Intro to Neural Networks

BA865 – Mohannad Elhamod

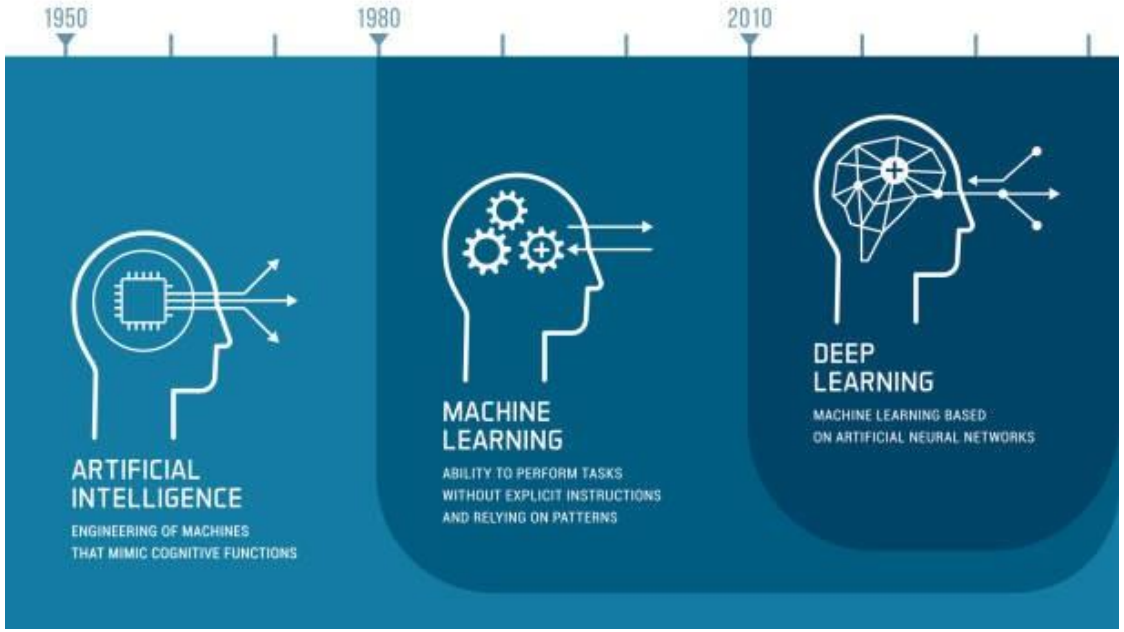
BOSTON UNIVERSITY

Boston University Questrom School of Business

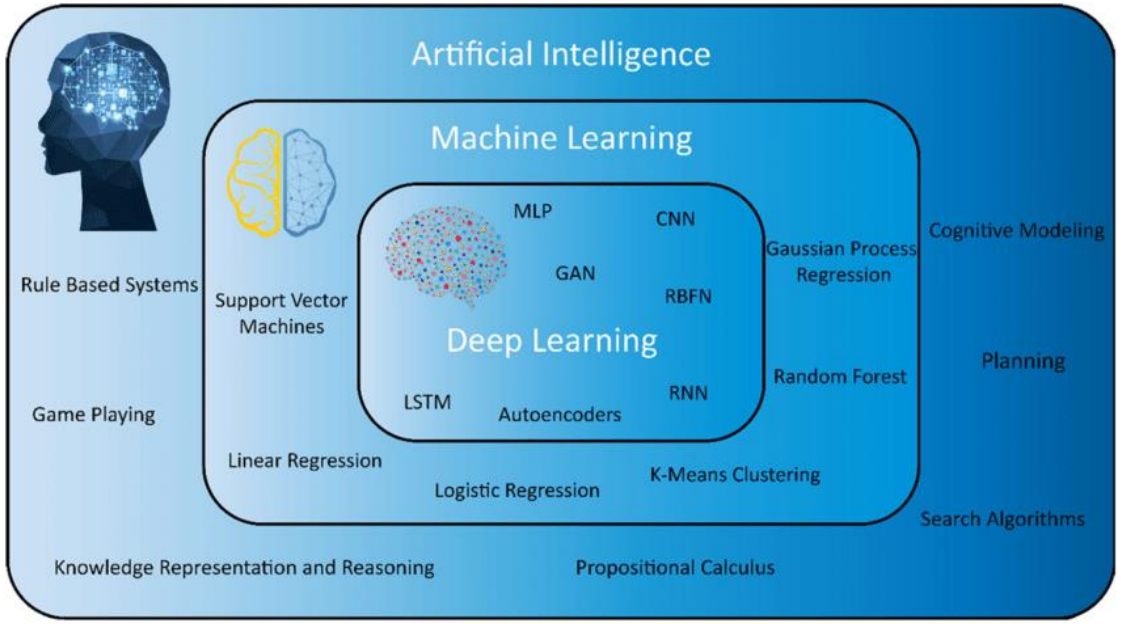
Motivation

BOSTON UNIVERSITY

What is Deep Learning?

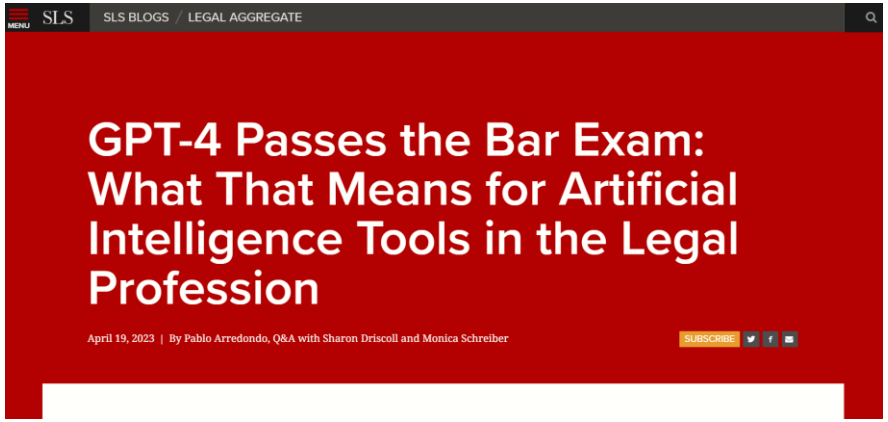


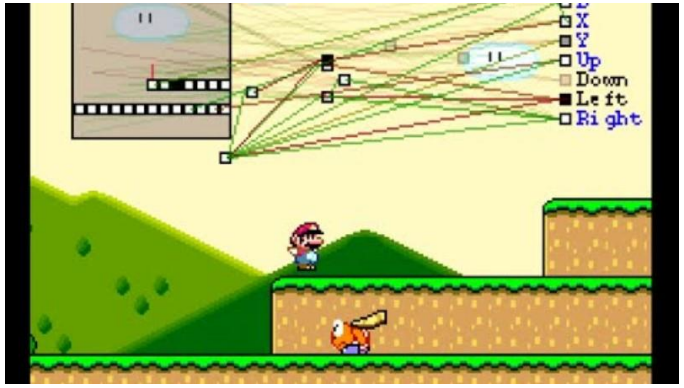
LinkedIn



ResearchGate

Behold The Almighty Deep Learning!



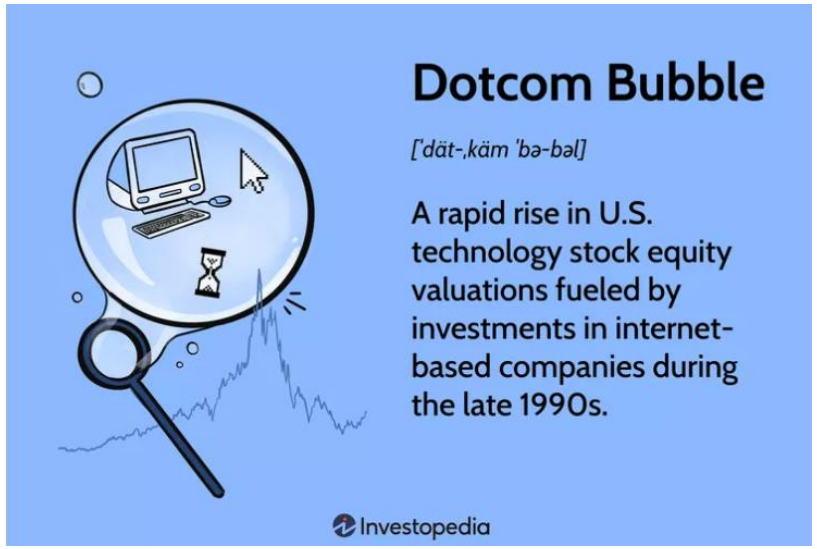


https://thispersondoesnotexist.com

Behold The Almighty Deep Learning!

