U D A C I T Y

Creditworthiness

Business and Data Understanding

CRITERIA	MEETS SPECIFICATIONS
Section Format	The section is written clearly and is concise. The section is written in less than 250 words.
Key Decisions	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

CRITERIA	MEETS SPECIFICATIONS
Section Format	The section is written clearly and is concise. The section is written in less than 100 words.
Justification	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.

Train your Classification Models

CRITERIA	MEETS SPECIFICATIONS
Section Format	The section is written clearly and is concise. The section is written in less than 500 words.

CRITERIA	MEETS SPECIFICATIONS
Complete	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

CRITERIA	MEETS SPECIFICATIONS
Section Format	The section is written clearly and is concise. The section is written in less than 250 words.

Complete	All questions have been answered:
	
	1. Which model did you choose to use? Please justify your decision using **only** of the following techniques:
	
	   ◦ Overall Accuracy against your Validation set
	   ◦ Accuracies within "Creditworthy" and "Non-Creditworthy" segments
	   ◦ ROC graph
	
	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?