

Agenda - Machine Learning (DSC 305)

Agenda - Machine Learning (DSC 305)

January 17, 2023

Week 1 - Course overview & ML overview



Photo by Katarzyna Pe at Unsplash.com

- ☒ Course overview - assignments, grading, topics, platforms
- ☒ Interesting article: "The Fourth Age of Programming"
- ☒ Interesting webinar: "Data Trends & Predictions 2023" report (PDF)
(Wed Jan 18, 10AM) - register/attend/watch the recording later
- ☐ Introduction to machine learning
- ☒ First DataCamp assignment "Understanding machine learning"

Week 2 - Introduction to ML & Review of R

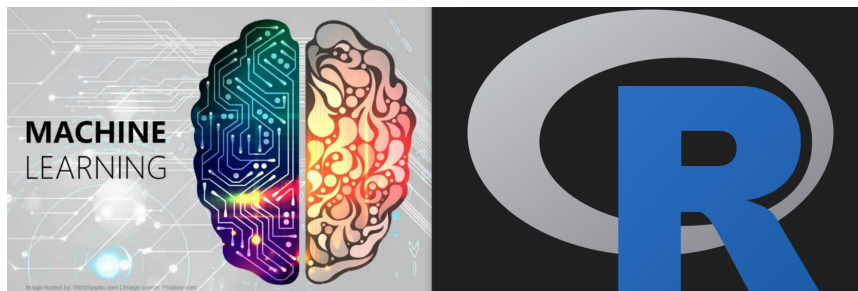


Figure 1: ML intro & R review in week 2 of DSC 305

Topics:

- ☐ Introduction to machine learning (continued)
- ☐ Review: R data structures, environment, exploration

Submissions:

- ☐ First multiple-choice test in class on Thursday, Jan-19
- ☐ First DataCamp home assignment "What is Machine Learning" Jan-19
- ☐ First in-class practice "Introduction to machine learning"

Recommendations:

- ☐ "Lyon Data & Analytics Nexus" on Discord (invitation only)
- ☐ Google search operators tutorial (video, 1 hr) => project
- ☐ ML for everybody by FreeCodeCamp (video, 4 hrs) => Python

Lyon data science nerd network (invitation only)

- Join the discord server to grow your professional network
- Lyon alumni, students, and companies are invited
- Share internship and job opportunities, and more

