

Task Overview

In this competition you will help to create an automation pipeline for consultants.

The task is to classify pairs – (web pages, criteria) to 3 classes: rejected_by_product / rejected_by_function / accepted.

We consider company's services / actions / activities as Function. For example, "manufacturing of cars" -> manufacturing – function, cars – product.

Firstly consultant checks whether the company can be rejected by product, if company is acceptable by product, consultant checks whether the company can be rejected by function.

Only after that, if company is acceptable by product and function, consultant marks it as accepted.



Data

id – unique id of sample.

html – list of html web pages related to this particular company.

text – list of extracted texts from the corresponded html.

keywords – list of relevant words from text (model based).

accepted_function – criteria for acceptance by function.

rejected_function – criteria for rejection by function.

accepted_product – criteria for acceptance by product.

rejected_product – criteria for rejection by product.

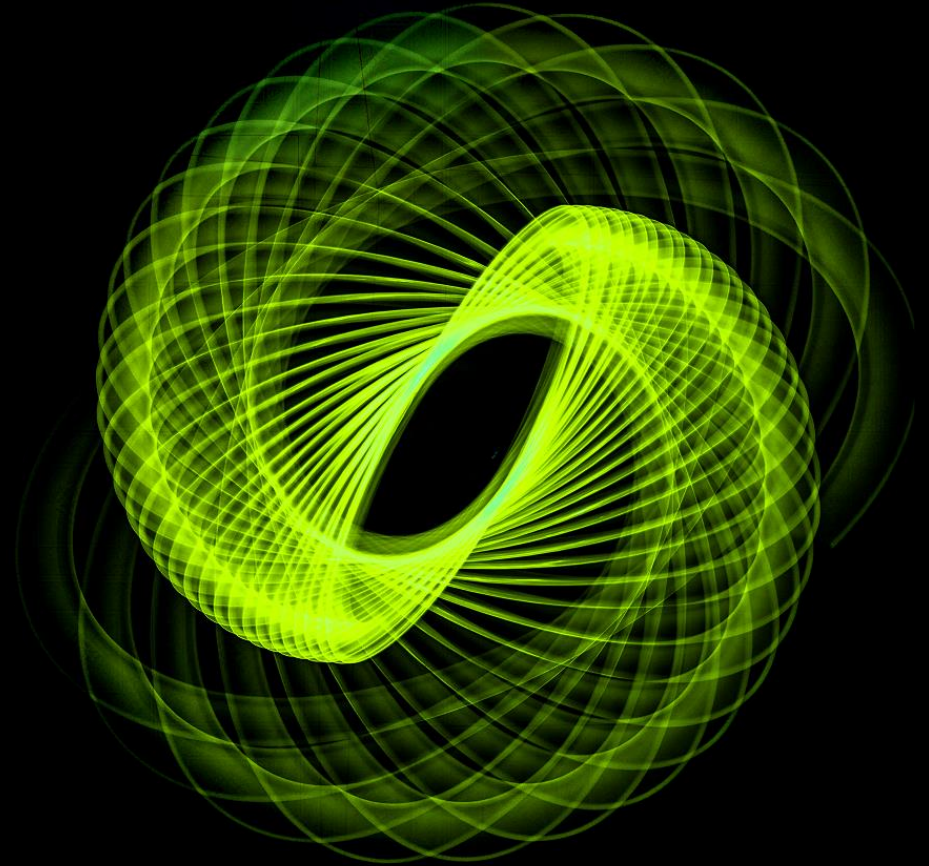
target – 0:rejected_by_product,
1:rejected_by_function, 2:accepted



