Assignment 2: Comparing Machine Learning Models

April 27, 2022

1 Tree-Based Models - 2 Points

Using the data sets provided in the Assignment 1, repeat the same steps using Tree-Based Models. Compare your results with the linear regression approach.

2 The Ising Model - 5 Points

Consider the 1D Ising Model with nearest-neighbour interactions:

$$\mathcal{H}[\vec{S}] = -J \sum_{j=1}^{L} S_j S_{j+1}, \tag{1}$$

on a chain of length L, with periodic boundary conditions and $S_j = \pm 1$ Ising spin variables. In one dimension, this paradigmatic model has no phase transition at finite temperature.

1. As a warm up, solve the following problems:
   - Compute the partition function in one dimension, at inverse temperature β in the thermodynamic limit.
   - Compute the model magnetization in canonical ensemble, again in the thermodynamic limit. How does it behaves with β?
2. Ising Model as a Regression Problem
   - Recast the same Ising Hamiltonian as a linear regression problem.
   - Simulate an Ising dataset with states and energy, in a lattice of size $L = 40$ and train a linear regression model using OLS, Lasso and Ridge. Evaluate the models.
   - Using the fitted models, and the same dataset for test, compare the analytical result and the results using the fitted models.
   - Provide an attempt for the last two step using Tree-Based Models. There were any improvement in the results?
3. Ising Model as a Classification Problem
   - Using the provided Ising Dataset(see attached Jupyter Notebook), build, train and evaluate a Logistic Regression Model to classify ths states in ordered and disordered.
   - Repeat the same problem using Tree-Based Model Classifiers.

3 The Fama-French-Carhart Model - 3 Points

After Hedonic Price Models, the most used model based on linear regression are the Factor Assets Pricing Model. From the seminal Capital Asset Price Model, to the most influential ones: The Nobel Laureate Fama-French Model and the Carhart Four Factor Model.

1. The Factor Models
   - Reproduce the results of the Nobel Laureate paper by Eugene Fama and Kenneth French. The required data can be downloaded from French's website.
   - Using the same dataset, extend the results to the Fama-French 5−Factor Model.

- Using the trained Model, predict the pricing behaviour of three stocks of your choice. I strongly suggest use stocks of different sectors.
- Provide the evaluation metrics for each model.
- Now, using the same variables repeat the steps using Tree-Based Regressors.