Project Based Learning:

Building a MLP Model

# Project Overview:

We have a binary classification datasets, the task is to categorize data into one of two categories/classes. These classes are typically labeled as **0** and **1.**

# Datasets:

* **Dataset – 1**
* It contains shape (100,3)
* The **U Shape** dataset is a binary classification dataset commonly used in machine learning research. It is based on a synthetic dataset designed to evaluate the performance of classification algorithms. The dataset exhibits a U-shaped decision boundary, hence the name "U Shape."
* **Dataset – 2**
* It contains shape (100,3)
* The **Concentric** dataset consists of two input features.
* The concentric circles dataset contains two classes. The data points are structured in concentric circles, where each circle represents a different class. Typically, the inner circle corresponds to one class (e.g., class 0), and the outer circle corresponds to another class (e.g., class 1)
* **Dataset – 3**
* It contains shape (500,3)
* Similar to the UShape dataset, the **concentric circles dataset** is synthetic, meaning it is artificially generated and not derived from real-world data.
* **Dataset – 4**
* It contains shape (100,3)
* **Linearly separable datasets** are commonly used to evaluate the performance of classification algorithms, particularly those that rely on linear decision boundaries.
* **Dataset – 5**
* It contains shape (600,3)
* An **outlier classification dataset** consists of input features that describe the characteristics or attributes of the data points. Outlier classification datasets present unique challenges compared to regular classification datasets.
* **Dataset – 6**
* It contains shape (100,3)

In **overlap classification datasets**, the decision boundaries that separate different classes are not linear or easily discernible. They often exhibit complex and non-linear patterns, making it challenging for traditional linear classifiers to accurately classify instances.

* **Dataset – 7**
* It contains shape (500,3)
* **The XOR dataset** is a well-known binary classification dataset that represents the exclusive OR (XOR) logical operation. The XOR operation takes two binary inputs and returns true (1) if the inputs are different and false (0) if they are the same.
* The XOR dataset is commonly used as a toy problem to demonstrate the limitations of linear classifiers, such as logistic regression or perceptron’s, which can only handle linearly separable datasets.
* **Dataset – 8**
* It contains shape (2000,3)
* The **Spiral dataset** contains two classes, often labeled as "0" and "1". The two spirals represent the different classes. Each data point is assigned a class label based on its position on the spirals.
* **Dataset – 9**
* It contains shape (400,3)
* **Random classification dataset** refers to a dataset generated with random values or patterns for classification tasks. These are commonly used to evaluate the baseline performance of classification algorithms or as a control group for comparative analysis.

# Steps involved in project:

* Importing Required libraries
  + **NumPy, Pandas, Seaborn & Matplotlib, Sci-kit learn, TensorFlow & Keras.**
* Loading the Dataset & Insights of the dataset and identifying the relationship between input and output.
* Segregate the data and split it into Train & Test.
* Building the Model
* Model compilation and Training.
* Evaluation of the Model and Accuracy of its performance.
* Plotting the Best Decision Surface Region.

# Summarize what you learned:

* In Deep learning we have to experiment with Activation Function Such as **ReLU,Tanh** and **Sigmoid** functions.
* By adding so many layers and using so many Neurons in the functions.
* Using some techniques like Dropout, Epochs and Batch size while compilation of models.
* Here the models always go into overfitting issue we have to train the model by Controlling Number of hidden layers and Regularization Techniques.
* So many initializers are there to initialize the weights.
  + **Glorot /XAVIER (**Uniform and Normal**)**
  + **Random (**Uniform and Normal**)**
  + **He (**Uniform and Normal**)**

# Conclusion:

# Tried so many Experiments to obtain best decision surface for all Datasets with good accuracy score.