University of Louisiana at Lafayette

School of Computing and Informatics

INFX-598 Machine Learning Applications

Homework # 05

Submitted By

Navid Yousuf

ULID: C00419219

Fall 2021

1. **What is the fundamental idea behind Support Vector Machines?**
   SVM or Support Vector Machine is a linear model for classification and regression problems. It can solve linear and non-linear problems and work well for many practical problems. The idea of SVM is simple: The algorithm creates a line or a hyperplane which separates the data into classes.
2. **What is a support vector?**
   Support vectors are the data points that lie closest to the decision surface (or hyperplane). They are the data points most difficult to classify. They have direct bearing on the optimum location of the decision surface.

3. **Why is it important to scale the inputs when using SVMs?**
   Because Support Vector Machine (SVM) optimization occurs by minimizing the decision vector w, the optimal hyperplane is influenced by the scale of the input features and it's therefore recommended that data be standardized (mean 0, var 1) prior to SVM model training.
4. **Can an SVM classifier output a confidence score when it classifies an instance? What about a probability?**
   It can output the distance between the test instance and the decision boundary, which can be used as the confidence score. This score can't be directly converted into a probability.
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 there are millions of instances, you should definitely use the primal form. This question only applies to linear SVMs. The computational complexity of the primal form of the SVM problem is proportional tot he number of training instances m, while the comp. complexity of the dual form is proportional to a number between m^2 and m^3.

6. **Say you've trained an SVM classifier with an RBF kernel, but it seems to underfit the training set. Should you increase or decrease γ (gamma)? What about C?**
   It's possible that there is too much regularization. To decrease it, increase gamma, C, or both.

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?**
   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,