# **1 Introduction**

When input and output are **not related linearly**, we use **Non-Linear Mapping Functions**. For this, we use the Scipy.optimize Python Library. Firstly, you will estimate the possible **Mapping Function** based on the dataset trend by displaying the data with plt.scatter(x,y). The mapping functions could be exponential, Power law, Logarithms, Cosine/Sine, or a combination of these. Then, apply the **curve fitting optimization algorithm** to estimate the mapping function coefficients.

In Mathematics **Optimizations** Means the Best Solution. Therefore, the optimized curve-fitting parameters will provide the best-fitting line.

# 2 How to perform a curve fitting with SciPy. Optimize?

**Here are the main steps:**

* **Import libraries**
  + import pandas as pd # **To read /import file**
  + import **matplotlib.pyplot** as plt # **To create plot**
  + from **scipy.optimize** import **curve\_fit**  **#To perform optimization** with curve\_fit
  + from **sklearn.metrics** import r2\_score # To perform model performance analysis
* **Import file**
* **Define** x and y variables.
* **Define the basic function** that describes the dataset behavior, takes x and y arguments and some number of arguments. The coefficients (or constants) (e.g. a, b, c) will be optimized by a nonlinear least square optimization process. (i.e. curve\_fit() **optimization algorithm**.)
* **Define a function by a name for example func (you can give any name.)**

def func(x, a, b):

return a \* x\*\*b

* **Curve\_fit**

popt, \_ = curve\_fit(func, x, y)

* Unpack optimal parameters for the basic function **#**

a, b, c = popt # ( **popt = parameter optimized)**

* Compute model-based y value, for a given x and optimized parameters.

y\_model = func(x, a, b)

* Perform model performance evaluation between **the True** (y) and the model predicted (y\_model)
  + r2\_score (y, y\_model) and display
* Plot results
  + plt.scatter(x, y)
  + plt.plot (y, y\_model)

# **3 Your Tasks**

* **Classwork (Practice).** In section 3.1, you will practice how **SciPy.optimize** model laboratory data with Linear and Non-linear mapping function models.
* **Project #3 Task #3:** In section 3.2, you will use your project file (**CleanedFeatureSelected.xlsx**) to do **Project #3-Task #3, which is to be handed in**.

## 3.1 Classwork #3: **Use SciPy.optimize** to model laboratory-measured viscosity Data. Your task is to model the viscosity dataset with three rheological models **and perform model accuracy analysis to select the best model that describes the dataset**.

What is Rheology? It is the study of fluid flow and deformation. The rheological parameters (i.e. viscosity) determine the fluid flow. It is applied in several sectors, where fluid flow occurs. For instance, in the petroleum industry, the rheological properties of the drilling fluids and cement determine the hydraulics. The rheological properties of the fluid are determined from the viscometer measurement data provided in Table 1 and the rheological models.

The **Rheological Models** are:

1. Power law : y = **a**x**b**
2. Herschel Bulkley: y = **a**x**b** + c
3. Robertson and Stiff: y = a\*(x + b)c

Table 1: Laboratory Measured Viscometer Dataset.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Shear rate, 1/s **(X)** | 5.1 | 10.2 | 170.3 | 340.6 | 510.9 | 1021.8 |
| Shear stress, lbf/100ft^2 (Y) | 8.5 | 9.6 | 19.2 | 25.6 | 32.1 | 45.9 |

1. **Your task** is to model the data with the four models and **perform model accuracy** analysis.
2. For each model, you will plot the **scatter (data)** and **line (model).**
3. Also **write equations** of the model.

## **Project #3-**Task #3. This is **to be Handed in. Guide!**

* **Use Your Project #3 – Task #1 File, which is the cleaned one**. (**CleanedFeatureSelected.xlsx**)
* Import Data and Display with a scatter plot (Vp vs DEN), (Vp vs Vs), (Vp vs NEU)
* Based on the dataset, estimate the Possible Mapping Function for the above pairs.
* Define the **mapping function** and, using SciPy, optimize the function. Estimate the optimized parameters (for example, a, b, and c as you have done above).
* Finally, **plot** the curve fitting model with the scatter plot providing the equations.
* Compare Co based on True data (Vp) vs **your Model**, **Gardner**  (Here, compute the R2 and Plot in a bar graph. In addition, your best model with **True Vp, Gardner**)