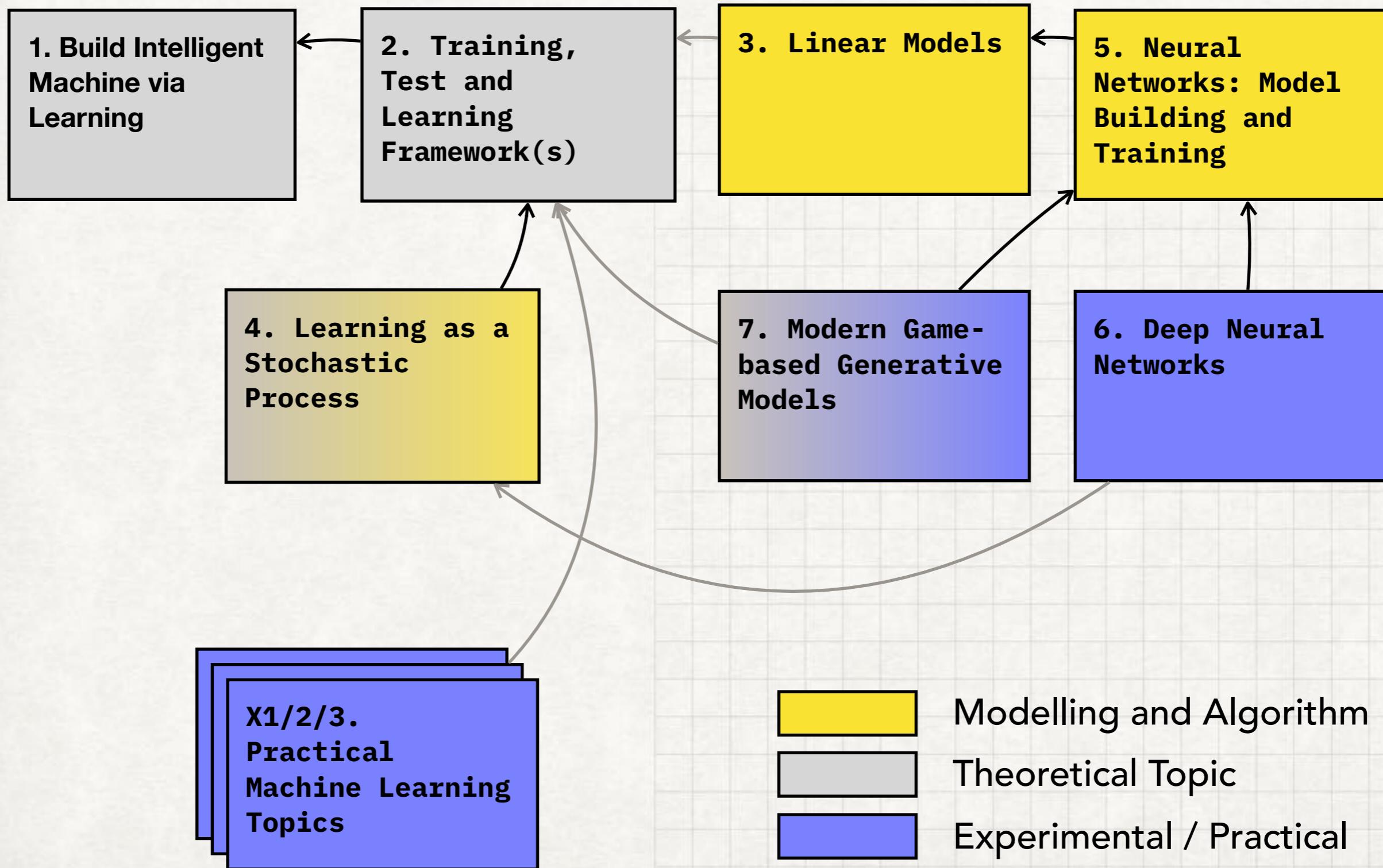


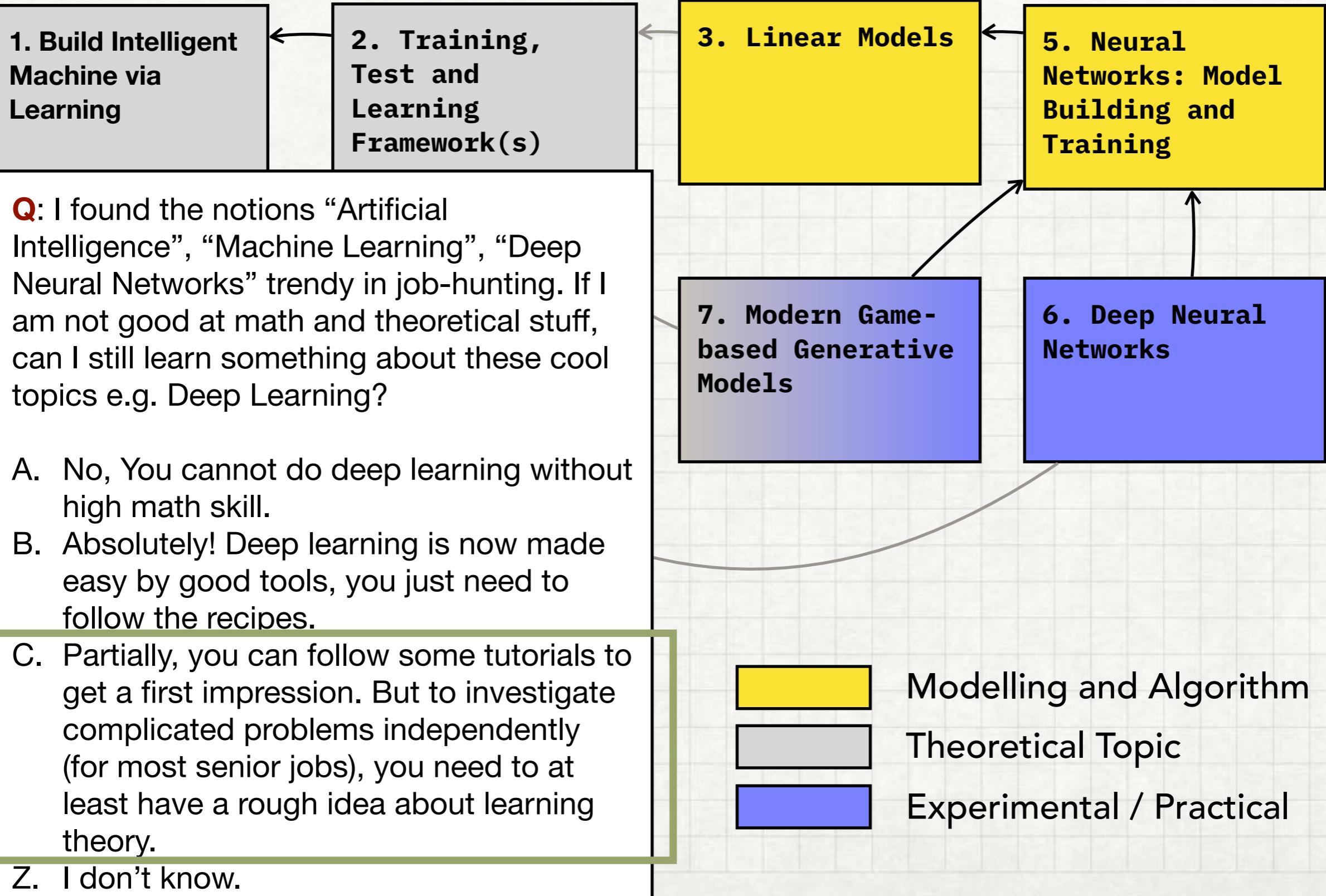
# MACHINE LEARNING

# **ABOUT THE SUBJECT**

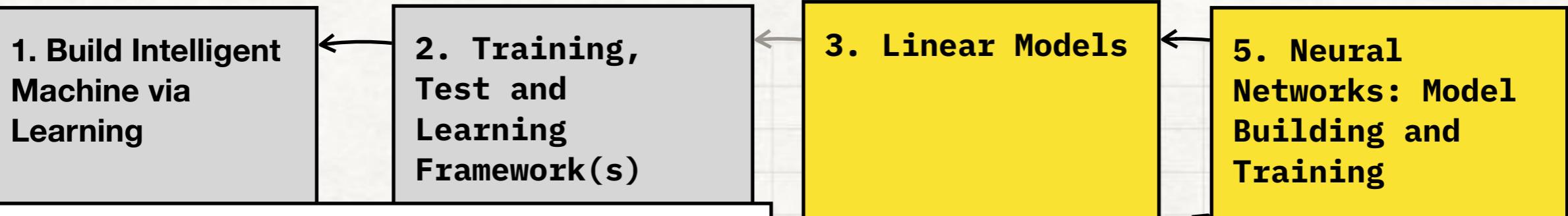
# 1. OVERVIEW



# 2. MY LEARNING PLAN

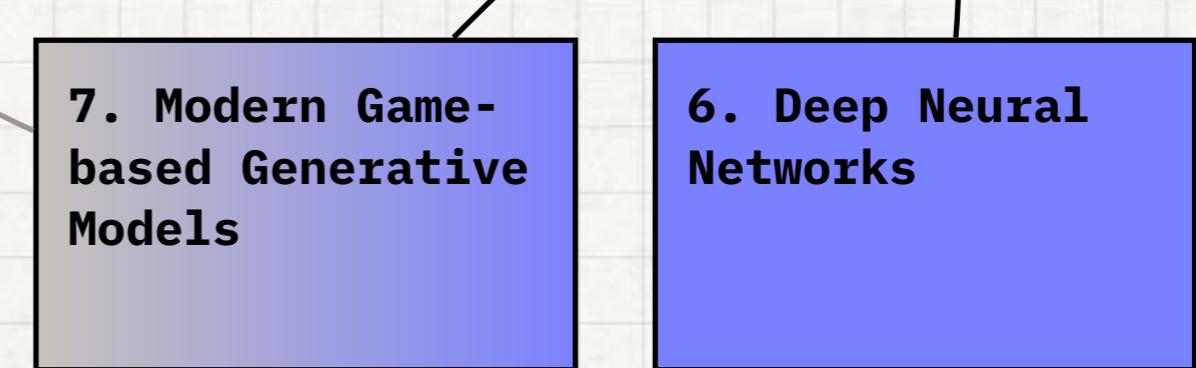


### 3. WHEN TO CODE



**Q:** From which part in this study do you think Python programming becomes necessary?

- A. From the very beginning
- B. When we start doing algorithms, in the part of “linear models”
- C. When we do deep neural networks
- D. No, it is not necessary. The requirement is just to make my life more difficult!
- Z. I don’t know.



