

Description

At ETS Factory we design quantitative algorithms to develop scientific solutions to the Asset and Risk Management Industry. In order to improve the quality of our solutions, we have developed a Synthetic Scenarios Generator for financial time series.

Given a set of series, are you able to distinguish the real ones from the synthetic ones?

Data

The data has been split into two groups:

- Training set (*TrainMyriad.csv*)
- Test set (*TestMyriad.csv*)

The training set should be used to build your machine learning models. For the training set, we provide the outcome (also known as the “ground truth” in the Class column) for each time series (rows): Class 1 for real series and Class 0 for the synthetic ones.

You are free to create as many new features as you want.

The test set should be used to see how well your model performs on unseen data. For the test set, we do not provide the ground truth for each row. It is your job to predict these outcomes. For each time series in the test set, use the model you trained to predict whether the time series is real or synthetic.

We also include *SubmissionMyriad.csv* as an example of what a submission file should look like.

Evaluation

Goal

It is your job to predict if time series is real or synthetic. For each row in the test set, you must predict the Class variable.

Metric

The performance is measured with the [ROC AUC](#). We need you to send the probability (between 0 and 1) of each series belonging to the positive class (1, real).

Submission File Format

You should submit a csv file with exactly 5000 entries plus a header row (Class). Your submission will show an error if you have extra columns or rows. The file should have exactly one column:

Class (contains your predictions)

myriad@etsfactory.com