

Deep Learning

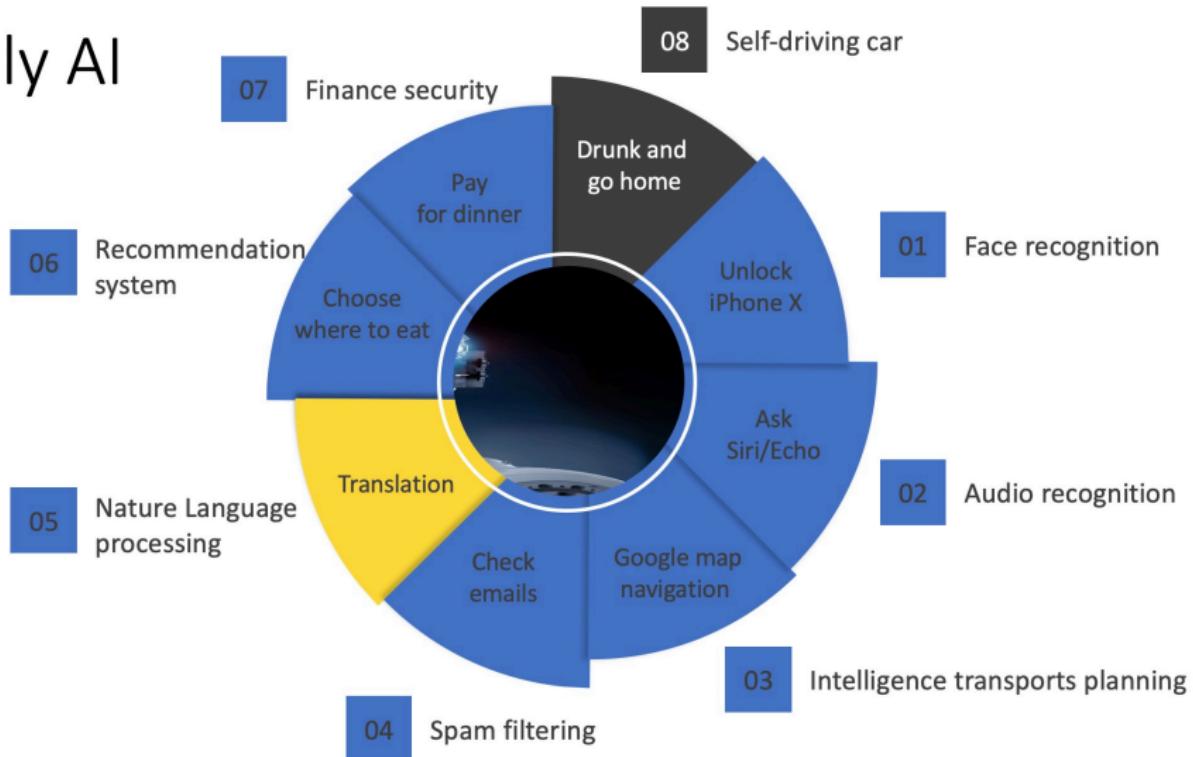
Vishnu Lokhande ¹

Department of Computer Science and Engineering
University at Buffalo, SUNY
`vishnulo@buffalo.edu`

January 23, 2025

¹ Slides adapted from the slides of Prof. Changyou Chen. All content is used with permission and credit to the original author.

Daily AI



Deep Learning in the Era of Big Data

Machine learning: next generation of artificial intelligence

Deep Learning for Real Applications

- Surpass human on
 - ▶ classification on ImageNet
 - ▶ AlphaGo
 - ▶ Atari game
- Interesting applications
 - ▶ self-driving cars
 - ▶ speech
 - ▶ image-style transformation
 - ▶ image captioning
 - ▶ ...

ES

Briefly explain chatGPT using a metaphor that would resonate with an audience of stock-market investors and avoiding AI jargon.



