UDACITY

Logout

Back to Machine Learning Engineer Nanodegree

Finding Donors for CharityML

REVIEW | CODE REVIEW | HISTORY

Meets Specifications

Congrats!! 🎉

You did a really good job here 👌, going through the project, selecting, optimizing and fitting classification algorithms. You have properly discussed the pros an cons of each algorithm and reasoned the reason behind the selection and also reviewed the impact of feature reduction.

I really hope I was able to help you out.

Enjoy the rest of the course!

Exploring the Data

Student's implementation correctly calculates the following:

- **Number of records**
- **Number of individuals with income >$50,000**
- **Number of individuals with income <=$50,000**
- **Percentage of individuals with income > $50,000**

This part was already fine!

Preparing the Data

Student correctly implements one-hot encoding for the feature and income data.

This part was already fine!

Evaluating Model Performance

Student correctly calculates the benchmark score of the naive predictor for both accuracy and F1 scores.

The accuracy and F-score calculated for the Naïve Bayes predictor are correct: `Naive Predictor: [Accuracy score: 0.2478, F-score: 0.2917]`.
Well done!

The pros and cons or application for each model is provided with reasonable justification why each model was chosen to be explored.

Please list all the references you use while listing out your pros and cons.

You have selected three classification algorithm and understood the strengths and weaknesses of the models, giving examples of real-world application. Also, a list of references is provided. 👍

Student successfully implements a pipeline in code that will train and predict on the supervised learning algorithm given.

This part was already fine!

Student correctly implements three supervised learning models and produces a performance visualization.

This part was already fine!

Improving Results

Justification is provided for which model appears to be the best to use given computational cost, model performance, and the characteristics of the data.

This part was already fine!

Student is able to clearly and concisely describe how the optimal model works in layman's terms to someone who is not familiar with machine learning nor has a technical background.

Well done here!

I also like when examples are used. Take a look this Quora discussion for inspiration, one of the first answers have a quite good explanation about comparing Gradient Boosting classifier to a group of teams. Also, this blog has a pretty good example of random forests making an analogy with a movie recommender.

The final model chosen is correctly tuned using grid search with at least one parameter using at least three settings. If the model does not need any parameter tuning it is explicitly stated with reasonable justification

need any parameter tuning if it is explicitly stated with reasonable justification.

Well done here!

Student reports the accuracy and F1 score of the optimized, unoptimized, models correctly in the table provided. Student compares the final model results to previous results obtained.

Well done here!

Feature Importance

Student ranks five features which they believe to be the most relevant for predicting an individual's' income. Discussion is provided for why these features were chosen.

Already correct!

Student correctly implements a supervised learning model that makes use of the `feature_importances_` attribute. Additionally, student discusses the differences or similarities between the features they considered relevant and the reported relevant features.

Well done here!

Student analyzes the final model's performance when only the top 5 features are used and compares this performance to the optimized model from Question 5.

Well done here!

DOWNLOAD PROJECT

RETURN TO PATH