Travel Insurance Prediction App Report

Abstract

A Tour & Travels Company Is Offering Travel Insurance Package To Their Customers.The New Insurance Package Also Includes Covid Cover. The Company Requires To Know The Which Customers Would Be Interested To Buy It Based On Its Database History.The Insurance Was Offered To Some Of The Customers In 2019 And The Given Data Has Been Extracted From The Performance/Sales Of The Package During That Period.The Data Is Provided For Almost 2000 Of Its Previous Customers And You Are Required To Build An Intelligent Model That Can Predict If The Customer Will Be Interested To Buy The Travel Insurance Package Based On Certain Parameters.

Dataset Details:

1) Numeric Datatypes:

--> "Index", "Age", "AnnualIncome", "FamilyMembers", and "TravelInsurance" columns have numeric data types (int64).

2) Categorical Datatypes:

--> "Employment Type", "GraduateOrNot", "FrequentFlyer", and "EverTravelledAbroad" columns have categorical data (object)

3) Target:

--> TravelInsurance (represented as 0 or 1)

Introduction

Travel insurance is a critical component for mitigating risks associated with travel. The decision to purchase travel insurance is influenced by various factors such as age, income, travel history, and personal circumstances. The Travel Insurance Prediction App leverages these factors to offer predictions on an individual's propensity to invest in travel insurance.

Technical Logic

The core of the Travel Insurance Prediction App is a Random Forest Classifier, a machine learning model known for its high accuracy and ability to handle non-linear data. The model is trained on a dataset containing attributes related to personal demographics, travel history, and socio-economic factors. An OrdinalEncoder translates categorical variables into numerical values for model processing. Calibration with cross-validation refines the prediction probabilities, enhancing the reliability of the results.

Usage

To use the app:

- Provide personal and travel-related details, such as age, employment type, graduate status, frequency of travel, history of international travel, annual income, family size, and chronic diseases which are present on left pane.
- Click on "Predict Insurance Need" to get a prediction.

Two sample results can be generated:

- Likely to purchase insurance if age > 30, income > $500,000, and family members > 4 because those features are best features for predicting the results.
- Not likely to purchase insurance for other input features combination.

Visualization Logic

- ROC Curve: Illustrates the diagnostic ability of the classifier by plotting the true positive rate against the false positive rate at various threshold settings. The area under the curve (AUC) serves as a measure of the model's accuracy.
- Feature Importance: Reveals the relative importance of each feature in making accurate predictions. Higher bars indicate more significant predictors.
- Feature Correlation Heatmap: Displays the correlation coefficients between variables. Darker colors represent stronger relationships, aiding in understanding the interdependencies among features.

Summary

The app is an exemplar of applied machine learning, assisting users in understanding the likelihood of purchasing travel insurance. Its predictions are based on a calibrated Random Forest Classifier, with visual aids for users to interpret the model's performance and the significance of their input data.