One way to explain chatGPT to stock-market investors is to compare it to a financial advisor who uses past performance data and market trends to make predictions about future investments. Just like a skilled financial advisor, chatGPT uses large amounts of training data to generate intelligent and informed responses to questions, providing valuable insights and suggestions to users.

Text-to-Image Generation



(a) A photo of a light bulb in outer space traveling the galaxy with a sailing boat inside the light bulb.



(b) A high-resolution photo of a chicken working out in a gym.



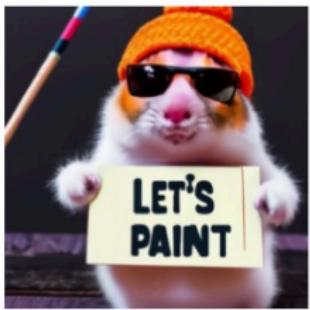
(c) New York Skyline with 'Hello World' written with fireworks on the sky.



(d) A television made of water that displays an image of a cityscape at night.



(e) A photo of llama wearing sunglasses standing on the deck of a spaceship with the Earth in the background.



(f) A high contrast portrait photo of a fluffy hamster wearing an orange beanie and sunglasses holding a sign that says "Let's PAINT!".



(g) A photo of an astronaut riding a horse in the forest. There is a river in front of them with water lilies.



(h) A portrait photo of a kangaroo wearing an orange hoodie and blue sunglasses standing on the grass in front of the Sydney Opera House holding a sign on the chest that says Wel-

ChatGPT/GPT-4/GPT-4V

GPT-4 visual input example, Solving École Polytechnique's Physics Exam:

User Answer question I.I.a. Think step-by-step.

I. Principe de la détection de rayonnement avec un bolomètre

Comme illustré sur la figure 1, un bolomètre est constitué d'un absorbeur qui吸收 le rayonnement que l'on désire détecter. Sa température T suppose uniforme, et mesure à l'aide d'un thermosondeur. La variation de température ΔT entre deux mesures successives est proportionnelle à dT . Si l'on connaît la constante de proportionnalité $\alpha = \frac{\Delta T}{dT}$, alors la température T est caractérisée par la relation $\alpha = \frac{d\theta}{dT}$. L'émissivité ϵ donne la puissance thermique G .

Un barreau conducteur thermique, homogène, de longueur L , de section S et de conductivité thermique λ a une réchange thermique latente, alors le barreau a un thermostat à la température T_0 fixe.

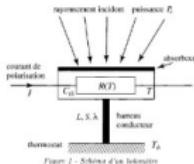


Figure 1 - Schéma d'un bolomètre

E.I. Dans cette première partie, on étudie le comportement du bolomètre en régime stationnaire, à la température de fonctionnement $T = T_0$.

E.I.a) Expliquez la température $T(x)$ en tout point du barreau conducteur reçue par son abscisse x ; on prendra $T(0) = T_0$ et $T(L) = T_b$.

GPT-4 To answer question I.I.a, we need to express the temperature $T(x)$ in every point of the conductive bar, which is characterized by the abscissa x . We are given that $T(0) = T_0$ and $T(L) = T_b$.

For a stationary regime, the heat equation for the one-dimensional heat conduction in the bar is:
$$d^2T/dx^2 = 0$$

Integrating once, we get:
$$dT/dx = C_1$$

Integrating a second time:
$$T(x) = C_1 x + C_2$$

To determine the constants C_1 and C_2 , we use the boundary conditions:
 $T(0) = T_0$, which gives $C_2 = T_0$
 $T(L) = T_b$, which gives $C_1 = (T_b - T_0)/L$

Finally, the expression for $T(x)$ is:

$$T(x) = (T_b - T_0) * (x/L) + T_0$$

This represents a linear temperature profile along the conductive bar, with a slope of $(T_b - T_0)/L$.

GPT-4 visual input example, Moar Layers:

User Can you explain why this is funny. Think about it step-by-step.



GPT-4 The comic is satirizing the difference in approaches to improving model performance between statistical learning and neural networks.

