Final Remarks

PHYS591000 2022.05.18

Outline

- Topics that are important yet we don’t get to discuss:
    - Reinforcement Learning
    - Anomaly detection
    - Explainable AI

  Ref: Prof. Hung-Yi Lee (NTU) lectures (2020)

- Final project presentation grading policy
- Attendance requirement for the guest lecture next week

Supervised vs Unsupervised Learning

Supervised Learning

Goal: Learn function to map
x (data) to y (label)

Examples:

- Classification
- Regression
- Image captioning
- …

Unsupervised Learning

Goal: Learn the hidden or underlying structure of the data without labels.

Examples:

- Clustering
- Dimensionality reduction
- Generative models
- ...

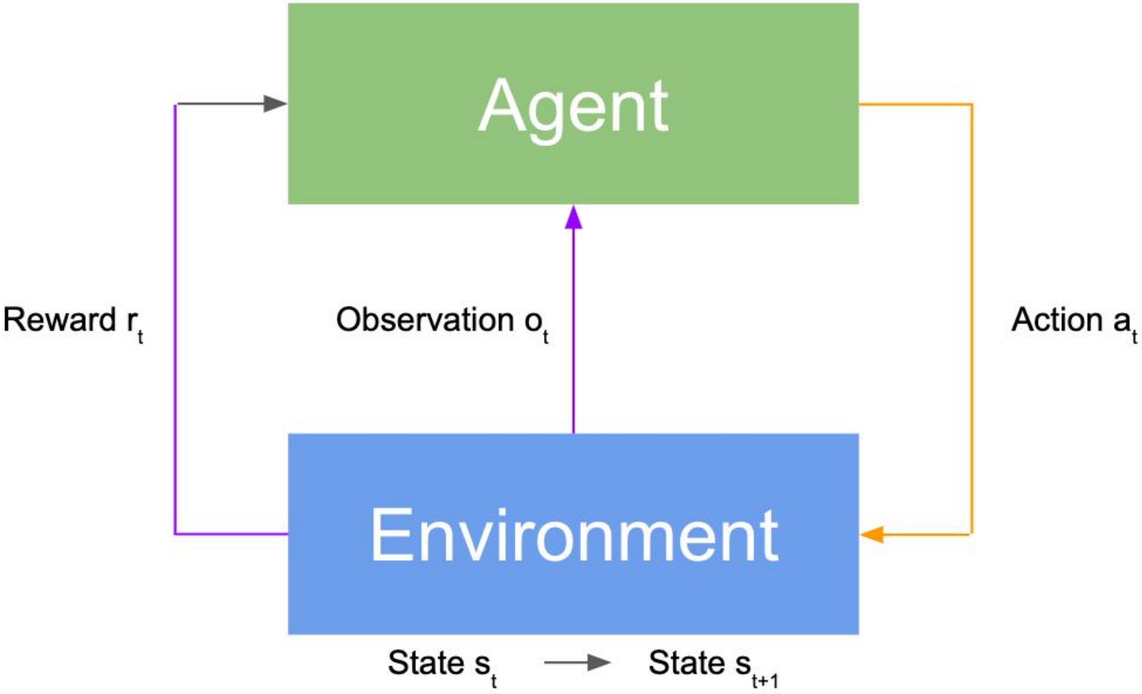
Reinforcement Learning

- Learn from the Environment: An **Agent** (program) interacting with the **Environment**:

    - Agent takes an **action** based on the current **state** (what it sees) of the Environment
    - Environment offers a **reward** (a number) based on the outcome of the action (usually changes the state)
    - Agent takes the next action → Reward → Action → …

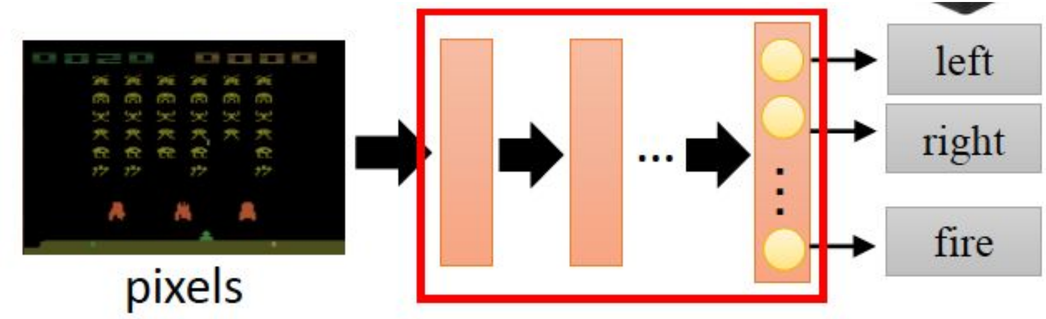
- Goal: Learn to take actions (decision making) to maximize the *total reward* during a process.

