Name: Soh Eng Khiong

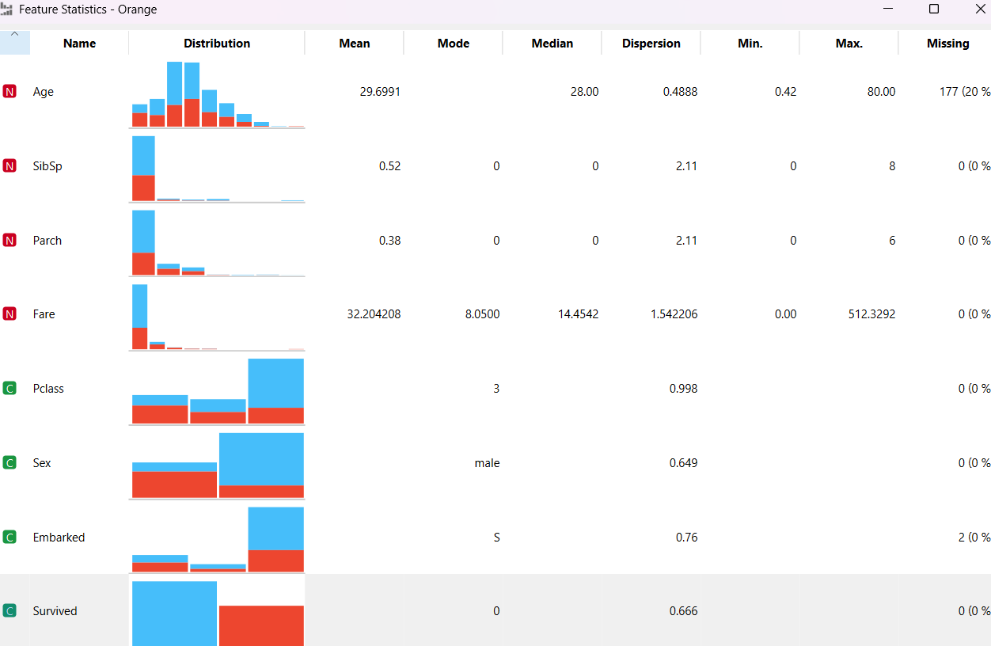
Student ID: 5129923N

Class NSDAI2 D

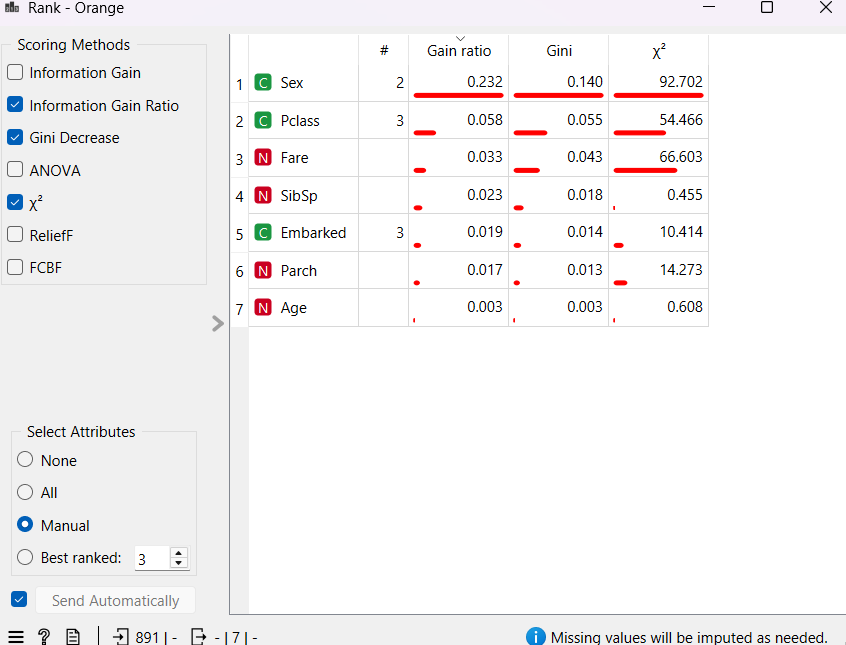
Using Orange software for supervised learning of the Titanic dataset to model the survival responses.

