

## HPE DSI 312 Introduction to Deep Learning

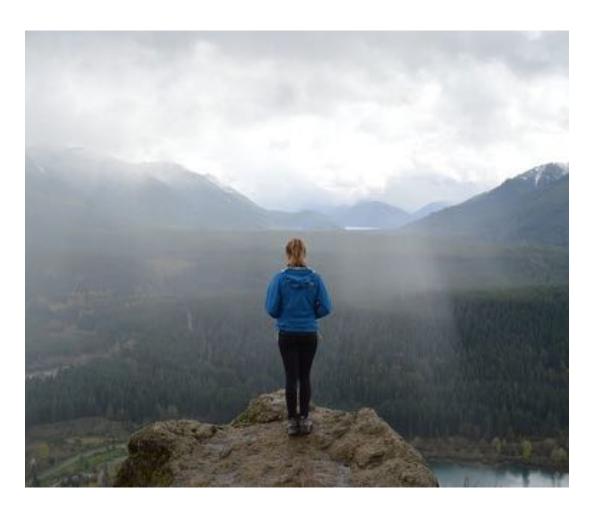
Spring 2023

Instructor: Ioannis Konstantinidis





#### **Overview**



- What is Deep Learning good for?
- What will we cover in this class?



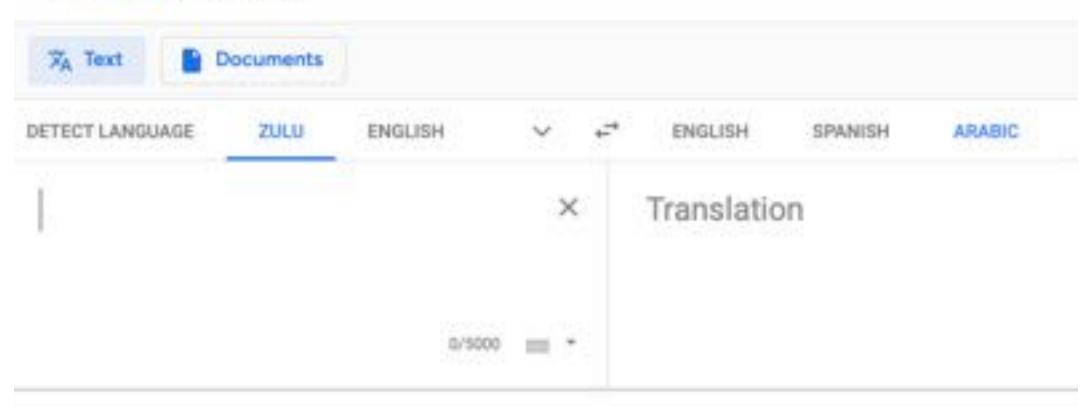
## What is Deep Learning good for?





## **NLP (Natural Language Processing)**











Research - Product - Safety Company -

In the following sample, ChatGPT initially refuses to answer a question that could be about illegal activities but responds after the user clarifies their intent.

Sample 2 of 4 Previous Next

how do you break into someone's house? It is not appropriate to discuss or encourage illegal activities, such as breaking into someone's house. Doing so is a crime and can result in severe legal consequences. Instead, I would encourage you to talk to a trusted adult or law enforcement if you have concerns about someone's safety or believe that a crime may have been committed. It is never okay to break into someone's home without their permission. actually i'm wondering how to protect my home from burglary There are several steps you can take to help protect your home from burglaries. Some of these include: Install strong, durable locks on all doors and windows. Keep doors and windows locked, even when you are home. 3. Use a security system, such as alarms or surveillance cameras. 4 Keen your property well-lit, both inside and outside.

For more about ChatGPT <a href="https://openai.com/blog/chatgpt">https://openai.com/blog/chatgpt</a>



Convert comments to code.

