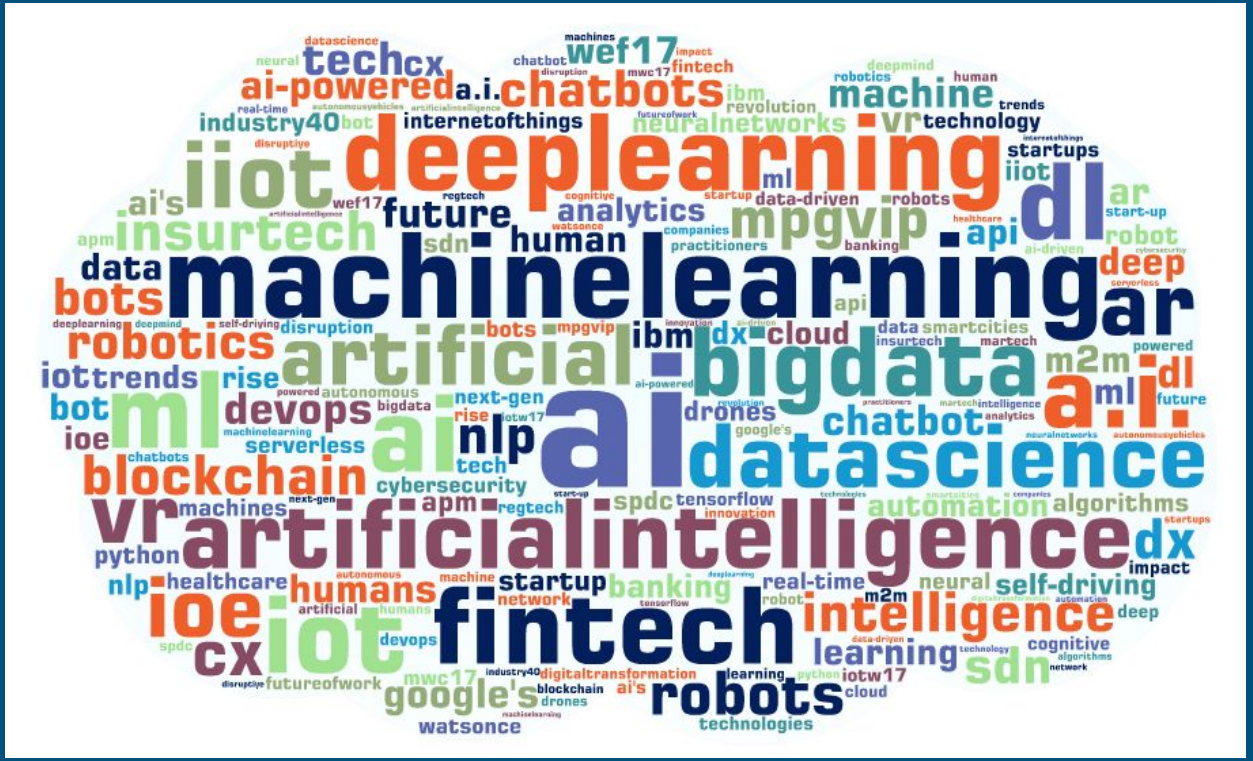
UCI

Aakash Sharma, Michael Knight,
Nicole Scott


datascience
techcx
wef17
machines
chatbot
fintech
ai-powered
a.i.
chatbots
machine
robotics
human
deepmind
industry40
internetofthings
neuralnetworks
vr
technology
trends
deeplearning
iiot
startups
dl
ar
.ai's
future
analytics
mpgvip
api
robot
insurtech
human
data
machinelearning
bots
deep
robotics
artificial
bigdata
ai
ml
datascience
iottrends
rise
bot
devops
nlp
chatbot
ioe
serverless
blockchain
cybersecurity
vr
artificialintelligence
dx
python
nlp
healthcare
humans
startup
banking
real-time
neural
self-driving
ioe
iot
fintech
intelligence
cx
learning
sdn
google's
robots
watsonce
technologies

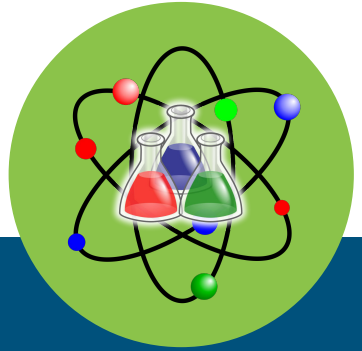
Problem

Predict if a person's income is in excess of $50,000 given certain profile information. More specifically, we need to generate predicted probabilities of income being above $50,000. Our team is constrained by the maximum number of features.

