MGMT 638

Session 6

Kerry Back

Fall 2025

Agenda

1. Machine learning and asset pricing paper
2. More random forests
3. Creating Quality Minus Junk characteristics

Machine Learning and Asset Pricing

- PDF: *Asset Pricing Via Machine Learning*, Gu, Kelly, and Xiu, 2020
- Video: *Asset Pricing Via Machine Learning*, Gu, Kelly, and Xiu, 2020
- NotebookLM: Asset Pricing Via Machine Learning

More Random Forests

- Ask Claude to get the California housing price dataset and to describe it.
- Ask Claude to do a train/test split and fit a random forest on the training data to predict house prices.
- Ask Claude to evaluate performance on the test data.
- Ask Claude to vary the maximum depth and to display performance on the training and test data.

Create Quality Minus Junk Variables

- See Session1/Quality Minus Junk for link to paper
- See Session3/Workflow for procedure to follow
- Important: create all growth rates of ratios or other fundamental variables before merging with returns