

HACKATHON PROBLEM STATEMENT

Invest Smart using AI&ML

A mutual fund is a professionally managed investment fund that pools money from many investors to purchase securities. These investors may be retail or institutional in nature. Mutual fund ratings is one of the most influential and prominent decision making factors that is used by investors in making a decision with regards to selecting a mutual fund.



Great Stone Rating is a star based ranking system.

These ratings are based on the performance of a mutual fund with adjustments for risks and costs as compared to other funds in the same category. The rating ranges from 0 to 5.

Goal: The goal of this hackathon is to predict GreatStone's rating of a mutual fund. In order to help investors decide on which mutual fund to pick for an investment, the task is to build a model that can predict the rating of a mutual fund. The various attributes that define a mutual fund can be used for building the model.



Dataset Information

This dataset comprises information of 25000 mutual funds in the United states. Various attributes related to the mutual fund have been described and these attributes will be used for making decisions on the rating of the mutual fund by GreatStone which is a top mutual fund rating agency.

The following files are provided in the form of CSVs. These files contain various attributes related to the mutual fund. Please find the following files for the same:

bond_ratings, fund_allocations, fund_config, fund_ratios, fund_specs, other_specs, return_3year, return_5year, return_10year.

Files Description:

bond_ratings consists of 12 columns which provide information on the bond rating percentage allocation of the mutual funds

fund_allocations consists of 12 columns which provide information on the sector wise percentage allocation of the mutual funds

fund_config comprises of 4 columns which comprise the metadata of the mutual funds

fund_ratios consists of 8 columns which provides information on various fundamental ratios that describe the mutual funds

fund_specs contains 9 columns which give information about the specifications of the mutual funds

other_specs contains 43 columns which give information of the other aspects of the mutual funds

return_3years contains 17 columns which give information about 3 year return and ratios

return_5years contains 17 columns which give information about 5 year return and ratios

return_10years contains 17 columns which give information about 10 year return and ratios

sample_submission contains the fund ids for which you need to provide the ratings for the submission file. Please maintain the order of the fund ids as shown in this file.

The tag column is a unique identifier and is also the same as the id.(i.e tag = id)

*Please note that the details regarding what each column signifies has not been provided and it a part of the hackathon to read and figure out what each column represents.

Train and Test Data

The train & test data are both provided in the CSVs described above as part of the same file. You need to segregate the training & test data based on where the greatstone ratings are provided. Go through the files carefully to understand how you can segregate both the datas. Please maintain the ordering of the test data. You can use the sample submission file in order to get ID of the test data.(check sample_submission.csv)



Evaluation Metric



Mean Precision Value - Mean of precision of all the classes = $P1+P2+...+P6/6$ Here P1 is Precision of Class 1 and P2 is Precision of Class2 and so on and so forth.

Submission Format

You need to submit the csv format file with 2 columns in the order - **fund_id** and **greatstone_rating** using the test dataset.

Please make sure the format of your file is exactly the same as mentioned above (order of ids is maintained, letter format etc).

