User Training

Notice: In addition to our training programs, there are many other computational groups at HMS that provide training, including at least:
* [http://bioinformatics.hms.harvard.edu/](http://bioinformatics.hms.harvard.edu/)
* [https://datamanagement.hms.harvard.edu/news-events](https://datamanagement.hms.harvard.edu/news-events)

Please directly contact them if you are interested to attend their trainings.

Summer 2019 Registrations Open!

You must have a Harvard University ID (HUID) to be able to register for a class. This is required to access the Harvard Training Portal.

Those members of the HMS community who do not have HUIDs - such as employees at affiliate hospitals, or collaborators from other institutions - may self-register for one as a "Person of Interest" with their faculty member's sponsorship. The form may take several days to process, so we encourage non-HUID users to fill out this form ASAP in preparation for upcoming training: [http://iam.harvard.edu/files/iam/files/authorized-identity-request-form.pdf](http://iam.harvard.edu/files/iam/files/authorized-identity-request-form.pdf) (please read through the form for submission details)

Current Schedule

<table>
<thead>
<tr>
<th>Class</th>
<th>Date</th>
<th>Time</th>
<th>Room</th>
<th>Page</th>
<th>Registration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Analysis and Visualization in MATLAB</td>
<td>06/04/2019</td>
<td>1-3p</td>
<td>TMEC L-008</td>
<td></td>
<td>Please email: Alyssa Silverman &lt; <a href="mailto:asilverm@mathworks.com">asilverm@mathworks.com</a>&gt;</td>
</tr>
<tr>
<td>Intro to R/Bioconductor</td>
<td>06/19/2019</td>
<td>3-5p</td>
<td>Countway 403</td>
<td>R Class Files</td>
<td>Register Here</td>
</tr>
<tr>
<td>Demystifying deep learning: A practical approach in MATLAB</td>
<td>06/25/2019</td>
<td>1-3p</td>
<td>TMEC106</td>
<td></td>
<td>Please email: Alyssa Silverman &lt; <a href="mailto:asilverm@mathworks.com">asilverm@mathworks.com</a>&gt;</td>
</tr>
<tr>
<td>Intro to O2</td>
<td>06/26/2019</td>
<td>3-5p</td>
<td>TMEC106</td>
<td>User Training github</td>
<td>Register Here</td>
</tr>
<tr>
<td>From Pixels to Features to Models: Image Processing, Computer Vision, and Machine- and Deep Learning with Images in MATLAB</td>
<td>07/16/2019</td>
<td>1-3p</td>
<td>TMEC209</td>
<td></td>
<td>Please email: Alyssa Silverman &lt; <a href="mailto:asilverm@mathworks.com">asilverm@mathworks.com</a>&gt;</td>
</tr>
<tr>
<td>Intro to Python</td>
<td>07/17/2019</td>
<td>3-5p</td>
<td>TMEC306</td>
<td>User Training github</td>
<td>Register Here</td>
</tr>
<tr>
<td>Intro to O2</td>
<td>07/31/2019</td>
<td>3-5p</td>
<td>TMEC306</td>
<td>User Training github</td>
<td>Register Here</td>
</tr>
<tr>
<td>Parallel Computing with MATLAB and Scaling to the O2 Cluster</td>
<td>08/13/2019</td>
<td>1-4p</td>
<td>Countway 403</td>
<td></td>
<td>Please email: Alyssa Silverman &lt; <a href="mailto:asilverm@mathworks.com">asilverm@mathworks.com</a>&gt;</td>
</tr>
</tbody>
</table>

Classes offered:

Intro to O2
O2 for New Users addresses the needs of users who have very little Linux experience, and are just getting started with HPC. More time will be devoted to covering Linux basics, and the concepts of schedulers and jobs, and data management best practices. The lecture portion of this class is one hour, the second hour will be spent clinic-style with HMS RC staff to address workflow-specific questions and help convert commands to O2 SLURM syntax.

Intro to Python

Python is a popular scripting language for scientific computing and available across all computer platforms. The course will introduce you to some of the basics of the Python language as well as some of the nuances involved with its use specific to the O2 environment. The goal is to provide users with a foundational level of familiarity. Topics covered include basic data types and declaration, flow control (if/else), loops, a brief introduction to constructing a script, and a briefer introduction to modules. The course will be taught on O2, but general concepts are easily translatable to desktop and local installations.

