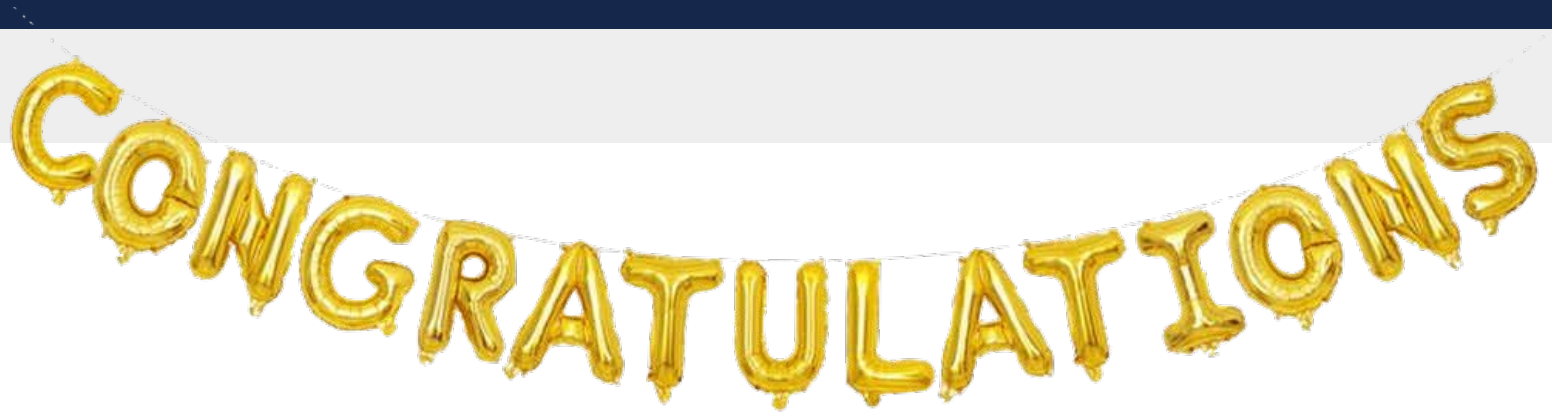
Berkeley Engineering | BerkeleyHaas

PROFESSIONAL CERTIFICATE IN MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

