ORION

PRACTICE PHASE LAUNCH

A Machine Learning Modeling Contest

April 2019

By AI Guild Team

Congratulations for registration to Orion…

Welcome aboard !

CONTENTS

1. CASE

Business Problem:

Mr. Bryan who is a CRO (Chief Risk Officer) of **ABC corporation**, a top US based NBFC (Non Banking Financial institution) is concerned about increasing losses in its newly acquired accounts for unsecured loan portfolio. He wants to identify characteristics and factors that are indicative of customers who are likely to charge-off/default on loans. He is concerned that the current model and strategy in place is inadequate to filter good and bad accounts.

Thus, He asks his modeling head to develop a credit scoring model that examine the credit worthiness of its newly acquired customers.

Goal:

As a modeling head of **ABC corporation**, your goal is to develop a robust modeling solution that can help Mr. Bryan predict the likelihood that the loan would be paid-off in full or it will be charged off/default.

2. DATA OVERVIEW

Data Accessibility:

Data can be accessed through personal laptops and is stored in the following location

http://bit.do/Orion_ML

Dataset Details:

- **Unique Identifier** - ID
- Total 3 samples will be used throughout this phase
  - ***Train*** – Dataset available for model development (***Release date*** – *8th April , 2019*)
  - ***Test1*** – First test sample to be used for scoring and evaluation (***Release date*** – *22nd April , 2019*)
  - ***Test2*** – Final test sample to used for scoring and evaluation (***Release date*** – *24th May, 2019*)
- Train dataset details:
  - ***# Observations*** – 152,514
  - ***# Events*** – 27,871
  - ***# Fields*** – 49
- Data dictionary is also available in the data download link (**http://bit.do/Orion_ML**)
- Participants are encouraged to merge any third party data for modeling/analytics
- Template submission form would be shared along with the test data on 22nd April, 2019

3. EVALUATION CRITERIA

There are 3 major criteria for evaluation:

Accuracy

- ✓ **Root Mean Square Error (RMSE)** will be used as an accuracy evaluation criteria
- ✓ **Least RMSE** model across test samples will be ranked higher