Intro to R/Bioconductor

Intro to using R and Bioconductor. R is a powerful, open-source, highly adaptable statistical language useful for crunching numbers to datasets like those produced by next-gen sequencing. This class covers R basics and learning to think like/understand R. Users will learn how to set up personal R libraries on O2, and use O2 R for its high memory allocations and parallelization. Topics include how to install packages, learn about variables, data types, data manipulation, flow control, and functions, perform simple statistical tests, and create a variety of plots. Laptops are encouraged.

Class Files Here

MathWorks/HMS Summer Seminar Series 2019

June 4th, 1 – 3pm

Data Analysis and Visualization in MATLAB

In this session we introduce interactive tools to analyze, visualize, explore and model data, and show how to create an executable notebook using Live Editor. We demonstrate how to fit curves and surfaces to data, perform statistical analysis and how to apply statistics, machine learning and visualization tools to data that does not fit in memory. We also show how to build and publish customized reports, and how to create desktop and web apps to share your work on repositories like GitHub and MATLAB Central.

June 25th, 1 – 3pm

Demystifying deep learning: A practical approach in MATLAB

Are you new to deep learning and want to learn how to use it in your work? Deep learning can achieve state-of-the-art accuracy in many humanlike tasks such as naming objects in a scene or recognizing optimal paths in an environment. The main tasks are to assemble large data sets, create a neural network, to train, visualize, and evaluate different models, using specialized hardware - often requiring unique programming knowledge. These tasks are frequently even more challenging because of the complex theory behind them. In this seminar, we'll demonstrate new MATLAB features that simplify these tasks and eliminate the low-level programming. In doing so, we'll decipher practical knowledge of the domain of deep learning. We'll build and train neural networks that recognize handwriting, classify food in a scene, and figure out the drivable area in a city environment.

Along the way, you'll see MATLAB features that make it easy to:

- Manage extremely large sets of images
- Visualize networks and gain insight into the black box nature of deep networks
- Perform classification and pixel-level semantic segmentation on images
- Import training data sets from networks such as GoogLeNet and ResNet
- Import and use pre-trained models from TensorFlow and Caffe
- Speed up network training with parallel computing on a cluster
- Automate manual effort required to label ground truth
• Automatically convert a model to CUDA to run on GPUs

July 16th, 1-3pm

From Pixels to Features to Models: Image Processing, Computer Vision, and Machine- and Deep Learning with Images in MATLAB

Transitioning image models from pixel-based to feature-based allows us to extract information from images and video at a high level, to detect, classify, and track objects, co-register images, or understand a real-world scene. Using collections of features, we can train computers to recognize objects, with user-specified or automatically determined features. In this session, you will learn about capabilities for image processing and computer vision with the MATLAB product family. We will also describe approaches to implementing different models for machine learning, including using deep convolutional neural networks (CNNs).

Demos include:
  • Perform automatic image registrations using feature-based approaches
  • Detect, recognize, and track objects in images
  • Ground truth images and video, and train new classifiers
  • Using pre-built networks to create new recognizers in a “transfer-learning-based workflow”

Tuesday, August 13, 2019, 1:00 PM - 4:00 PM Countway 403

Title

Parallel Computing with MATLAB and Scaling to the O2 Cluster

During this hands-on, self-paced workshop, a MathWorks engineer will introduce you to parallel and distributed computing in MATLAB with a focus on speeding up your application code and offloading compute. By working through common scenarios and workflows using hands-on demos, you will gain an understanding of the parallel constructs in MATLAB, their capabilities, and some of the common hurdles that you’ll encounter when using them. You’ll also learn how to run your MATLAB code on local HPC resources.

Agenda
01:00pm Introductions and Overview of Workshop
01:05pm Overview of HPC cluster and support
01:10pm Overview of Parallel Computing with MATLAB
01:35pm Hands-on, self-paced exercises
03:00pm Question and Answers
03:15pm Running MATLAB on the cluster and submitting MATLAB parallel jobs
04:00pm Wrap-up

Prerequisites
For those not familiar with MATLAB it is recommended to taking the free self-paced MATLAB Onramp course (2 hours)
https://www.mathworks.com/learn/tutorials/matlab-onramp.html

System Requirements
• Laptop with power adapter
• MATLAB R2016a or newer
• Parallel Computing Toolbox
• SSH client
• Account on HPC cluster