## Overview of the Machine Learning Process Flow.

## Exploratory Data Analysis:



Ranking of the predictors



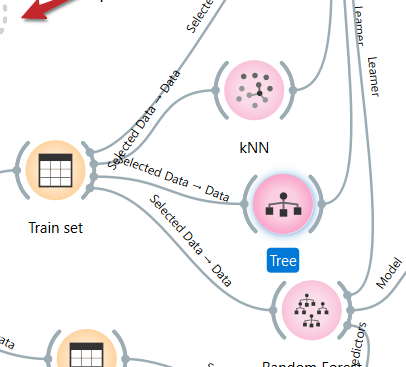
The rankings indicate that Sex, Pclass, and Fare significantly impact the model's survival response.

## Preprocess Data

Data splitting

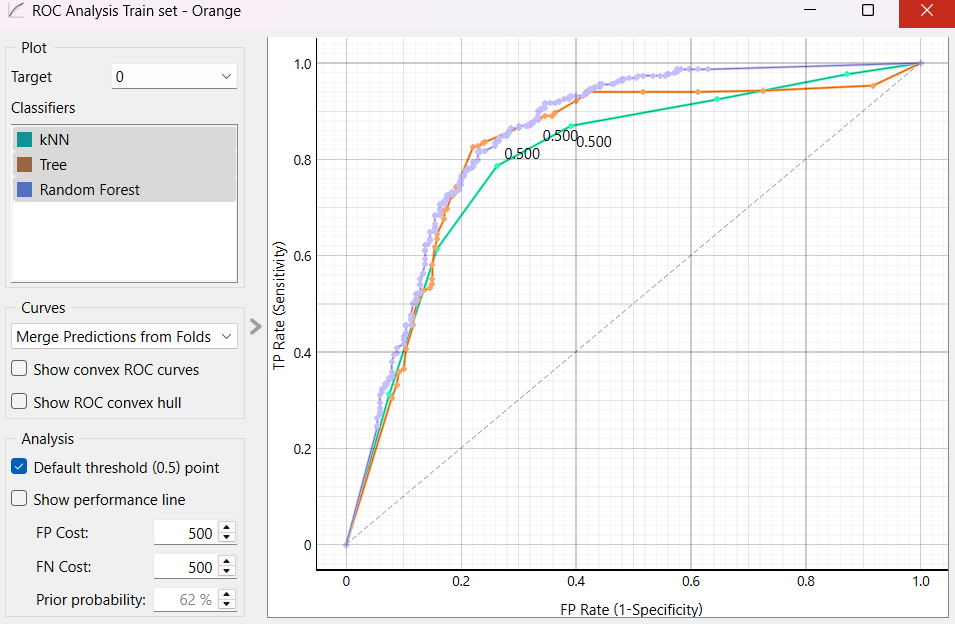
The data is split into train and test set with 70:30 ratio.

