Adult Census

Generated by Keita Onabuta on 07/20/2022

Model Summary

Purpose

classification to predict house income

This is a classification model

How the model is evaluated

This model is evaluated on a test set with 1000 datapoints.

Target values

Here are your defined target values:

- accuracy_score: >= 0.8
- Top important features: 6

Model Performance

Observe evidence of your model performance here:

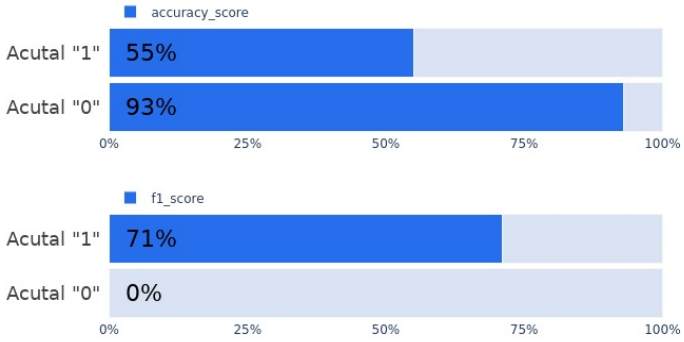
84% Accuracy

84% of data points have the correct prediction.

Accuracy = correct predictions / all predictions
= (133 + 713) / 1000

None

	Acutal "1"	Acutal "0"
Predicted "1"	133 correct prediction	46
Predicted "0"	108	713 correct prediction



Data Explorer

Evaluate your dataset to assess representation of identified subgroups:

Age

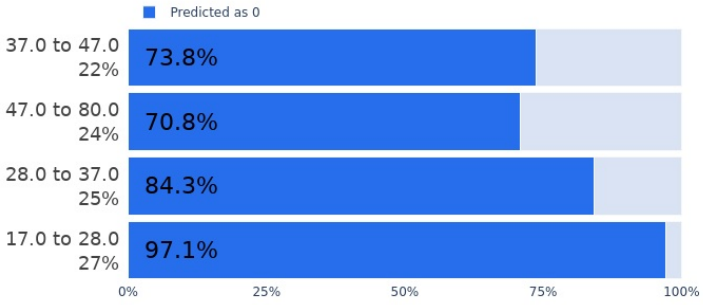
"17.0 to 28.0" have 96.7% accuracy_score

"28.0 to 37.0" have 80.7% accuracy_score

"47.0 to 80.0" have 77.4% accuracy_score

"37.0 to 47.0" have 82.1% accuracy_score

Predicted classification output of the different subgroups are as follows:

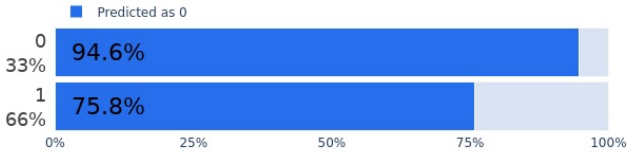


Sex

"1" have 81.0% accuracy_score

"0" have 91.7% accuracy_score

Predicted classification output of the different subgroups are as follows:



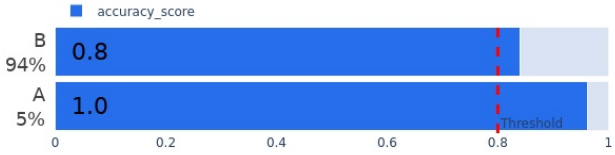
Cohorts

Observe evidence of model performance across your passed cohorts:

My Cohorts: accuracy_score

- A: Young
- B: Not Young

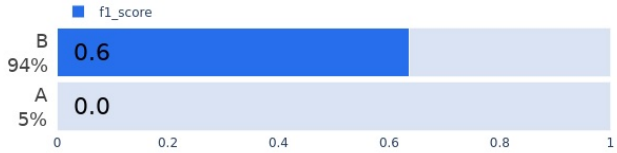
My prebuilt dataset cohorts: accuracy_score



My Cohorts: f1_score

- A: Young
- B: Not Young

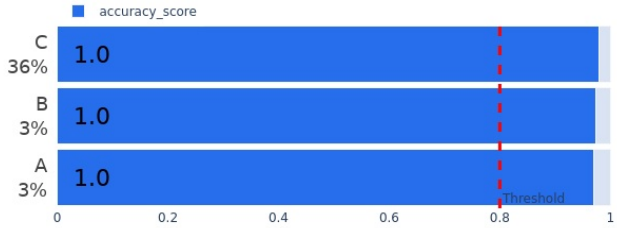
My prebuilt dataset cohorts: f1_score



Highest ranked cohorts: accuracy_score

- A: Capital Gain > 4518.00
  AND
  Relationship > 3.50
- B: Age <= 30.50
  AND
  Hours per week > 45.50
  AND
  Relationship <= 3.50
- C: Age <= 44.50
  AND
  Hours per week <= 45.50
  AND
  Relationship <= 3.50

