

Machine Learning

Exercise 3: Usage of different regression algorithms in python

Prof. Dr. Thomas Kopinski

November 2, 2022

Abstract

The goal of this exercise is to use python and scikit-learn to perform logistic and polynomial regression on different datasets.

Task 1: Getting familiar with the regression algorithms

- Further information and examples about the implementation of logistic regression can be found in [this](#) jupyter notebook.
- Additional information about polynomial regression in python can again be found [here](#).

Task 2: Logistic Regression

- Download the MNIST Dataset with the scikit-learn function:
`sklearn.datasets.fetch_mldata("MNIST original")`
- look into the data and analyze the data structure
- plot some samples of the data with matplotlib
- split the data into a training and test dataset
- classify the MINST dataset with a logistic regression model
- evaluate the model performance using the accuracy
- display a sample of the wrong predictions to understand the models shortfalls

Task 3: Polynomial Regression

- read the data
`x = np.arange(0, 30)`
`y = [3, 4, 5, 7, 10, 8, 9, 10, 10, 23, 27, 44, 50, 63, 67, 60, 62, 70, 75, 88, 81, 87, 95, 100, 108, 135, 151, 160, 169, 179]`
- use matplotlib to display the data
- make polynomial features of degree 2 from the input data using the `sklearn.preprocessing.PolynomialFeatures` function
- use a linear regression model to fit to the transformed data
- plot the model together with the data and evaluate the models performance
- use the `numpy.polyfit` function to fit a model to the data and compare it to the first model