AI HACKATHON DOCUMENTATION

This guide provides essential information to help you get started with resources and tools you might need or use throughout the hackathon.

Platforms and Resources

1. MyCSC Portal

Access and Registration:

- URL: https://my.csc.fi
- Sign in with your (HAKA login) or CSC credentials.
- After login, access your project resources and manage your CSC account details.
- Here connect to the Hackathon project and then after login in to the platform (Lumi/Puhti/ALLAS)

2. Puhti

Access Puhti:

- SSH login thru terminal
- or Login through OpenonDemand GUI

File System:

- **Home directory:** /users/username
- **Project directory:** /projappl/project_xxx
  Shared space for project data and results. Good for collaboration and medium-to-large datasets.
- **Scratch directory:** /scratch/project_xxx
  High-performance, temporary storage. Not backed up. Best used for large intermediate data or temporary files created during computations.

3. LUMI

Access LUMI:

- SSH login thru terminal:
- or Login though the OpenonDemand GUI

File System:

- **Home directory**: /users/username
- **Project directory**: /projappl/project_xxx
  Shared space for project data and results. Good for collaboration and medium-to-large datasets.
- **Scratch directory**: /scratch/project_xxx
  Fast, non-backed up temporary storage for large files, training datasets, or intermediate computation outputs.

Important Guidelines:

- Do **not** use the home directory for heavy I/O operations. Avoid using it for intensive computation or large file storage.
- Do **not** store critical data long-term in the scratch directory, as it may be purged regularly.
- Run heavy computational jobs from the scratch directory to optimize I/O performance.

- Backup essential results to the project directory. (Will be deleted on the expiry of the Project)

Slurm Job Scheduler

It efficiently allocates and manages computational resources, schedules tasks, and monitors resource utilization. With Slurm, you will submit, monitor, and manage batch jobs across a distributed computing environment.

Capabilities of Slurm:

- Allocate resources like CPUs, GPUs, memory, and nodes.
- Schedule jobs with detailed parameters (e.g., time limits, priority, resource specifications).
- Automate execution of complex workflows.
- Enable parallel job execution for optimal performance.

Sample Job Script:

```
#!/bin/bash
#SBATCH --job-name=myjob
#SBATCH --account=project_xxx
#SBATCH --partition=standard
#SBATCH --time=01:00:00
#SBATCH --mem=8G
#SBATCH --cpus-per-task=4

module load python
srun python myscript.py
```

Now in the terminal:
Submit Job:
sbatch myjob.sh

Monitor Jobs:

- List your jobs:
  squeue -u username

- Cancel a job:
  scancel JOB_ID

Documentation: Puhti: Puhti example scripts - Docs CSC
Lumi: https://docs.lumi-supercomputer.eu/runjobs/scheduled-jobs/batch-job/

Virtual Environments and Wrappers

Some reasons to use Virtual Environments:

- Avoid conflicts between project dependencies.
- Ensure reproducibility of your software environment.
- Easily manage and switch between different setups.

Python Virtual Environments (venv)

- **Create environment:**
  python -m venv myenv

- **Activate environment:**
  source myenv/bin/activate

- **Install dependencies:**
  pip install -r requirements.txt

Conda Environment (Anaconda)

- **Create environment:**
  conda create -n myenv python=3.10

- **Activate environment:**
  conda activate myenv

Wrappers (Singularity)

Wrappers (containers) like Singularity - Apptainer, encapsulate software and its dependencies into a self-contained environment. They provide a standardized way to package software, making it portable and ensuring consistent execution across various platforms (Puhti ,LUMI, etc).

Reasons to use Wrappers:

- Facilitate software portability and reproducibility.
- Easily deploy software across different computing environments.
- Manage complex dependencies within a single container.

Documentation: Puhti: **Containers - Docs CSC**
LUMI: **LUMI container wrapper - Documentation**

Further Support

For support and troubleshooting:

- Visit: https://docs.csc.fi

ICT Solutions for Brilliant Minds

CSC – TIETEEN TIETOTEKNIIKAN KESKUS OY
Keilaranta 14, PL 405, 02101 Espoo, puh. 09 457 2001,
Y-tunnus 0920632-0, www.csc.fi

CSC – IT CENTER FOR SCIENCE LTD.
Keilaranta 14, P.O. BOX 405, FI-02101 Espoo, Finland, Tel. +358 9 457 2001
VAT number FI09206320, www.csc.fi

ICT Solutions for Brilliant Minds

CSC – TIETEEN TIETOTEKNIIKAN KESKUS OY
Keilaranta 14, PL 405, 02101 Espoo, puh. 09 457 2001,
Y-tunnus 0920632-0, www.csc.fi

CSC – IT CENTER FOR SCIENCE LTD.
Keilaranta 14, P.O. BOX 405, FI-02101 Espoo, Finland, Tel. +358 9 457 2001
VAT number FI09206320, www.csc.fi