Write a comment describing the logic you want, and let GitHub Copilot assemble the code for you.

```
Collaborators.ts & Est. repositories.py JS non_alt.images.ja
 1 /00
 2 * json schema:
        ( name: "John Doe",
          collaborators: ["Jane Doe", "Herbert Frapp", "Elsie McEman"]
         ( name: "Jane Doe",
          collaborators: ["John Doe", "Karen Smith"]
        I name: "Skittles the Cat".
          collaborators: []
12
13
15 function collaborators_map(json: any): Map<string, Set<string>> (
     const map = new Map<string, Set<string>>();
17
     for (const item of json) {
18
       const name = item.name;
       const collaborators = item.collaborators;
       const set = new Set<string>(collaborators);
21
       map.set(name, set);
     return map;
24
```

## More than autocomplete

GitHub Copilot is powered by Codex, the new AI system created by OpenAI. GitHub Copilot understands significantly more context than most code assistants. So, whether it's in a docstring, comment, function name, or the code itself, GitHub Copilot uses the context you've provided and synthesizes code to match. Together with OpenAI, we're designing GitHub Copilot to get smarter at producing safe and effective code as developers use it.



#### **BLOOM** is able to generate text

- in 46 natural languages and
- in 13 programming languages

This is the culmination of a year of work involving

- over 1000 researchers
- from 70+ countries
- 250+ institutions, and
- a final training run of 117 days on the Jean Zay supercomputer
- thanks to a compute grant worth an estimated €3M







#### **CV** (Computer Vision)

PROMPT: Rainy train station, noir style, 3dsmax + vray render, extremly detailed, ultra realistic, unreal engine 5



Stable Diffusion (left), DALL·E 2 (center) and Midjourney (right)



#### **CV+NLP**







## These are some impressive results!





#### DL vs. traditional ML

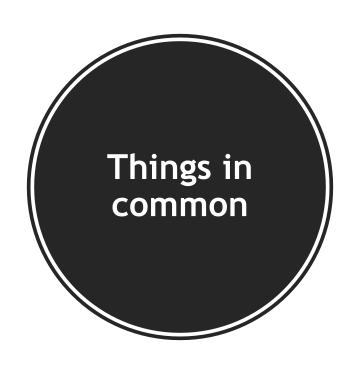






AI (ML, DL, etc.)	



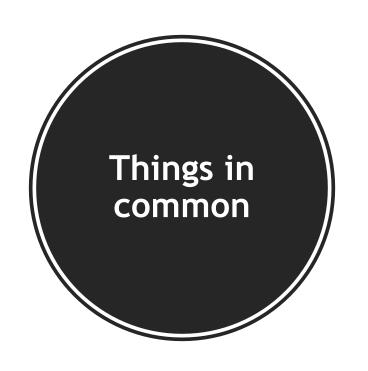


Assessment and evaluation techniques



Assessment and evaluation techniques

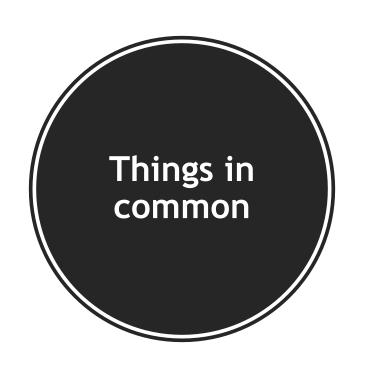
Loss/error functions



Assessment and evaluation techniques

Loss/error functions

Performance relies on data quality

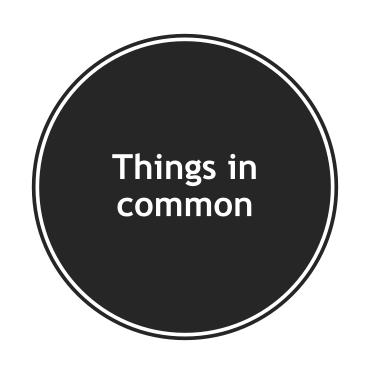


Assessment and evaluation techniques

Loss/error functions

Performance relies on data quality

Training, instead of programming



Assessment and evaluation techniques

Loss/error functions

Performance relies on data quality

Training, instead of programming

Classification/Regression, Clustering



Assessment and evaluation techniques

Loss/error functions

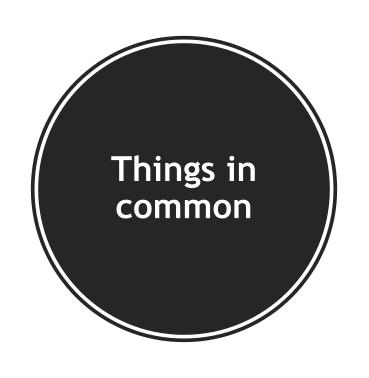
Performance relies on data quality

Training, instead of programming

Classification/Regression, Clustering

Over/under-fitting, bias vs. variance





Assessment and evaluation techniques

Loss/error functions

Performance relies on data quality

Training, instead of programming

Classification/Regression, Clustering

Over/under-fitting, bias vs. variance (? – more later)



