

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

# Creditworthiness

## REVIEW

## HISTORY

### Meets Specifications

Congratulations on completing the project. All your answers are correct and well explained. You have met all the rubric requirements to pass the project. Good Work!!!

Keep up the good work and all the best with future projects and submissions. Please check the annotation for details.

Happy Learning!!!

### Business and Data Understanding

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



**Awesome:** Indeed the section is written clearly and is concise and the word limit has been respected as per the rubric requirement. Amazing Work!!!

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?



**Awesome:** Indeed you have answered all the 3 questions correctly. Amazing Work!!!

## Building the Training Set

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



**Awesome:** Indeed the section is written clearly and is concise and the word limit has been respected as per the rubric requirement. Amazing Work!!!

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.



**Awesome:** You have removed and imputed all the required columns from the data with the correct reasoning as per the rubric requirement. Amazing Work!!!

## Train your Classification Models

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



**Awesome:** Indeed the section is written clearly and is concise and the word limit has been respected as per the rubric requirement. Amazing Work!!!

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.



**Awesome:** You have correctly trained all the models and identified the significant variables as per the rubric requirement. Also, the overall accuracy and the confusion matrix calculated by you are absolutely correct and within the expected range. Amazing Work!!!

## Writeup

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



**Awesome:** Indeed the section is written clearly and is concise and the word limit has been respected as per the rubric requirement. Amazing Work!!!

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?



**Awesome:** You have correctly identified that the Forest Model should be used here and correctly justified doing the same using the Overall accuracy, confusion matrix, ROC curve, and by discussing the bias in the model as per the rubric requirement. Also, the number of creditworthy applicants calculated by you is within the expected range of 405-420. Amazing Work!!!

 [DOWNLOAD PROJECT](#)

[RETURN TO PATH](#)