

Parallel and Distributed Computing

Lab1

DataBricks Registration

In this lab you will setup the development environment we will use throughout the rest of the course.

We will be using the community edition of DataBricks. DataBricks is an enterprise quality version of Spark with integrated Notebook and Cluster management. There are editions of DataBricks that cost money, but we are using the free community edition.

To login or sign-up for the community edition use the following link. To sign-up, there is a small link under the sign-in button. Please follow that link and fill in the registration form. I suggest using your DU as your workplace and your DU email as your email. You should receive an email that allows you to complete the registration.

<https://community.cloud.databricks.com/login.html>

Once you reset your password, you should try to log in and make sure it works. Note there is nothing to download as the system is web based. The features and use of DataBricks will be explained in upcoming weeks.

You can also read the community edition FAQ:

<https://databricks.com/product/faq/community-edition>

There is nothing official to turn in for this lab. Everyone will get full credit. Please do ask questions if you cannot get signed up.