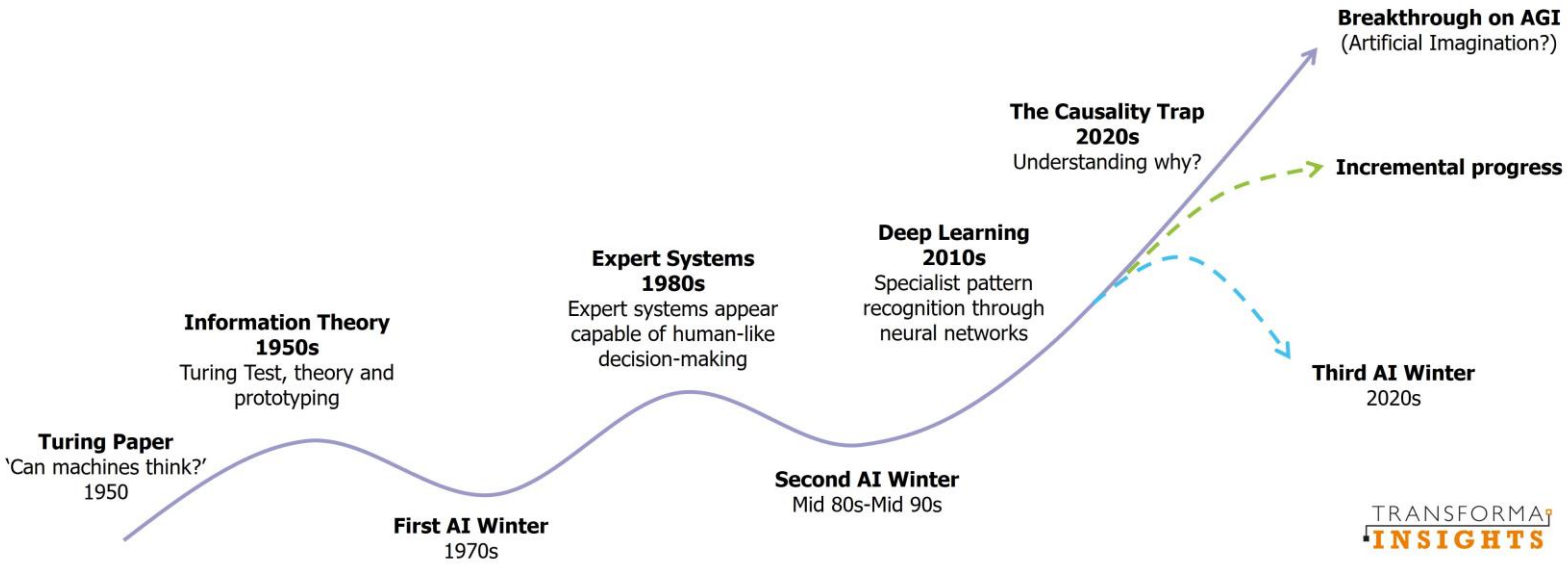
- Generate music
- Classify images
- Classify images (2)
- Dimensionality Reduction

Beware The Hype!



Beware The Hype!

- It hasn’t always looked rosy…There are some "hypes" along the way.



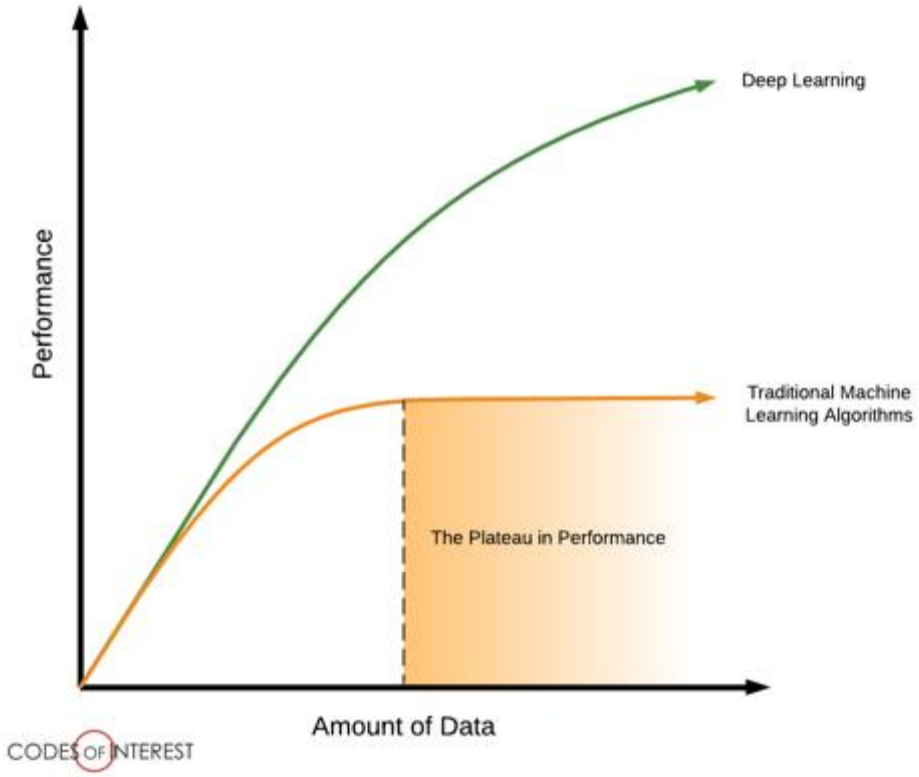
Algorithms Come and Go...

- Different algorithms gain popularity at different periods of time.

medium.com

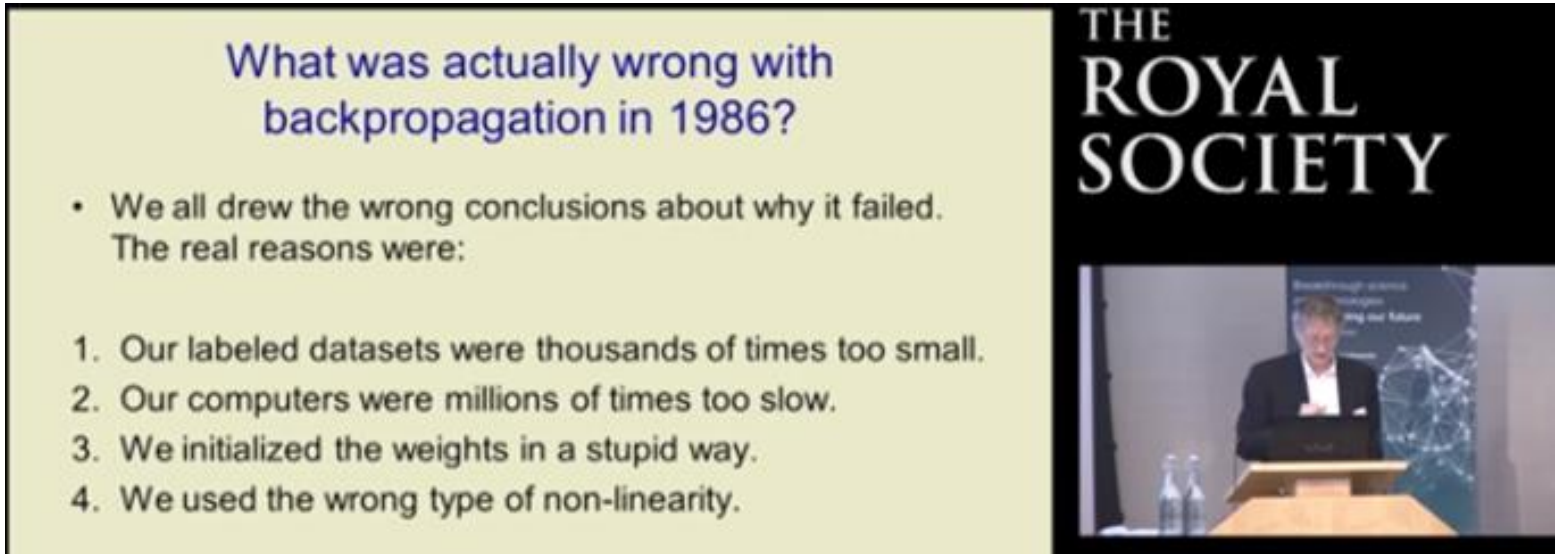
Why Deep Learning?

- Traditional ML was found to plateau in performance in the “Big Data” realm.
- Deep Learning has shown a much better capacity for learning from large amounts of data.



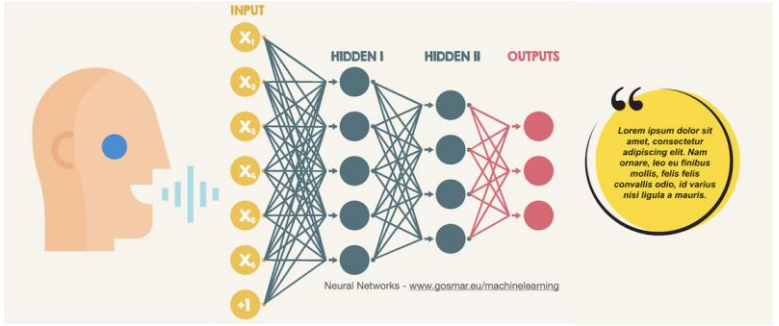
Why Now?

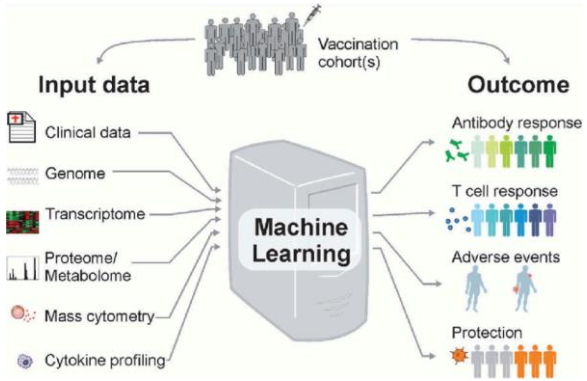
Geoffrey Hinton, Godfather of AI.



Where Can Deep Learning Be Applied?

- Almost anywhere where you have data to learn from and make decisions based on.





