Learning With Kernel Machines   
Map Your Own Problem

Instructions:

In this part of the course project, you will map your own problem taking the ERM approach. You will specifically describe the particulars of the problem you would like to solve. Answer the following questions within the provided textboxes.

What is the problem space/context?

|  |
| --- |
| We have several features regarding customer behavior that we want to input to a machine learning model to predict customer actions going forward. Specifically, we would like to determine the probability that a customer will churn in 6 months’ time based on the features we have. |

What do you want to learn?

|  |
| --- |
| We want to classify customers into two groups:   * Customers who will likely churn in 6 months. * Customers who will NOT likely churn in 6 months.   Hence, this is a classification problem. We would like to calculate the probability that a customer will churn. Prediction probabilities >= 0.5 will be considered ‘churn’ and prediction probabilities < 0.5 will be considered ‘no churn’. |

What are the constraints/limitations of your data?

|  |
| --- |
| The limitations that we encounter are the typical challenges for almost every data problem:   * We have a limited amount of data to train our model. * We must be selective about which customer features (i.e., which ones have the most influence on churn?) we incorporate into our train data set. |

What algorithm, loss function, and regularizer will you use? Why?

|  |
| --- |
| Since this is a classification problem where we are looking for probabilities, we will select a logistic regression algorithm for our model. Our loss function will be the logistic (binary cross-entropy) loss. Moreover, we will add L1 regularization to our loss function to avoid overfitting on the training data. An additional advantage of L1 regularization is that it can identify or eliminate any features that do not have direct impact on the model output. |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*To submit this assignment, please refer to the instructions in the course.*