

Final Project

Advanced Machine Learning (SS 2022)

Default Prediction for Peer-to-Peer Lending

Download Data From: <https://www.kaggle.com/wordsforthewise/lending-club>

Submission Deadline: *Su., 07.08.2022 23:59Uhr*

Submission Format: *Your codes and outputs (Jupyter Notebook or PDF)*

Email: *nazemi@kit.edu*

Tasks:

- Explain dataset and business problems in two paragraphs. (10 points)
- Do an explanatory analysis for this dataset. (10 points)
- Use logistic regression (in-sample) for default prediction. Interpret your results for logistic regression. (20 points)
- Compare logistic regression, classification tree, random forest, bagging, support vector machine (with two kernel functions), and neural network for default prediction based on the out-of-sample performance measures (at least three different measures). (35 points)
- Use LASSO for variable selection and explain results. (15 points)
- Which model do you suggest to the peer-to-peer company? Why? (10 points)