Machine Learning:

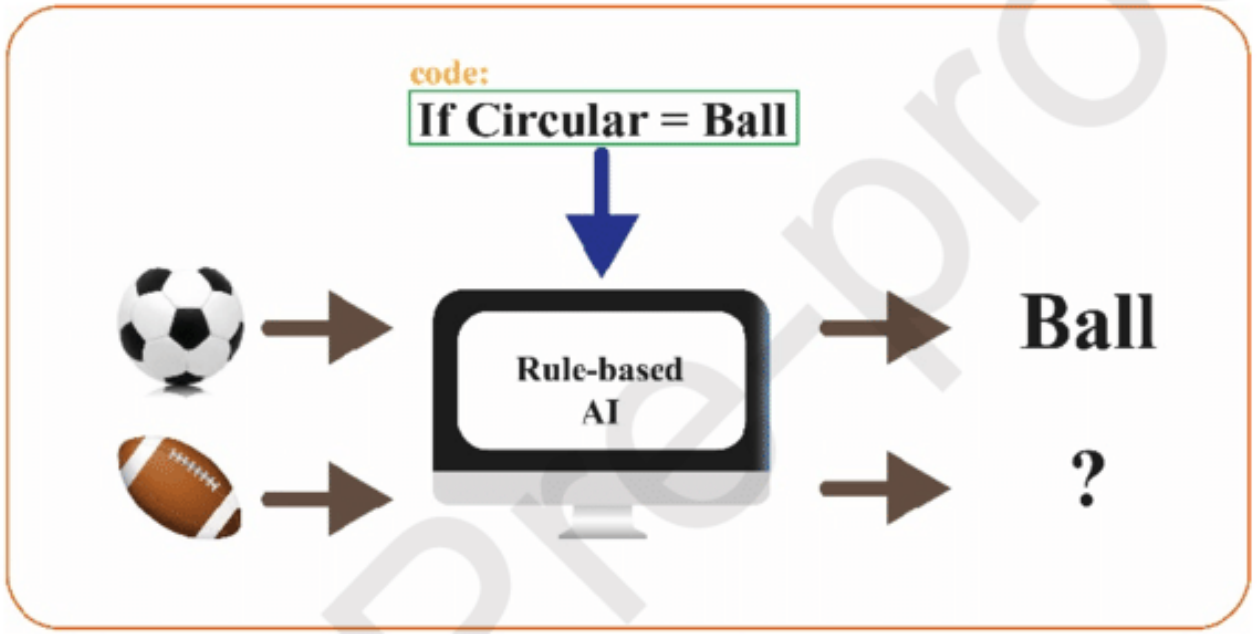
An Essential Review!

BOSTON UNIVERSITY

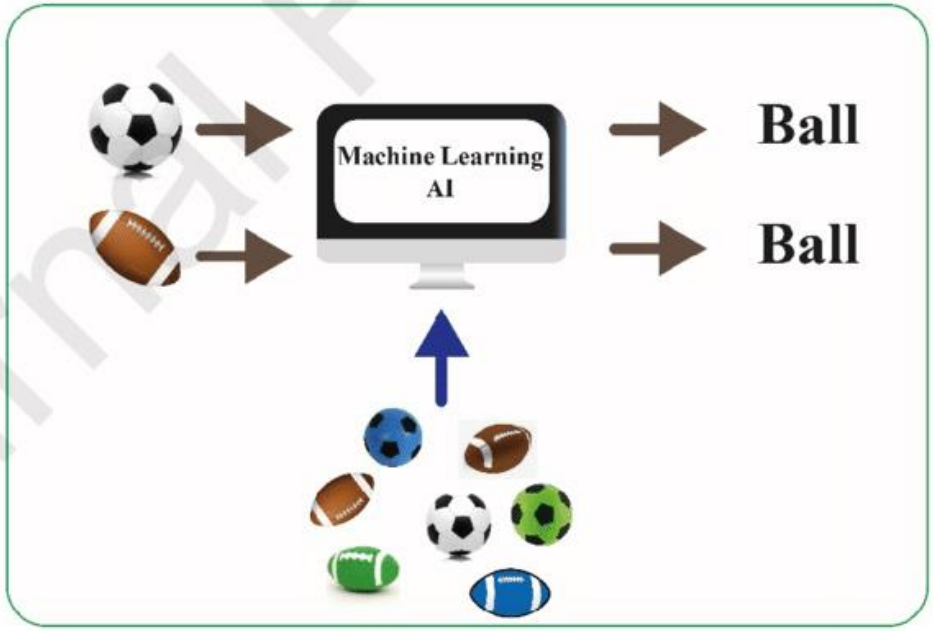
Boston University Questrom School of Business

What is Machine Learning?

Elbadawi, Moe & McCoubrey, Laura & Gavins, Francesca & Ong, Jun Jie & Goyanes, Alvaro & Gaisford, Simon & Basit, Abdul. (2021). Harnessing Artificial Intelligence for the Next Generation of 3D Printed Medicines. Advanced Drug Delivery Reviews. 175. 10.1016/j.addr.2021.05.015.



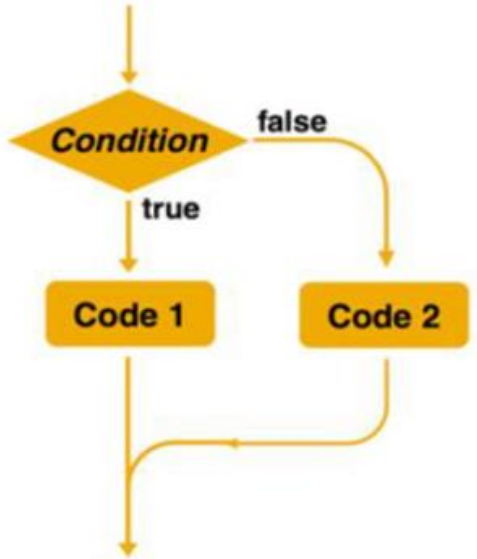
Knowledge-based modeling



Data-driven modeling

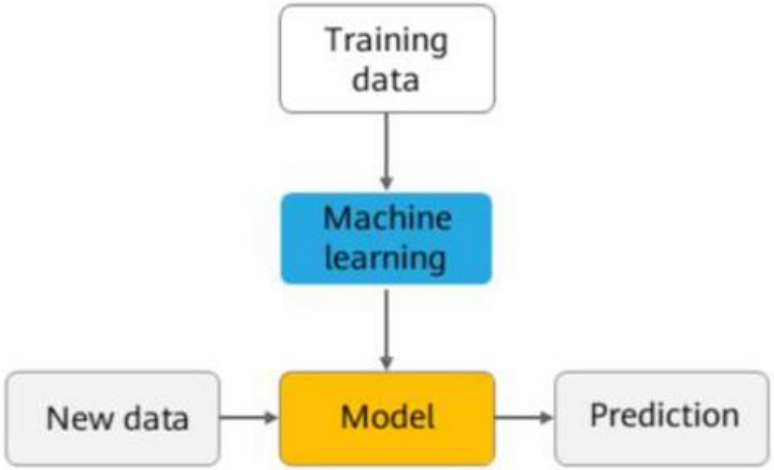
What is Machine Learning?

https://www.wangzze.top



- Explicit programming is used to solve problems.
- Rules can be manually specified.

Knowledge-based modeling

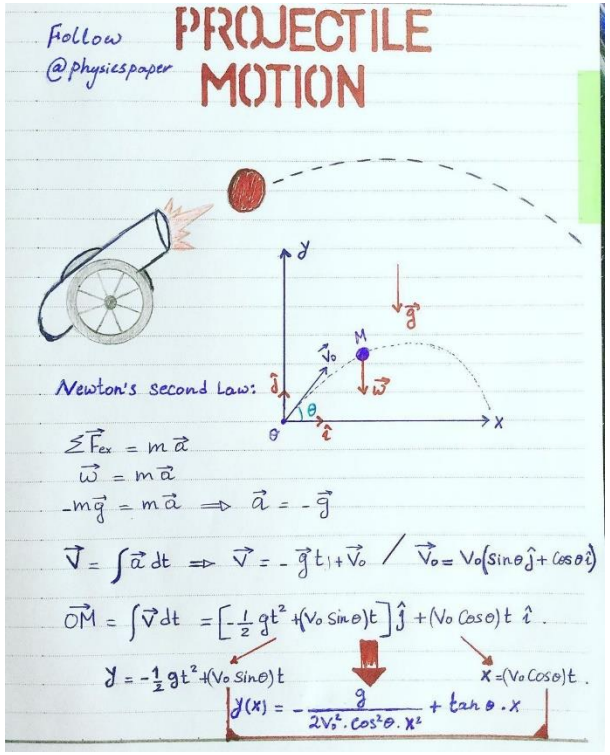


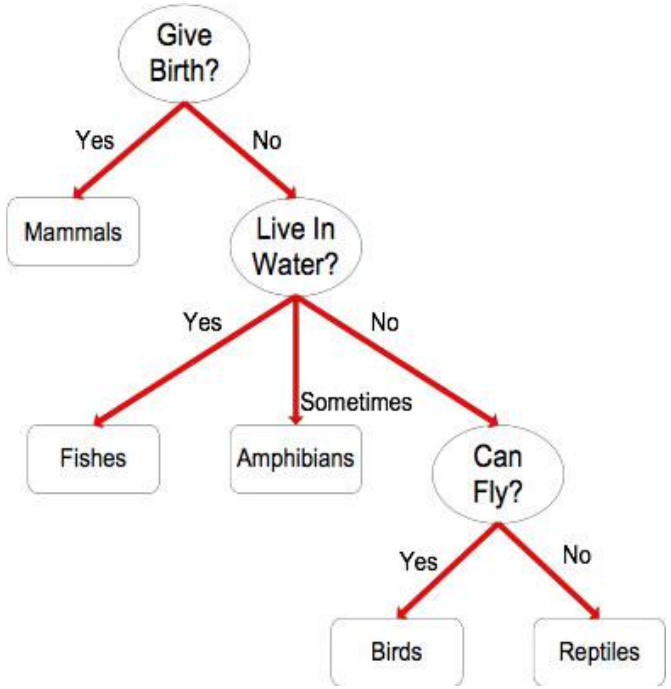
- Samples are used for training.
- The decision-making rules are complex or difficult to describe.
- Rules are automatically learned by machines.

Data-driven modeling

Knowledge-based Modeling

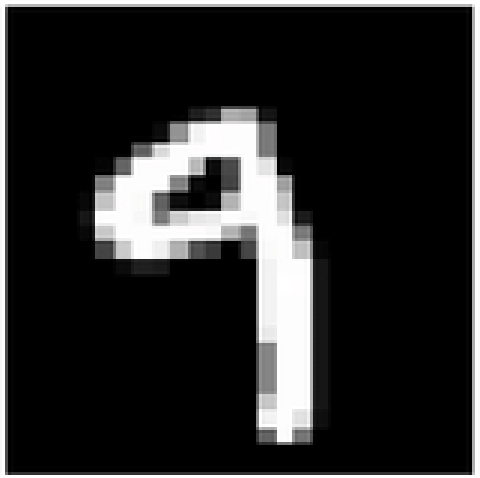
- What challenges you find with this approach?



