**ML Lab (CS360)**

**Assignment 3**

**Total Marks: 20**

**Logistic Regression**

Q1: (a) Split the Diabetes Prediction dataset( [link to download](https://www.kaggle.com/datasets/iammustafatz/diabetes-prediction-dataset?resource=download)) into training (80%) and testing (20%) sets. (5 marks)

(b) Write a code to implement logistic regression for without using the in-built library.

(10 marks)

Q2: Implement the logistic regression model on Iris dataset ( [link for dataset](https://www.kaggle.com/datasets/himanshunakrani/iris-dataset)) by using in-built library. Divide the dataset into training and testing sets in the ratio of 90:10.

(5 marks)

----------------------------------------------------------------------------------------------------------------------------

1. Implement logistic regression algorithm (Multi-class classification) without using the inbuilt library using the **Iris** dataset and perform the following:

*(Hint to load the dataset*

*from sklearn import datasets*

*iris = datasets.load\_iris())*

1. Divide the dataset into train, validation and test with the dataset division ratio of your choice.
2. Hyperparameter tuning on the validation set is required to be done.
3. Check for overfitting on the validation set by plotting graph during training of the finalized model.
4. Report the classification accuracy (on test set and train set) by varying the percentage of training samples.

**Please do not copy anyone's code. Try to learn to code.**

Thank you

Happy learning