Evaluation

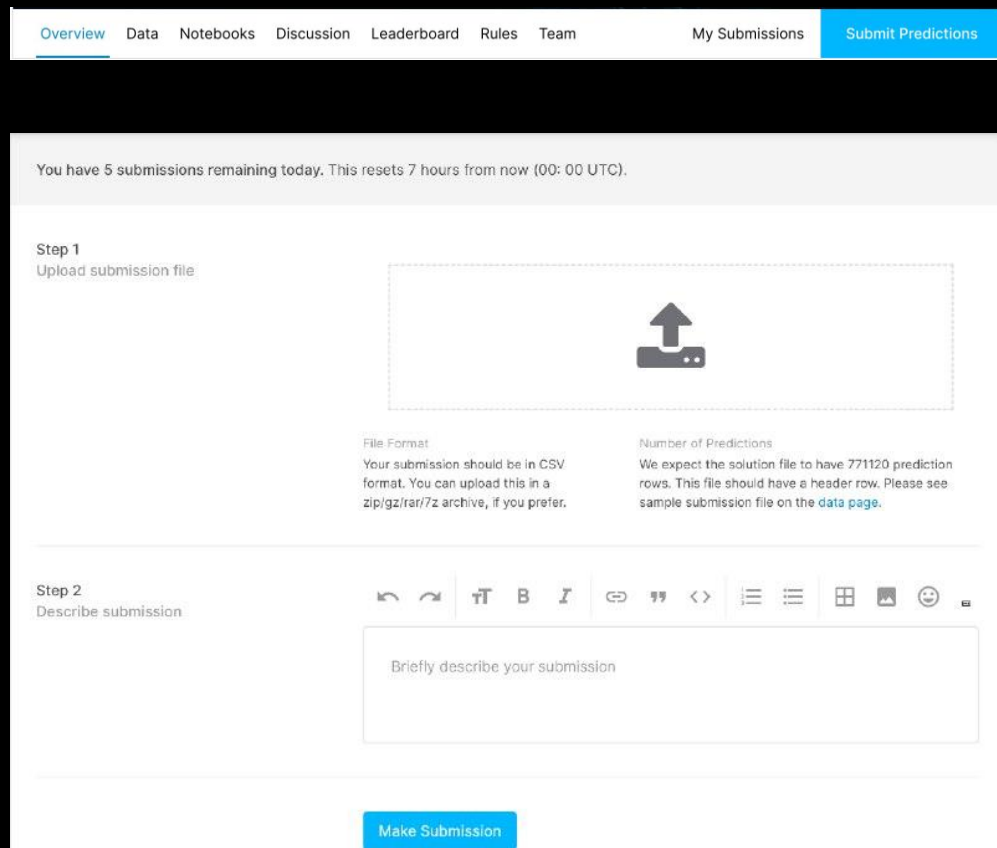
Metric – weighted accuracy.

Weight of each sample depends on its target:
{rejected_by_product:1,
rejected_by_function:1, accepted:2}



How to submit

Instruction




The screenshot shows a submission interface with a top navigation bar and a main content area. The navigation bar includes links for Overview, Data, Notebooks, Discussion, Leaderboard, Rules, Team, My Submissions, and Submit Predictions. The main content area has a status bar at the top indicating 5 submissions remaining. It is divided into two steps: Step 1 (Upload submission file) and Step 2 (Describe submission). Step 1 includes a file upload area, file format instructions (CSV or zip/gz/rar/7z), and prediction row requirements (771120 rows with a header). Step 2 includes a rich text editor for describing the submission and a 'Make Submission' button at the bottom.

Overview Data Notebooks Discussion Leaderboard Rules Team My Submissions Submit Predictions

You have 5 submissions remaining today. This resets 7 hours from now (00: 00 UTC).


Step 1
Upload submission file



File Format
Your submission should be in CSV format. You can upload this in a zip/gz/rar/7z archive, if you prefer.

Number of Predictions
We expect the solution file to have 771120 prediction rows. This file should have a header row. Please see [sample submission file on the data page](#).

Step 2
Describe submission



Briefly describe your submission

Make Submission

#1 Step

Go to the "Submit Predictions" tab

#2 Step

Select and download a csv structure file similar to `sample_submission.csv` from the organizers

#3 Step

Optionally write a comment of the file (model, validation, etc.)

#4 Step

Click on «Make Submissions»