BVRIT HYDERABAD

College of Engineering for Women

SEAWISE GIANT

Team No: 07

Team Members:

B.Amulya-20WH1A1294
B.Harshitha-20WH1A1295
K.Sri Nithya-20WH1A1296
M.Manjitha-20WH1A1297
D.Kalpana-20WH1A1298

VISHNU
UNIVERSAL LEARNING

Problem Statement

- Your task is to predict whether or not a passenger survived the sinking of the Seawise Giant (a synthetic, much larger dataset based on the actual Titanic dataset). For each PasengerId row in the test set, you must predict a 0 or 1 value for the Survived target.

Python Packages Used

- Numpy
- Pandas
- matplotlib
- sklearn
- seaborn

Algorithm

- Random Forest.
- Decision Tree.
- Logistic Regreesion.
- KNN.

VISHNU
UNIVERSAL LEARNING

Random Forest

- This algorithm works by randomly selecting a subset of features from the input data and constructing a decision tree based on this subset of features. This process is repeated multiple times to create multiple decision trees, each based on a different subset of features.
- When making a prediction for a new data point, the random forest algorithm aggregates the predictions of all the individual decision trees and returns the most common prediction as the final output. This helps to reduce the impact of overfitting and noise in the input data, leading to more accurate predictions.

Decision Tree

- A decision tree is a structure that is used to predict the value of a target value based on certain input variables.Each node of the tree represents a decision that will affect the outcome of the target value.One advantage of using a decision tree to predict behavior is that decision trees are easy to understand and follow.It is easy to follow the conditional logic that decision trees use.

VISHNU
UNIVERSAL LEARNING

Logistic Regression

- Logistic regression is a classification algorithm.It predicts the probability of an input belonging to a certain set by separating data into two regions. Logistic regression is used when the response variable will be binary, for example, pass/fail.

VISHNU
UNIVERSAL LEARNING

KNN

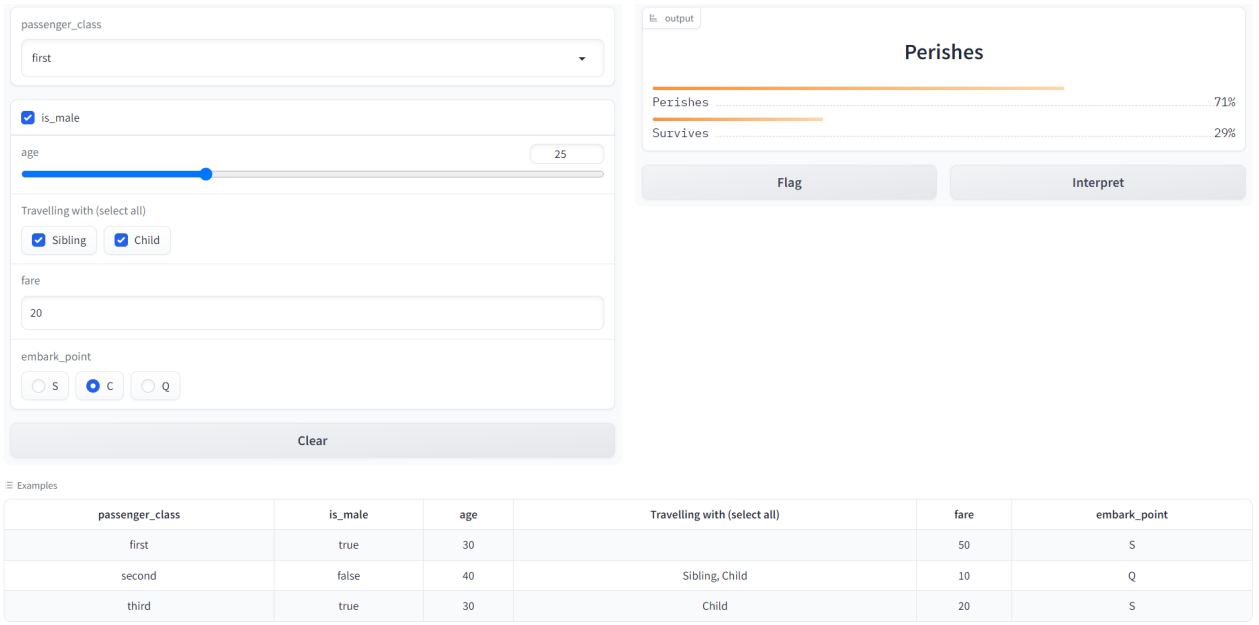
- K-Nearest Neighbors is an algorithm that can be used for both classification and regression. It works by taking in training data and then seeing which data points are close to a data point and then classifying that data point as part of the same class as the majority of the k-nearest data points. An advantage of k-nearest neighbors is that it is usually pretty accurate and works well for non-linear data. A disadvantage is that it has to store all the training data which can lead to memory and runtime issues.

