CPE 486/586: Machine Learning for Engineers

01 Tools for Machine Learning

Fall 2025

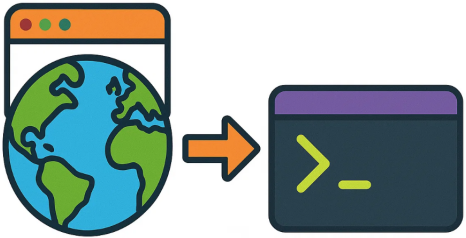
Rahul Bhadani

Outline

Command Line Tools and Linux

Why Command Line Tools and Linux

1. Download data from another location, webpage or server

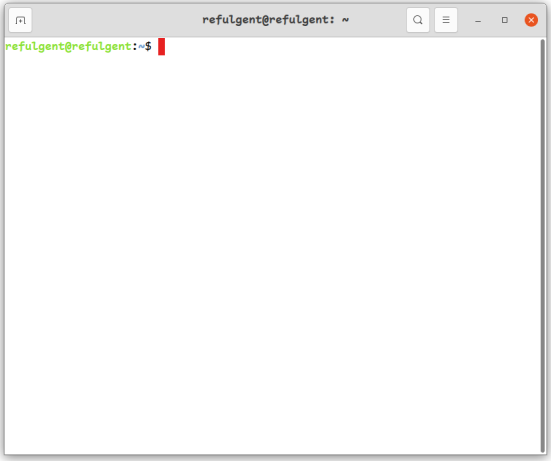


Why Command Line Tools and Linux

2. File operation, renaming, copying, editing, etc, in bulk.

3. Logging into remote server
4. To integrate with other languages and tools.
5. Command Line provides scalability, extensibility, and can automate the entire pipeline for machine learning and data science.

Command Line



1. Terminals in Linux and Macbook
2. Windows. Several Options:
   1. Git Bash. Install guide: https://youtu.be/SsdpuprzRE0
   2. Cygwin. Install guide: https://youtu.be/_j0Prs7aggo
   3. WSL.Install guide: 1. https://youtu.be/GMhV5Uqd8R8, 2. https://youtu.be/NPuIUT_6NeM?si=g0N39WfsBf3kIKCx

Common Linux Commands

My recommendation is to use Windows Subsystem for Linux, as it provides full linux funtionality to Windows.

A good tutorial to get started is **Getting started with Linux and Bash** at
https://learn.microsoft.com/en-us/windows/wsl/tutorials/linux

1. Installing Software:

   ```
   sudo apt-get install vim git
   ```

2. Get the working directory path:

   ```
   pwd
   ```

3. Change directory:

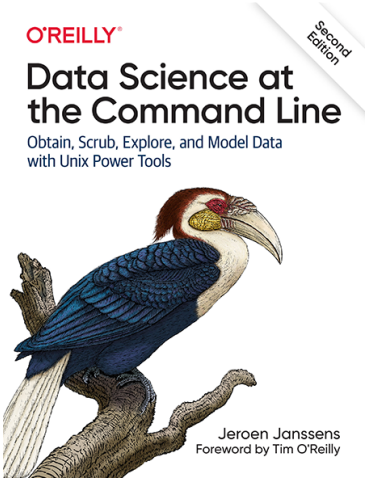
   ```
   cd /data/ch02
   ```

4. Create a directory or folder:

   ```
   mkdir data
   ```

An in-depth tutorial is at

https://jeroenjanssens.com/dsatcl/chapter-1-introduction



Remote Logging

If you do not have a powerful machine, you can remotely logging into UAH Engineering Linux Server using your UAH ID.

```
ssh rkb0022@blackhawk.ece.uah.edu
```

Windows users should use WSL Terminals.

Python

Installing Python

Most machine learning libraries and codebase are written Python. Python is supported by a strong open-source community.
Windows users download from https://www.python.org/downloads/
Mac users download from https://www.python.org/downloads/macos/
Linux users install from command line

```
sudo apt-get install python3.12
```

or

```
sudo apt-get install python3.12-full
```

Recommended version of Python for this course is Python 3.12.

Git and GitHub

Version Controlling your Development

It is must that any of your machine learning project, and in general any coding project should use Git.
You should already have Git installed in your system through Git, otherwise install.

1. Create a GitHub account and a user name.
2. Create a new repository.
3. Clone to local machine:

```
git clone https://github.com/rahulbhadani/TestRepo TestRep
```

Commonly used Git Commands I

1. Initialize an existing local folder as a git repo.

```
git init
```

2. Add remote to your local git folder (Only needs to do once per repo).

```
git remote add origin https://github.com/rahulbhadani/TestRepo.git
```

3. Add files for commit.

```
git add . # add all files
git add somefile.txt # Add specific file(s)
```

4. Commit changes with a message before you push to the remote.

```
git commit -m "Added some files" # add a meaninful message
```

Commonly used Git Commands II

5. Push to the remote.

```
git push  # shortcut command
git push -u origin # push to a specific remote named origin
git push -u origin2 HEAD:master # push to specific origin's specific branch
```

6. Pull updates from a remote.

```
git pull  # shortcut command
git pull origin master # pull from specific remote
```

7. Git undo add before commit

```
git reset filename.txt
```

8. Git undo add after commit but before push and keep your changes

Commonly used Git Commands III

9.
```
git reset --soft HEAD~1
```

10. Git undo add after commit but before push and unstage

11.
```
git reset HEAD~1
```

12. Git undo add after commit but before push and discard every changes

13.
```
git reset --hard HEAD~1 # use n for undoing n multiple commits
```

Development Environment

Integrated Development Environment (IDE)

Why use IDEs?

1. Provides integrated view of editor, file explorer, command lines.
2. Some IDEs provide intellisense for autocomplete.
3. IDEs are also integrate with compiler/intepreter – one click to run your code.
4. IDEs are beautiful, rich colors and fonts over boring plain text editors.
5. Many IDEs add support for additional features such as visualization, remote log in, data wranling, etc.

VS Code

Many IDEs are available such PyCharm, Atom, VS Code.

For this course, I recommend VS Code.

Visual Studio Code

VS Code supports several extensions for Python, and other necessary tools.

Packages and Package Manager

Jupyter Notebook

Markdown

Latex

Other Notable Tools

The End