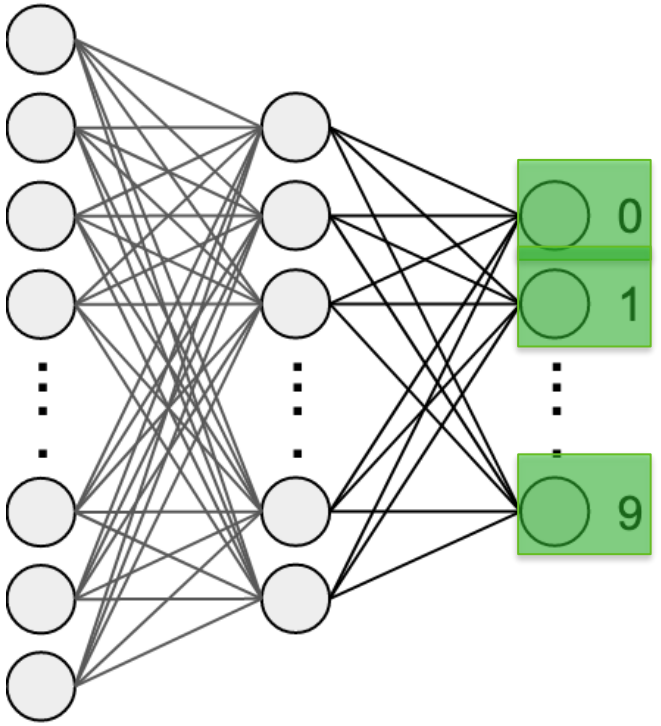


http://jcsites.juniata.edu/faculty/rhodes/ml/rulebasedClass.htm
https://www.pinterest.com/pin/598626975454754006/

