The Tunisian Company of Electricity and Gas (STEG) is a public and a non-administrative company, it is responsible for delivering electricity and gas across Tunisia. The company suffered tremendous losses in the order of 200 million Tunisian Dinars due to fraudulent manipulations of meters by consumers.

Using the client’s billing history, the aim of the challenge is to detect and recognize clients involved in fraudulent activities.

The solution will enhance the company’s revenues and reduce the losses caused by such fraudulent activities.

**Evaluation**

The metric used for this challenge is [Area Under the Curve.](https://zindi.africa/learn/zindi-error-metric-series-how-to-use-area-under-curve-auc-as-an-evaluation-metric-for-machine-learning)

Then the submission file should be as follows:

client\_id target

test\_Client\_0 0.986

test\_Client\_1 0.011

test\_Client\_10 0.734

**About**

The data provided by STEG is composed of two files. The first one is comprised of client data and the second one contains billing history from 2005 to 2019.

There are 2 .zip files for download, train.zip, and test.zip and a SampleSubmission.csv. In each .zip file you will find a client and invoice file.

**Variable definitions**

**Client:**

* Client\_id: Unique id for client
* District: District where the client is
* Client\_catg: Category client belongs to
* Region: Area where the client is
* Creation\_date: Date client joined
* Target: fraud:1 , not fraud: 0

**Invoice data**

* Client\_id: Unique id for the client
* Invoice\_date: Date of the invoice
* Tarif\_type: Type of tax
* Counter\_number:
* Counter\_statue: takes up to 5 values such as working fine, not working, on hold statue, ect
* Counter\_code:
* Reading\_remarque: notes that the STEG agent takes during his visit to the client (e.g: If the counter shows something wrong, the agent gives a bad score)
* Counter\_coefficient: An additional coefficient to be added when standard consumption is exceeded
* Consommation\_level\_1: Consumption\_level\_1
* Consommation\_level\_2: Consumption\_level\_2
* Consommation\_level\_3: Consumption\_level\_3
* Consommation\_level\_4: Consumption\_level\_4
* Old\_index: Old index
* New\_index: New index
* Months\_number: Month number
* Counter\_type: Type of counter