How to Tune Hyperparameters for SVM

Initial: April 14, 2020

Hyperparameter tuning needs a lot of trials and thus could run very slowly. With many hyperparameters, SVM training and parameter tuning process might be very time consuming, especially when sample size is large. The following are several general approaches that can speed up the hyperparameter tuning.

1. Use a different R package

There are multiple R packages that implement the similar algorithms. For example, algorithms in the kernlab package usually run faster than those methods in the e1071 package. The kernlab packages support "svmLinear", "svmRadial", and "svmPoly" methods for training SVM with different kernels by using the unified train() method in the caret package.

For a detailed list of all methods supported by the caret package, refer to:

https://topepo.github.io/caret/available-models.html

2. Simplify the hyperparameters

There might be several hyperparameters in SVM kernels. You may want to reduce the number of combinations of parameters. For example, just set the cost or C parameter as 1. Do not set too many combinations of parameter values.

3. Simplify the degree for polynomial kernel

Polynomial kernel could become very complicated when degree is very large. To speed up the process, you can simply set degree = 2. This can greatly improve the performance of polynomial kernel. Especially when your polynomial kernel cannot converge to a solution, reducing degree might fix the problem.

4. Try parameter tuning on a smaller training set

You can try the parameter tuning on a smaller training set (for example, randomly select 1000 observations from the training data). This allows you to quickly tune the parameters. Then you can use the optimal values of the parameters to fit the final model to the whole large dataset. This may not output the "optimal" parameters, but better than waiting for the tuning process for ever. Essentially, even when the whole training dataset is used to tune hyperparameters, there is still no guarantee that the hyperparameters found are really "optimal."