

HW 5
MACHINE LEARNING PIPELINE
ADDITIONAL IMPROVEMENTS

Contents

1	Pipeline Improvements	1
1.1	Code Structure	1
1.2	Methodology	1
1.3	Models	1

Notes

- Representative code snippets are interspersed with analysis and explanations below; all code is available on GitHub: <https://github.com/satejsoman/capp30254/tree/master/hw5>.
- The pipeline library lives in the code/pipeline directory, while the sample application which imports the library is code/donors_choose.py.
- The generated table of hyperparameters is uploaded as a separate CSV; it also is available on GitHub as code/evaluations.csv.

1 Pipeline Improvements

The following improvements have been made to the pipeline library:

1.1 Code Structure

- model grid per pipeline
- discretize on each dataset
- metrics at k
- unit tests

1.2 Methodology

- The label for the analyzed DataChoose data set has been updated so it is 1
- set gaps

1.3 Models

- SVM
- meaningful params