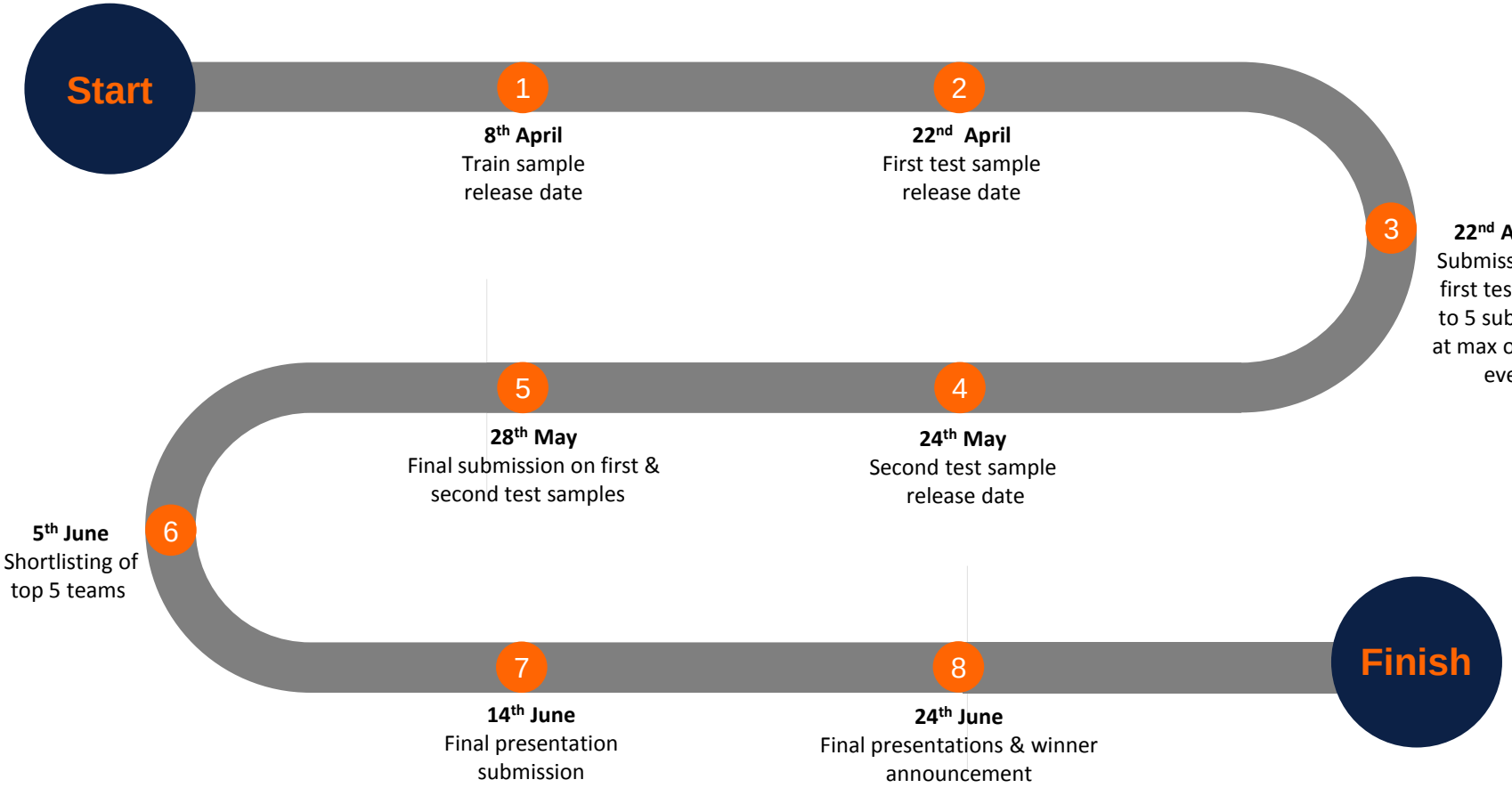
Innovation

- ✓ Modeling approach and methodology
- ✓ Feature Engineering

Final Model Interpretability & Quality of final recommendation

- ✓ Final model variables business justification and relation with target
- ✓ Final model results and insights
- ✓ Final Recommendation

4. TIMELINES


Start
1
8th April
Train sample release date
2
22nd April
First test sample release date
3
22nd Apr-20th May
Submission period for first test sample - (up to 5 submissions with at max one submission every week)
4
24th May
Second test sample release date
5
28th May
Final submission on first & second test samples
6
5th June
Shortlisting of top 5 teams
7
14th June
Final presentation submission
8
24th June
Final presentations & winner announcement
Finish

5. RULES

DO'S

- ✓ The Score file should be named as "TeamName_DDMM_v1 (submission No.)". For example file name for team ABC 2nd submission on 20th April will be "ABC_2004_v2".
- ✓ All Score file shared should be in csv format only
- ✓ The columns in score file should be id and Score_TeamName. For example score file should have 2 columns "id" and "Score_ABC"
- ✓ The Email subject for sharing score file should be "Orion Submission Team Name Week 1". Example "Orion Submission ABC Week 1"

DON'T'S

- ☒ No other file formats other than csv would be accepted
- ☒ Only 1 submission for score would be accepted per week. If multiple submission are made, first one will be considered for evaluation
- ☒ Entry without Team Name will not be considered
- ☒ For Score file, only keep Id and Score columns in the file. Please do not include all attributes
- ☒ Do not delete any observations from the test samples while providing scores

6. RESOURCES

Software Support:

R & Python installation guide and links are available in the below mentioned link. This should be accessed from personal laptops

http://bit.do/Orion_ML

Training Refresher

- Create an account on https://cognitiveclass.ai/. Sign Up using ***LinkedIn or Email*** so that the lectures are accessible on office computers as well.
- There are 4 sets of modules:

Modeling Basics – Pick any 1

Introduction to Data Science
Cognitive Class / Fireside Analytics Inc. **DS0101EN**

Predictive Modeling Fundamentals I
Cognitive Class **PA0101EN**

Machine Learning – Pick any 1

Machine Learning with R
Cognitive Class **ML0151EN**

Machine Learning with Python
Cognitive Class **ML0101ENv3**

REWARDS

** Maximum of 2 awards per team*

CONTACT US

AI.Guild@exlservice.com

Happy Learning !

EXL (NASDAQ: EXLS) is a leading operations management and analytics company that designs and enables agile, customer-centric operating models to help clients improve their revenue growth and profitability. Our delivery model provides market-leading business outcomes using EXL's proprietary Business EXLerator Framework®, cutting-edge analytics, digital transformation and domain expertise. At EXL, we look deeper to help companies improve global operations, enhance data-driven insights, increase customer satisfaction, and manage risk and compliance. EXL serves the insurance, healthcare, banking and financial services, utilities, travel, transportation and logistics industries. Headquartered in New York, New York, EXL has more than 27,000 professionals in locations throughout the United States, Europe, Asia (primarily India and Philippines), South America, Australia and South Africa.

EXLservice.com

GLOBAL HEADQUARTERS

280 Park Avenue, 38th Floor
New York, New York 10017
T +1 212.277.7100 **F** +1 212.771.7111

United States • United Kingdom • Czech Republic • Romania • Bulgaria • India • Philippines • Colombia • South Africa