SUMMARY REPORT

INTRODUCTION AND DISCUSSION OF PROJECT

1. Identification and Survey paper

This Diabetes Dataset Identifies Supervised Learning. Requirements are very clear. Both input and output is present.

In Supervised Learning it is a Classification Data.

This paper category includes an excessive survey for any machine learning domain.I Prefered Healthcare to start with survey on that.The Website I choosen Kaggle to get a free dataset Because I can use various Machine Learning Algorithms and conclude the best Algorithm for the choosen problem.

2. Paper with just proof of concept

This category of paper requires in-depth knowledge of the selected area. This paper showcases the brief, logical, and technical proof of the proposed new algorithm.

3. Developing new machine learning algorithms

Machine learning is still an emerging field. However, there are many application areas of machine learning algorithms like agriculture, health, social media, computer vision, image processing, NLP, sentimental analysis, recommender system, prediction, business analytics, and almost all the fields can directly or indirectly use machine learning in one or another way.

4. Comparison of various machine learning algorithms

The paper title for such category includes, “DIABETES DATASET: Survey of various machine learning algorithms”.

5. Analysis with Data visualization

Finally Analysis the Project with Multiple Algorithm and Make it with Visualization Using Google Studio and select Best Algorithm.

SUMMARY

In Diabetes Dataset, the various tests checks for the women are who have pregnancy and non pregnancy too. But Diabetes is varying for all the women's but some women's have same condition of Test result. To know the comparison and variance of test result and variation depends upon Age can Applicable In this ML Algorithms.

While Analysis with Visualization of Data using Google Studio can see the Multiple Information at the same time. It can Display and show the Results Using Many Type of charts and controls.

I Compare with Multiple Models, **Logistic, Support Vector Machine, NavieBayes, KNN, Random Forest, Decision Tree. Using of Advanced ML Feature Selection, K Best, RFE,KernelPCA, Principal Component Analysis and Linear Descriminant Analysis.**

Using Principal Component Analysis, by Dimensionality Reduction, PCA of **Logistic Regression-Classification using Grid Method** is Good Model and Accuracy. The Traditional way also given the accuracy report.