Supervised Learning

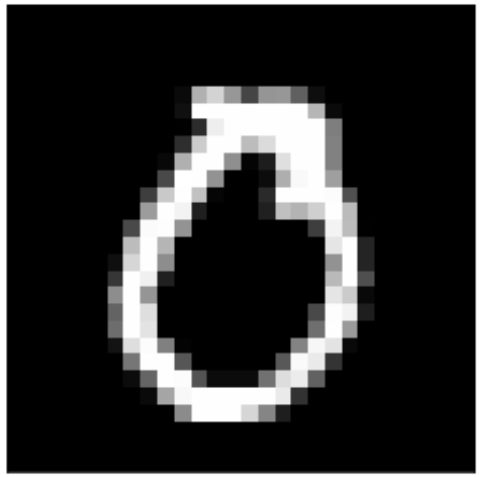


MNIST Dataset

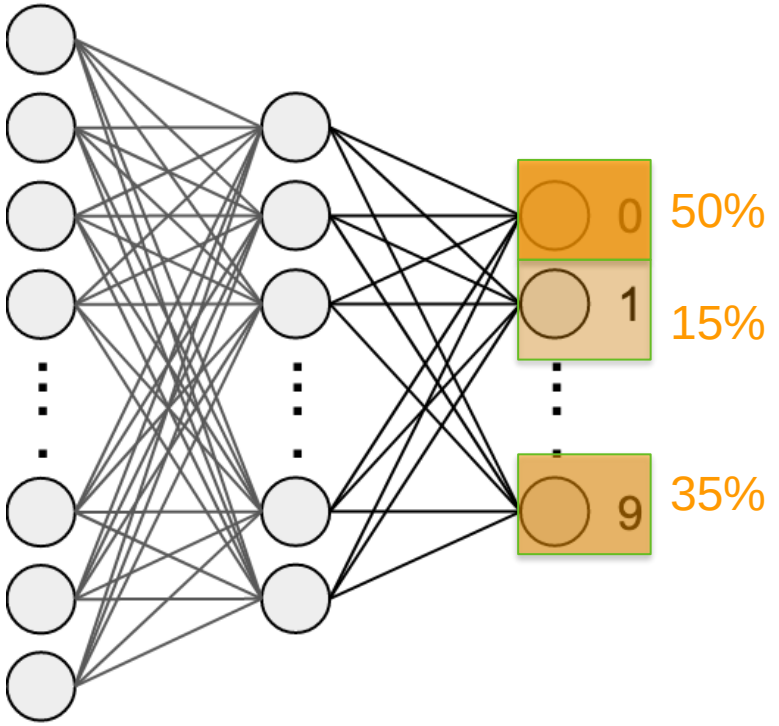


Phase 1: Training

Supervised Learning

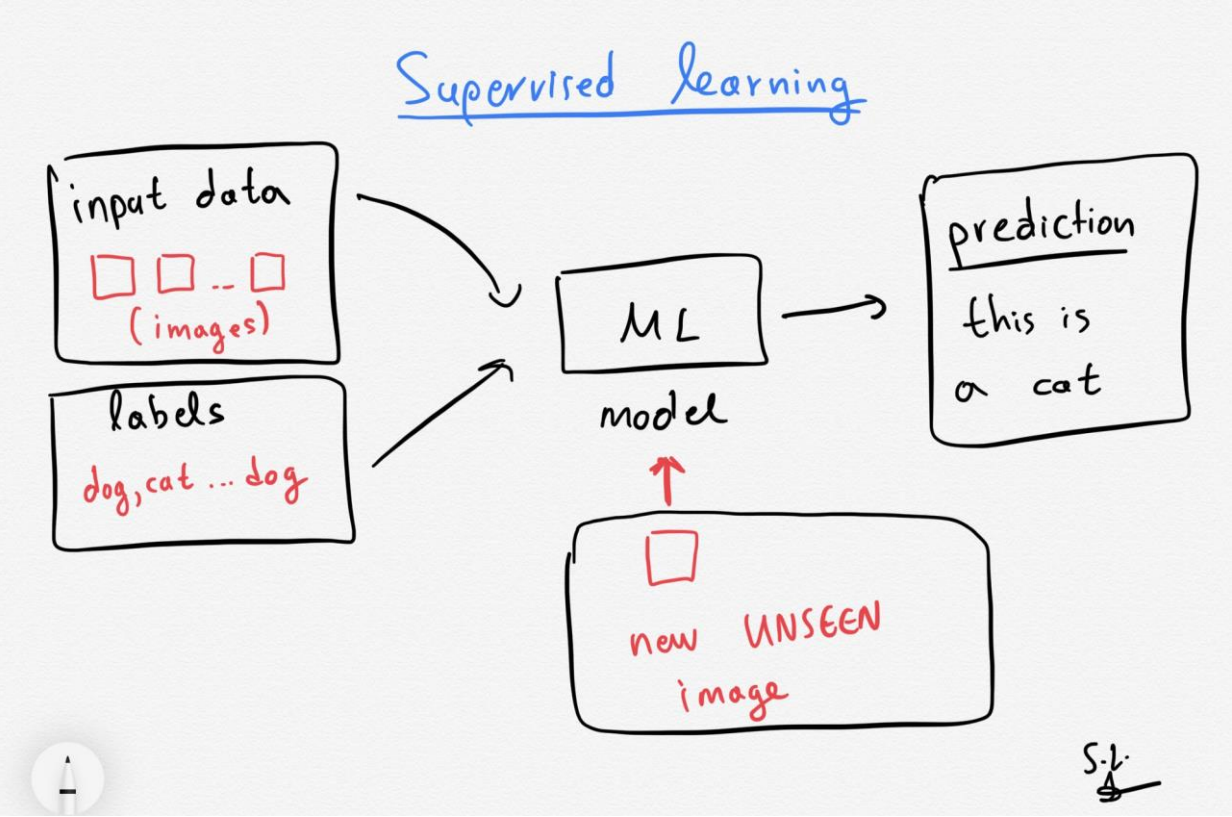


MNIST Dataset



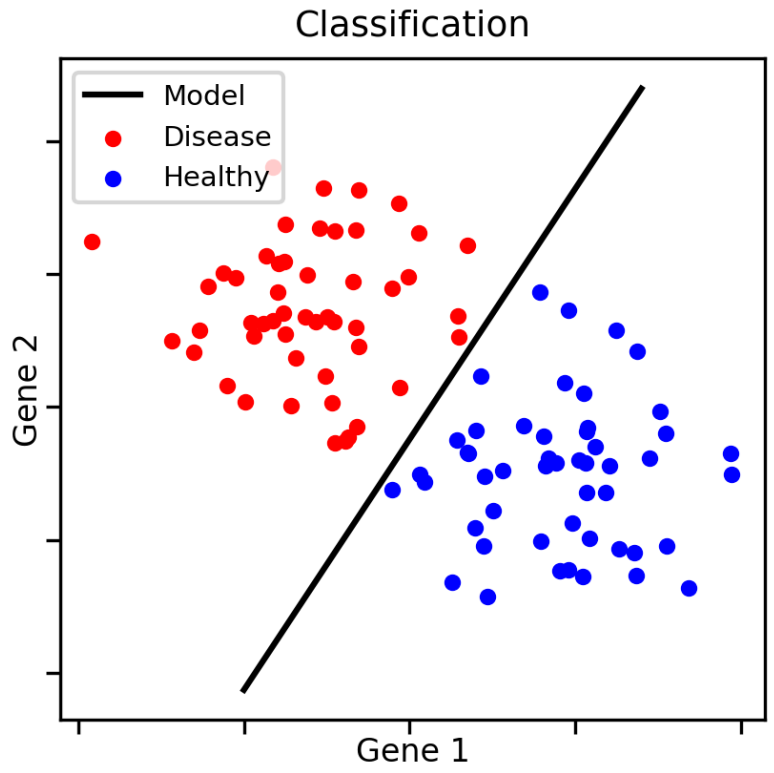
Phase 2: Validation/Testing

Supervised Learning



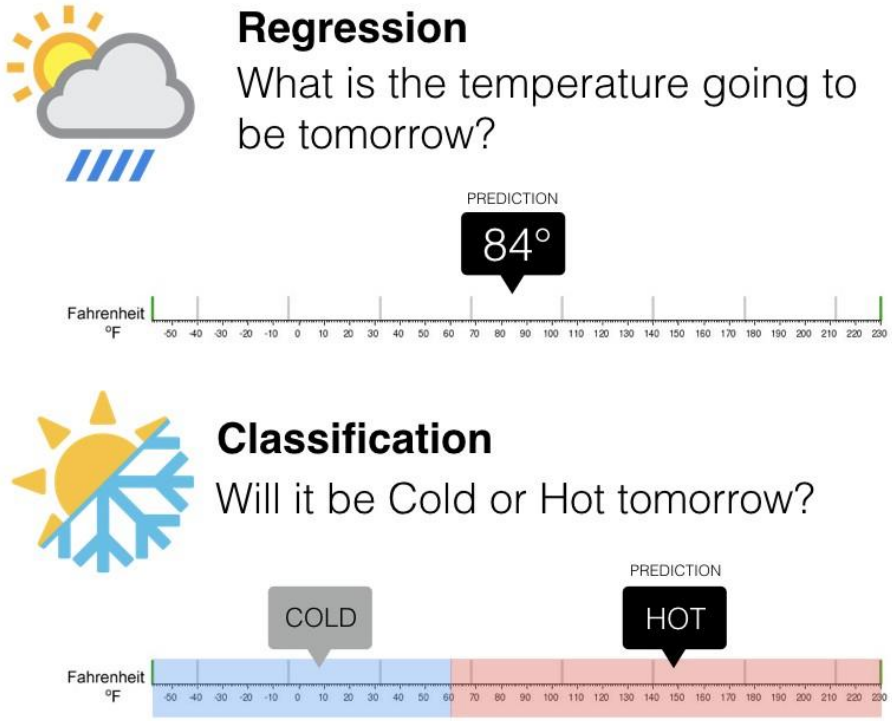
https://towardsdatascience.com/what-is-machine-learning-a-short-note-on-supervised-unsupervised-semi-supervised-and-aed1573ae9bb

Supervised Learning - Model Types



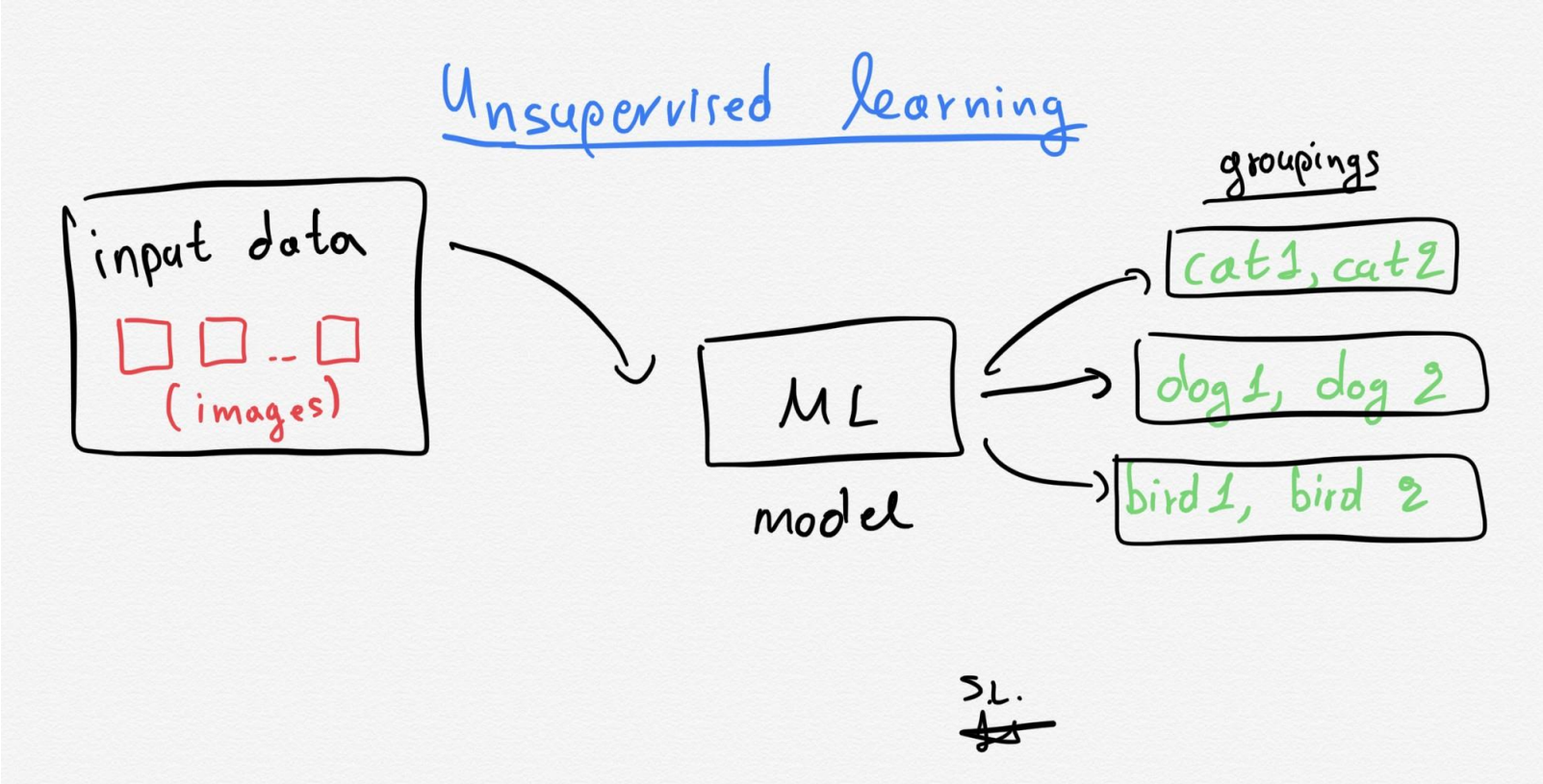
https://aldro61.github.io/microbiome-summer-school-2017/sections/basics/

Supervised Learning - Model Types



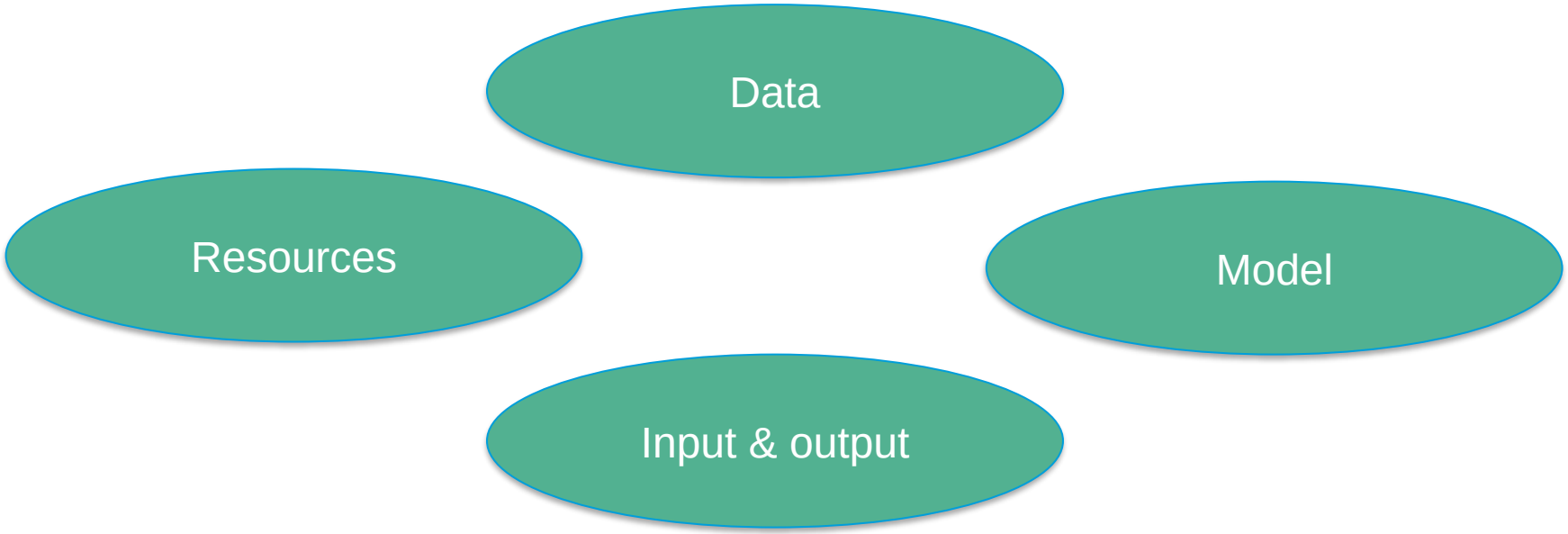
https://medium.com/@ali_88273/regression-vs-classification-87c224350d69