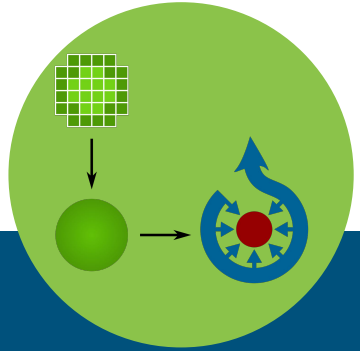
Team Features Constraint:

- Your choice of algorithm
- Limited to a maximum of 20 features
- Your choice of samples

Steps



EDA/Cleaning



Data Conversion
& Encoded

Feature Engineer
& Models

Our Best Features

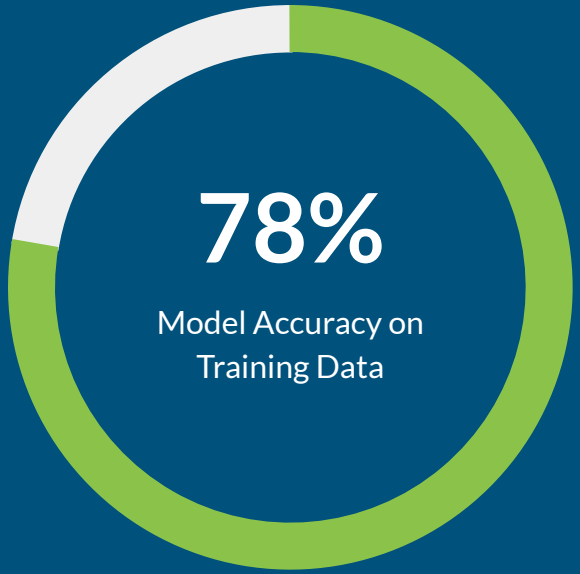
Strongest Positive Correlation

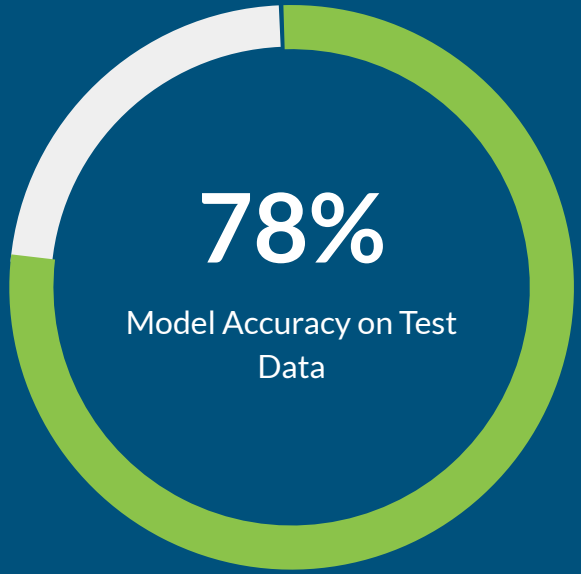
- 'Age'
- 'Hours per Week'
- ‘Education Number’
- 'Marital Status Married Civ Spouse'
-

Strongest Negative Correlation

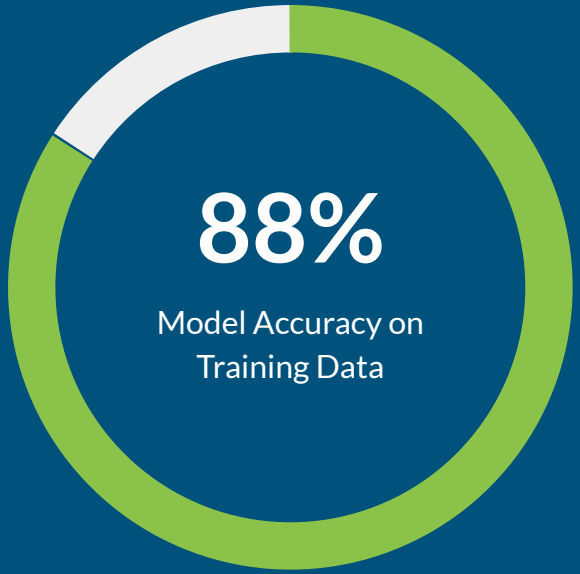
- 'Marital Status Never Married'
- 'Relationship Own Child'
- ‘Relationship Not In Family'
- ‘Occupation Other Service'
- ‘Relationship Unmarried'