Modelling and Algorithm



Theoretical Topic



Experimental / Practical

## 4. GOOD BOOKS

- Introduction to Statistical Learning (with App in R) [amazon](#) (do the R part only if you've already in)
- Machine Learning (Tom's) ([amazon](#), 0)
- Deep Learning (Modern) ([amazon](#), +)
- Learning from Data: A short course ([amazon](#), \*)
- Pattern Recognition and Machine Learning ([amazon](#))
- Information Theory, Inference and Learning Algorithms ([amazon](#), \*\*)

## 5. ONLINE LECTURES

- Andrew Ng's courses (start from the Stanford one if you want a solid basis and don't know where to start).
- Udacity (engineering oriented)
- Udemy (easy to get your hands dirty)
- fast.ai (toolset close to ours)

## 6. GETTING HELP (GENERAL)

- Make a team
    - Answering questions helps yourself understand greatly
      - Try to explain to the guy next to you what is information gain you might think you had learned in the “introduction to data analytics”
  - Get your hands dirty
  - Come to the consultation before it gets too crowded.
    - Consultation arrangement
- **Q:** When to book a face-to-face QA.
    - A. I don't know the deadline of Assignment 2.
    - B. When I import a library in Python, the computer complains “No module named ‘XXX’”
    - C. Things just don't work. I don't know how to describe the problem.
    - D. My ID3 decision tree works on 70% training subset of MNIST data, but not on the test subset. I tried to find answers on the Internet but failed. I prepared the experiment configuration, including data and source code.
    - E. You said simpler models are generally better. I checked on the Internet and watched tutorial by Andrew. But I am still confused.
    - Z. I don't know.

# **HOW TO PASS (WITH A DESIRABLE GRADE)**

# 1. ASSIGNMENTS OVERVIEW

- No Final Exam!
- 3 Assignments
- Need to code in A2!

		A1	A2	A3
31005	Weight	30%	40%	30%
	Deliverable Content	Literature Review Report	Project Report (including implementation) + Presentation	Take-home exam
	Time Budget	4 Weeks	4 Weeks	2 Weeks (Error)
	Time Cost (Estimate)	2 hr x 12 Deep Working	2 hr x 24 Deep Working	2 hr x 6 Deep Working
32513	Weight	30%	30%	40%
	Deliverable Content	Literature Review Report	Project Report (including implementation)	Proposal + Presentation
	Time Budget	4 Weeks	4 Weeks	2 Weeks
	Time Cost (Estimate)	2 hr x 12 Deep Working	2 hr x 22 Deep Working	2 hr x 8 Deep Working
	Submission Format	Link to a Notebook @ your_github_repo + PDF submission to Turnitin		

## 2. A1 - READING LITERATURE

- A list of works will be provided next week.
- What to submit: a critical review report.
- Time: 4 Weeks. Bad news: no examples. Good news: you will be guided
- What's in the report
  - What attracts you.
  - Setting up the task: input and output
  - Motivation and background
  - The Method/Technique (you can be forgiven is making mistakes here)
  - Conclusion (Taking home points).
  - Presentation
  - Your own comments on each part.

- **Q:** You will lose marks for
  - A. Clearly stating the I/O of the proposed algorithm/processing model
  - B. Stating that Principle Component Analysis is used to reduce the number of attributes of the data from 256 to 16, but I don't understand PCA.
  - C. Writing a piece of pseudo-code contains mistaking min-max normalisation with zero-mean-unit-var normalisation
  - D. Identifying the essential contribution of the paper by finding the key statement by the authors in the abstract of the article.
  - Z. I don't know and don't want to pretend to know.

### 3. A2(VER.A) - PRACTICAL PROJECT

- Like A3 in Introduction/Fundamental Data Analytics
- What to submit: an project implementation report.  
**(31005: With a Video Pitch)**
- Time: 4 Weeks. Good/Bad news: same as above
- What's in the report
  - Why you find this project worthy doing (significance and challenge)
  - Background: existing approach and why yours is different. (You CAN copy approaches to new problem!)
  - The Proposed Method/Technique and why you have proposed them
  - Implementation and Evaluation
  - Conclusion: (Lesson Learned)
- **Q: Over-simple tasks (multiple)**
  - A. Get NOAA data and UN FAO data to check if climate change has direct effect on food production
  - B. Iris classification.
  - C. Predict control (right/left push) on a cart so a pole is standing.
  - D. Recognise hand-written digits.
  - E. Predict Titanic survivors.

## 4. A3(VER.B 31005 ONLY) - ALGORITHM IMPLEMENTATION

- What to submit: a implementation report of an algorithm  
+ Video Pitch
- Time: 4 Weeks. Bad news: same as above; Good news:  
lots of code examples/tutorials on the Internet
- What's in the report
  - Algorithm Brief (Input, output, intermediate data structure design)
  - Implementation **with comments**
  - Testing
  - Reflect on challenging part (e.g. improvement of memory usage etc.)

## 5. A3 (31005) - TAKE-HOME EXAM

- Open Problems
- 2 Weeks (Error on Subject Outline, but on your side!)
- What's in the report
  - Your understanding of the key challenge in the problem
  - Your proposed solution
  - Attempt is more important than success

## 6. A3 (32513) - PROPOSAL

- Identify an open problem
- 2 Weeks
- **Video Pitch**
- What's in the report
  - Choose a problem with passion and interest ()
  - Your proposed method
  - Administrative aspects (schedule / personnel / budget etc.)

## 7. PREPARE A GITHUB ACCOUNT

- UTSOnline  
+ **github**

**Create a new repository**

A repository contains all project files, including the revision history. Already have a project repository elsewhere? [Import a repository](#).

---

Owner      Repository name \*

 junjy007 ▾ / UTS\_ML2019\_ ✓

Great repository names are short and memorable. Need inspiration? How about [literate-octo-couscous](#)?