Output

- RANDOM FOREST CLASSIFIER

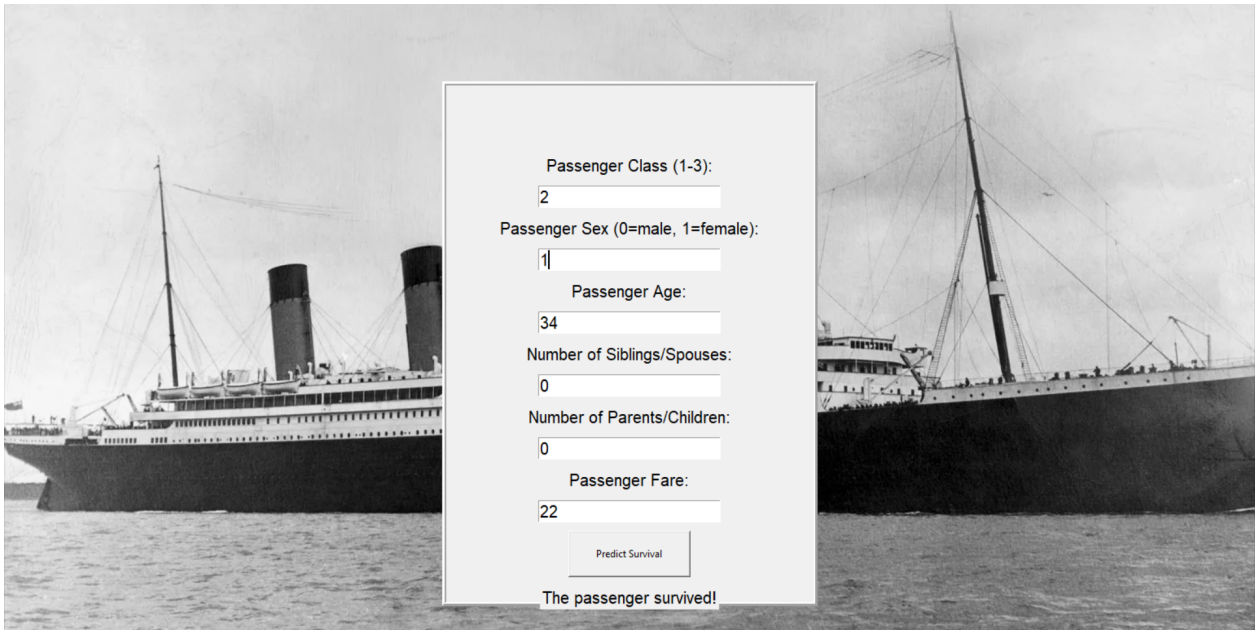
1	Passenger	Survived
2	100000	0
3	100001	0
4	100002	1
5	100003	0
6	100004	1
7	100005	0
8	100006	0
9	100007	0
10	100008	0
11	100009	0
12	100010	1
13	100011	0
14	100012	0
15	100013	0

Gradio



passenger_class	is_male	age	Travelling with (select all)	fare	embark_point
first	true	30		50	S
second	false	40	Sibling, Child	10	Q
third	true	30	Child	20	S

GUI


Passenger Class (1-3):
2
Passenger Sex (0=male, 1=female):
1
Passenger Age:
34
Number of Siblings/Spouses:
0
Number of Parents/Children:
0
Passenger Fare:
22
Predict Survival
The passenger survived!

VISHNU
UNIVERSAL LEARNING

Comparison Table

Score	
80.73	Random Forest
80.73	Decision Tree
76.00	Logistic Regression
75.47	KNN

VISHNU
UNIVERSAL LEARNING

THANK YOU