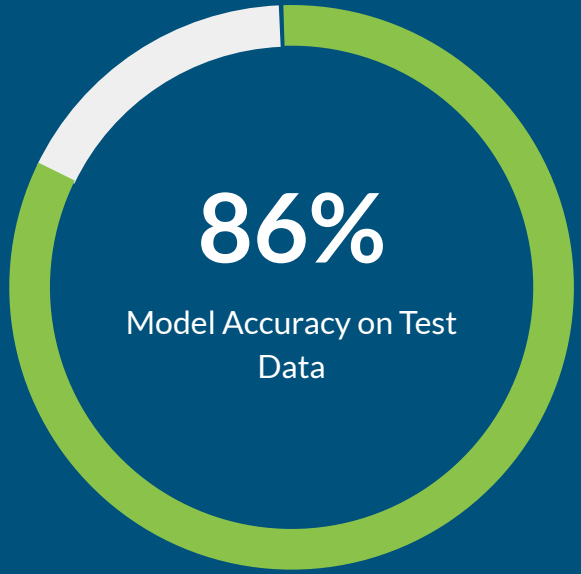
Supporting **Information: Naive Bayes**

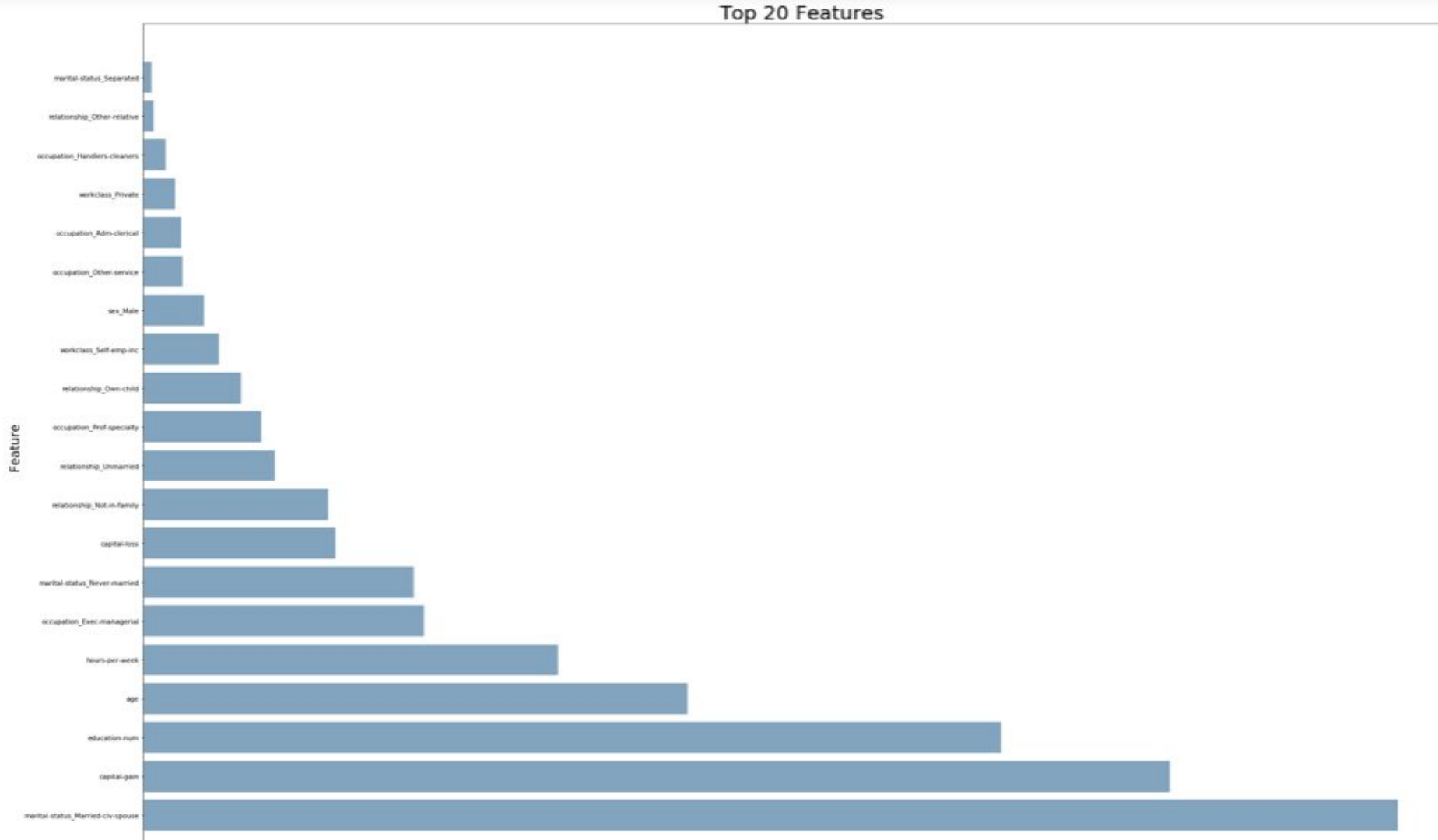


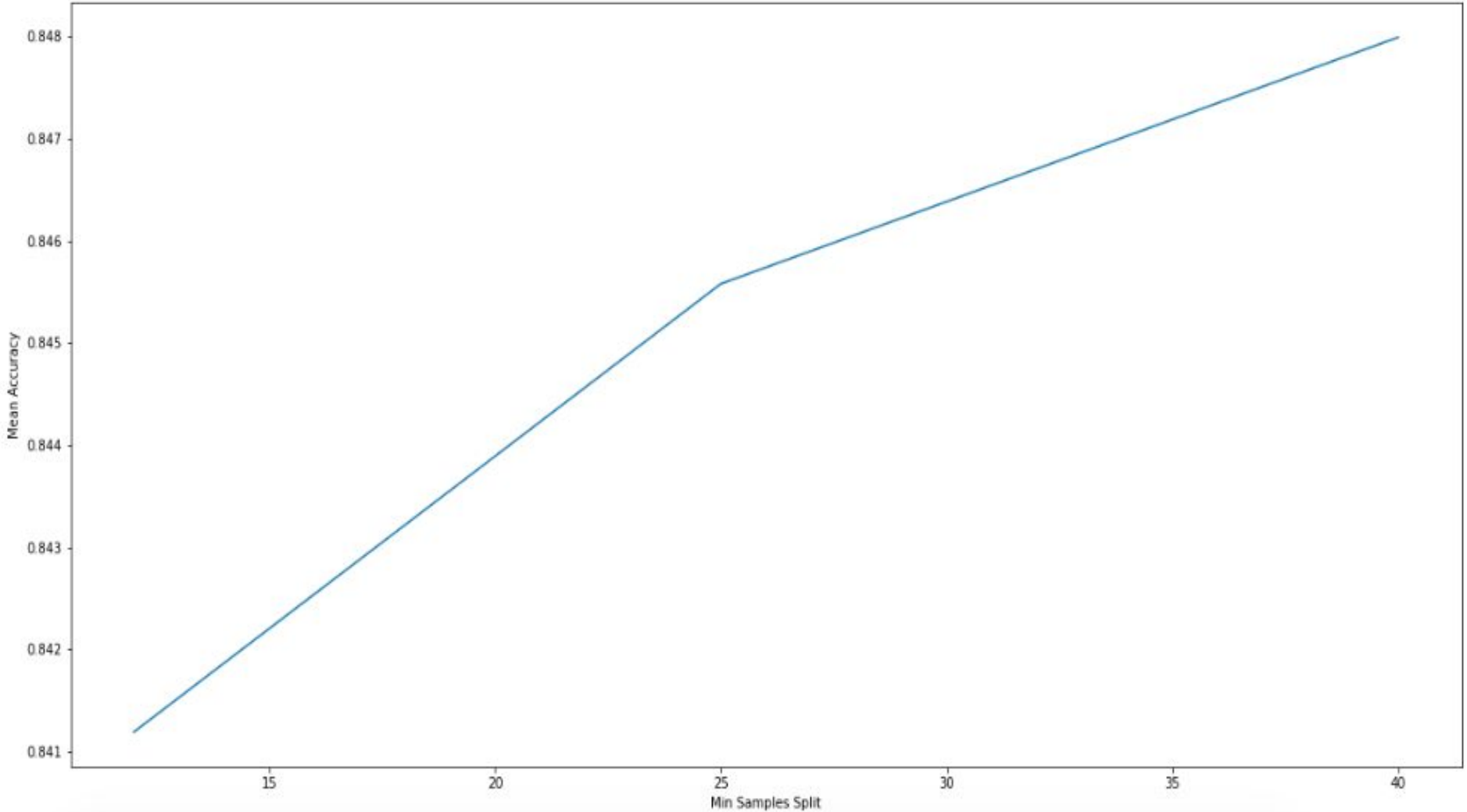


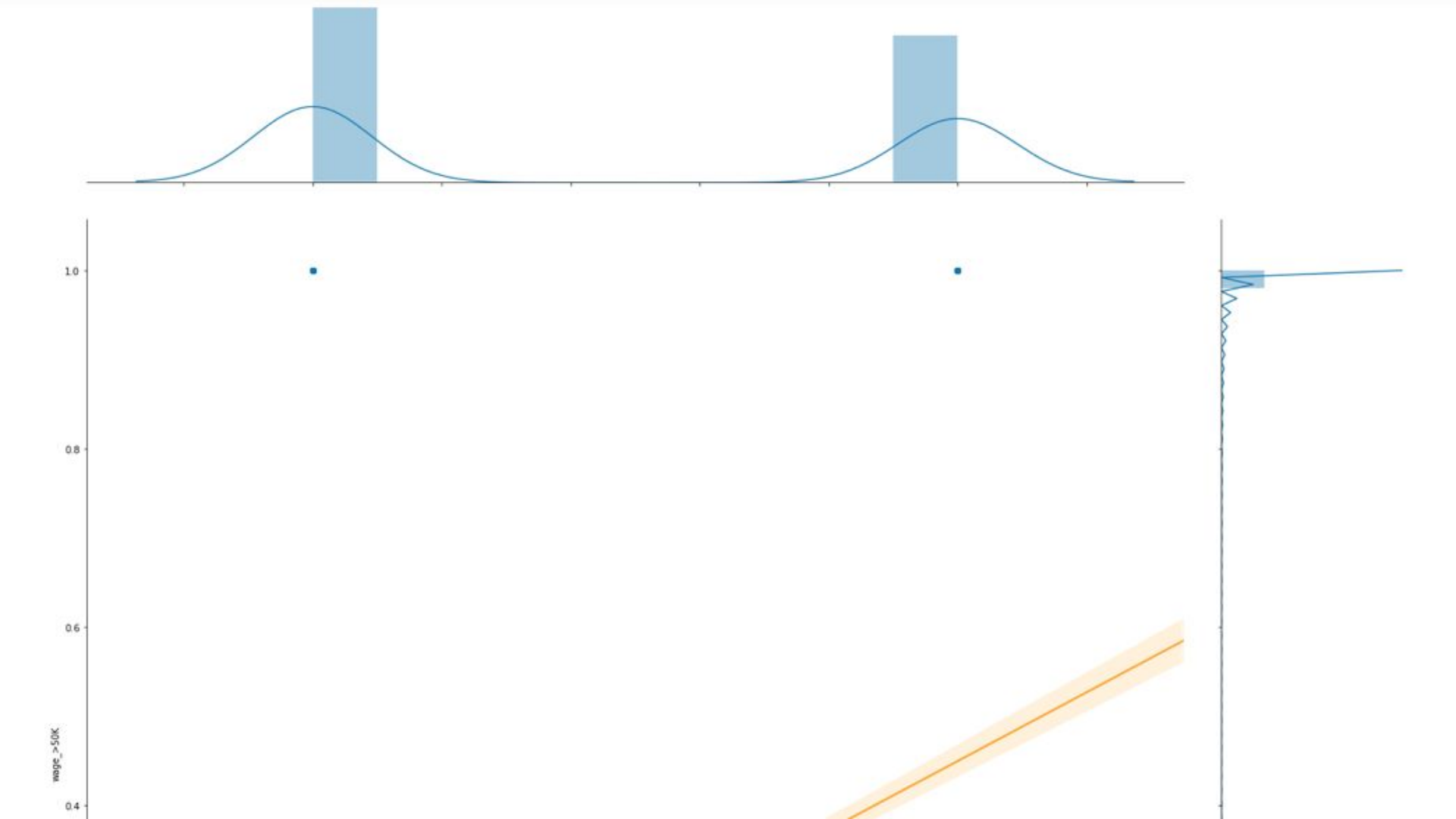
Supporting **Information: Random Forest**






Top 20 Features
Feature
marital-status_Separated
relationship_Other-relative
occupation_Handlers-cleaners
workclass_Private
occupation_Adm-clerical
occupation_Other-service
sex_Male
workclass_Self-emp-inc
relationship_Own-child
occupation_Prof-specialty
relationship_Unmarried
relationship_Not-in-family
capital-loss
marital-status_Never-married
occupation_Exec-managerial
hours-per-week
age
education-num
capital-gain
marital-status_Married-civ-spouse


0.848
0.847
0.846
0.845
0.844
0.843
0.842
0.841
Mean Accuracy
15
20
25
30
35
40
Min Samples Split


1.0
0.8
0.6
0.4
wage_>50K

Questions