# Universidade da Beira Interior
# Departamento de Informática
# Inteligência Artificial

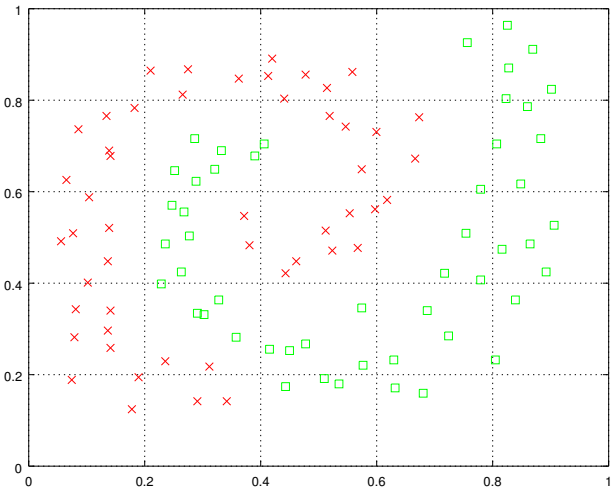
Practical exercises 7

Ano letivo 2025-26

## Exercises

1. We have received the data file `size_price.csv` that contains house prices as a function of their sizes. We want to know what is the expect price of a house with $350m^2$. Solve this problem implementing the linear regression predictor. The line equations are on the theoretical class slides.

2. In this exercise we will use an SVM from the scikit-learn library. There are several implementations, we want to use the SVC. Check here for examples. Create the SVC with the following parameter `svm.SVC(C=1000)` instead of using the default parameter as appears in the examples. Use the SVC to classify the problem of the two spirals:




Train your classifier on the data from file `espiral_train.csv` and estimate the generalization error on the data from the `espiral_test.csv`.

3. Study the impact of the $C$ parameter in the accuracy of the SVM classifier in the previous problem. Vary the $C$ in $[1, 5, 10, 50, 100, 500, 1000]$ and create a plot where the X-axis represents the $C$ value and the Y-axis represents the accuracy in the test set of the SVM classifier. What can you conclude?

4. Adapt the code in the theoretical class slides to implement a naive Bayes classifier to solve the spirals problem.

5. Implement the $k$-NN classifier (**do not** use the implementation from the scikit-learn library). Use it to solve the spirals problem with $k = 1$ and $k = 7$.

6. Complete the following table with the test error rates you obtained in exercises 2 to 4. You should also compute the training set error rates for those exercises to complete the table. What do you conclude from this table?

| Classifier | Train | Test |
|---|---|---|
| Naive Bayes | | |
| SVM | | |
| 1-NN | | |
| 7-NN | | |