Predict the product type

You are working as a data scientist in a B2B company. The company provides its services to various clients. The management of the company has decided to analyze the information about its clients, to provide them with the best offerings.

Task

As a part of the data and analytics team, your task is to use the information related to each company from its info page (HTML file) and identify the product type of that company.

Dataset description

The dataset contains the following files:

- train(folder): Contains 13452 HTML files
- test(folder): Contains 2375 HTML files
- *train.csv*: 13452 x 2
- *test.csv*: 2375 x 1
- *sample_submission.csv*: 5 x 2

The columns provided in the dataset are as follows:

Column	Description
File_Name	Represents the name of the file.
Product	Represents the type of the Product.

Evaluation metric

```
score = 100*f1_score(actual, predicted, average = 'weighted')
```

Result submission guidelines

- The index is "File_Name" and the target is the "Product" column.
- The submission file must be submitted in *.csv* format only.
- The size of this submission file must be 2375 x 2.

Note: Ensure that your submission file contains the following:

- Correct index values as per the *test.csv* file
- Correct names of columns as provided in the *sample_submission.csv* file

Instructions:

- Click *Download dataset* to download the dataset.
- Solve the problem in your local environment.
- Save the submission in a *.csv* file.
- Click *Upload File* (under the *Upload File* section) to upload your prediction file (*.csv*).
- Add any instructions or comments in the *Your Answer* section.
- Click *Submit*.

Download dataset