Machine Learning Methods in Credit Risk

Capstone Timeline

Problem Statement
The main interest was to help bank determine whether to grant loan depending on the risk of the mortgage. Our goal was to develop a robust model to predict default using available data at the time of the house mortgage application.

Definition of Default

Data Sources

Fannie Mae® Housing Price Index
Mortgage Loan with terms 10-30 years, acquired by Fannie Mae 2004-2013

Survival Analysis in Default Prediction

Classification (Predict the state: Event/no event/censored)
Regression (Predict time to default: Time to event/no event/censor)

Data Processing And Performance

Methodology, Data Processing

Business Impact

Key Conclusions
Future Directions

- Machine learning models especially survival models add value and valuable insights
- The results of this study will be used to build comprehensive and accurate credit risk models for future customers
- Choose cutoffs via portfolio loss optimization
- Real-time portfolio risk monitor with time-varying covariates (through Deep Learning algorithms)

Economic Impact via Optimization

Dynamic Time to Default Estimation