IRIS DATA

NAME :-Akash Sharma
ROLL NO :- 19223005
CLASS:- MCA(3rd sem)

- ## 1.Introduction

The purpose of this project is to investigate "The Iris Flower Data Set" and prove my understanding of the data set and explain its importance to machine learning.Iris data set is the famous smaller databases for easier visualization and analysis techniques. The IRIS flower data set contains the the physical parameters of three species flower Versicolor, Setosa and Virginica. In this data we will be predicting the species of the flowers based on these parameters, we will be building a machine learning project to determine the species of the flower

- ## 2.Problem Statement

The data set contains 3 classes of 50 instances each, where each class refers to a type of iris plant. One class is linearly separable from the other two , the latter are not linearly separable from each other. The data base contains the following attributes:

1). sepal length in cm
2). sepal width in cm
3). petal length in cm
4). petal width in cm
5). class:
- Iris Setosa , Iris Versicolour, Iris Virginica

3. Solution overview

LOGISTIC REGRESSION ALGORITHM:

Logistic Regression is a type of regression . Logistic Regression uses the logistic function to find a model that fits with the data points. . Logistic Regression can then model events better than linear regression, as it shows the probability for y being 1 for a given x value. Logistic Regression is used in statistics and machine learning to predict values of an input from previous test data. Supervised learning consists in learning the link between two datasets: the observed data X and an external variable y that we are trying to predict, usually called “target” or “labels”. Most often, y is a 1D array of length n_samples. All supervised estimators in scikitlearn implement a fit(X, y) method to fit the model and a predict(X) method that, given unlabeled observations X, returns the predicted labels y

- ## 4. Conclusion

The primary goal of supervised learning is to build a model that “generalizes”. Here in this project we make predictions on unseen data which is the data not used to train the model hence the machine learning model built should accurately predicts the species of future flowers rather than accurately predicting the label of already trained data