Programming Assignment Overview

Help Center

Overview

The course includes an optional programming assignment where you can gain hands-on experience with some of the important text mining concepts and algorithms. The assignment will be out in Week 2 and will cover the following topics:

1. Overview of some NLP Tools
2. Word Association Mining
3. Topic Modeling
4. Text Mining Competition

To complete the assignment, you will use MeTA, a data sciences toolkit that facilitates mining big text data. You can learn more about MeTA and its functionalities by visiting MeTA's website. Since MeTA is written in C++, you are going to use C++ to complete the programming assignment. If you do not have prior knowledge of C++ but have experience with any other object-oriented programming language, you are highly encouraged to complete the assignment as the programming tasks will barely rely on the language's specific syntax.

Below we provide instructions to install the software prerequisites. Make sure to have them installed during Week 1 so you are ready to start working on the assignment when it is out during Week 2.

If you have questions about this installation process, use the **Programming Assignments Forum**. This is a great place to ask questions and also help your fellow classmates.

Software Installation

There are two software prerequisites you must install in order to complete the assignment:

1. MeTA 1.3.6
2. Python 2.7 (used only for uploading your assignment, not for coding)

MeTA officially supports Linux operating systems and Mac OS X 10.6 or higher but does not support Windows (it is expected to start supporting it very soon). If you have an officially supported operating system and some experience in updating and installing packages, go through the steps in **Option 1** below. If your operating system is not supported or you want to get started with the programming assignment in minimum time without having to install packages, proceed with **Option 2** where we provide a virtual image of Ubuntu with MeTA pre-installed.

Option 1 (Supported OS)

Go to the MeTA Setup Guide and follow the steps specific to your operating system. Note: There is one extra command you should execute in the installation process that is not mentioned in the MeTA Setup Guide. After executing the following two commands:

```
git clone https://github.com/meta-toolkit/meta.git
cd meta/
```

You should, in the same termimal, execute:

```
git reset --hard v1.3.6
```

This ensures that v1.3.6 of MeTA, the version on which the assignment is based, will be installed. Then proceed with the rest of the installation instructions in the MeTA Setup Guide. After installing MeTA, you should make sure that you have Python 2.7 installed. If it is not installed, check Python's Download page for the installation instructions specific to your OS.

Option 2 (Easier Installation/Non-supported OS)

In case your operating system is not supported or you prefer not to install the packages needed by MeTA, use this option. We have created a virtual image of Ubuntu with MeTA and Python 2.7 pre-installed. The image is around 2 GB, so plan to install it early on if you have a slow Internet connection. You will need a virtualization software such as VirtualBox to run the virtual image. Below are the installation steps:

1. Install the latest version of VirtualBox.
2. Download the virtual image. You have two options to download the image: either using a BitTorrent (highly recommended) or a direct download:
   - Torrent
   - Direct Download
3. Open VirtualBox. Click on the **File** menu and select **Import Appliance**. A wizard should open; point it to the location of the image you downloaded in Step 2. See a graphical illustration of these steps.

Now, a new image called **Text Mining** should show up in the VM list. Click on **Text Mining** and start it. Ubuntu should load and might ask you for the username and password. Use "coursera" (without quotes) for both the username and password. After logging in, you should find a folder called **meta** on the desktop. This folder contains a pre-installed version of MeTA which you are going to use in the upcoming assignment. If you are interested in exploring MeTA early on, you can go through the tutorials on MeTA's website.

Created Thu 4 Jun 2015 10:47 AM PDT

Last Modified Mon 8 Jun 2015 7:00 AM PDT