NL4OPT Q&A Sessions

Session 1 August 9, 2022



NL4OPT Q&A Sessions

1 Evaluation Details

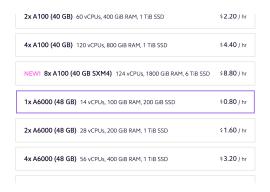
2 Competition Updates

3 General Q&A

Evaluation Details

Evaluation Server

We will be using the Lambda Labs for evaluating the submissions.



We encourage you to test your script on the same instance (see above) to ensure that the model runs smoothly.

Submission Folder

Since the submissions are evaluated using a bash script, your submission folder should follow the following structure:

- subtask1
 - submission folder 1
 - ...
 - evaluate.sh
 - submission folder 2
 - •
 - submission folder 3
 - ...
- subtask2
 - ...

Note

Please ensure that exact names are used for the subtask folder and bash script, and at most 3 submission folders for each subtask

Saving Results

The corresponding test set will be copied into each submission folder, your script should produce a text file "results.out" with a single line for the evaluation metric

- subtask1
 - submission folder 1
 - ...
 - test.txt
 - evaluate.sh
 - results.out
 -baaala
- subtask2
 - ...

Examples

Please refer to our example submissions for both subtasks on Google Drive

Things to Consider

CUDA Version

 Please ensure your installation supports Compute Capability 8.6 for A6000, the local CUDA version is 11.6

Runtime

 Due to the number of registered teams, please ensure that your evaluation scripts finish in a reasonable amount of time (< 10 min)

Storage Capacity

- Please refrain from storing overly large files in your submission folders (> 10GB).
- You can store parameters of overly large models in a remote location before copying them to the evaluation server

Using Conda in Shell Scripts

 Please include the following line to activate Conda environments inside your shell script source ~/miniconda3/etc/profile.d/conda.sh



Competition Updates

Updates to Subtask-1 Data

(Discussion #8 & #15) We have removed white space tokens from the NER dataset for Subtask-1.

```
Bold
                                         Bold
                                         Tycoon
                 B-LIMIT
                                                           B-LIMIT
work
and
                                         and
decides
                                         decides
                                                           B-CONST_DIR
                 B-CONST DIR
money
                 0
                                         monev
                                                           0
GICs
                 B-VAR
                                         GICs
                                                           B-VAR
                                         and _
index
                 R-VAR
                                         index
                                                           R-VAR
GICs -
                 B-VAR
                                         GICs
                                                           B-VAR
                                         vields
vields
```

- (a) Data before update (b) Data after update

The held-out test set has also been updated to the same format.

Dev Data for Training

(Discussion #24) After some considerations, we have decided against using the development set for training the model.

Note

Top submissions will be required to submit training code, where we will attempt to reproduce your approach without the development set.

General Q&A

Thank you for Joining!

Please continue to monitor our Website and Github Discussion Forum for announcements and updates.