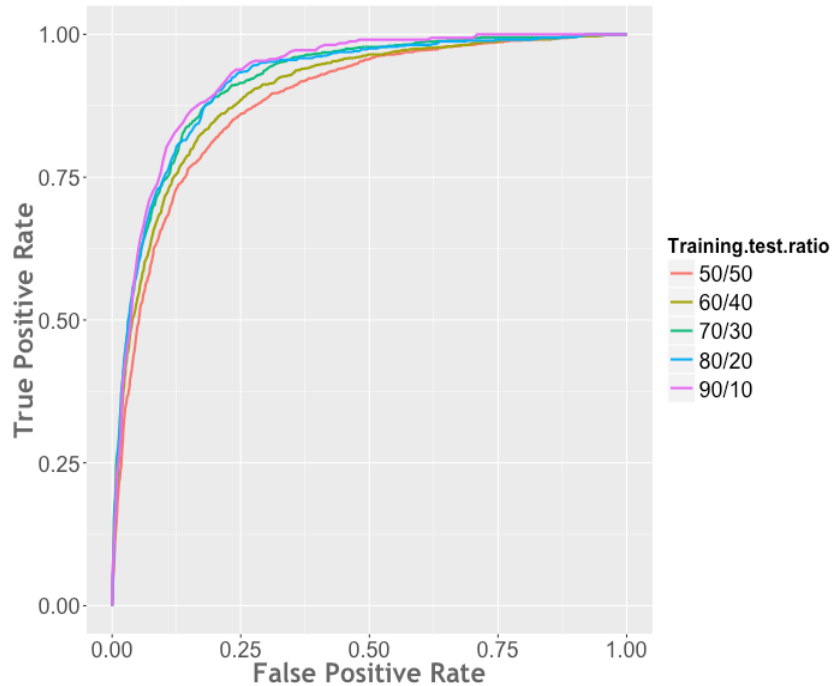
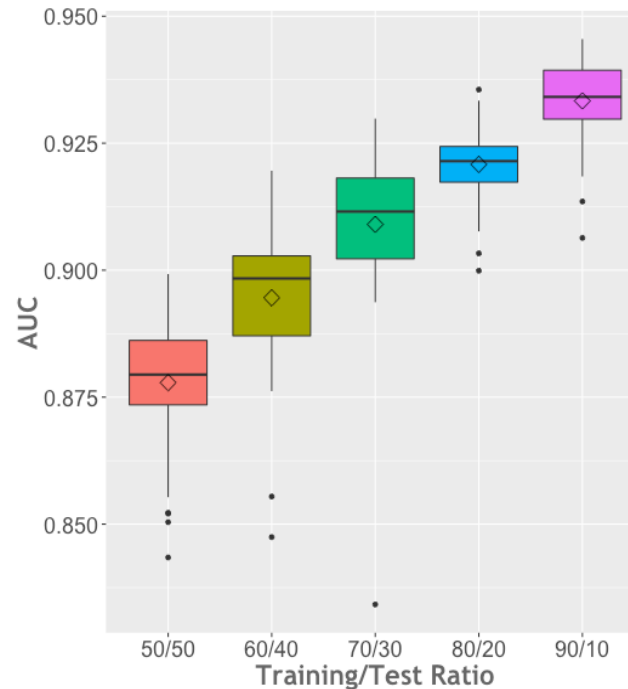


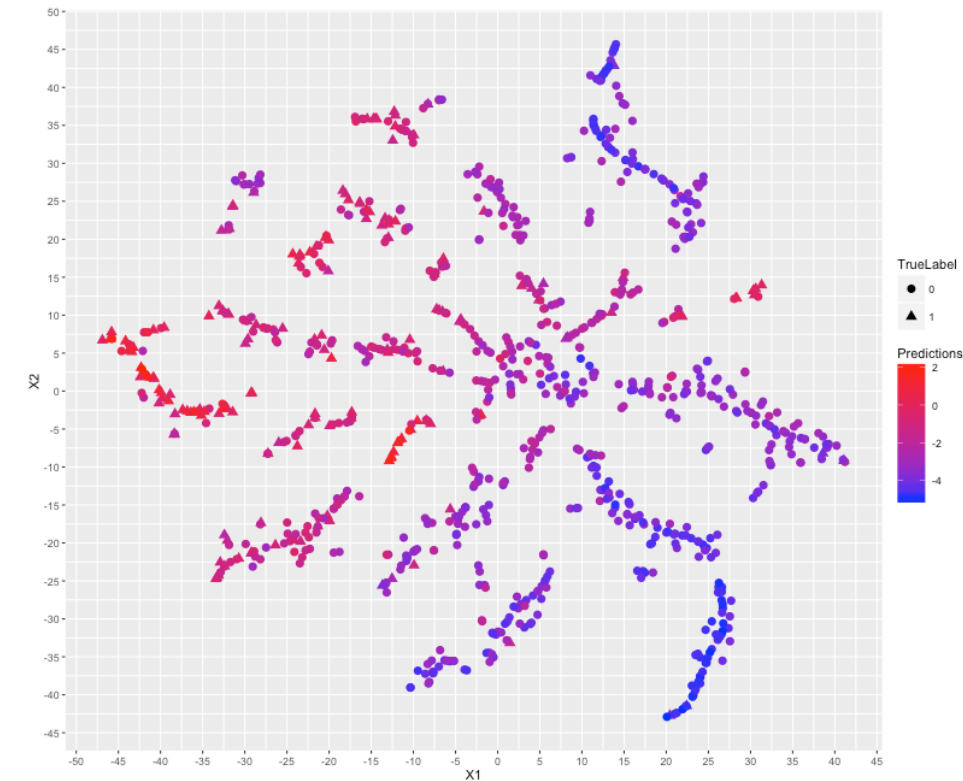
Predictive Performance and Visualization



ROC curves with different training / test ratios



AUC vs. training / test ratio for 50 random sampling of control groups



Two dimensional embedding by t-SNE (van der Maaten and Hinton, 2008)

▲ Pancreatic cancer diagnosis
● No pancreatic cancer diagnosis
Colors: predicted score

Summary

- Applied the supervised LDA model to predict pancreatic cancer in Duke EMR data
- Learned interpretable representations with excellent predictive performance
- Identified essential hypertension as a possible early contributor
- Found individuals with EMR highly similar to pancreatic cancer but without a diagnosis