

Machine Learning

COMP 5630/ COMP 6630/ COMP 6630 - D01

Instructor: Dr. Shubhra (“Santu”) Karmaker

TA 1: Dongji Feng

TA 2: Souvika Sarkar

Department of Computer Science and Software Engineering

Auburn University

Fall, 2022

October 12, 2022

Assignment #6

Kernels and Support Vector Machines (SVM)

Submission Instructions

This assignment is due Tuesday, October 25, 2022, at 11:59 pm. Please submit your solutions via Canvas (<https://auburn.instructure.com/>). You should submit your assignment as a PDF file. Please do not include blurry scanned/photographed equations as they are difficult for us to grade.

Late Submission Policy

The late submission policy for assignments will be as follows unless otherwise specified:

1. 75% credit within 0-48 hours after the submission deadline.
2. 50% credit within 48-96 hours after the submission deadline.
3. 0% credit beyond 96 hours after the submission deadline.

Tasks

1 Kernel computation cost [30 pts]

1. [10 Points] Consider we have a two-dimensional input space such that the input

vector is $x = (x_1, x_2)^T$. Define the feature mapping $\phi(x) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)^T$. What is the corresponding kernel function, i.e., $K(x, z)$? Do not leave $\phi(x)$ in your final answer.

2. [20 Points] Suppose we want to compute the value of the kernel function $K(x, z)$ from the previous question, on two vectors $x, z \in \mathbb{R}^2$. How many additions and multiplications are needed if you
 - (a) [10 Points] Map the input vector to the feature space and then perform the dot product on the mapped features?
 - (b) [10 Points] Compute through the kernel function you derived in question 1?

2 Kernel functions [30 pts]

Consider the following kernel function:

$$K(x, x') = \begin{cases} 1, & \text{if } x = x' \\ 0, & \text{otherwise} \end{cases}$$

1. [10 Points] Prove this is a legal kernel. That is, describe an implicit mapping $\phi : X \rightarrow \mathcal{R}^m$ such that $K(x, x') = \phi(x) \cdot \phi(x')$. (You may assume the instance space X is finite.)
2. [10 Points] In this kernel space, any labeling of points in X will be linearly separable. Justify this claim.
3. [10 Points] Since all labelings are linearly separable; this kernel seems perfect for learning any target function. Why is this a bad idea?

3 Linear SVM Implementation [30 pts]

In this assignment you will implement a simple linear SVM classifier and run them on the following dataset:

- **Mushroom dataset:** a simple categorical binary classification dataset. Please note that the labels in the dataset are 0/1, as opposed to -1/1 as in the lectures, so you may have to change either the labels or the derivations of parameter update rules accordingly.

The goal of this assignment is to help you understand the fundamentals of the linear SVM classifier and become trained in implementing SVM using Python. You will also get experience in hyperparameter tuning and using proper train/validation/test data splits.

Download the starting code from the package (“SVM Programming.zip”) provided to you.

The top-level notebook (“SVM.ipynb”) will guide you through all of the steps. Setup instructions are below. The format of this assignment is inspired by the Stanford CS231n assignments, and we have borrowed some of their data loading and instructions in our assignment IPython notebook.

None of the parts of this assignment require the use of a machine with a GPU. You should be able to complete the assignment using your local machine.

Environment Setup (Local): You will need a Python environment set up with the appropriate packages.

IPython: The assignment is given to you in the SVM.ipynb file. Ensure that IPython is installed (<https://ipython.org/install.html>). You may then navigate to the assignment directory in the terminal and start a local IPython server using the Jupyter notebook command.

Reporting: Describe the hyperparameter tuning you tried for learning rate, number of epochs, and Regularization constant. Report the optimal hyperparameter setting you found in the list below. Also report your training, validation, and testing accuracy with your optimal hyperparameter setting.

- **Optimal hyperparameters:**
- **Training accuracy:**
- **Validation accuracy:**
- **Test accuracy:**

Also, create two plots as follows:

1. Fix the number of epochs and Regularization constant to their optimal values, then create a plot where the x-axis varies the learning rate and the y-axis plots the training, validation, and testing accuracy.
2. Fix the number of epochs and learning rate to their optimal values, then create a plot where the x-axis varies the Regularization constant and the y-axis plots the training, validation, and testing accuracy.

Finally, submit “SVM.ipynb” on Canvas as well.

Disclaimers: This assignment re-uses some materials from the publicly available website: [CMU Introduction to Machine Learning Course, 10-315, Spring 2019](#). I personally thank Prof. Maria-Florina Balcan for sharing her teaching materials publicly. This assignment is exclusively used for instructional purposes.