

Predictors of Survival for Titanic Passengers

Alden Chen, Birinder Singh

2018-11-17

Introduction

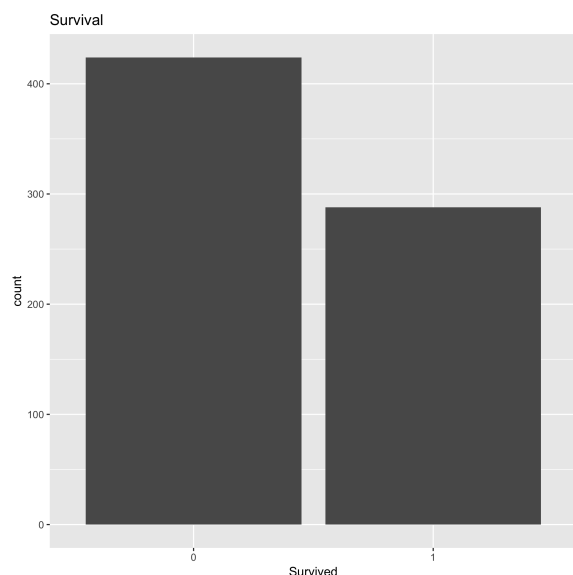
In this analysis, we look at data on passengers from the Titanic. Using seven features, we fit a decision tree model to find out which two features are the best predictors of survival for Titanic passengers.

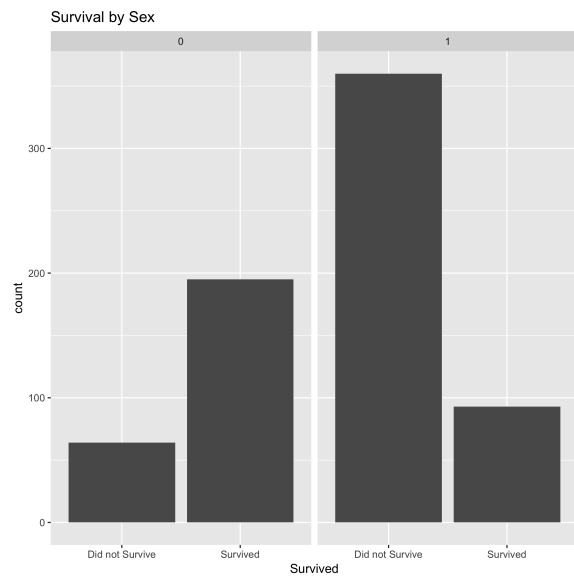
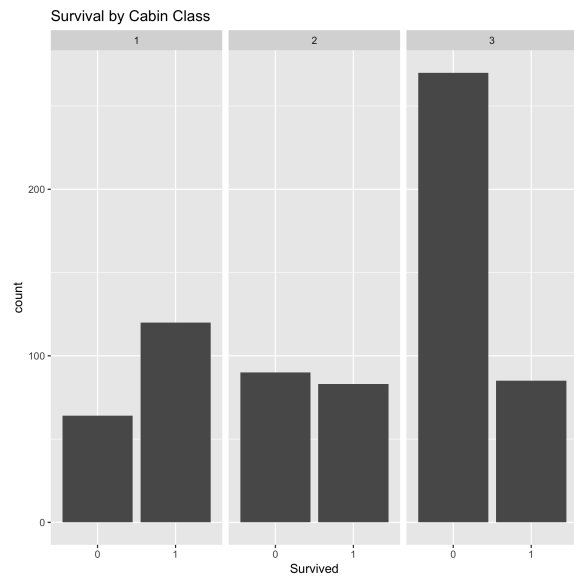
Exploratory Data Analysis:

