\section{Comparing classifiers}

In this section, I used two classic datasets, the divorce classification dataset and the MNIST handwritten digit dataset, to compare the quality and effectiveness of several different classifier models. I focus on classifiers: Naïve Bayes, Logistic Regression, k-Nearest Neighbors (KNN), Support Vectors Machine (SVM), and Neural Networks.

\subsection{Divorce classification/prediction}

I use the divorce dataset, consisting of 170 participants who documented personal information, which includes 54 predictors in the attempt to predict if the subject got divorced. This dataset does have added noise such that it is not identical to the publicly available version and I have the ground truth values for each data point. For this dataset I compare three classifiers: Naive Bayes, Logistic Regression, and KNN. I first train the three classifiers using the first 80% of the data and used the trained models to predict the labels for the remaining 20% of the data (test set). I compare these model predictions to the ground truth values provided to obtain the overall accuracy of each classifier. It is worth noting that the training set consists of data points from each class but the test set consists As an addi