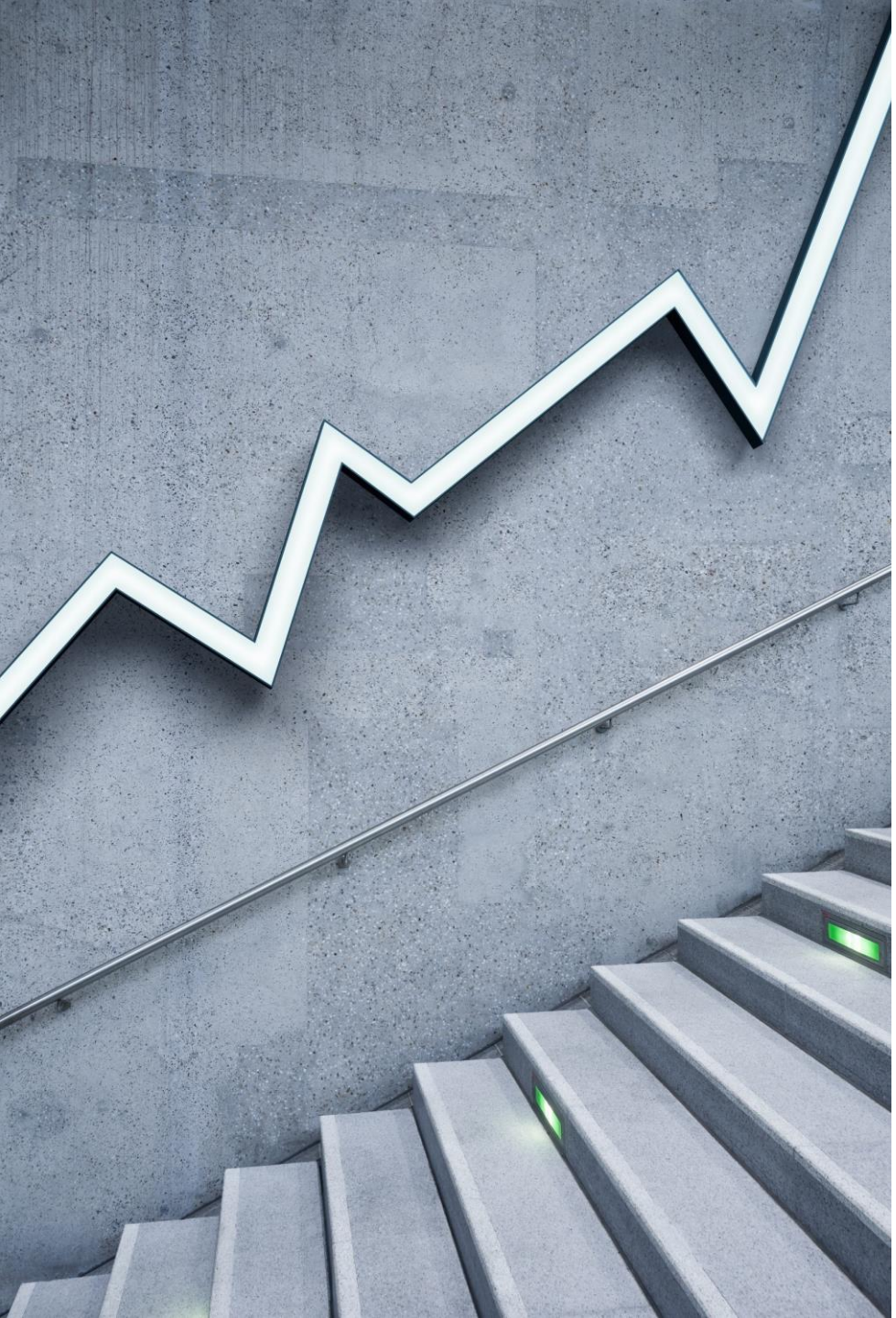


New Coder Model

“An investment in knowledge pays the best interest.”

- BENJAMIN FRANKLIN

New Coder Bootcamp Information

Who is most likely to complete a New Coder Bootcamp based on the data set retrieved from Github Free Code Camp?

Who participated?

More than 31,000 new coders responded to this survey. These are people who have been coding for less than 5 years.

Some of the Demographics included in the Survey

Here are some high level facts about them:

67% of them live outside the United States

Their median age is about 30 years old

21% are women

One quarter identifies as being part of an ethnic minority in their country

8% are military veterans

They've been coding for an average of 21 months

17% have already landed their first developer job

Learner goals and approaches

40% of them would like to either freelance or start their own business.