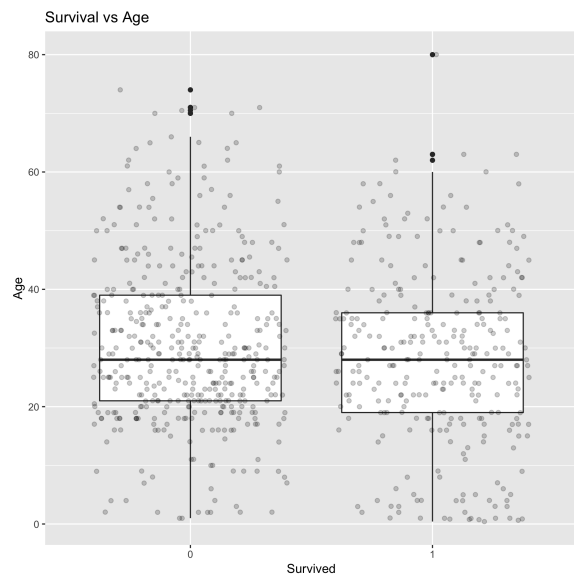
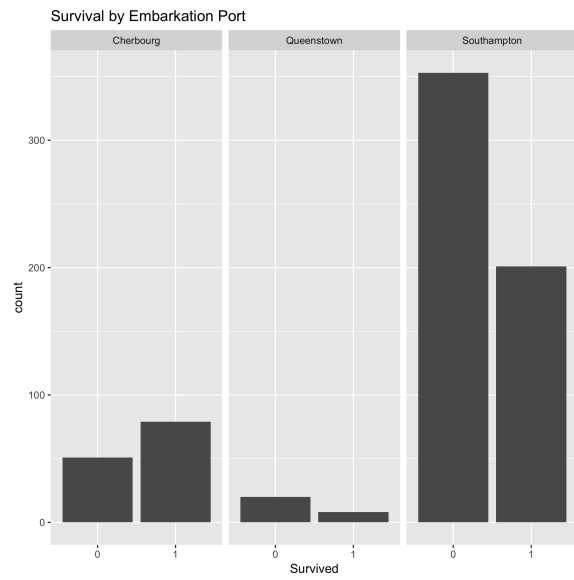
Our target variable is **Survived**, a dummy variable that takes the value 0 if a passenger did not survive and 1 if he or she did survive. We have data on 712 passengers for these seven features

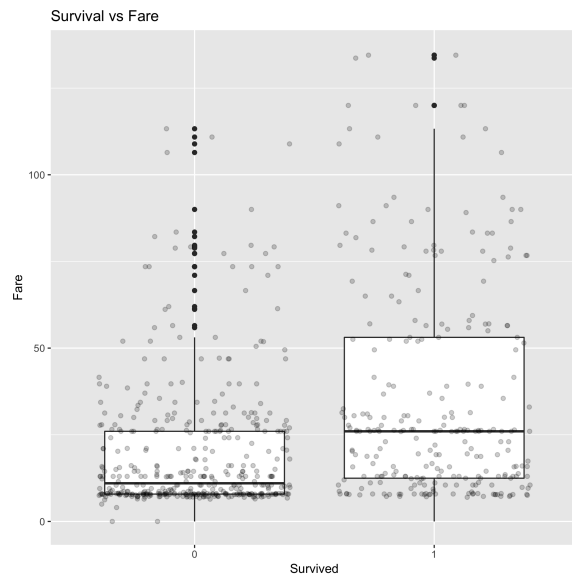
Feature	Description
Passenger Class	First, Second or Third Class
Sex	Male or Female
Age	Age of each passenger years
Siblings/Spouses	Number of siblings and spouses aboard the Titanic
Parents/Children	Number of arents and children aboard the Titanic
Fare	The amount each passenger paid for a his or her ticket
Port of Embarkation	Cherbourg, Queenstown, or Southampton

In this section, we look at some plots of the data and discuss their implications for prediction.



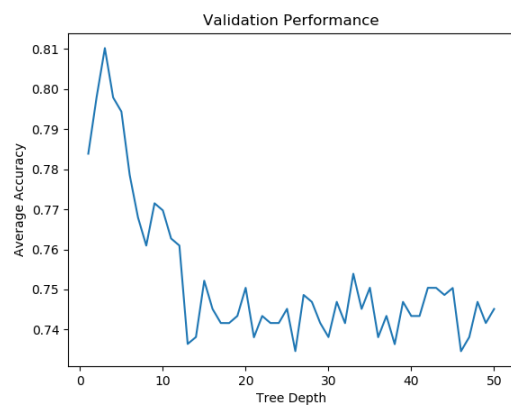






Model Development

- Discussion of cross-validation to select hyperparameter and fitting the model



Results

