Assignment #6

Course: *Machine learning*
Date: *December 15th, 2023*

Assignment

In this assignment, you will learn about kernel methods.

Download the Breast Cancer Wisconsin dataset `https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data` and perform all the necessary data pre-processing.

Use the Scikit-learn implementation of SVM (the SVC() function) to predict if the tumor is malignant or not. Try different kernels.

Try different:

- kernels
- regularization parameters
- different σ values when using the RBF kernel (in Scikit-learn this is the γ parameter).

  Find the best values for the parameters and do a grid search. Tip: take a look at the GridSearchCV() function.

Test your models.

Be careful not to use the same data for any stage of training and testing your models.

Implement kernel ridge regression to model the data from Učilnica.

Implement different kernels. Train and test your kernel regression model. Which is the most suitable kernel?

Plot the model fit to your train data and to the predictions.

Ridge regression is a linear model. What kind of predictions did you get? Comment.