Unsupervised Learning



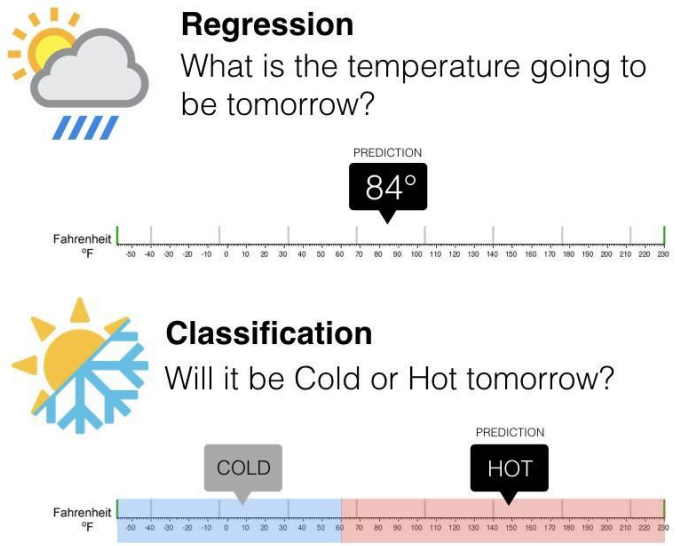
https://towardsdatascience.com/what-is-machine-learning-a-short-note-on-supervised-unsupervised-semi-supervised-and-aed1573ae9bb

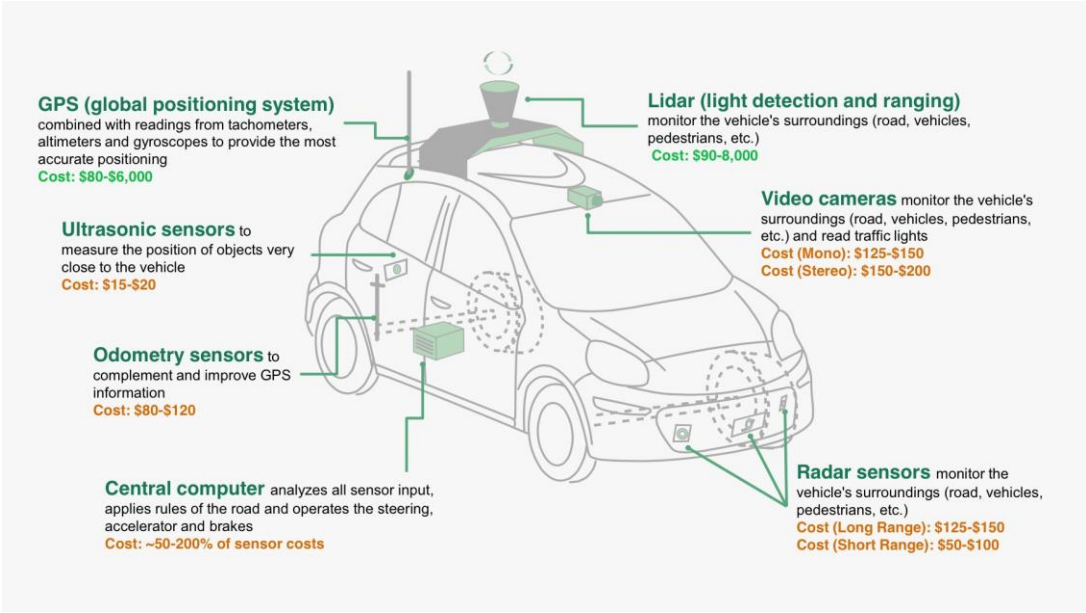
The Pillars of ML



Input & Output

- What exactly are you trying to model?





https://www.wired.com/2015/04/cost-of-sensors-autonomous-cars
https://medium.com/@ali_88273/regression-vs-classification-87c224350d69

Data

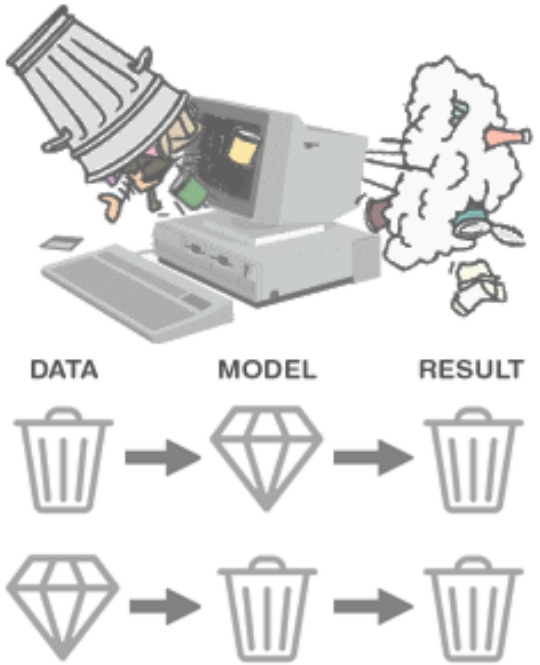
- Is there enough of it?
- Does it need clean-up?

Damaged specimen

Missing Features

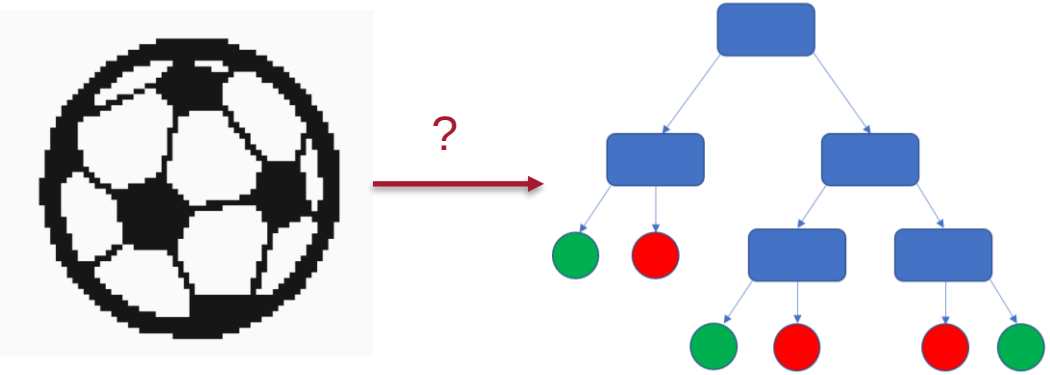
Occluded Features

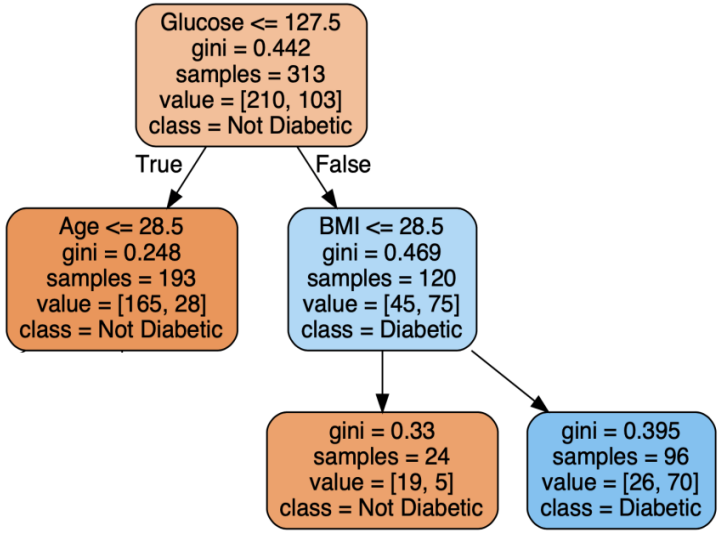
Elhamod, M., Diamond, K. M., Murat Maga, A., Bakis, Y., Bart, H. L., Mabee, P., Dahdul, W., Leipzig, J., Greenberg, J., Avants, B., & Karpatne, A. (2022). Hierarchy-guided neural network for species classification. *Methods in Ecology and Evolution*, 13, 642–652. https://doi.org/10.1111/2041-210X.13768



Model

- What kind of model would be sufficient/suitable for modeling your data?

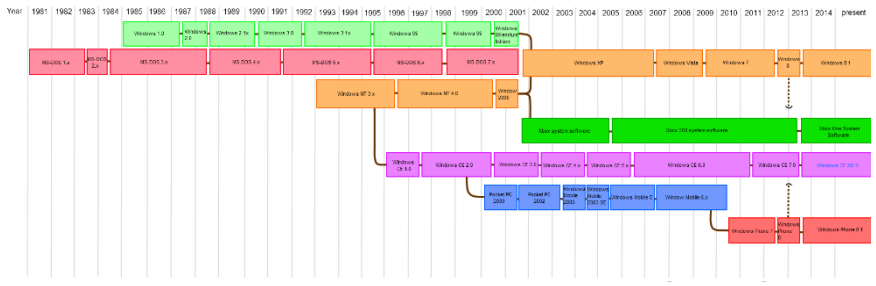




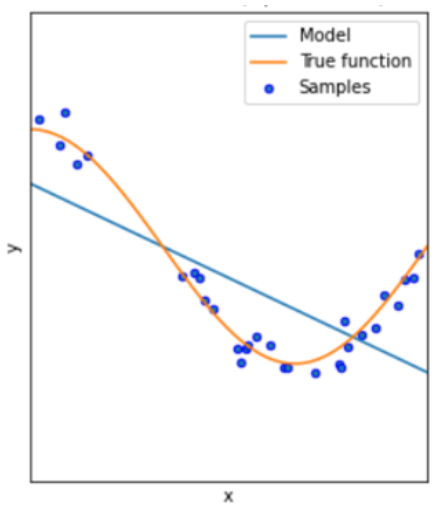
https://statisticallyrelevant.com/decision-trees-in-python-predicting-diabetes/

Resources

- What compute resources are available?
- What is the project's timeline?