Reinforcement Learning



Reinforcement Learning

- Examples:
  - Playing video games: Actions taken based on the current image (pixels) on the screen

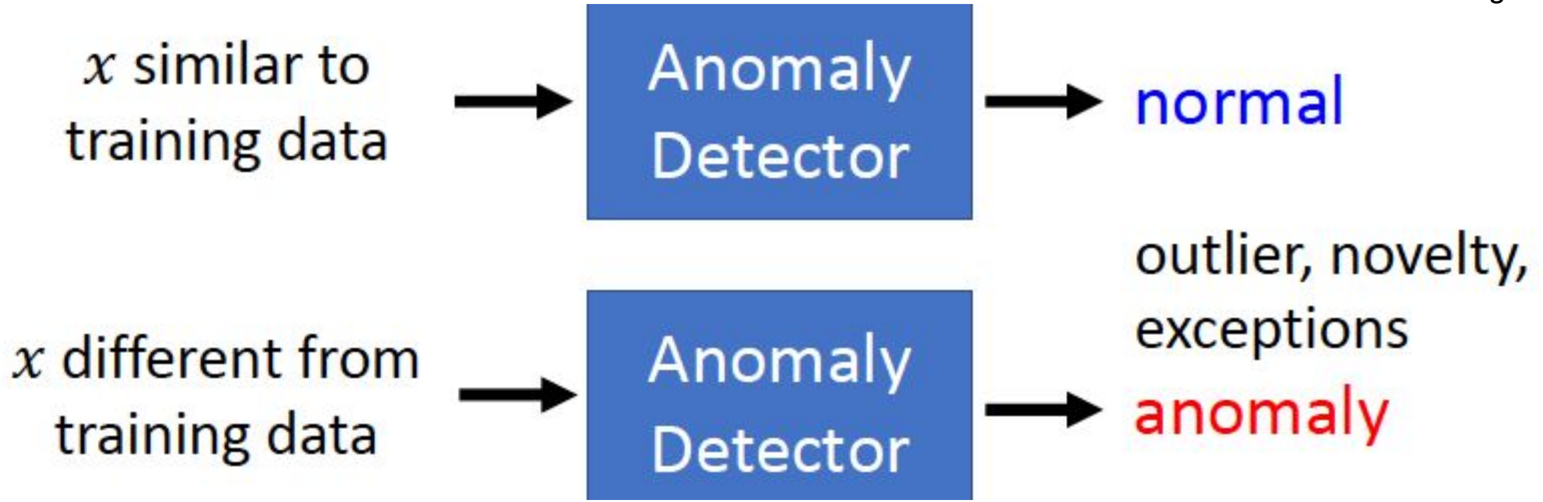


  - AlphaGo: The first computer Go program which had beaten professional human players.

Anomaly Detecton

- Teach the program to know ‘it doesn’t know’: Classify whether the input sample looks like the training data.

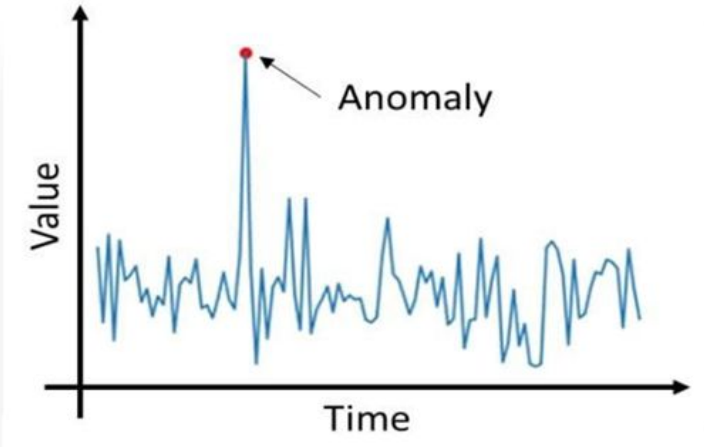
Lecture from Prof. Hung-Yi Lee (NTU)

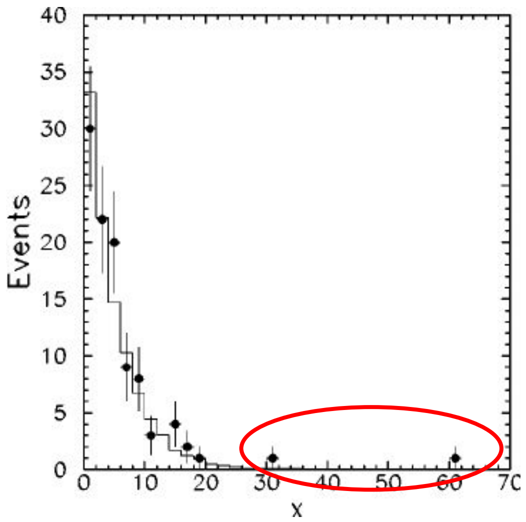


Anomaly Detecton

- Applications:
  - Fraud detection
    (Is this transaction unauthorized?)
  - Network intrusion detection
    (Is someone trying to hack the server?)
  - Cancer detection

