Data that assumes a non-linear relationship can be handled through non-parametric regression, which will help accurately find the conditional expectation between two random variables. Kernel regression is a popular form of non-parametric regression. Specifically, Kernel regression uses a kernel smoothing method to estimate a curve for non-parametric data. Kernel estimator or smoothing is done by taking the weighted average of training samples that surround a targeted y value. Inside this smoothing parameter is distance-based kernel function, which helps attach weights to the observations surrounding the y value. In essence, the curve is estimated by a kernel estimator of the conditional mean. There is also a scale factor h, known as the bandwidth, which determines area of points surrounding the response value that will be used for the weighted local averaging. The development and critique of non-parametric kernel regression estimation is ongoing.

The history of kernel estimators continues to be researched. Nadaraya-Watson Kernel Estimator, which was founded in 1964, has been considered a very effective approach in accurately modeling real world data. In 1998, Richard Blundell implements kernel methods in microeconomics data, including analysis on food and alcohol expenditure for households in the United Kingdom(Blundell). In 2017, this approach was used to predict monthly electricity demand production, which aids with appropriately scheduling upkeeping as providing important information during contract negotiations between energy companies and customers (Dudek). However, latest works indicate kernel regression can provide meaningful results in classification scenarios. Support Vector Machines (SVM) is a type of supervised machine learning method that has been shown to have better performance measurements (sensitivity, specificity, Area Under the Curve, etc.) than other machine learning methods such as the Decision Tree and Neural Network models. SVMs solve binary classification problems by defining a maximum margin separating hyperplane. This hyperplane is then used to classify training and testing datapoints. To construct the best hyperplane, SVMs need support vectors, or datapoints around the boundary. For non-linear data, kernels SVMs allow the hyperplane to be mapped on a higher dimensional space(Awad).

To achieve a classifier, kernel SVM standard model selection practice requires us to select a regularization parameter, C, and a parameter governing the sensitivity of the kernel, γ. Since we do not know the true values of γ and c, we can only achieve the optimal estimates through resampling procedures such as k-fold cross validation. It should also be k-fold cross validation is best suited when both the training and testing datasets are large, and the number of k-folds used ensure that the variance and computational costs tradeoff is minimized (Wainer).

This report presents a two-part study. The first section addresses how to determine the optimal parameters for Nadaryan-Watson Kernel estimator. We will examine the bias-variance tradeoff and conclude the optimal bandwidth using a commonly used cars dataset in R. The second section focuses on kernel SVM, answering questions about how to fine tune the regularization and sensitivity parameters discussed above. We will conclude with a comparative analysis of the four different kernel functions provided in R.

Awad, M., & Khanna, R. (2015). Support vector machines for classification. In *Efficient learning machines* (pp. 39-66). Apress, Berkeley, CA.

Dudek, G., & Pelka, P. (2017, May). Medium-term electric energy demand forecasting using Nadaraya-Watson estimator. In *2017 18th International Scientific Conference on Electric Power Engineering (EPE)* (pp. 1-6). IEEE.

Blundell, R., & Duncan, A. (1998). Kernel regression in empirical microeconomics. *Journal of Human Resources*, 62-87.

Wainer, J., & Cawley, G. (2017). Empirical evaluation of resampling procedures for optimising SVM hyperparameters. *Journal of Machine Learning Research*, *18*(15), 1-35.

Demir, S., & Toktamiş, Ö. (2010). On the adaptive Nadaraya-Watson kernel regression estimators. *Hacettepe Journal of Mathematics and Statistics*, *39*(3), 429-437.