Description (optional)

---

 **Public**  
Anyone can see this repository. You choose who can commit.

 **Private**  
You choose who can see and commit to this repository.

---

Skip this step if you're importing an existing repository.

**Initialize this repository with a README**  
This will let you immediately clone the repository to your computer.

---

Add .gitignore: **None** | Add a license: **None** ⓘ

---

**Create repository**

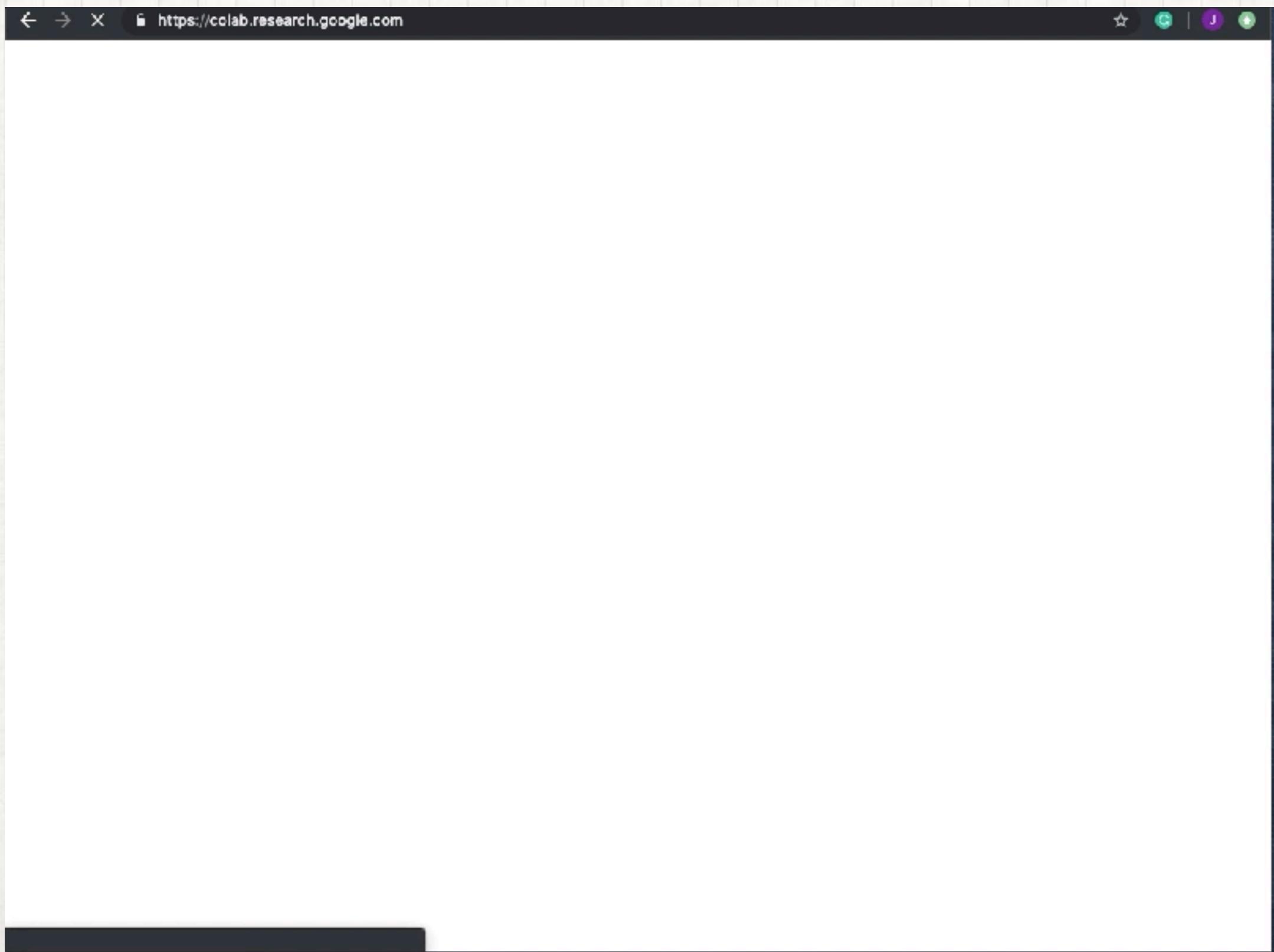
## 8. SUBMISSION

- UTSONline+ **github**

The assignments are designed to (especially A2) encourage you to take challenging tasks. EFFORT and HARD TRYING win marks.