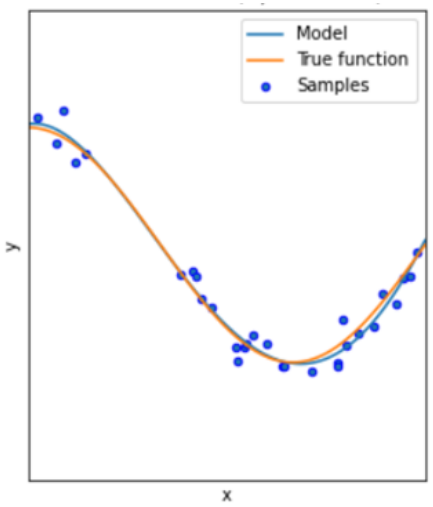


Model Generalization



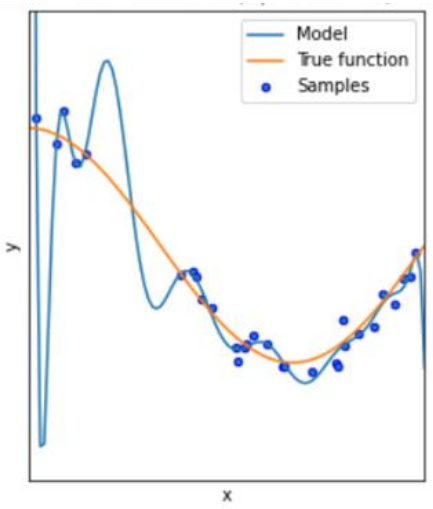
Underfitting

$\theta_0 + \theta_1 x$



Well-fitted

$\theta_0 + \theta_1 x + \theta_2 x^2$



Overfitting

$\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$

https://datascience.foundation/sciencewhitepaper/underfitting-and-overfitting-in-machine-learning

What is Model Complexity?

- It is the degrees of freedom, the “flexibility” of the model to fit the data.
- Bigger datasets or datasets with more attributes need more complex models, and vice versa.

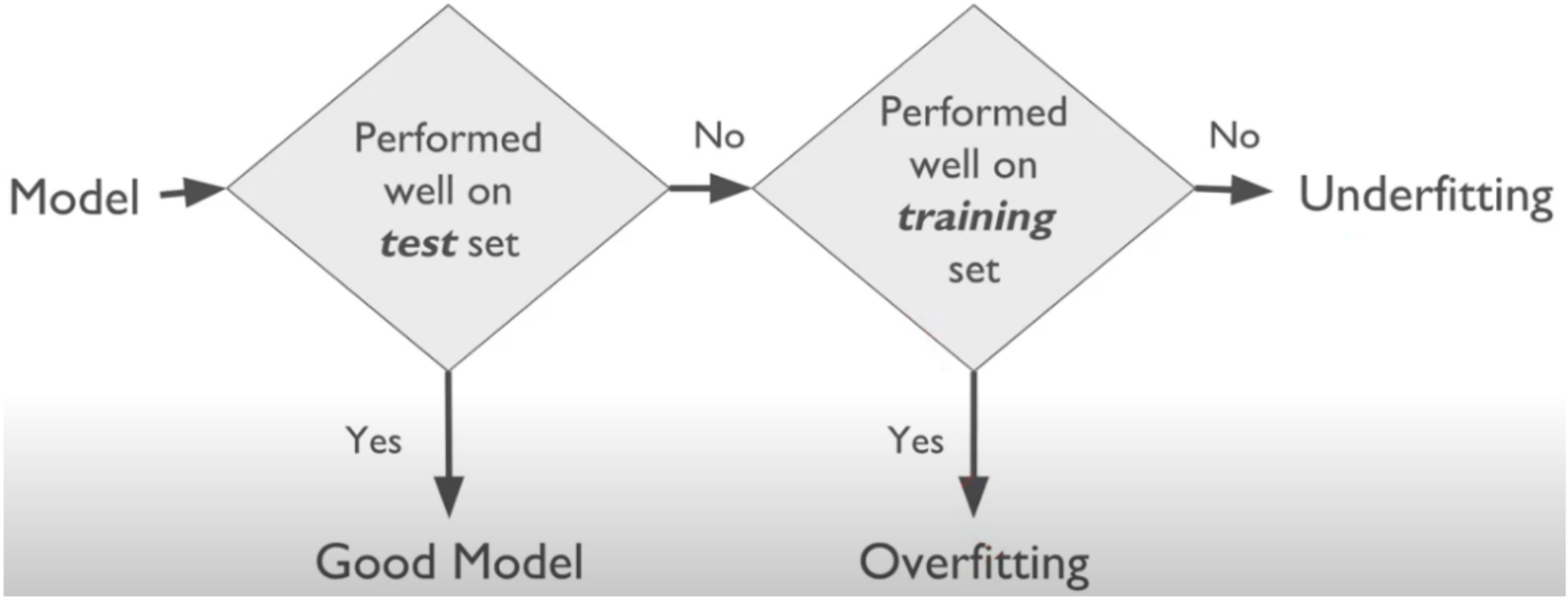
- Try it.

Demo

Put Your Own Subtitle In This Box

BOSTON UNIVERSITY

Are We Overfitting or UnderFitting?

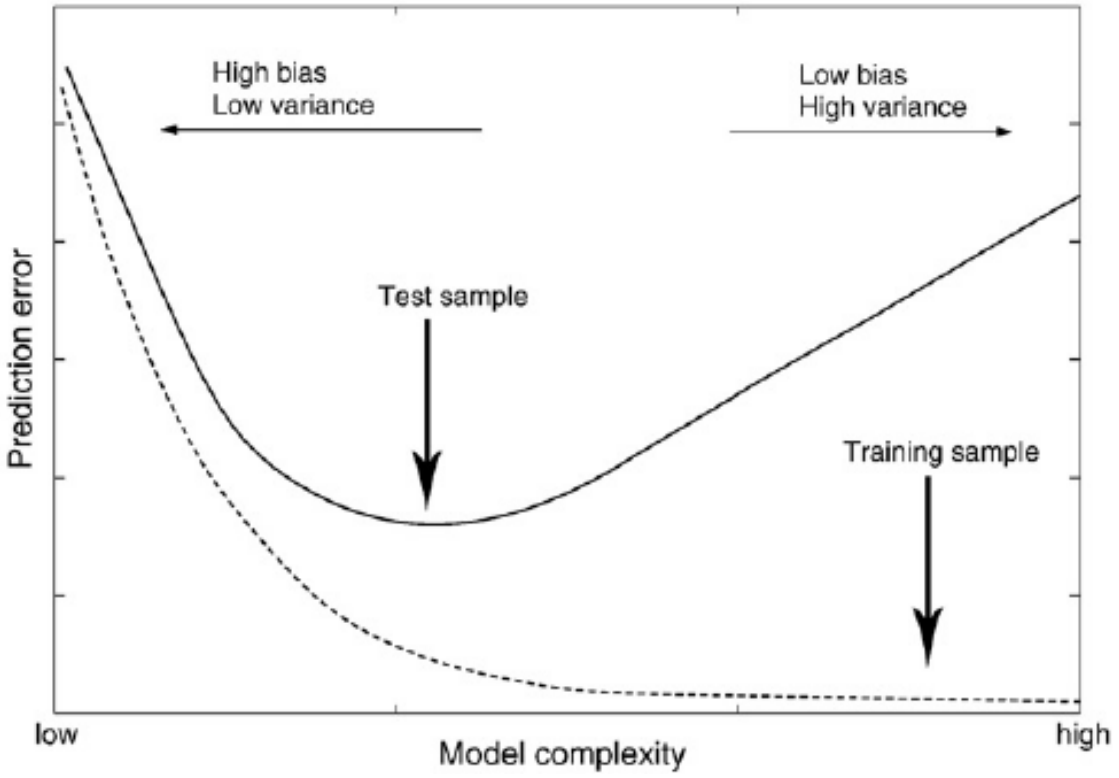


So, What If We Overfit?!

- Without validation/testing, you could overfit
  - In other words, you could learn the wrong patterns in the data!

https://medium.com/trusted-ai/explaining-ai-model-behaviour-with-ibm-watson-openscale-86515702c177

Training vs. Test Errors



www.researchgate.net

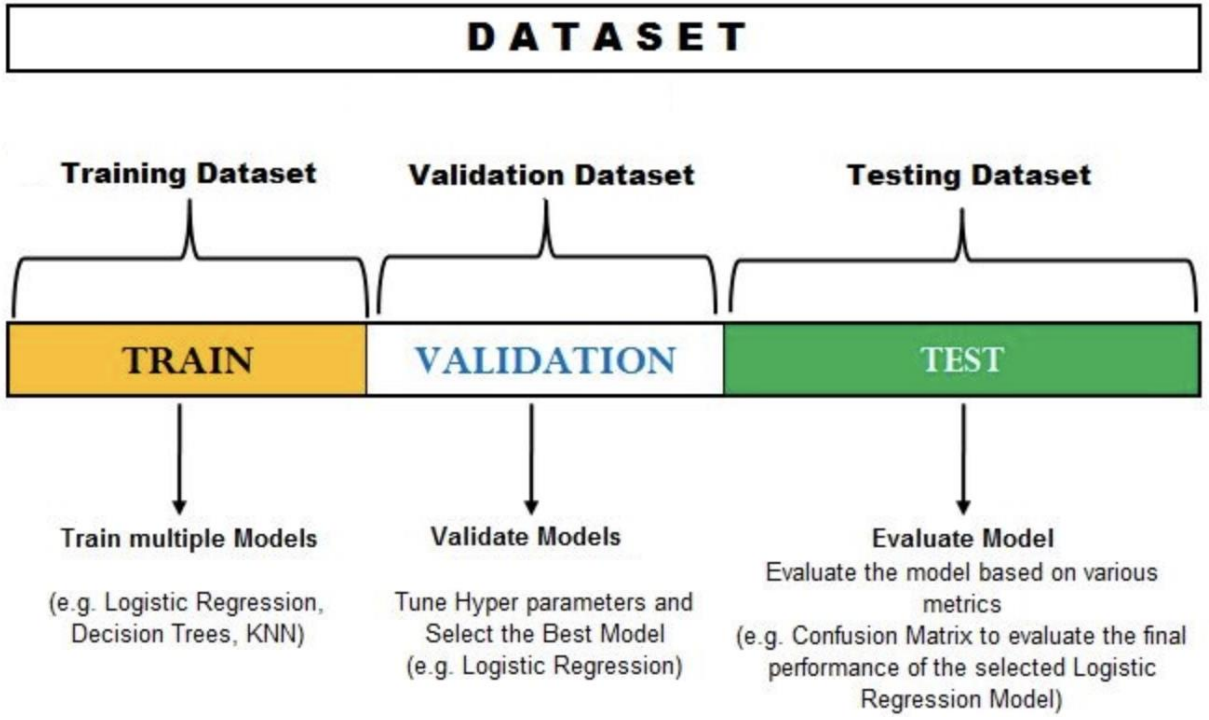
Training vs Test vs. ….?

