**Labwork 1**

Sample Dataset:

COIL20: <https://academictorrents.com/details/1d16994c70b7fff8bfe917f83c397b1193daee7f>

ORL Dataset: <https://www.kaggle.com/datasets/tavarez/the-orl-database-for-training-and-testing>

**Without Feature Selection**

1. Process all images in a dataset and transform them into a matrix.
2. Build a classifier (Random Forest, SVM, Decision Tree, etc.) on all features.
3. Evaluate the performance of the model using quality measures (accuracy, precision, recall, F1-score, …)

**With Feature Selection filters**

1. Apply filter techniques to select relevant features.
2. Re-run the classifier and compare the performances before and after feature selection.

**With wrappers**

1. Remove the least significant features from the initial features and verify the improvement in performance.
2. Repeat the process until a certain number of iterations has been reached or the performance measure has reached a threshold.

Discussion:

1. How did feature selection affect model accuracy? Did the model improve or decrease after removing features?
2. Did feature selection reduce overfitting? Check the difference between training and test accuracy.