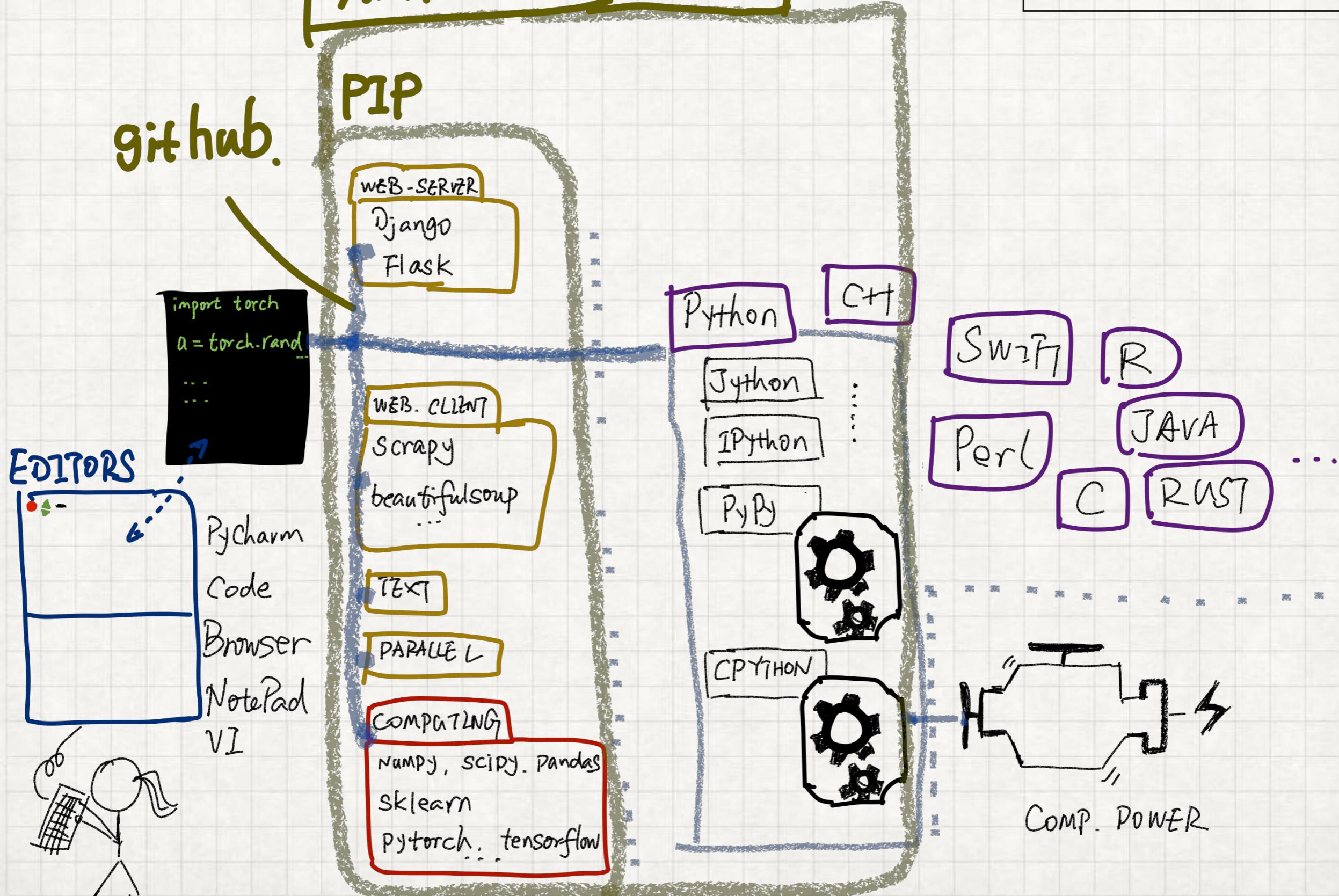
- **Q:** What will my marker need to assess my assignments (multiple)
  - A. A PDF report submitted to Turnitin task on UTSONline
  - B. A link within the PDF file pointing to the repository holding the draft of my report
  - C. A link within the PDF file pointing to my github account
  - D. My source code submitted to Turnitin as a zip file.

