11
[[-0.19077385 2.60705662 3.18773849 0.01280459 1.60616982 -0.86082337]]

Degree3



0.11760207984434126 0.3104088355947956 0.06 0.1
[[0.09810307 1.14542978 2.47878121 -0.61252193 1.85580144 -1.3286611
1.37466788 0.77834139 0.89519369 0.73945932]]

Degree4



0.11087100824609009 0.3200373514760757 0.05 0.11
[[-0.0669957 1.0081198 2.46194036 -0.30964605 0.91749538 -0.42963131
1.5745728 0.6024594 0.55439898 1.36104911 0.06127144 0.51359281
-0.70175256 1.05701703 -0.47898942]]

Exercice 2 Confusion Matrix and Performance Measures

1. Write a function to take classification decisions on such outputs.
    a. See confusion_matrix_ex-stud.ipynb
2. What is the overall error rate of the system ?
    a. 0.10729999999999995
3. Compute and report the confusion matrix of the system A.
    a. See confusion_matrix_ex-stud.ipynb
4. What are the worst and best classes in terms of precision and sensitivity (recall) ?
    a. best: [nr 1] with 0.979
    b. worst: [nr 8] with 0.792
5. In file ex1-system-b.csv you find the output of a second system B. What is the best system between A and B in terms of error rate and F1.
    a. System A
        recall : 0.8927
        precision : 0.8927
        accuracy : 0.8927
        f1-score : 0.8927
    b. System B
        recall : 0.9613
        precision : 0.9613
        accuracy : 0.9613
        f1-score : 0.9613