- Model-independent way for searching new physics





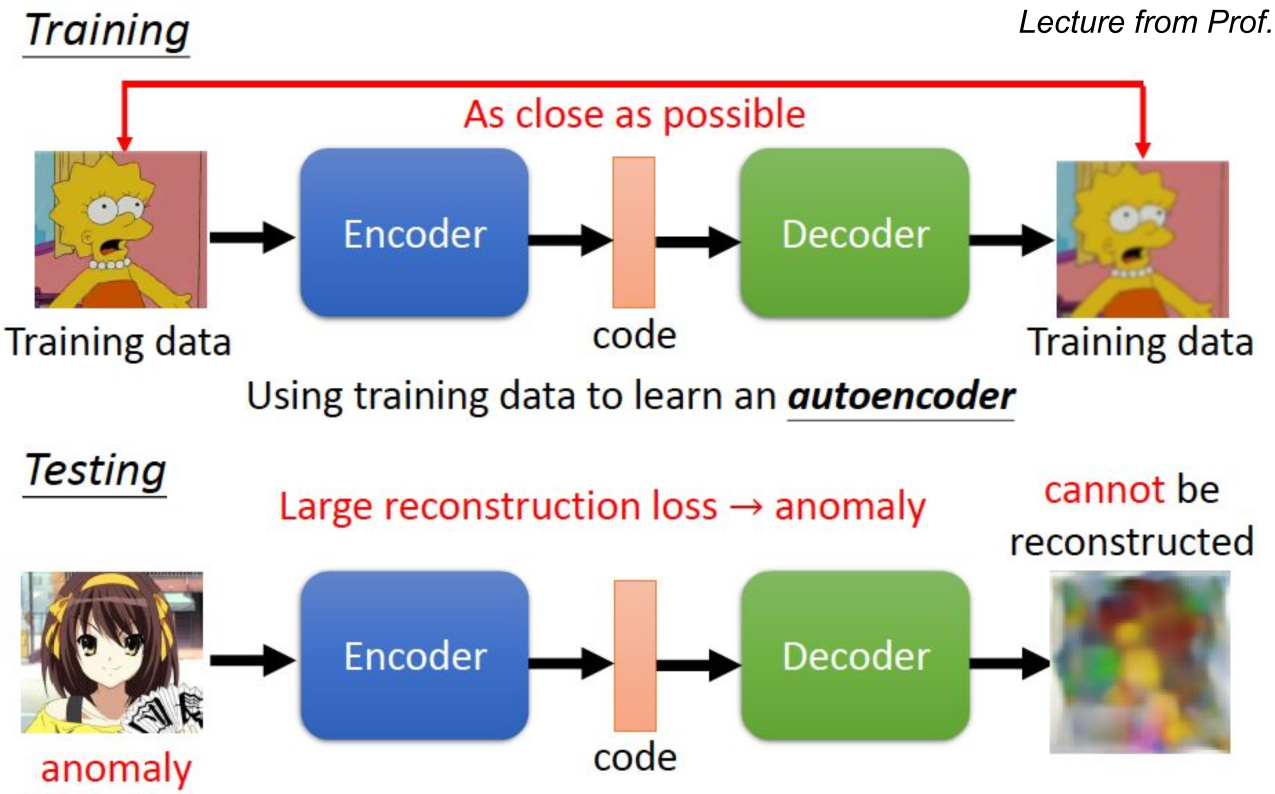
Anomaly Detecton

- Usually there are not much ‘ready-to-use’ anomaly data for training/testing the model.
  → Use Generative models to produce anomaly data!

- Alternatively, one can train an Autoencoder to detect anomaly:
  – Input belong to the same class as training data can be recontstructed (small loss).
  – Anomaly data cannot be reconstructed (large loss).

Anomaly Detecton

Lecture from Prof. Hung-Yi Lee (NTU)



Explainable AI

- AI (machine learning, especially NN) are powerful, yet the models often work like black boxes.
  – We know it works, but don’t know why or how it works.

- Can we explain (to ourselves) exactly how the model makes a decision (output)?
  – It may not be what we think it should be.

For example, a 2017 system tasked with image recognition learned to "cheat" by looking for a copyright tag that happened to be associated with horse pictures, rather than learning how to tell if a horse was actually picture. (Wikipedia)

Explainable AI

- If we can explain how our AI model works, we can improve/optimize it more effectively.
- The AI may in turn provides new insights in explaining the data.
  → Search for new physics.

