

- [Hardware Utilization Prediction](#)
 - [Overview](#)
 - [Installation](#)
 - [Installation of the Environment](#)
 - [Installation and Usage of Git Large File Storage](#)
 - [Datasets](#)
 - [Timestamp Sorted Datasets](#)
 - [Machine Sorted Datasets](#)
 - [How to use](#)
 - [Timestamp Prediction](#)
 - [Machine Prediction](#)
 - [Resource Wastage Calculation](#)

Hardware Utilization Prediction

Overview

This subrepository uses the GPU machine traces collected by Alibaba to predict future hardware utilization and correct the user allocated resource parameters to values closer to the actual values. This is done with a **Long-Short Term Memory** machine learning model.

Installation

In this section, the installation steps are described in order to execute code inside this folder.

Installation of the Environment

To install the environment, it is recommended to use the Python dependency framework [anaconda/miniconda](#).

If [anaconda/miniconda](#) is not installed on your system, please visit [anaconda installation](#) for further instructions before continuing.

Once anaconda/miniconda is installed, open your preferred CLI and navigate to the **prediction** root folder this **README.md** resides in.

Execute the command **conda env create --file=environment.yml**. The CLI will ask you to accept installing the environment, please accept it and after doing so, the environment will be installed on your system.

After the environment is installed on your system, you can activate it with the command **conda activate lstm_prediction**. Then the python files with the file extension **.py** should be executable.

The Jupyter dependency is also included in this environment, but the Python kernel **lstm_prediction** has to be chosen inside a Jupyter notebook, in order to execute the code inside the notebook.

Installation and Usage of Git Large File Storage

In order to be able to download and use the datasets, *Git Large File Storage* is required.

If the git extension is not installed on your system, please follow the instructions of [git-lfs installation](#).

After successfully installing this extension, execute the command **git lfs pull** inside a CLI. This will download all required datasets.

Note: Since this download includes many monitored traces and evaluation data, it is recommended to use an internet connection that has no data cap.

Datasets

This section briefly describes the different datasets used in this repository.

Timestamp Sorted Datasets

The dataset **preprocessed_worker_new.csv** is sorted by timestamp (index **start_time**), meaning tasks are sorted by arrival time on the GPU servers, regardless

of which machine they will be sent to. If two tasks arrive at the same time, the tasks will be sorted by the **machine** column in the dataset.

This dataset is used in the Jupyter notebook [lstm_timestamp_prediction.ipynb](#) and the approach is further described in section **Timestamp Prediction**.

Machine Sorted Datasets

[df_machine_sorted.csv](#), [machine_indices.csv](#)

How to use

Timestamp Prediction

Machine Prediction

Resource Wastage Calculation

$$p := \{x \mid i \in N \wedge x_i \in pred \wedge y_i \in actual, x_i \geq y_i\}$$

$$wastage(pred, alloc) := 1 - (\frac{pred_{mean}}{alloc_{mean}})$$