**Titanic Survival Prediction 🚢**

**📌 Project Overview**

This project predicts whether a passenger survived the Titanic disaster using the **Logistic Regression** algorithm.  
The dataset is preprocessed to handle missing values, encode categorical variables, and prepare it for model training.

**📂 Dataset**

* **Source:** Kaggle Titanic Dataset
* **Features:** Pclass, Sex, Age, SibSp, Parch, Fare, Embarked
* **Target:** Survival (1 = Survived, 0 = Did not survive)

**⚙️ Technologies Used**

* Python
* Pandas, NumPy
* Scikit-learn

**🔍 Model Approach**

1. **Data Cleaning:**
   * Identified and handled missing values in Age and Embarked columns.
   * Converted categorical variables into numerical form using one-hot encoding.
2. **Modelling:**
   * Applied Logistic Regression.
   * Split dataset into **training** and **validation** sets.
3. **Evaluation:**
   * **Training Accuracy:** 81.18%
   * **Validation Accuracy:** 75.98%

**💡 Insights**

* **Passenger Class:** Individuals from **upper-class (Pclass 1)** had a significantly higher chance of survival compared to those from lower classes.
* **Gender:** Female passengers had nearly **double the survival rate** of male passengers.
* **Age Factor:** Younger passengers showed a slightly higher survival probability than older ones.

**📈 Model Performance**

The model shows good predictive capability with Logistic Regression, highlighting key survival factors such as passenger class, gender, and age.

**▶️ Open in Google Colab**

Click the button below to open and run this notebook in Google Colab:

<https://colab.research.google.com/github/Pooja-Pj205/Titanic-Survival-prediction/blob/main/Titanic.ipynb>