



In the Game of Thrones you either win or you die...

Supervised Learning

Supervised learning is a data mining task in which a machine is trained based on a number of labeled observations. By logging the outcome of every possible scenario, supervised learning allows a machine to make prediction based on precedent.

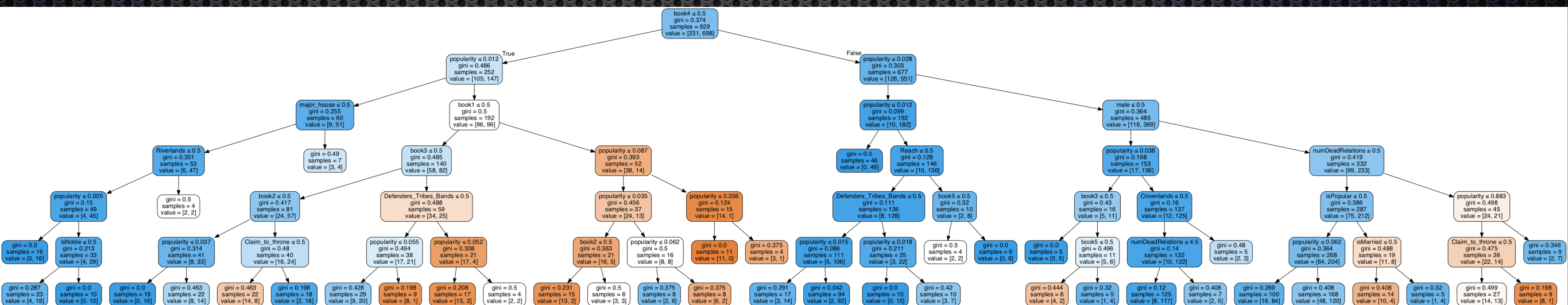
I will use supervised learning to train my model to make predictions on whether a character in the popular TV show, Game of Thrones, survives

The Predictors

- Gender: Male or Not
- House : Which house they were born in
- Book: Which book do they make an appearance in
- Married: Are they married?
- Nobility: Are they noble?
- Number of Dead relatives
- Are they popular? How popular?

Decision Tree

- ✦ The decision tree model is a flow chart that uses the observed behavior of different features to make predictions on our target.
- ✦ This decision tree was able to predict with 79% Accuracy and 87% accuracy whether a character will survive or not



Random Forest

- ✦ This model makes use of several decision trees to make a prediction on a target
- ✦ The random forest model that I created is able to predict with 83% accuracy and 90% precision on whether a character dies or not

XGB Boost

- ✦ The model I ultimately chose however, is the XGB Boost classification model.
- ✦ XGB is a high performing machine learning method which implements gradient boosting to make accurate predictions
- ✦ My XGB model is able to predict with 83% accuracy and 90% precision on whether a character survives or not. It was able to do this by using half the amount of features needed in the first two models

Conclusion

- ✦ Decision classification uses a flow-chart observing all features provided to make predictions on a target
- ✦ Random forest classification uses several different decision trees to weigh in on which features hold more weight and uses it's observations to make predictions on a target
- ✦ XGB classification proves to be the most efficient model using the least amount of features and shortest time to make predictions that are on par with the above mentioned models.