

CSE343/ ECE363/ ECE563: Machine Learning W2022
Assignment (Optional): SVMs
Max Marks: Programming: 50, Theory: 10
Due Date: 16/5/2022, 11:59 PM

Instructions

- This is a individual assignment.
 - Try to attempt all questions.
 - The theory questions should be your individual effort. Copying/Plagiarism will be dealt with strictly.
 - This is an optional assignment. **No extension days** will be permitted.
 - Your submission would be a single .zip file (rollno_HW1.zip) file, that would contain two items (codes + .pdf file). Zip your saved models and include them in the codes folder. You have to include all your plots, results, analysis, conclusion, and solutions for the theory questions in the pdf file. **Code should be well documented. Use comments for python scripts or markdown in Jupyter Notebooks.**
 - Anything not written in the report will fetch 0 marks.
 - It is preferred that you write LaTeX reports.
 - Remember to **Turn in** after uploading on Google Classroom. No excuses or issues would be taken regarding this after the deadline.
 - Start the assignment early. Resolve all your doubts from TAs in their office hours **at least two days before the deadline.**
-

PROGRAMMING QUESTIONS

1. (50 points) **SVM for image classification.** Use the [CIFAR-10](#) dataset to implement this question. Use sklearn for 5 fold cross-validation and choose the hyper-parameters in systematic manner.
 1. (10) Perform PCA using sklearn on the dataset such that 90 % variance is retained.
 2. (10) Use Grid search CV with 5 folds to find best hyperparameters for SVMs with linear (C) and RBF kernels (C, *gamma*).
 3. (15) Use SVM with the best hyperparameters from above to classify this dataset through one vs one (**OVO**) approach. Report the accuracy over five folds along with mean accuracy. Also, report mean class accuracy.

4. (15) Use SVM with best hyperparameters from above to classify this dataset through one vs rest (**OVR**) approach. Report the accuracy over five folds along with mean accuracy. Also, report mean class accuracy.

THEORY QUESTION

1. (10 points) Using a Kernel $K(x, y) = \phi(x)^T \phi(y)$ we can transform data into a higher dimension for classifying inseparable data. ($\phi(x)$ is a feature mapping function for input x). Given two points $x = (x_1, x_2)$ and $y = (y_1, y_2)$, what would be the feature mapping function ϕ for the kernel $K(x, y) = (x^T y)^2$.