- How do we test the model's performance?
- How do we assess the model's fine-tuning?

https://www.machinecurve.com/index.php/2020/11/16/how-to-easily-create-a-train-test-split-for-your-machine-learning-model/

Training vs Test vs. Validation?

- The test set should only be used once our model is ready, it should not be used to “tune” the model.
- Hence, we need a third “validation” dataset, which is a split of the training set.



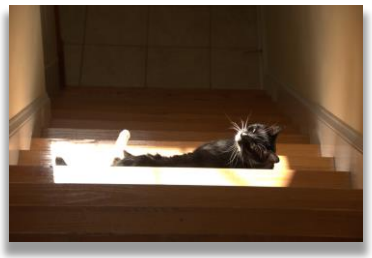
https://vitalflux.com/hold-out-method-for-training-machine-learning-model/

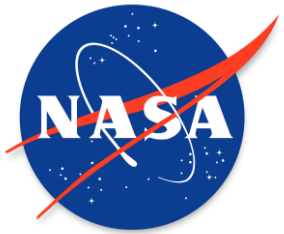
Logistics

Put Your Own Subtitle In This Box

BOSTON UNIVERSITY

Who is your professor?





Why are you here?

- I am sure it eventually leads to money….
- Not a required class. So, hopefully based on interest!
- Nonetheless, understanding how things work is essential for decision making and innovation.

> Someone who had begun to [study] geometry asked Euclid, 'What shall I get by learning these things?' Euclid called his slave and said, 'Give him [some money], since he must make gain out of what he learns'.
>
> (Heath, 1981, loc. 8625)

Euclid

Fundamentals are important!

- The more foundational knowledge you skip, the more fundamental errors you will make.
- Work hard. **Be patient!**

Zen Speaks: Shouts of Nothingness

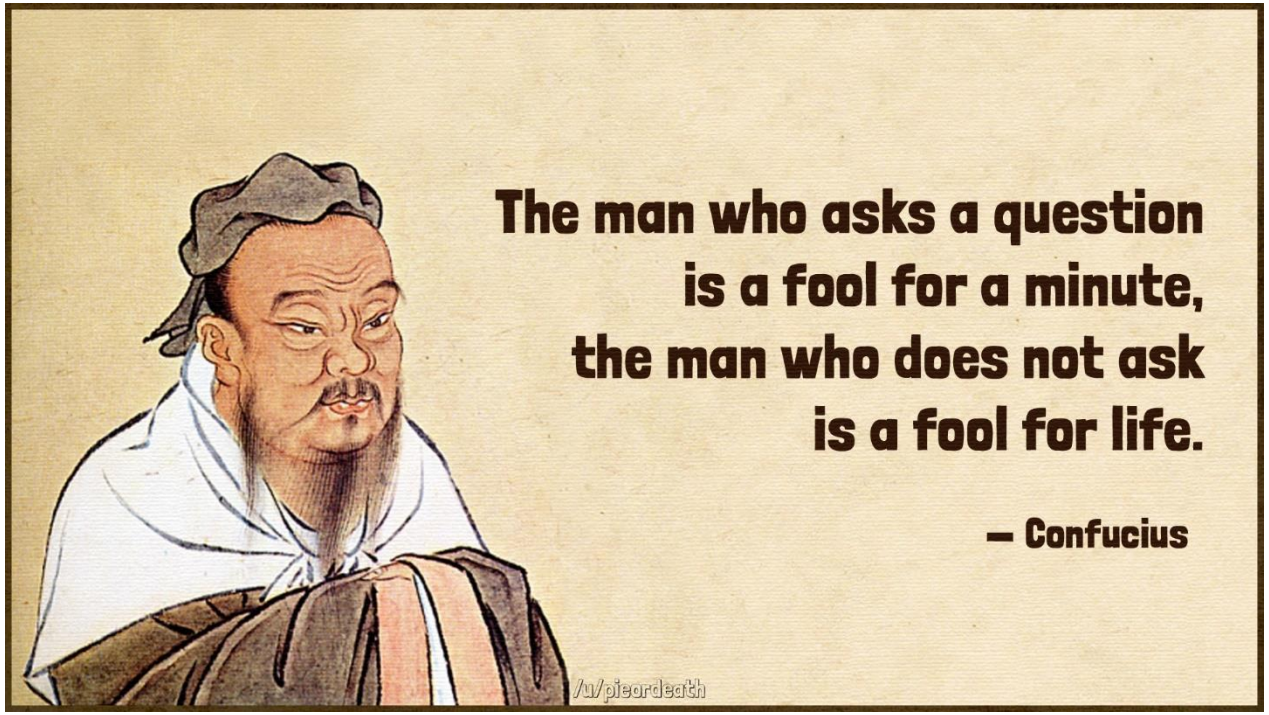
Your professor is not a God

- Deep Learning is a fast-growing domain.
- The internet nowadays has all kinds of learning material.

- Your professor is **NOT** here as a walking encyclopedia. He is here to guide your learning experience and build you a solid foundation, so you could continue learning on your own later.

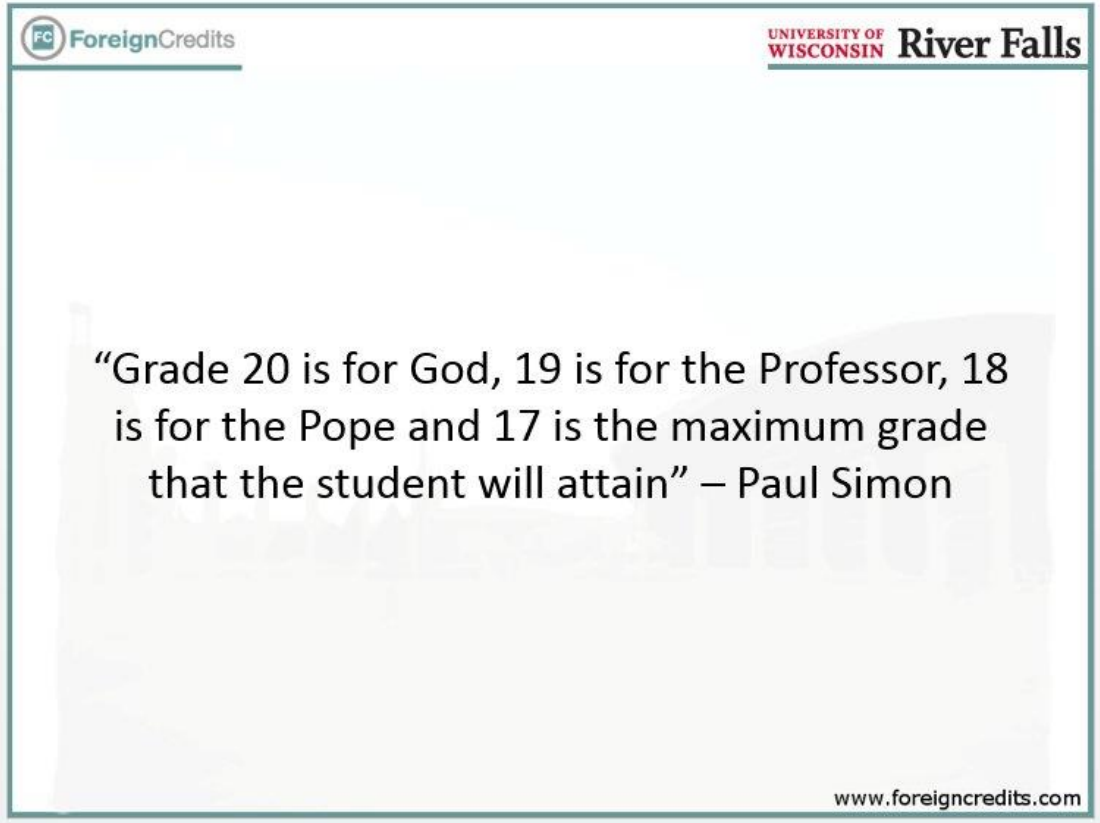
No Question is Foolish

- Other might have the same question.
- Knowledge is hierarchical.



Can I get an A?

- Questrom policy
- But you will get your fair chance.



Participation is Essential!

- 7.5% of your grade!
- The instructor reserves the right to cold-call.

Office Hours

- They are for you to take advantage of!
- However, to make the best of your and the TA's/instructor's time, do your homework before dropping in:
  - If you have a question about your project, make sure you have synched with your colleague in advance.
  - If you have question about assignments, make sure you have done your best and that your question is specific rather than "How do I solve this?".

Syllabus

- Here
  - Pay attention to Piazza!
  - Attendance and participation.
  - The class is somewhat more technical than BA820 and BA810.
  - Zero tolerance for cheating (AI or peers).
  - 5%-20% Penalties for deviating from assessment instructions!

Shared Computing Cluster (SCC)

- BU's computing cluster with access to CPUs and GPUs.
- We have a limited number of compute hours. Use your allocations wisely!
- When facing issues, ask the professor or TA for help first, not IT support!
- **Don't leave things till last minute!**

BU TechWeb

Coding Environments

- https://www.simplilearn.com/keras-vs-tensorflow-vs-pytorch-article
  - Keras is easier to use for non-CS users. It is good at an introductory level.
  - PyTorch is most commonly used for research as it provides powerful low-level API.
  - We will also use some more advanced packages that help us diagnose and use more advanced models.