Topic: Predict if a customer is likely to buy travel insurance using R/Python

Problem Statement:

This project aims to predict whether a customer will purchase travel insurance based on their demographic and travel information. To train and evaluate eight machine learning models. To compare their performance in terms of accuracy and other relevant metrics.

Data Source: Toffee

Approach:

1. Data Loading and Preprocessing
2. Exploratory Data Analysis (EDA)
3. Model Selection and Training
4. Model Evaluation

Algorithms Used:
1. Logistic Regression
2. Decision Tree
3. Random Forest
4. K-Nearest Neighbors (KNN)
5. Naive Bayes (Gaussian NB, Bernoulli NB, Multinomial NB)
6. SVC

Data Preprocessing:
1. Missing Values
2. Encoding(One hot Encoding)
3. Standardization

Model Evaluation:
The models were evaluated using 10-fold cross-validation, and the accuracies were compared using a boxplot, Scatterplot, Pie Chart
Results:
The boxplot below shows the accuracy distribution of different classification models.

Conclusion:

Sr. No.	Models	Accuracy
1	BernoulliNB()	71%
2	GaussianNB()	71%
3	K-Nearest Neighbors (KNN)	65%
4	MultinomialNB()	65%
5	Support Vector Machines (SVM)	63%
6	Random Forest Classifier	57%
7	DecisionTreeClassifier	52%
8	LogisticRegression	29%