Learning Resources Used by the Courses

Some of the resources that were used were:

freeCodeCamp

Stack Overflow

Codeacademy

Udemy Mozilla Developer Network (MDN)

Khan Academy

Coursera

Less than half attended In-Person-Coding –Related Events

New Coder Bootcamp Information

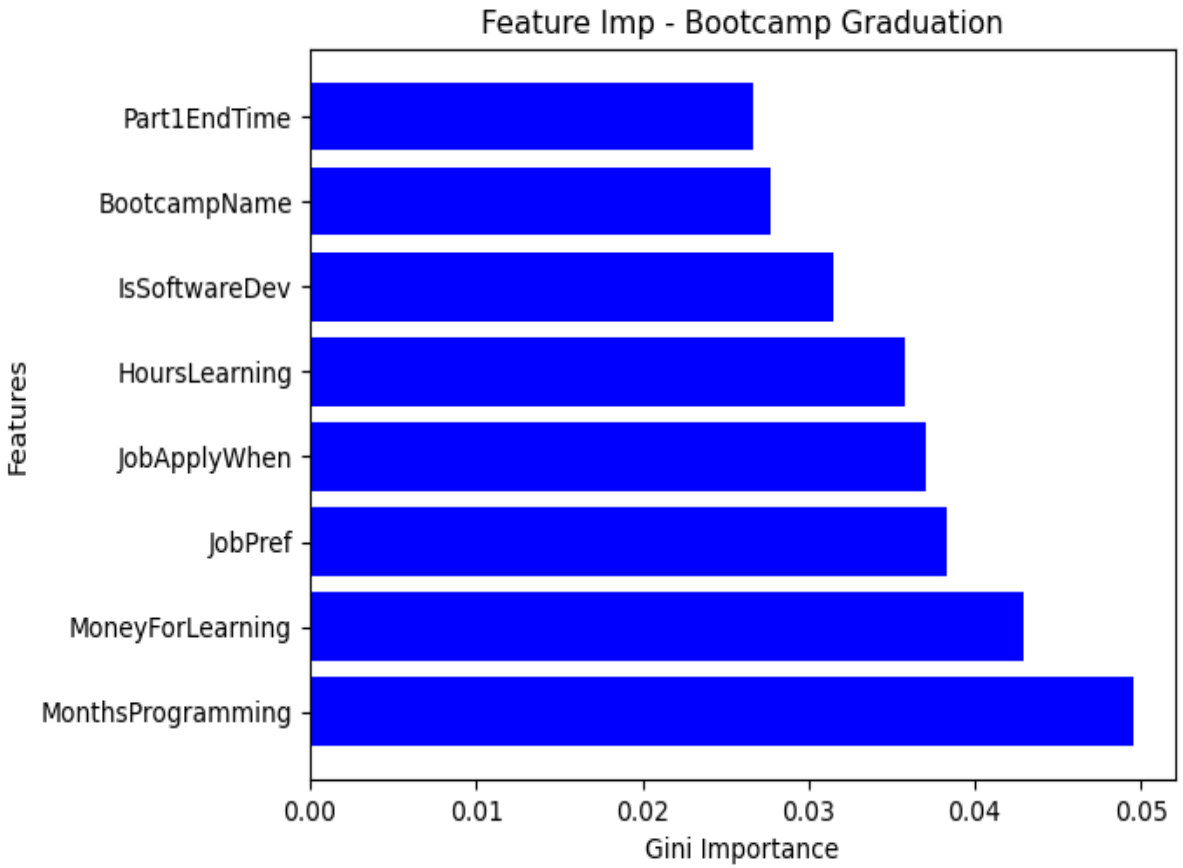
- Roberta Chandler used the random Forest model.
- Phil Klimkewicz Tried the XGBoost &Light GBM Model.
- Ayana tried the Logistics Regression & Decision Trees Model.
- Dominique Davis tried the SVM and KNN Models.

New Coder Bootcamp Information

Our Models used all the columns.

The reason all the columns were needed was there were so may variables.

The top eight most important features were months programming, money for learning, job preference, How soon after the course they were planning to apply for jobs, or if they already had applied for jobs, Hours learning, are already software developers, something related to the bootcamp names, what time of day the courses started, when we think based on the data points may be the length of the course, time the course takes and the


Feature Imp - Bootcamp Graduation
Part1EndTime
BootcampName
IsSoftwareDev
HoursLearning
JobApplyWhen
JobPref
MoneyForLearning
MonthsProgramming
Features
0.00
0.01
0.02
0.03
0.04
0.05
Gini Importance

Training Scores for Models Used

XGBoostClassifier produced the highest with a score of 81.6%

While Random Forest model produced a score of 81.3 percent.

So the the XGBoost Model produced the best result at a depth of one. Though it was higher there was no statistically significant difference between the two models that produced data.