

Online Hackathon Machine Learning

5th May - 21st May

24 day and 23:45:08

[Dashboard \(https://www.techgig.com/codegladiators/dashboard\)](https://www.techgig.com/codegladiators/dashboard)

[Problem](#)

[Submission](#)

[My Submissions](#)

[Need Help?](#)

Risk for Redeeming

OVERVIEW

Invesco Ltd. is an independent investment management firm with more than 6,500 employees dedicated to delivering an investment experience that helps people get more out of life. We understand that every client has a unique set of investment objectives that can be achieved in a variety of ways, and our comprehensive range of investment capabilities has been constructed over many years to help clients achieve these objectives. Our high conviction fundamental and factor based methodologies aim to deliver client outcomes that go beyond the limitations of traditional passive investing and benchmark centric active management, and we believe they can help investors build portfolios in more precise and impactful ways. We are privileged to manage more than \$830 billion in equity, fixed income, money market, balanced and alternative assets on behalf of retail and institutional clients around the world. (As of December 31, 2016.)

Invesco sells retail mutual funds across the United States via financial intermediaries. In order to maximize effectiveness of our sales staff, we would like to understand who is at risk for selling (redeeming) our mutual funds in the next month and which funds are at risk, so that we may take a proactive engagement approach targeted specifically at the risk areas.








Invesco has a wealth of internal data describing purchase and sale behaviour of our products with our clients. In addition to sales transactions, we collect asset holding balances, product investment experience information, and financial intermediary activity information. Our clients (financial intermediaries) are already segmented by behavioural criteria.

You are a **Senior Data Scientist** within our strategy and innovation team, and are assigned the task to analyse the submitted data. **Please write an algorithm using the prediction and test data sets, and provide a data output file that should generate a sample_output.csv file.**

Objective: **Predict which investment id of different investment advisors will be redeemed in 2017- 01?**

The details to support this effort are in 4 data domains: transactions, assets, product investment experience, and activity engagement. Below are the file names and details.

DATA FILES (7 files)- Download All Files (https://www.techgig.com/files/contest_upload_files/invesco/CodeGladiators-Invesco-DataSet.zip)

File Name	Description	Format	Size
CodeGladiator-DataDictionary.xlsx	4 tabs that give the fields and descriptions of all data provided in remaining files.	.xlsx	0.03 (mb)  (https://www.techgig.com/files/contest_upload_files/invesco/CodeGladiator-DataDictionary.xlsx)
sample_output.csv	Example of the desired outputs	.csv	0.002 (mb)  (https://www.techgig.com/files/contest_upload_files/invesco/sample_output.csv)
Code-Gladiators-InvestmentExperience.csv	the prediction data set for investment experience	.csv	2 (mb)  (https://www.techgig.com/files/contest_upload_files/invesco/Code-Gladiators-InvestmentExperience.csv)
Code-Gladiators-Transaction.csv	the prediction data set for transactions	.zip	1 (mb)  (https://www.techgig.com/files/contest_upload_files/invesco/Code-Gladiators-Transaction.zip)
Code-Gladiators-AUM.csv	the prediction data set for assets under management	.zip	7 (mb)  (https://www.techgig.com/files/contest_upload_files/invesco/Code-Gladiators-AUM.zip)
Code-Gladiators-Activity.csv	The prediction data set for engagement activity	.csv	1 (mb)  (https://www.techgig.com/files/contest_upload_files/invesco/Code-Gladiators-Activity.csv)
test_data.csv	The test data set for which Propensity_Score and Redeem_Status need to be predicted.	.csv	0.11 (mb)  (https://www.techgig.com/files/contest_upload_files/invesco/test_data.csv)

FILE FORMATS

Please see Code Gladiator File Spec.xlsx for file layouts and contents. Note multiple tabs.

EVALUATION CRITERIA

- Implementation



Submissions would be evaluated based on accuracy. Score is calculated with the following formula:
(<https://www.techgig.com/codegladiators>) Presented by

Deloitte

(<https://www.techgig.com/codegladiators/deloitte>)

$$\text{score} = \text{Number of correct predictions} / \text{Total records} * 100$$

- Presentation
 - Participants are expected to share a short presentation explaining their model.

GUIDELINES

1. You must upload your output on the Submission page in the format given in problem statement. In addition to your output, you will also have to submit your source and other files in .zip or .tar compressed archive.
2. You can use any tools or libraries for building your solution. There is no restriction on the tools that you can use.
3. You have unlimited number of attempts to make submission. But only your last submission will be considered for evaluation.
4. Privately sharing code or data outside of your team is not permitted. It's okay to share code if made available to all participants on the forums.

Create Team (<https://www.techgig.com/codegladiators/leader/9105>)

Prizes and Recognition

Win Prizes worth Rs 3 Crores

Participate and score to win Rs.150 Dineout voucher and one month subscription of gaana.com. Hurry Up!

Champion **Rs 2,00,000**

1st Runner-up **Rs 1,00,000**

Top 200 Performers get Rs. 700 Dineout Vouchers & Three months Subscription of gaana.com

You Must Participate Because

It is India's biggest coding Contest

You can Earn title of best coder of India

Get recognised by IT companies