Module 5

Practical Applications

Office Hours with Viviana Márquez
May 2, 2024

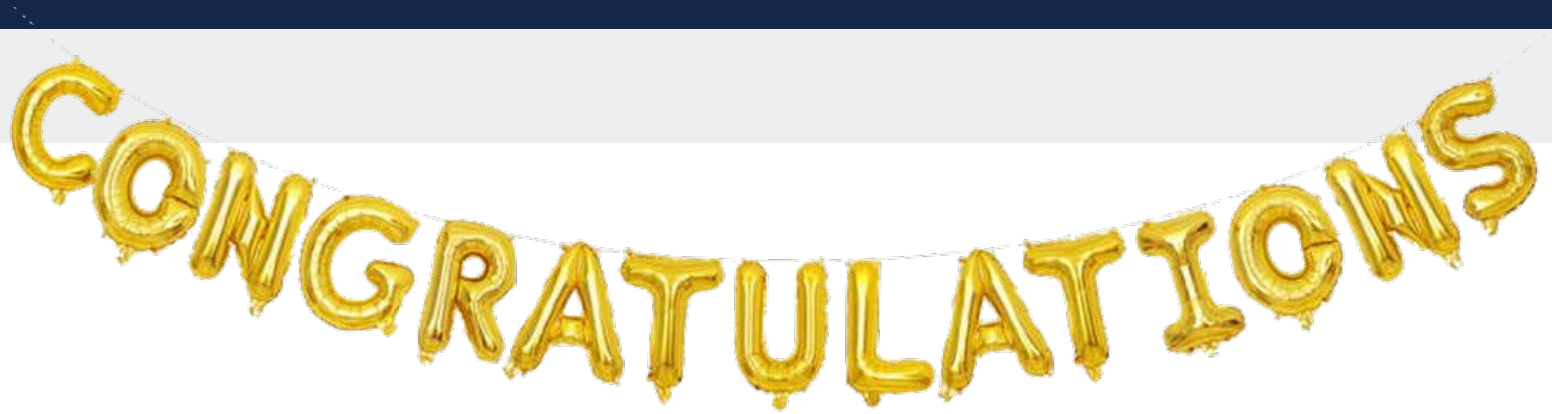


on finishing

Section 1: Foundations of ML/AI

Break week: May 8 - May 14

Please enjoy this short break before the launch of the next module. You can use time to catch up with assignments, review videos or peruse resources you may not have had the chance to look through. No new activity or content is released this week.



on finishing

Section 1: Foundations of ML/AI

Break week: May 8 - May 14

Please enjoy this short break before the launch of the next module. You can use time to catch up with assignments, review videos or peruse resources you may not have had the chance to look through. No new activity or content is released this week.

Next Section

Section 2: ML/AI Techniques

AGENDA

- Required activities for Module 5
- Content review Module 5: Practical Applications I
- Questions

Required Activities for Module 5

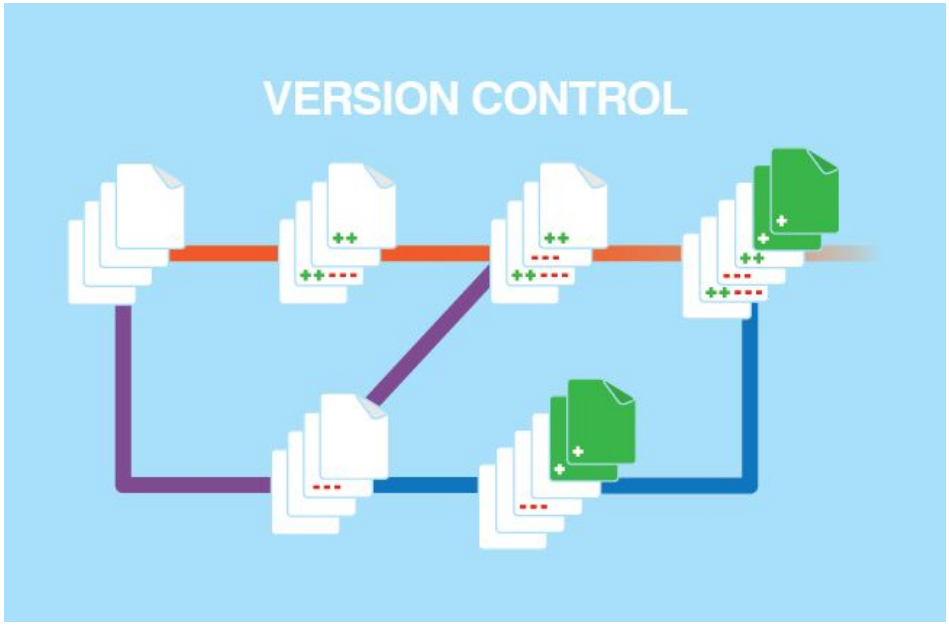
- 👩‍💻 Practical Application Assignment 5.1: Will the Customer Accept the Coupon?
- 🤍 Activity 5.1: Setting up your GitHub Portfolio

https://classroom.emeritus.org/courses/6063/assignments/212728?module_item_id=1513343

Content review Module 5: Practical Applications I

- Version Control
  - Git
  - GitHub

The importance of Version Control



- System to track changes in files over time during software development
- Allows multiple people to work on the same project without stepping on each other's toes
- Facilitates easy rollback, collaboration, and parallel work