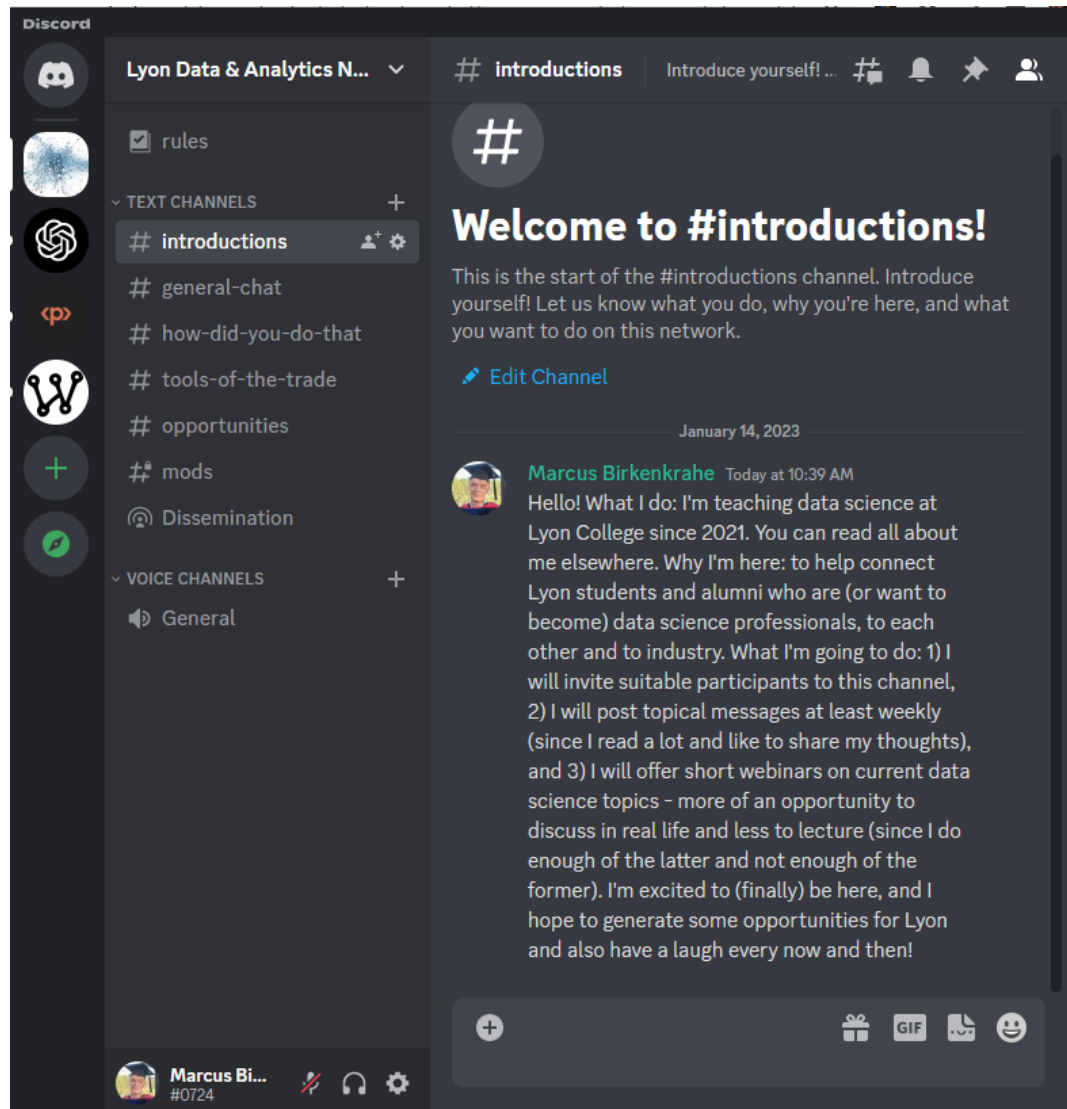


Figure 2: Lyon Data & Analytics Nexus Discord server dashboard

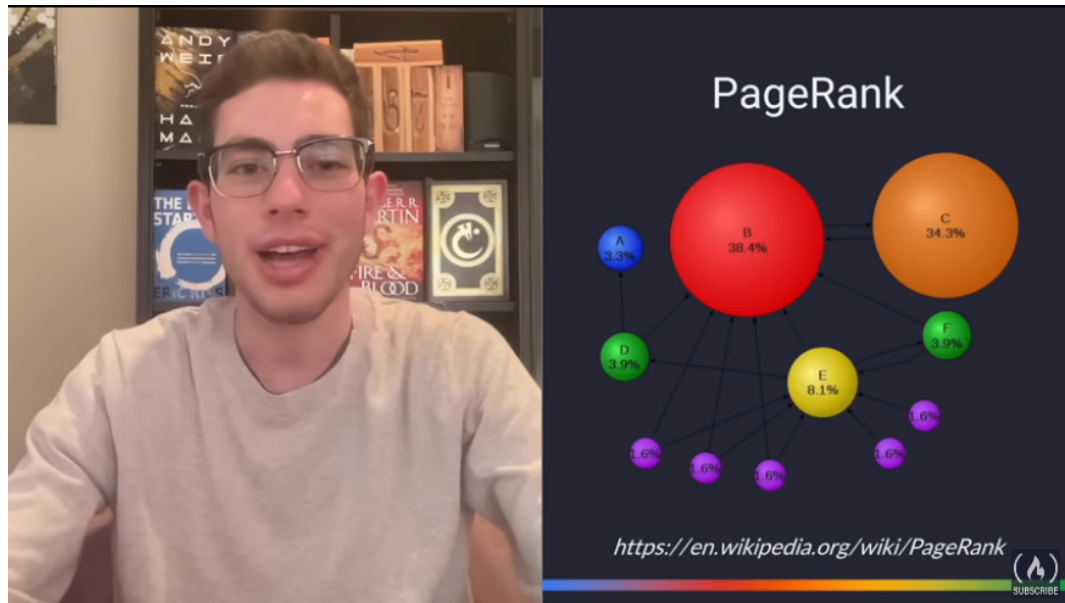


Figure 3: Google search tutorial by Seth Goldin (freeCodeCamp)

Google search for programmers (Goldin, 2023)

Waste of time?

1. Searching is not finding
2. Finding is not understanding
3. Understanding is not using

Google's business model = data + algorithm + infrastructure + brand

1. data are offered by users for free
2. algorithm is largely known and public
3. infrastructure is proprietary and hard to duplicate
4. brand is unique and highly valuable (but also vulnerable)

ChatGPT's answer to "how do I search using Google effectively?"

Here are a few tips for searching effectively using Google:

1. Use quotation marks to search for an exact phrase. For example, searching for `"chocolate chip cookies"` will return pages that include that exact phrase, rather than pages that include the words `"chocolate"`, `"chip"`, and `"cookies"` separately.
2. Use a minus sign to exclude certain words from your search. For example, searching for `chocolate chip cookies -recipe` will return pages that include `chocolate chip cookies` but exclude any pages that include the word `"recipe"`.
3. Use the `site:` operator to search within a specific website. For example, searching for `site:nytimes.com chocolate chip cookies` will return pages on the New York Times website that include the words `"chocolate chip cookies."`
4. Use the `filetype:` operator to search for specific file types. For example, searching for `chocolate chip cookies filetype:pdf` will return PDF documents that include the words `"chocolate chip cookies"`.
5. Use the `related:` operator to find websites that are related to a specific website. For example, searching for `related:nytimes.com` will return websites that are related to the New York Times website.
6. Use the Google advanced search page for more complex queries.

Keep in mind that Google's search algorithm is constantly evolving, so these tips may not always return the most accurate results.

Caveat: professional programmers use aggregators instead

- `stackoverflow.com` with sub-groups like `chess.stackexchange.com`
- Internet Relay Chat - in Emacs (`M-x erc`), e.g. `libera.chat`
- Usenet, mailing lists, Slack, Discord, reddit, quora etc.

Rapid review - intro to ML

1. Most popular OS for ML?
2. Which tools are we using in this course?



Figure 4: Photo by Benjamin Davies on Unsplash

3. What are the deliverables for your project?
4. Where do you get the topic for your term project?
5. What are the steps for a supervised learning process?
6. What is the general ML process?

References

- Goldin, Seth @freeCodeCamp.org (Sep 9, 2023). Google Like a Pro – All Advanced Search Operators Tutorial [2023 Tips]. Online: youtube.com.
- Ying, K. @freeCodeCamp.org (Sep 26, 2022). Machine Learning for Everybody - Full Course. Online: youtube.com.
- Stokes, Jon (Jan 4, 2023). The Fourth Age Of Programming [Blog]. URL: blog.repolit.com