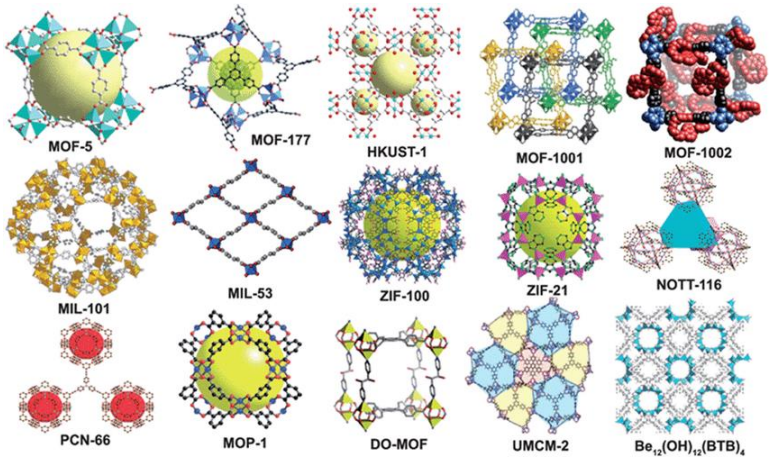
ChE 197/297: Intro to AI/ML for Chemical Engineers
Case Studies in ChemE

Instructions: Answer each problem then create a solution using Python code via Jupyter Notebook.

Problem 1: Fitting Langmuir Isotherms to Adsorption Data



Metal-organic frameworks (MOFs) are promising new materials for a wide range of applications due to their porosity, modularity, crystallinity, and tunability. Their structure consists of a metal ion bonded to organic ligands. MOFs have the highest reported surface area among materials, hence they are viewed as next-generation catalyts, adsorbents, and separation agents. To combat climate change, researchers are now searching for the best MOFs that can capture CO_2 from the atmosphere.

You are given a data set of MOF isotherms (**mofDatabase.csv**): adorption loading of the MOFs for CO_2 at different pressures but same temperature. The data set contains >130,000 MOFs along with their adsorption loadings at 0.01, 0.05, 0.1, 0.5, and 2.5 bar at 298 K. Load the data then do the following:

1. Take only 20 random MOFs then make a scatter plot of their isotherm: loading (q) vs. pressure (p).

2. For each of the 20 MOFs, fit the following Langmuir isotherm by finding a and b.

$$q = \frac{ap}{1 + bp}$$

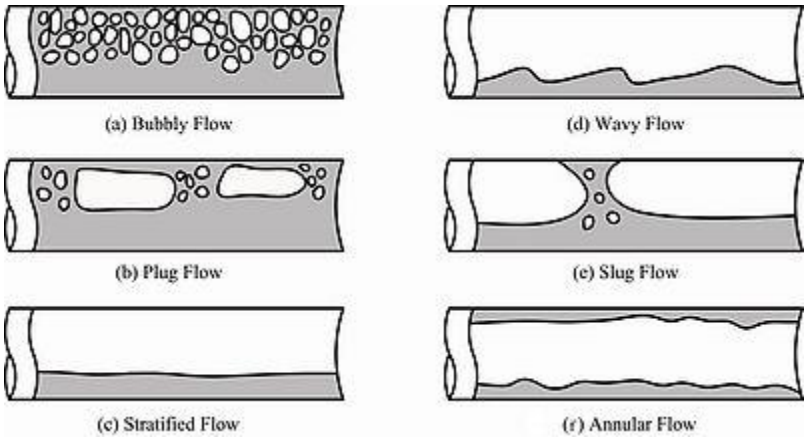
   a. Use the following linearized equation, then apply linear regression. Report a and b.

$$\frac{p}{q} = \frac{b}{a}p + \frac{1}{a}$$

   b. Report the accuracy of the fits in terms of MSE and R^2.

Note: This problem is for practice only. Model testing is not required for now.

Problem 2: Slug vs. Stratified vs. Annular Flow Classification



You are given a data set of different scenarios of two-phase flow in a pipe (**flow_regime.csv**). Each scenario corresponds to a specific gas and liquid velocity (flow rate) setting in the pipe and the governing flow regime at that setting. Your task is to find the decision boundary between the flow regimes using Logistic Regression.

1. Use 70% of the data for training and 30% for testing. Pre-process the gas-liquid velocities by taking their logarithms, as typically done in flow regime maps.

2. Train a Logistic Regression model to perform three-class classifiction on all flow regimes. Visualize the decision boundary between the three classes.

3. Report the accuracy of the model on training and testing data. Also, report the model parameters (`coef_` and `intercept_`) describing each boundary.

END OF EXERCISE