Statistical Machine Learning – Week 12

Sascha Jecklin

HS 2022/23

1 Task

Support Vector Machines (SVM) is a good off-the-shelf classifier which performs well even if the predictors are very high dimensional. SVM had their best days in the 90s before the beginning of the deep learning age. Even today they can still be used for advanced tasks such as face recognition. However, with such tasks it is difficult to grasp the concept behind this method, so we will apply SVM to much simpler data in this lab.

We will first take a look at Maximum Margin Classifier and Support Vector Classifier before we get to the Support Vector Machine.

- Open `week12_exercise.py` and try to get an overview of the provided code. In the main function three artificial datasets are generated. Try to get an understanding of the data distribution by creating a scatter-plot. For which dataset are the classes linearly separable?
- Complete all Todo's in the function `exercise_1(x, y)`. Instantiate and fit a SVC object with a linear kernel. Display a scatter-plot of the training dataset and illustrate the decision boundary, margin and support vectors by calling the `plot_svc_decision_function` function. How many support vector do you expect?
  Hints:
  - Make use of sklearn's SVC class and set the parameter `kernel` to `linear`.
- Complete all Todo's in the function `exercise_2(x, y)`. Fit a linear SVC model for different values of the hyperparameter `C` and display the training dataset, decision boundary, margin and support vectors for each value by calling the `plot_svc_decision_function` function. What do you notice? How does the value of C affect the model flexibility or the number of support vectors? How does the model bias change as the number of support vectors increases?
  Hints:
  - Make use of sklearn's SVC class and set the parameter `kernel` and `C` accordingly.
- Finish exercise 3 which is prepared in the function `exercise_3(x, y)`. In many cases the data is not separable by a line / plane. However, by appropriately increasing the dimension of the observations, a dataset can be made linearly separable. Try to make the dataset of exercise 3 linearly separable by adding a third dimension to the given two dimension. Display the 3D dataset with the provided `Axes3DSubplot` object. Hints:
  - Test the following functions as a third dimension: $z = x^2 + y^2$ and $z = e^{-x^2-y^2}$
  - You only need to call the `scatter` method of the `ax1` object to display a three-dimensional scatter-plot.

- Use function `exercise_4(x, y)` for exercise 4. Instantiate and fit a SVC object with a RBF kernel and a polynomial kernel of degree 2. For both models plot the dataset, decision boundary, margin and support vectors by calling the `plot_svc_decision_function` function. Play with the hyperparameters `C` and `gamma`. How does each hyperparameter affect the decision boundary?
  Hints:
  - Make use of sklearn's SVC class and set the parameter `kernel` to `rbf` or `poly`.

Comments

- Take a look at sklearn's documentation of the SVC class.
- This Stackoverflow answer and this short PDF may give you a better understanding of the kernel method.