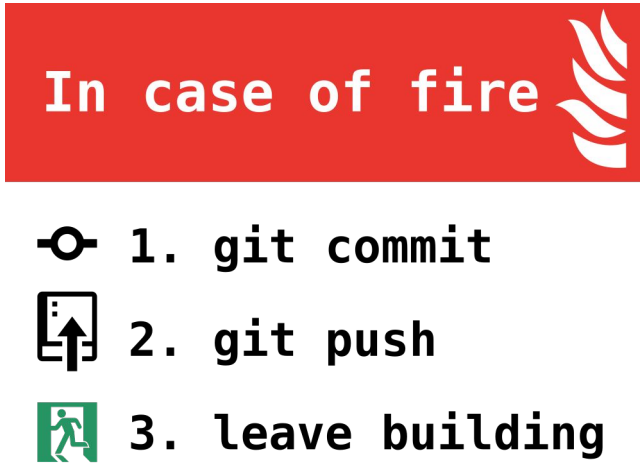
Git

- Git is the most popular Version Control system nowadays
- Specifically, it's a Distributed Version Control System (DVCS), meaning every working copy of the codebase is a complete repository with full history
- A command-line tool (though there are GUI versions available) that manages source code history
- Can be used independently of any online platform

GitHub

- GitHub is a platform that provides hosting for software development version control using Git. It allows multiple people to work on projects simultaneously, without interfering with each other's work
- It's the world's leading software development platform with millions of developers hosting and reviewing code, managing projects, and building software
- Offers a user-friendly interface and additional features like pull requests, issue tracking, and integrations with other developer tools
- While it's built around Git, using GitHub also means using Git, but the reverse isn't necessarily true. One can use Git without ever touching GitHub

GitHub – Importance in Data Science



- **Collaboration**: Data scientists often work on teams and need a platform to collaborate, share code, and maintain versions of their scripts, notebooks, and data
- **Reproducibility**: GitHub ensures that there's a record of code changes, allowing for transparency and reproducibility in data science tasks
- **Portfolio Building**: For budding data scientists, having a GitHub profile can serve as a portfolio of their projects, demonstrating their coding and analytical skills to potential employers

GitHub – Basic terminology

- **Repository (Repo):** A directory or storage space where your project lives
- **Commit:** A saved change to your repo
- **Branch:** A parallel version of a repository
- **Pull Request (PR):** Proposing your changes and requesting that someone review and pull in your contribution
- **Merge:** Merging your changes back to the main (master) branch
- **Fork:** A personal copy of another user's repository that lives on your account
- **Clone:** A copy of a repo that exists on your local computer

GitHub Pages

- GitHub Pages is a static site hosting service offered by GitHub. It allows users to transform their GitHub repositories into websites
- By default, the URL of your GitHub Pages site will be in the format `username.github.io/repository-name` . However, you can also set up a custom domain if you want a more professional or personalized URL
- GitHub Pages is free for public repositories, but there are some limitations
- Documentation: https://docs.github.com/en/pages/getting-started-with-github-pages/creating-a-github-pages-site
- Examples: https://github.com/collections/github-pages-examples

GitHub – Safety and Etiquette

- **Do not upload sensitive information:** Never commit passwords, API keys, or any other sensitive information to public repos
- **READMEs and Licensing:** Create clear README files for every repo and the understanding of software licensing
- **Respectful Collaboration:** Be kind, understanding, and patient when collaborating on shared projects

GitHub – Resources

- Command Line: MAC vs Windows
  https://tracer.gitbook.io/manual/support/command-line-mac-vs.-windows

- Is a Mac or Windows PC Better for Programming?
  https://medium.com/geekculture/is-a-mac-or-windows-pc-better-for-programming-d5556bf06f1#:~:text=The%20Operating%20System%3A%20macOS%20vs,benefit%20of%20macOS%20is%20security

- Installing Git
  https://git-scm.com/downloads

GitHub – Getting started!

- Sign in or create an account at: https://github.com/
- Configure Git with your username and email (important for commit messages):

```
git config --global user.name "Your Name"
git config --global user.email "youremail@example.com"
```

- Create a new repository

```
git clone https://github.com/username/repository-name.git
```

- Make Changes and Commit

```
git add .
git commit -m "A descriptive message about the changes."
git push origin main
```

- Fetches Changes

```
git pull
```

QUESTIONS?