

Problem Statement:

- This is the age of Data as it is the lifeblood of all business. Data-driven decisions increasingly make the difference between keeping up with competition or falling further behind.
- Machine learning can be the key to unlocking the value of data and enacting decisions that keep a company ahead of the competition.
- Is it necessary to know all the machine learning development stages to get insights from the data. If we didn't perform all the stages in the correct manner, we might get wrong results and this will leads to wrong decision making.
- So, there is a need for a **general purpose dashboard** where **all machine learning stages** can be **performed in a structured manner** to get the correct results.

Proposed Design:



Feature Engineering and Model Building

- We are working on various several datasets from different domains
 - Flower Classification
 - House Price Prediction
 - Placement Prediction
 - Diabetes Prediction
 - Forest Fire Data
- After performing EDA and Data visualization, we select independent features and dependent features.
- Splitting Data into Training and Testing set with `test_size=0.20`
- Applied various Supervised machine learning algorithms [Classification and Regression]
- Evaluation Metrics: Finding best scores using hyperparameter tuning

Supervised Algorithms:

Classification Algorithms

- Logistic Regression
- K-Nearest Neighbours
- Support Vector Classifier
- Naive Bayes Classifier
- Decision Tree Classifier
- Random Forest Classifier'

Regression Algorithms

- Linear Regression
- Lasso Regression(L1 Regularizer)
- Support Vector Regressor
- Random Forest Regressor

RESULT

Classification Output

✓ Evaluation Metrics

Accuracy of Decision Tree Classifier classifier is 0.7857142857142857

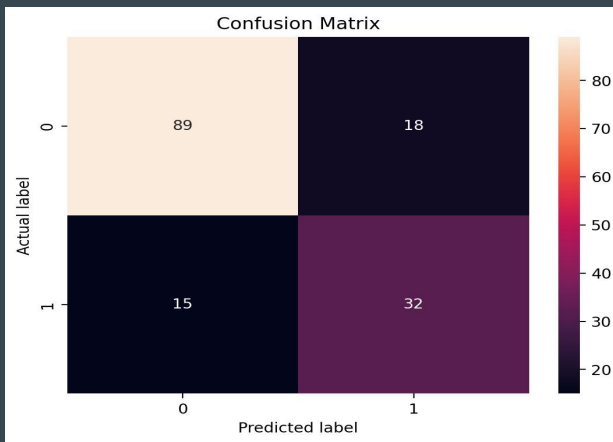
Precision of Decision Tree Classifier classifier is

	0
0	0.8558
1	0.6400

Recall of Decision Tree Classifier classifier is

	0
0	0.8318
1	0.6809

F1_Score of Decision Tree Classifier classifier is 0.7516978550837934



Regression Output

✓ Evaluation Metrics

RMSE of Linear Regression is 5.783509315085134

MAE of Linear Regression is 3.8429092204445023

R2 Score of Linear Regression is 0.589222384918251

It's considered to be Good Model

Conclusion:

- We developed a dashboard where no requirement of coding from the user to get insights from the data for decision making.
- Users can upload data in .csv format
- Performed Data Preprocessing
- Implemented various Supervised Machine Learning Algorithms.
- Finally, we performed Metrics Evaluation to find the best algorithms.

Future Scope:

- When data is highly imbalanced then our model is not performing good. So, we need an algorithm that will perfectly remove the biases from the data.
- We can add unsupervised learning algorithms to the dashboard.