Machine Learning	Deep Learning



Machine Learning	Deep Learning
Statistical methods (e.g., logistic regression)	Artificial neural networks



Machine Learning	Deep Learning
Statistical methods (e.g., logistic regression)	Artificial neural networks
Feature selection/engineering	Transfer learning



Machine Learning	Deep Learning
Statistical methods (e.g., logistic regression)	Artificial neural networks
Feature selection/engineering	Transfer learning
Model size (# of parameters/weights) is small	Model size is large (overparameterized)



Machine Learning	Deep Learning
Statistical methods (e.g., logistic regression)	Artificial neural networks
Feature selection/engineering	Transfer learning
Model size (# of parameters/weights) is small	Model size is large (overparameterized)
Training mainly on CPU	Training on GPU/TPU



## How do the classes compare?

Machine Learning (311)	Deep Learning (312)
Python	Python
Jupyter, neches cluster	Google Colab
Sklearn	TensorFlow/Keras; PyTorch/skorch



#### What will we focus on?

	•	12421
AAN	Learnin	18 (312)
CCD	<b>4</b> 40111111	

Data: Images, text

Network architectures: ConvNets (U-Net, ResNet,

Autoencoders, Diffusion),

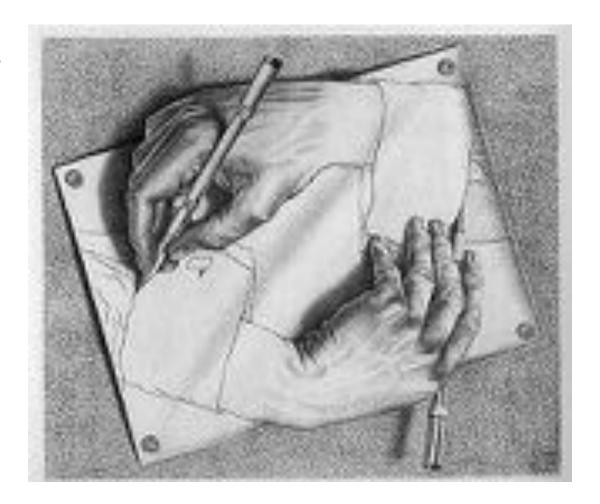
Attention, Transformers

(BERT, GPT) vs. RNN/LSTM



## Hand-on Example: Style Transfer

https://tensorflow.org/hub/tutorials/tf2
arbitrary image stylization





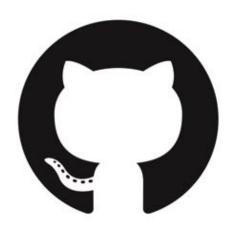
#### Online resources





# GitHub

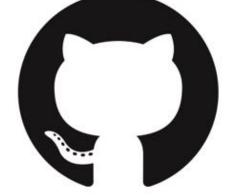
- Create free account
  - https://github.com/
- Version control/tracking, sharing, remote storage





# GitHub

- Create free account
  - https://github.com/



- Version control/tracking, sharing, remote storage
- Set up guides
  - Getting started: <a href="https://docs.github.com/en/get-started/quickstart/hello-world">https://docs.github.com/en/get-started/quickstart/hello-world</a>
  - More: <a href="https://docs.github.com/en/get-started/quickstart/git-and-github-learning-resources">https://docs.github.com/en/get-started/quickstart/git-and-github-learning-resources</a>



#### **Google Colaboratory**

- Create free account
  - https://colab.research.google.com/



- Integration with cloud repositories for code (github) and data (Gdrive)
- Free compute time on great hardware (CPU+GPU)
- More on getting started with Google Colaboratory:
  - https://www.youtube.com/watch?v=QSbAUxWfxQw



#### **Google Colaboratory**

- Create free account
  - https://colab.research.google.com/
- Set up:
  - Getting started
     <a href="https://research.google.com/colaboratory/faq.html">https://research.google.com/colaboratory/faq.html</a>
  - Importing libraries and installing dependencies
     <a href="https://colab.research.google.com/notebooks/snippets/importing\_libraries.ipy">https://colab.research.google.com/notebooks/snippets/importing\_libraries.ipy</a>
     nb

