Dataset Description

The dataset is provided by the Georgia Tech ChBE Unit Control Lab, which is obtained during the two-phase flow correlation experiment. It consists of 8 parameters: air flow rates, water flow rates and 6 pressures along the pipeline. We will use the pressure-6 as our “y”, while the rest 7 features are the inputs. We have 8057 data points available, which is a good sample for this system, considering the number of features. After we vitalize the data, we find that the distribution of each features are all close to normal distribution, but the order of magnitude vary. Thus we will rescale the features for our following steps.