## Model selection

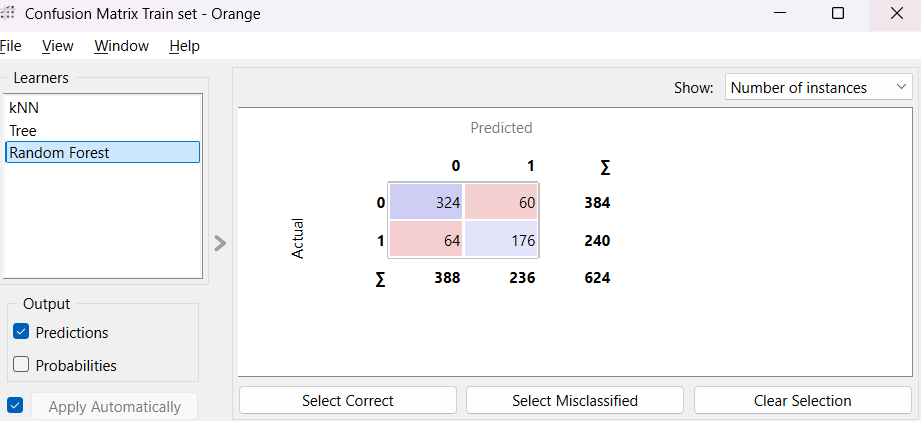
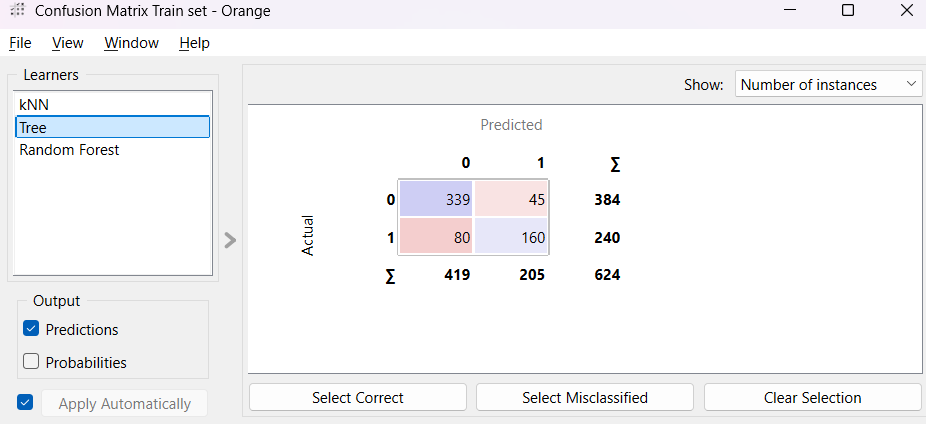
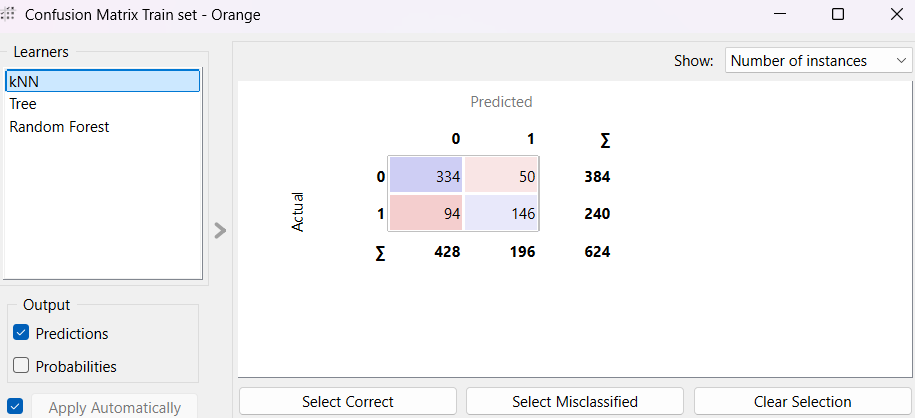


I used KNN, Decision Tree, and Random Forest models to predict survival. The training set trains the model, and the test set validates it.

## Evaluate the model's performance

ROC analysis of the training dataset using the kNN model.

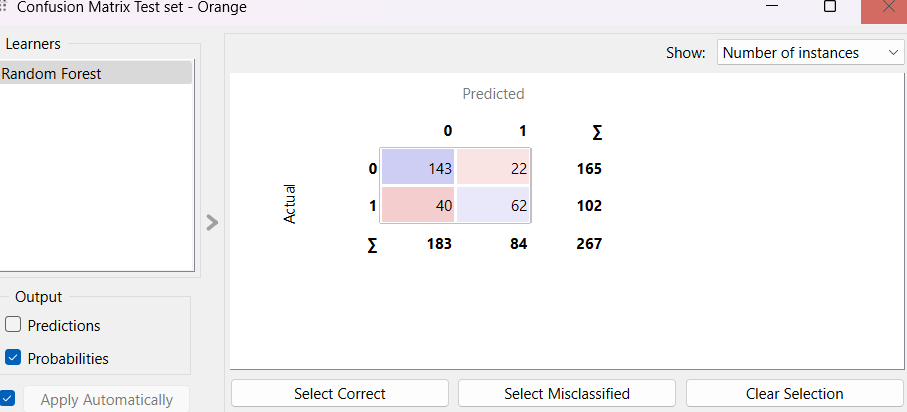
Confusion Matrix



Based on the confusion matrix of the model, the Random Forest algorithm demonstrates superior performance.

## Validating with Test data

Confusion matrix for the test data



The confusion matrix demonstrates that the Random Forest model accurately classifies survival responses.