Machine Learning Course 2024 Spring: Experiment 2

March 26, 2024

1   Overview

In this experiment, you will implement linear/non-linear Support Vector Machine (SVM) models on given data set, and deepen the knowledge and understanding of the hyperparameters and kernel functions involved through visualization.

2   Dataset Description

The datasets are ready for you and are stored in the file *'T1_data.csv'* and *'T2_data.csv'*. Each row in the files is a sample, with the last column being the label and the other columns being the features.

3   Task Decription

3.1   Task 1

Given the *T1* dataset, implement the **linear SVM** model. Visualize the decision boundary of the model. Tune the value of hyperparameter C (regularization parameter), and observe the change of decision boundary.

3.2   Task 2

- Given the *T2* dataset, implement the **non-linear SVM** model with **Gaussian kernel**. Visualize the decision boundary of the model. Tune the value of hyperparameter C and γ, and observe the change of decision boundary.

- Tune the hyperparameter C and γ of the SVM model using a grid search. Use 21 values for C and γ in $[2^{-10}, 2^{-9}, \ldots, 2^9, 2^{10}]$ on a log scale. Report accuracy under 5-fold cross validation and the optimal hyperparameter values.

3.3 Task 3 (Optional)

Try to implement SVM without using existing machine learning libraries (**sklearn**, **LIBSVM**, **LIBLINEAR**, etc.).

4 Requirements

- Implement learning algorithms with **Python**. You can use integrated libraries such as **sklearn**, **LIBSVM**, **LIBLINEAR**, etc.

- Record the procedure of experiments in **Jupyter Notebook**. It is recommended to use **Markdown** in Notebook to properly explain the procedure.

- Your submission includes:

  - source code: python notebook (.ipynb);

  - 'report.pdf' file : your report. Refer to the report template for the format.