Final Presentation Grading Policy

Final Project Grading Policy

- Final Project - 30%
  - **Oral presentation: 15%**
  - Final project results: 10% (whether your model and/or fit results make sense)
  - Kaggle competition: 5% (1st or 2nd place)

- **Deadline for Kaggle submission: 5PM Tuesday May 31**

Final Project Oral Presentations

- There are two parts of oral presentations for each group:
    - On Week 16 (June 1): A 10-min ‘intro’ of your task and your model from each group before we reveal the Kaggle competition results LIVE in class.
    - The speaker will stay for another 5 mins to discuss the competition results.

Final Project Oral Presentations

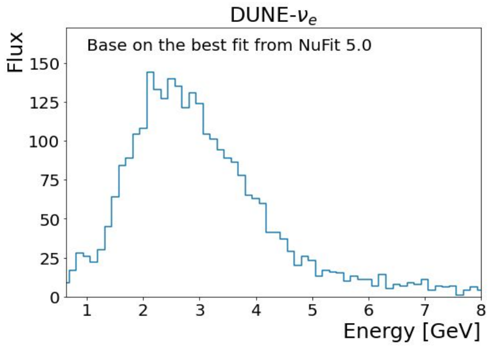
- The ‘intro’ presentation will be evaluated based on:
  - Physics/Science motivation: 3%
  - Data description (Structure/shape of data, visualization) and Task description (explain what you plan to do): 3%
  - Model description/architecture: 3%
  - Presentation skills (English correctness, font and plot size, etc.): 3%
  - Another 3% for performance during the 5-min discussion

Example: Neutrino Oscillation

- Physics motivation: Explain what is neutrino oscillation

$$\underbrace{\begin{bmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{bmatrix}}_{\text{Flavor Eigenstates}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{bmatrix} \begin{bmatrix} c_{13} & 0 & s_{13} e^{-i\delta_{\mathrm{CP}}} \\ 0 & 1 & 0 \\ -s_{13} e^{i\delta_{\mathrm{CP}}} & 0 & c_{13} \end{bmatrix} \begin{bmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{PMNS Matrix}} \underbrace{\begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix}}_{\text{Mass Eigenstates}}$$

- Data description: What are in the dataset? How do they look like?
- Task: Regression of parameters
- What model do you use for the task?



Final Project Oral Presentations

- There are two parts of oral presentations for each group:
  - On Week 17 (June 8): A 15-min presentation on the whole story, including domain knowlege, data structure and visualization, model architecture, final results, and what you’ve learned from comparison to the other group on the same project.

Final Project Oral Presentations

- The ‘full’ presentation will be evaluated based on:
  - Recap of motivation, data structure/visualization, and task description: 3%
  - Model description/architecture/how you optimize it: 3%
    – E.g. Have you tried different ways/architectures and choose the best one? Tell us about it!
  - Results and discussion/interpretation: 3%
  - What’s learned from the other team and Summary: 3%
  - Presentation skills: 3%

Final Project Oral Presentations

- For the group with 3 members, there will be an extra 15-min talk (on June 8) on whether your model can detect the ‘fake’ distribution generated by your competitor.
  - Recap of the ‘intro’ talk: 4%
  - Recap of the models from both teams: 4%
  - Report on the performance of your discriminator, and discuss/summarize: 4%
  - Presentation skills: 3%

Final Project Oral Presentations

- Please inform the TA’s the names of speakers for each talk by **5PM next Wednesday (May 25)**, so that we can setup the agenda.

  Please fill this google survey:
  https://docs.google.com/forms/d/e/1FAIpQLSeISLniIDBI9C-h_dtbdpVfmU91db4Z6RvQ3yZUcdIDWT3B-g/viewform

- You should be able to upload your talk to the agenda page (will inform you to try when it’s setup). If not, send a link to your slides to the TA by **5PM on Tuesday** before your talk.

Final Project Oral Presentations

- On the day of the presentations, the speaker should turn on the camera/video so that we can see you talking!

Final Project Grading Policy

- Final Project - 30%
  - Oral presentation: 15% (Focus of today)
  - **Final project results: 10%**
    - Notebook readiblity (well-commented): 3%
    - Consistency with oral presentations: 3%
    - Results reproducibility: 4%
  - Kaggle competition: 5% (1st or 2nd place)

Attendance Requirement

- Next Wednesday (May 25) we will have a guest lecture from Prof. Daw-Wei Wang. Please connect to Google meet at 10:10 **on time**!
  Topic: Some Applications of Machine Learning in Astronomy, Neurosciences and Condensed Matter Physics.
- Since attendance is strictly required, at the beginning we will ask you to turn on your camera and say hi :)
  (you can turn it off afterwards).
- Please check if your camera is working *now*!