Tuning Fork Quick Start Guide

“Your Cloud data Science Lab”

Tuning Fork is your data science lab. It enables Enterprise data science departments to run experiments on the cloud. Using the REST API, data scientists can run standard tests and procedures in Tuning Fork enabling them to fine-tune models for performance.

Enterprise customers can define custom tests and procedures which can be reused in multiple projects.

Benefits

- **Is Machine Learning Feasible with your data?:** Caltech Prof. Yaser Abu Mostafa's 3-criteria test: Is there a pattern, to learn from? Are you unable to pin down the pattern mathematically? Does data that represents the pattern exist? Traditionally this process takes weeks. Tuning Fork shrinks this process to hours.
- **Design Experiments quickly:** “I have the advantage of having found out, how hard it is to really get to know something, how careful you have to be about checking the experiments, how easy it is to make mistakes and fool yourself." - Richard Feynman. Design and run lots of experiments on your data leaving out the guesswork.
- **Fine-tune models for performance:** While modeling machine learning solutions is only half of the battle. The other half is fine-tuning the model for performance which is an arduous and often expensive

process. The tools provided by Tuning Fork enables you to fine-tune models for performance effortlessly.

- **Rapid Development Development:** Rapid Model development is not about providing automated tools to create models, but tools to understand and transform data. The tests and procedures of Tuning Fork enable the data scientist to develop and deploy machine learning model rapidly in days, instead of months.

Features

- **Tests to learn about data:** 60% of data science is about collecting and organizing data. Get visualizations about your data such as normality, 7-number summary, outliers instantly without writing code.
- **Procedures to transform data:** Transform data instantly using procedures such as normalizing, standardizing, splitting data without writing code. Enterprise Customers can write their own custom procedures.
- **Weakly label data:** "If big data is the new oil, Labeled data is the 'new new' oil". Write functions to label lots of data automatically utilizing the scarce domain experts for manually labeling a small dataset.
- **Augment Data:** Machine learning and especially deep learning needs lots and lots of data. Tuning Fork can automatically generate synthetic data enabling you to do machine learning with a much smaller dataset.

Tuning Fork REST API

The dead-simple REST API lets data scientists and engineers, interact with Tuning Fork, to create tests and procedures, query the results of the tests and experiments using POST and GET HTTP requests.

This Quick Start Guide is intended to be an introduction to run tests and procedures rapidly, in a matter of hours, rather than months.

Ingest Content

The first thing, you need to do is to do is ingest content via HTTPS POST. You could use curl for accomplishing the ingest.

curl -F 'unlabeled=@<location of the unlabeled data file>' POST https://api.tuningfork.ai/ingest

Here is an example:

curl -F 'unlabeled=@/home/alice/data.csv' POST https://api.tuningfork.ai/ingest

There are two high functions one can accomplish, namely, tests and procedure.

The POST API responds with a list of URLs which one can do further, i.e., tests or procedures. These are the tests and procedures allowed to perform in that data.

For instance, if the ingest to an unlabeled data was successful, it will return the following list of standard tests, i) Normality, ii) Outliers, iii) 7-number summary statistics, iv) linearly separable tests and v) enumeration of all the categorical variables.

It will also list the following standard procedures, i) weakly label which will label the data given labeling functions as input and ii) normalize procedure which will normalize the data and iii) standardize procedure which will standardize the data.

The tests are simply GETs while the procedures are POSTs.

```
$ # Ingest the unlabeled CSV file and get a new Project ID
$ curl -F 'unlabeled=@/home/alice/data.csv' POST https://api.tuningfork.ai/ingest
https://api.tuningfork.ai/a17766c6/test/normality
https://api.tuningfork.ai/a17766c6/test/outliers
https://api.tuningfork.ai/a17766c6/test/stats
https://api.tuningfork.ai/a17766c6/test/separable
https://api.tuningfork.ai/a17766c6/test/enumerate/<text_column_name>
https://api.tuningfork.ai/a17766c6/procedure/weakly_label
https://api.tuningfork.ai/a17766c6/procedure/normalize
https://api.tuningfork.ai/a17766c6/procedure/standardize

$ # tests are HTTP GETs and procedures are HTTP POSTs
```

If the ingest was an unlabeled data, it will return the following list of standard procedures, i) augment which will create artificial data and ii) split which will split the dataset automatically into Training, Test and Dev. Tests.

```
$ # Ingest the labeled CSV file and get a new Project ID
$ curl -F 'labeled=@/home/alice/labeled_data.csv' POST https://api.tuningfork.ai/ingest
https://api.tuningfork.ai/b28877b5/procedure/augment
https://api.tuningfork.ai/b28877b5/procedure/split

$ # procedures are HTTP POSTs
```

If the ingest was a model, i.e., either a scikit-learn or keras, it will return the following standard procedure, learn, i.e., given data, it will generate a model with weights.

```
$ # Ingest the Model file and get a new Project ID
$ curl -F 'model=@/home/alice/model.json' POST https://api.tuningfork.ai/ingest
https://api.tuningfork.ai/e40044g3/procedure/learn

$ # tests are HTTP GETs and procedures are POSTs
```

If the ingest was a model with weights, i.e., either a scikit-learn or keras, it will return the following standard test, metrics, which will return a confusion matrix comprising of performance metrics and a standard procedure, predict, i.e., given data, it will predict the outcome.

```
$ # Ingest the Model with weights file and get a new Project ID
$ curl -F 'model=@/home/alice/model.h5' POST https://api.tuningfork.ai/ingest
https://api.tuningfork.ai/c39933f2/test/metrics
https://api.tuningfork.ai/c39933f2/procedure/predict

$ # tests are HTTP GETs and procedures are POSTs
```

Enterprise Support

If you need Tuning Fork for data over 10,000 rows and/or for custom tests and procedures, please contact us for Enterprise plans.

Contact

Website: https://tuningfork.ai

Mail: sales@tuningfork.ai

Twitter: https://twitter.com/TuningForkAI

Github: https://github.com/vasusrini/TuningForkAIDocs