## 9. PROGRESSIVE SUBMISSION

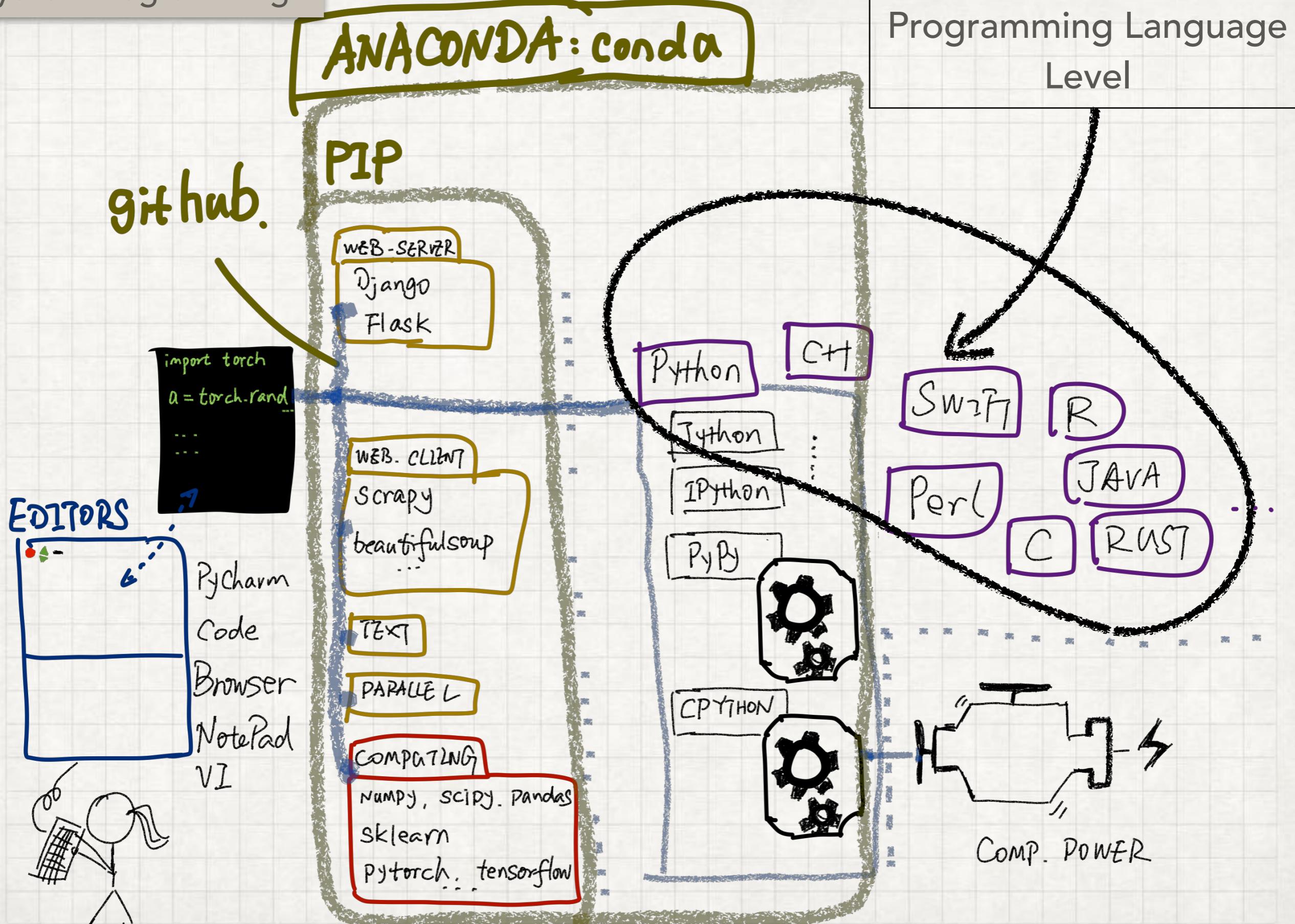


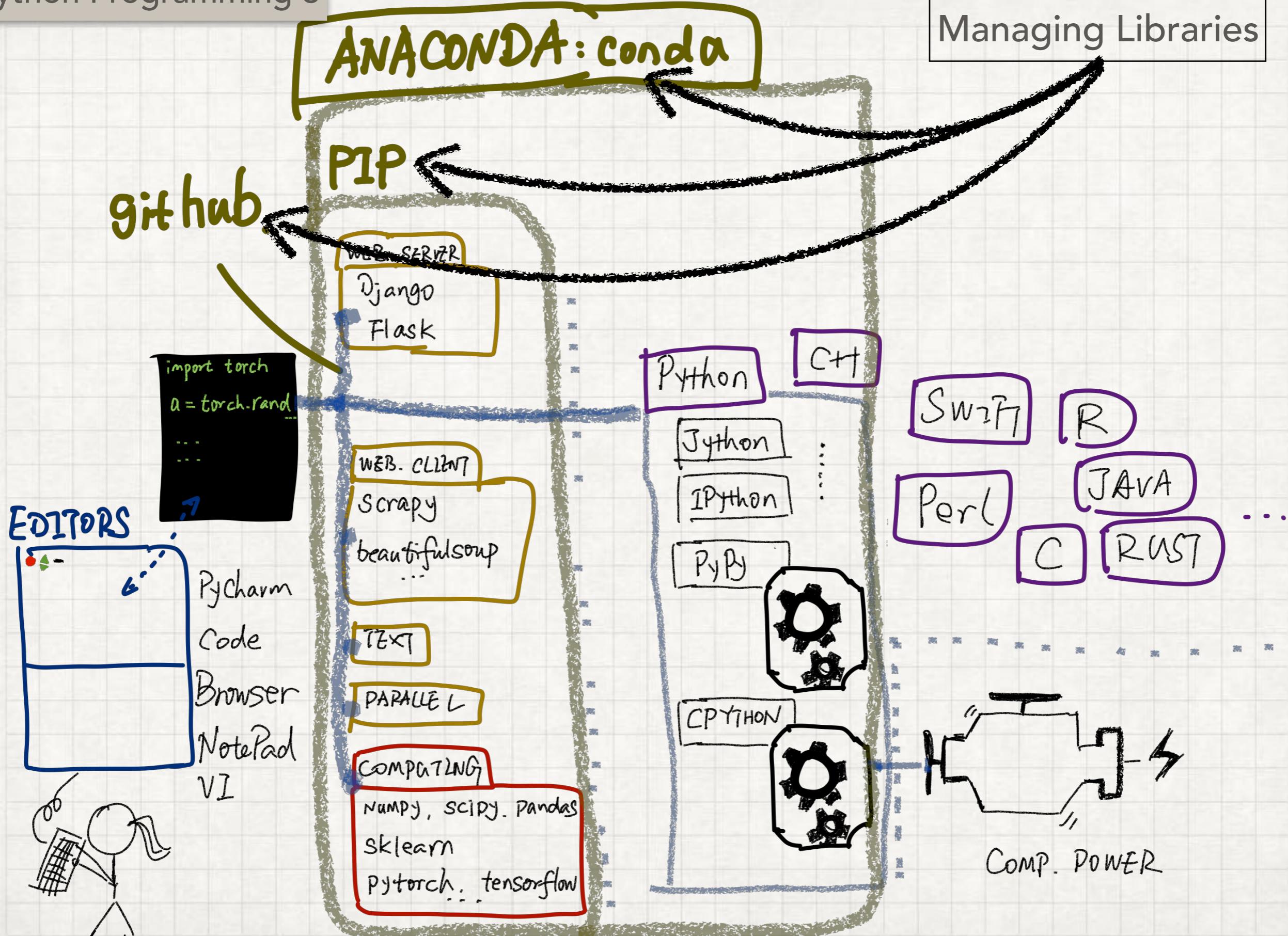
# PYTHON PROGRAMMING

## ANACONDA: conda



An overall Picture

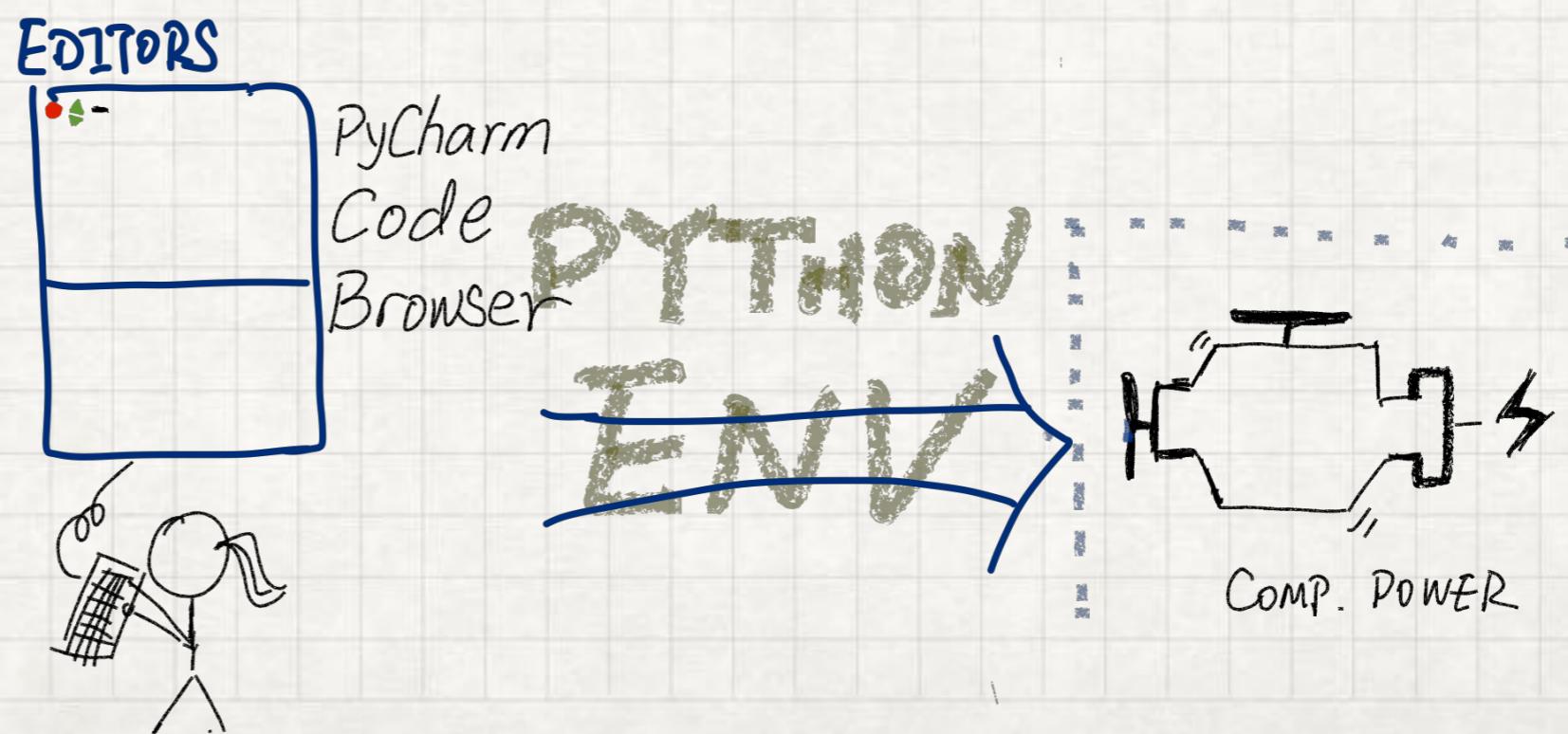




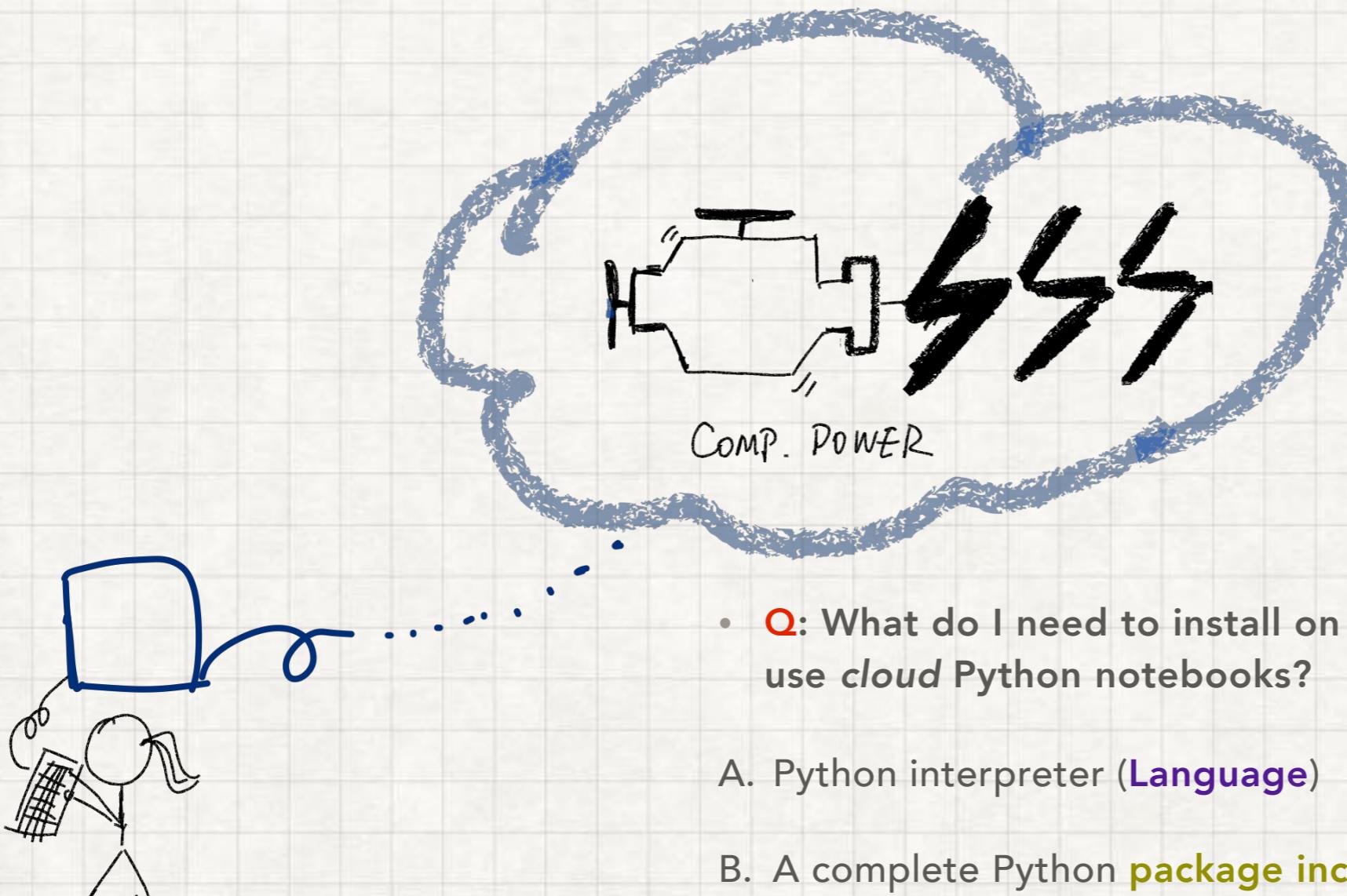
- Anaconda: package, wholesale solution
    - conda install
  - PIP: standard installer
    - pip install
  - source-code: 3rd party library
    - git clone
    - make
    - make install / python setup.py
- **Q:** If I need a library `xyz`, and I tried different methods, all of them *completed without any error*. However, in which of the following cases I am *least* sure that my library is working?