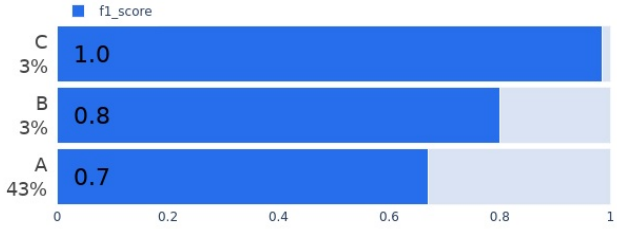
Highest ranked cohorts: accuracy_score



Highest ranked cohorts: f1_score

- A: Relationship > 3.50
- B: Age <= 30.50
  AND
  Hours per week > 45.50
  AND
  Relationship <= 3.50
- C: Capital Gain > 4518.00
  AND
  Relationship > 3.50

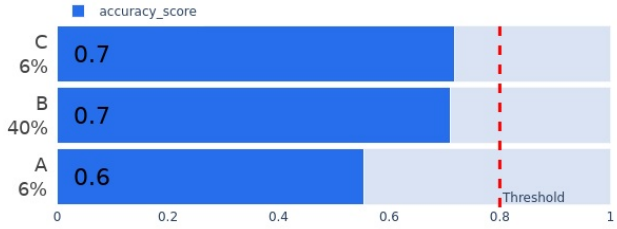
Highest ranked cohorts: f1_score



Lowest ranked cohorts: accuracy_score

- A: Hours per week > 53.00
  AND
  Capital Gain <= 4518.00
  AND
  Relationship > 3.50
- B: Capital Gain <= 4518.00
  AND
  Relationship > 3.50
- C: Age > 30.50
  AND
  Hours per week > 45.50
  AND
  Relationship <= 3.50

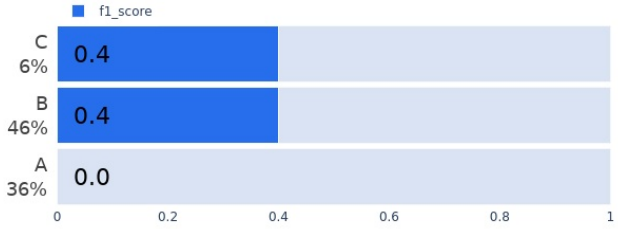
Lowest ranked cohorts: accuracy_score



Lowest ranked cohorts: f1_score

- A: Age <= 44.50
  AND
  Hours per week <= 45.50
  AND
  Relationship <= 3.50
- B: Hours per week <= 45.50
  AND
  Relationship <= 3.50
- C: Age > 30.50
  AND
  Hours per week > 45.50
  AND
  Relationship <= 3.50

Lowest ranked cohorts: f1_score

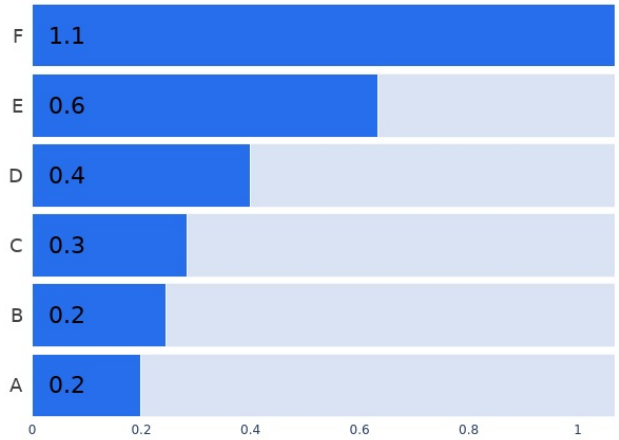


Feature relevance (explanability)

Understand factors that have impacted your model predictions the most. These are factors that may account for performance levels and differences.

- A: Marital Status
- B: Hours per week
- C: Age
- D: Capital Gain
- E: Education-Num
- F: Relationship

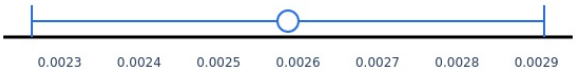
Feature Importance



Causal analysis answers real world what if questions about how changes of treatments would impact a real world outcome.

Age

On average, increasing "Age" by 1 unit increases the prediction outcome by 0.003



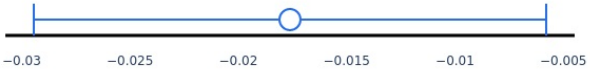
Top data points responding the most to treatment on "Age":

Data points which experience the largest treatment impact (causal effect) when adjusting "Age"

Index	Current Value	Recommended Treatment	Effect Estimate
999	34.0	increase	0.01
366	39.0	increase	0.01
382	34.0	increase	0.01

Sex

On average, increasing "Sex" by 1 unit increases the prediction outcome by -0.018



Top data points responding the most to treatment on "Sex":

Data points which experience the largest treatment impact (causal effect) when adjusting "Sex"

Index	Current Value	Recommended Treatment	Effect Estimate
960	0	1	0.12
276	0	1	0.12
530	0	1	0.12