Project 1: Model Evaluation

The aim of this project is to evaluate the model's performance on the testing dataset (Titanic dataset). Logistic Regression is applied to predict survival outcomes, and the model's performance is assessed using appropriate metrices such as accuracy, precision, recall, and F1-score. Results are further visualized with a confusion matrix and classification report.

Dataset Description:-

-Dataset: Titanic Dataset (Titanic Dataset.csv).
- Selected Features: Pclass, Sex, Age, Fare.
-Target Variable: Survived.
- Missing values handled using median.
-Categorical variables encoded numerically.

Methodology:

1. Data Preprocessing.
2. Splitting data into train (80%) and test (20%).
3. Training Logistic Regression model.
4. Model evaluation using Accuracy, Precision, Recall, and F1-score.
5. Visualization using Confusion Matrix and Classification report.

Results:

1. Accuracy: 1.0
2. Precision: 1.0
3. Recall: 1.0
4. F1 Score: 1.0

Classification Report:

-Class 0 → Precision: 1.00, Recall: 1.00, F1: 1.00
-Class 1 → Precision: 1.00, Recall: 1.00, F1: 1.00
-Overall Accuracy: 1.00

Discussion:-

-The model achieved 100% accuracy.

- Possible reasons include data leakage, a simplified dataset, or a lucky train-test split.

-Real-world datasets rarely produce perfect accuracy.

Conclusion:-

-Logistic Regression was successfully applied to the Titanic dataset for survival prediction.

-The model achieved 100% accuracy, precision, recall, and F1-score on the test set.

-The project clearly demonstrates the importance of model evaluation using metrics and confusion matrix.