Introduction to Pattern Recognition Assignment 4 Report

Andrés Ponce(彭思安) 0616110

June 12, 2020

1 Coding

1.1 K-fold data partition: Implement the K-fold cross-validation function. Your function should take K as an argument and return a list of lists which contains K elements. Each element in a list contains two parts, the first part contains the index of all training folds. The second part contains the indeces of the validation fold.

The `kfold_data` array contains the indices of the training and validation sets that will be used during grid search. To do this, we divide the number of elements in our training set by k. We create a `KFold` model named `kf`, and add the index and validation indices returned by `kf.split(x)` to our `kfold_data` list.

1.2 Use `sklearn.svm.SVC` to train a classifier on teh provided train set and conduct the grid search of "C" and "gamma", "kernel=rbf" to find the best hyperparameters by cross-validation. Print the best hyperparameters you found.

To find the best hyperparameters using grid search, we need to evaluate all the possible combinations of C and gamma. We then train the model on a subset of the original testing data. During this training phase, we reserve one of the K folds to perform validation on our estimated dataset. Our main algorithm consists of trying each of the K folds with each of the different C and gamma combinations, for a combined complexity of $O(\text{gamma x C x K})$.

2 Questions

2.1 Given a valid kernel $k_1(x, x')$, prove that 1) $k(x, x') = ck_1(x, x')$ and 2) $k(x, x') = f(x)k_1(x, x')f(x)$ are valid kernels, where $c > 0$ is a positive constant and $f(\cdot)$ is any real-valued function.

A valid kernel function $k_1(x, x')$ has to have a positive semi-definite **Gram matrix** $\mathbf{K} = [k(\mathbf{x}_n, \mathbf{x}_m)]_{nm}$. A positive semi-definite matrix is one whose eigenvalues are all positive. We are given that $k_1(x, x')$ is a valid kernel, so at the start it meets this condition. We thus have to show that multiplying k_1 by a constant also produces a valid kernel, i.e. multiplying the kernel by a constant also produces a positive semi-definite Gram matrix.