# SKLEARN

-MUKESH KUMAR

# AGENDA

- Sk-Learn
- Seed
- Train_Test_split

# The name Scikit-learn or sk-learn

- The name **scikit-learn** has an interesting origin that reflects its history and purpose:
  - **Scikit:**
  - **"Sci"**: This prefix stands for **"Scientific"** and is a nod to the scientific computing community in Python, particularly the SciPy ecosystem. SciPy is a core library used for scientific and technical computing in Python.
  - **"kit"**: This suffix stands for **"toolkit"**, indicating that scikit-learn is a toolkit or a collection of tools for a specific purpose—in this case, machine learning.
- **Sklearn:**
  - **"sklearn"** is the shorthand or module name used in Python when importing the library. It's a concise way to reference scikit-learn in code.

# Key Features of Scikit-learn:

- **Machine Learning Algorithms**:
  - Scikit-learn includes a wide range of algorithms for supervised learning (like linear regression, decision trees, and support vector machines) and unsupervised learning (like k-means clustering and principal component analysis).
- **Data Preprocessing**:
  - It provides tools for preprocessing data, such as scaling features, handling missing values, encoding categorical variables, and splitting datasets into training and testing sets.
- **Model Evaluation**:
  - Scikit-learn has functions to evaluate the performance of models using metrics like accuracy, precision, recall, and mean squared error.
- **Model Selection**:
  - It includes methods for selecting the best model, such as cross-validation, hyperparameter tuning, and grid search.
- **Integration with Other Libraries**:
  - Scikit-learn integrates well with other Python libraries like NumPy, pandas, and Matplotlib, making it easy to work with data and visualize results.

# What is Seed?

- Open collab notebook

# What is train_test_split?

- Open collab notebook