

Submission Instructions

Prepared by the Leaders Prize scientific committee: Vikash Balasubramanian (University of Waterloo), Vikash Goel (University of Waterloo), Kashif Khan (University of Waterloo), Allison McConnell (Communitech), Neda Paryab (University of Waterloo), Pascal Poupart (University of Waterloo, RBC Borealis AI, Vector Institute), Gaurav Sahu (University of Waterloo)

This document describes the steps to submit a solution:

1. Prepare your submission
2. Upload your submission
3. View the score of your submission

1. Prepare your Submission

A submission consists of a Docker container that contains a program that reads an input file (dataset of claims with related articles) and outputs a prediction file (truth rating for each claim).

1.1 Docker Container

First, you will need Docker installed on your machine. See <https://docs.docker.com/install/> for instructions on how to install it.

A submission consists of a Docker container (see the sample submission as an example). The container will have the dataset mounted (read-only) inside of it. An input dataset consists of a json metadata file in the same format as `train.zip` (except that the “label” field is empty) and a folder of related articles in the same format as the folder `train_articles` (see the data documentation under the data tab of the DataCup website for more details). The metadata will be accessible at `/usr/local/dataset/metadata.json` and the related articles will be located at `/usr/local/dataset/articles`. The container is expected to produce a predictions file to `/usr/local/predictions.txt`.

The provided sample submission shows how to load the dataset using a Python script and produce a valid predictions file. All software installation needed to run the program should be specified in the Dockerfile. The Dockerfile ends with a command to launch the script. This runs when the container is started. For a reference on how to use Docker see <https://docs.docker.com/get-started/>.

To build the container, run

```
docker build -t image_name .
```

in the same directory as the Dockerfile. Then, to test running the submission locally, run this command:

```
docker run \
  -v $DATASET_PATH:/usr/local/dataset/:ro \
  --name container_name \
  image_name
```

where `$DATASET_PATH` is the folder with the dataset on your local machine. We recommend testing the submission on Google cloud (free credits available at <https://cloud.google.com/free/>) with the same type of machine (`n1-standard-4`, see hardware limits in Section 2) as what will be used to execute submissions to DataCup.

1.2 Predictions File Format

The Docker container is expected to produce a predictions file to `/usr/local/predictions.txt` in which each line contains two comma-separated values. The first value represents the claim id and the second value represents the prediction for that claim id (either 0, 1, or 2, corresponding to False, Partially True and True). Here is an example of a predictions file:

```
3, 2
7, 1
2, 0
5, 0
...
```

2. Upload your Submission

To submit your Docker container to the leaderboard, you will first need to save it as a tar file using this command:

```
docker save -o tar_file_name.tar image_name
```

Then, upload the tar file (max 4 GB) to the submission tab of the DataCup website. The following limits apply to each submission:

- Max number of submissions per day: 1
- Max Docker image size: 4 GB
- Docker container internet access: none
- Time limit for the execution of a Docker container: 30 min

- Hardware limits for the execution of a Docker container by DataCup on a Google cloud `n1-standard-4` machine:
 - 4 vCPUs
 - 1 T4 GPU
 - 15 GB RAM
 - 375 GB SSD

3. View the Score of your Submission

A predictions file is scored according to the macro F1 score described in the data documentation under the data tab of the DataCup website. The score is automatically posted on the leaderboard. In periods of high congestion, submissions will be queued and scored in order. As a result, a score may take several hours before appearing on the leaderboard.