C00535934

Rajesh Awal

Informatics

1. What is the fundamental idea behind Support Vector Machines?

   Support Vector Machines implements hyperplane to separates two classes. It is designed in a way that it manipulates data so that it can implement best separation functions. Its aim is to make hard margin so that there are very few miscalculations.

2. What is a support vector?

   Support vectors are reasonable features and is responsible for creating decision boundary. Support vectors are data points that are closer to the hyperplane and influence the position and orientation of hyperplane.

3. Why is it important to scale the inputs when using SVMs?

There should be scaling of features, to have fair prediction. Otherwise, data features containing very larger values will affects the result. SVM tend to neglect small features.

4. Can an SVM classifier output a confidence score when it classifies an instance? What about a probability?

   Yes, SVM classifies an instance with confidence score. It provides distance from separator for each output. No, it does not provide probability output. But these confidence score can be implemented to approximate probability using logistic regression function and training.

5. Should you use the primal or the dual form of the SVM problem to train a model on a training set with millions of instances and hundreds of features?

   If the number of instances outnumbered features, then the effective method will be primal Form.

6. Say you trained an SVM classifier with an RBF kernel. It seems to underfit the training set: should you increase or decrease γ (gamma)? What about C?

   Gama or C or both should be decreased in this situation. This is because underfitting in SVM classifier with an RBF kernel occurs when there is too much regularization.

7. How should you set the QP parameters (H, f, A, and b) to solve the soft margin linear SVM classifier problem using an off-the-shelf QP solver?

   Here, m is number of rows and n is number of columns.

   Let's call the QP parameters for the hard-margin problem H′, f′, A′ and b′. The QP parameters for the soft-margin problem have m additional parameters (np = n + 1 + m) and m additional constraints (nc = 2m). They can be defined like so:

   H is equal to H′, plus m columns of 0s on the right and m rows of 0s at the bottom.

   f is equal to f′ with m additional elements, all equal to the value of the hyperparameter C.

   b is equal to b′ with m additional elements, all equal to 0.

A is equal to A′, with an extra m × m identity matrix Im appended to the right.

8. Train a LinearSVC on a linearly separable dataset. Then train an SVC and a SGDClassifier on the same dataset. See if you can get them to produce roughly the same model.

   Solution is in google colab.

9. Train an SVM classifier on the MNIST dataset. Since SVM classifiers are binary classifiers, you will need to use one-versus-all to classify all 10 digits. You may want to tune the hyperparameters using small validation sets to speed up the process. What accuracy can you reach?

   Solution is in google colab.

10. Train an SVM regressor on the California housing dataset.

    Solution is in google colab.