Lab 2: Classification – KNN and Naive Bayes

Goals:

- KNN
- Naïve Bayes
- Accuracy and Confusion Matrix
- Training strategies

1. Load the _iris_ data. Train a classifier using the KNN algorithm. Keep all parameters with their default values Use percentage split with 70%. Keep class as the class attribute.

    a. What is the accuracy achieved?

    b. And the number FP and FN for Iris-virginica?

    c. And for Iris-setosa?

    d. Compare the results achieved through Cross-validation with 10 folds.

2. Repeat the process with naïve Bayes classifier.

    a. Answer the previous questions again.

    b. Compare the results achieved with KNN and Naïve Bayes.

3. Load the _glass_ data. Train a classifier using the KNN algorithm. Keep all parameters with their default values, but the number of neighbours (KNN). Keep type as the class attribute

    a. What is the accuracy with 1 neighbour?

    b. And with 5, 10, 15, 50 and 100 neighbours?

    c. How does the accuracy change?

    d. Compare the results achieved through Cross-validation with 10 folds.

    e. Is any of the models in overfitting?

4. Repeat the process with naïve Bayes classifier.

    a. Answer the previous questions again.

    b. Compare the results achieved with KNN and Naïve Bayes.