Big Data Platforms homework guide

Step 1: Download and Install Docker

1. **Windows**:
   - Visit the Docker for Windows page.
   - Follow the instructions to download and install Docker Desktop.
   - After installation, launch Docker from the Start menu.
2. **macOS**:
   - Visit the Docker for macOS page.
   - Follow the instructions to download and install Docker Desktop.
   - After installation, launch Docker from the Applications folder.
3. **Linux**:
   - For Linux, the installation steps depend on your distribution. Follow the instructions provided on the Docker for Linux page.

Step 2: Get released homework files

The homework files can be downloaded from Google Drive.

Step 3: Start Jupyter Notebook Container

1. Open a terminal or command prompt and navigate to the folder containing the downloaded homework files.
2. Pull the docker image with the following command: (You only need to do this once)

   ```
   docker pull hainingt/big_data_platforms:2024
   ```

3. Start a Jupyter Notebook container with the following command: (make sure to remove any line breaks when copying and pasting)

   ```
   docker run --name bdp_homework --rm -p 8888:8888 -v
   .:/home/jovyan/work hainingt/big_data_platforms:2024
   ```

Some notes about the above command:

- In case of you cannot find the local folder in the Jupyter Notebook, please check the permission of the folder. You can change the permission by running the following command before starting the container:

   ```
   chmod 777 -R .
   ```

- If you're using Docker Desktop on Windows or macOS, check your Docker Desktop settings. In some cases, you may need to explicitly grant Docker Desktop access to certain directories.
- Make sure there is no other Jupyter Notebook container running on port 8888.
- Make sure there is no illegal character in the path of the folder. You can also replace the aforementioned command with the absolute path of the folder.

4. Open the jupyter notebook in your browser by copying the URL with the token from the terminal or command prompt. The URL should look like this:

```
http://127.0.0.1:8888/?token=<TOKEN_VALUE>
```

Step 4: Access and Save Jupyter Notebook

1. Open a web browser and paste the copied URL with the token.

2. You will be directed to the Jupyter Notebook interface in your browser.

3. Create or open a notebook, work on it, and save the results within the Jupyter environment.

4. To save the notebook to your local machine:

   - Click on "File" in the Jupyter interface.
   - Select "Download as" and choose the desired format (e.g., Notebook (.ipynb)).

5. Please ensure that you only modify the cells marked with "# YOUR CODE HERE". We use Nbgrader for automatic grading, so Jupyter notebooks from other sources or those with unexpected modifications may be rejected by the grading system, meaning that the system will fail to give points for the exercise.

Step 5: Stop and Remove Jupyter Notebook Container

1. To stop the container, press `Ctrl+C` in the terminal or command prompt where you started the container.
2. If the container is running in the background, you can stop it with the following command:

```
docker stop bdp_homework
```

The container will be automatically removed when it is stopped.

Step 6: Upload Jupyter Notebook to Mooc Page

1. Go to the big-data-platforms-24 Mooc Page and sign in with your Mooc account.
2. The submission box is in the bottom of the page. Choose the assignment title and upload your notebook file.

Step 7: Receive Feedback

1. Every one hour the server will autograde your submission and generate the feedback result.

2. The feedback result will be sent to your email address (the one you used to register the Mooc account) from bigdata.platforms.feedback.gmail.com. Please check your spam folder if you cannot find the email.
3. You can submit multiple times and the system will only record the best score.
4. The feedback is auto-generated by a web crawler based on the output of your submitted Jupyter notebook, so the email layout may sometimes be incorrect. The grade displayed on the MOOC page is the official result and should be your highest score. (Click the blue icon in the bottom right corner to view the result.)