- A. `import xyz` completes without errors.
  - B. `conda install xyz` ...
  - C. `pip install xyz` ...
  - D. `git clone <https://github.com/xyzauthor/xyz.git>; cd xyz; make; make install`

- Use Anaconda when possible.
- Then pip.
- Check [here](#). (ignore the `pyenv` and version 2.7 part)

- Editors choice is important but subjective
  - MS Code is simple
  - PyCharm is also good
  - Notebook is the simplest (I always start developing in a Notebook when possible) — And you need to use it anyway for assignments.

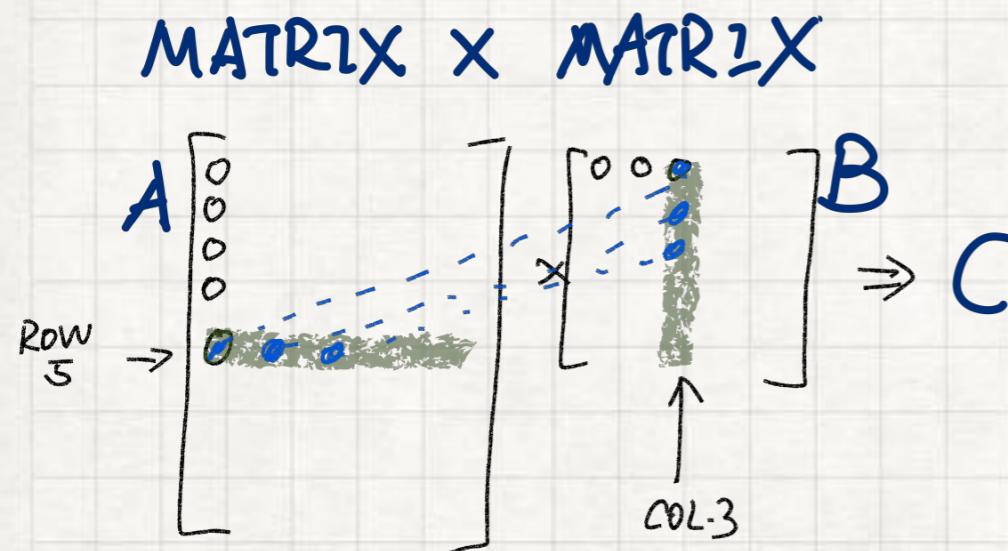


- Browser's advantage: connecting to a powerful machine!



- Q: What do I need to install on my computer to use *cloud* Python notebooks?
- A. Python interpreter (**Language**)
  - B. A complete Python **package including all libraries**
  - C. Nothing.
  - D. Python interpreter and a few packages

- Example: multiply and add numbers.



e.g.

$$C[5,3] = A_{5,1} \cdot B_{1,3} + A_{5,2} \cdot B_{2,3} + \dots + A_{5,k} \cdot B_{k,3}$$

for each row      for each column.

- Q: E.g. Say, A has 5 rows and k=10 columns and B has k=10 rows (must, why?) and 3 columns, computing  $C = A \times B$  requires  $5 \cdot 10 \cdot 3 = 150$  multiplications. Then how many multiplications is needed to perform if A is of (2000, 1000) and B is of (1000, 3000)?

A. 6 billion

B. 6000

C. 6 million

D. 5 million

**THANKS**