

[Return to Classroom](#)[DISCUSS ON STUDENT HUB](#)

Creditworthiness

REVIEW

HISTORY

Meets Specifications

Dear Excellent Student,

Thank you for your submission. I enjoyed reviewing your work because it was great. This submission meets all of our expectations. The Classification Model is not an easy course to grasp, however, the task was really understood. A lot of work has been done and you should be proud of yourself. Continue practicing on these projects and other projects of yours and you will become the best in your domain.

Keep up the good work and good luck in future projects!!

Business and Data Understanding



The section is written clearly and is concise. The section is written in less than 250 words.

Excellent work done!!

The section is written clearly and is concise, the word limit is very well respected.



All following questions have been answered:

1. What decisions need to be made?
2. What data is needed to inform those decisions?
3. What kind of model (Continuous, Binary, Non-Binary, Time-Series) do we need to use to help make these decisions?

Building the Training Set



The section is written clearly and is concise. The section is written in less than 100 words.

Awesome!!

The section is written clearly and is concise, the word limit is very well respected.



The following question has been answered:

1. In your cleanup process, which field(s) did you impute or remove?

Please justify why you imputed or removed these fields. Visualizations are encouraged.

The correct fields are removed or imputed.

Excellent work Done!! 🎉🎉

You have removed all the variables which need to be removed and imputed are correctly identified with a reasonable justification and the correct visualization.

I like the way, you have justified the reason for each of the removed variable, followed by the visualization.

Tips: If you would like to better understand - in a very intuitive way - why the median is a good value to impute the missing values in Age field, check this site out: [measures-central-tendency](#)

Train your Classification Models



The section is written clearly and is concise. The section is written in less than 500 words.



All questions have been answered for each of the four models built: Logistic, Decision Tree, Forest Model, Boosted Model

1. Which predictor variables are significant or the most important? Please show the p-values or variable

importance charts for all of your predictor variables.

2. Validate your model against the Validation set. What was the overall percent accuracy? Show the confusion matrix. Are there any bias seen in the model's predictions?

There should be 4 sets of questions answered.

Writeup



The section is written clearly and is concise. The section is written in less than 250 words.



All questions have been answered:

1. Which model did you choose to use? Please justify your decision using all of the following techniques.

Please only use these techniques to justify your decision:

- Overall Accuracy against your Validation set
- Accuracies within "Creditworthy" and "Non-Creditworthy" segments
- ROC graph
- Bias in the Confusion Matrices

Note: Your manager only cares about how accurate you can identify people who qualify and do not qualify for loans for this problem.

1. How many individuals are creditworthy?

 [DOWNLOAD PROJECT](#)

[RETURN TO PATH](#)

Rate this review

[START](#)