Brandenburgische
Technische Universität
Cottbus - Senftenberg

Exercises

Feature Importance and Impact

Exercises

Brandenburgische
Technische Universität
Cottbus - Senftenberg

- Read data *FBPS-ValidationData.csv*
- Why is it necessary to make a copy of the read data?
- Which are the output data?
- Change output into the following form:
  - If birthpos == 1 -> y = 1
  - If birthpos == lastchild -> y = 3
  - Else -> y = 2
- Select only probands with native language English

- Which columns can you delete?

Exercises

Brandenburgische
Technische Universität
Cottbus - Senftenberg

- Use the *RandomForestClassifier()* for the prediction
- Print accuracy and confusion matrix
- Reduce features using *LinearDiscriminantAnalysis*
- Compare the results without and with using LDA
- Is the dataset balanced? Why or why not?
- Compare *accuracy* and *balanced accuracy*

Exercises

Brandenburgische
Technische Universität
Cottbus - Senftenberg

- Create the 3 Tree-based models from the lecture
- Fit them with the training data
- Compute feature importance
- Rank it
- Print the cardinality of the features *gender* and *I sometimes ruin my jokes by laughing in the middle of them*

Exercises

Brandenburgische
Technische Universität
Cottbus - Senftenberg

- Compute coefficients of LDA
- Normalize the coefficients
- How much each class is represented in the model?
- Print weighted features sorted into a HTML table
- Who doesn’t insult people?
- Who doesn’t boss people around?
- What is the advantage of Permutation Feature Importance and why?

Exercises

Brandenburgische
Technische Universität
Cottbus - Senftenberg

- Read the fetch_california_housing data

```
from sklearn.datasets import fetch_california_housing
X, y = fetch_california_housing(return_X_y=True, as_frame=True)
```

- Use only 100 datasets
- Use RandomForestRegressor
- The target variable is the median house value
- Plot PDP of the feature average number of rooms per household
- Plot ICD of the feature average number of household members
- Plot the relationship between average number of bedrooms per household and median house age