**ADVANCED DATA ANALYTICS FOR MANAGERS**

**Total Marks: 50 Time duration: 48 Hours**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**QUESTIONS**

## Build and interpret a logistic regression model to identify good and bad customers – a good customer is one who has never defaulted, any customer with a single default is a bad customer (15 Marks)

*Note:*

*Please see the dataset associated with this case. Run a logistic regression model (with two category) using the defaulter variable (DefaulterFlag) as the dependent variable and the other mentioned variables (see the data) as independent variables. However, considering the importance of the some of the variables, the company recommended to consider the following variables in the logistic model. See Exhibit 1 (See the case) for more details about the variables.*

* *No. of dependents*
* *Monthly income in thousands*
* *Salary date fraction*
* *Tenor in years*
* *Down-payment fraction*
* *Professional or business flag*
* *Highest qualification is high school*
* *Highest qualification is post-graduation*
* *Sex*
* *Has provided post-dated check*
* *Has refrigerator*
* *Has washing machine*

*Apply two-fold cross validation for model validation. Apply suitable data pre-processing (if required).*

1. Build alternative logistic regression models & check the predictive accuracy of the models and construct the ROC and estimate the AUC (10 Marks)
2. Apply alternative machine learning algorithms (e.g., Decision Tree & Random Forest) and examine whether the use of these algorithms is suitable to add better predictive power over the logistic regression constructed to predict the creditworthiness of the customers (20 Marks [10+10]).

*Note: Using the given data, build Decision Tree & Random Forest. Analyze the accuracy of the models and build the ROC and estimate the AUC of these models. Compare the ROC and AUC across these ML algorithms.*

1. Based on the models constructed, suggest the necessary recommendations from your side as an analyst, so that the company can better assess the creditworthiness of its future/potential customers (5 Marks)