In statistical learning, the character is shown to be concerned with overfitting and suggests a series of complex and technical solutions, such as minimizing structural risk, reworking the loss function, and using a soft margin.

In contrast, the neural networks character simply suggests adding more layers to the model. This is often seen as a common solution to improving performance in neural networks, but it's also considered a simplistic and brute-force approach.

The humor comes from the contrast between the complexity and specificity of the statistical learning approach and the simplicity and generality of the neural network approach. The "But unironically" comment adds to the humor by implying that, despite being simplistic, the "stack more layers" approach is often effective in practice.

DALL-E 3: Instruction-Following Image Generation

The sidewalks bustling with **pedestrians** enjoying the nightlife.



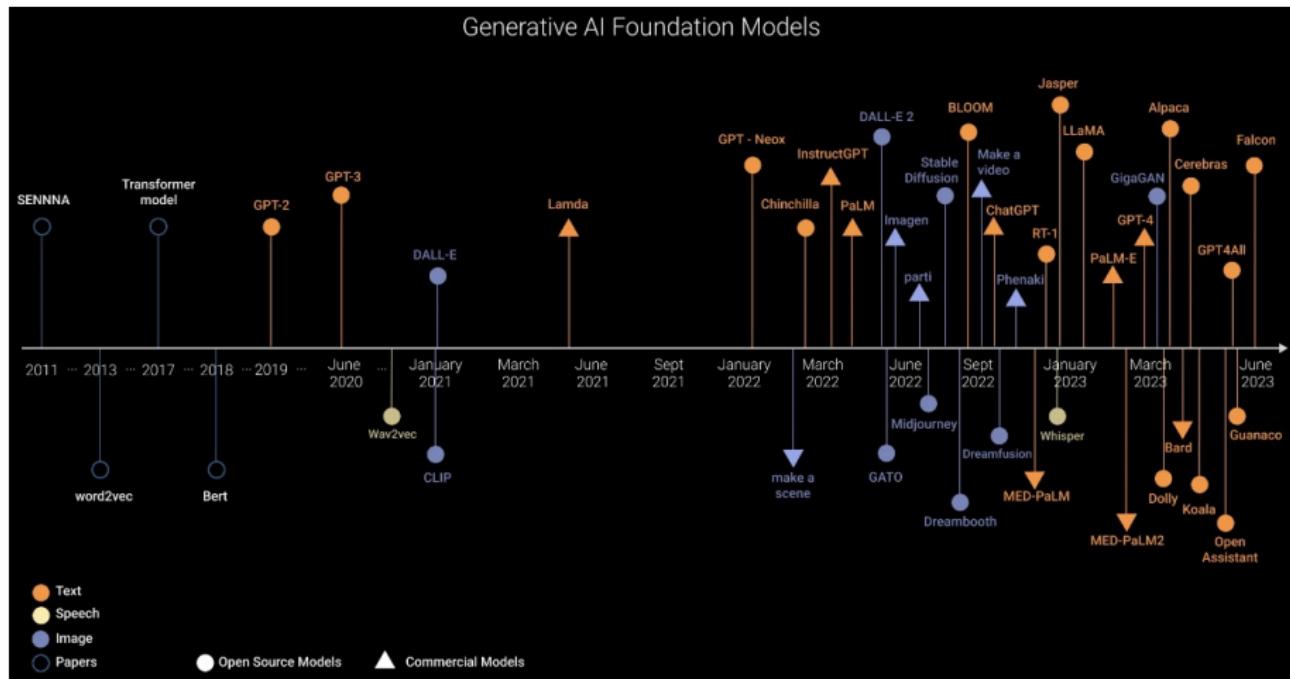
At the corner stall, a **young woman** with fiery red hair, dressed in a signature velvet cloak, is **haggling with the grumpy old vendor**.

A bustling city street under the shine of a **full moon**.



The grumpy vendor, a **tall, sophisticated man**, is wearing a sharp suit, sports a **noteworthy moustache** and is animatedly conversing on his **steampunk telephone**.

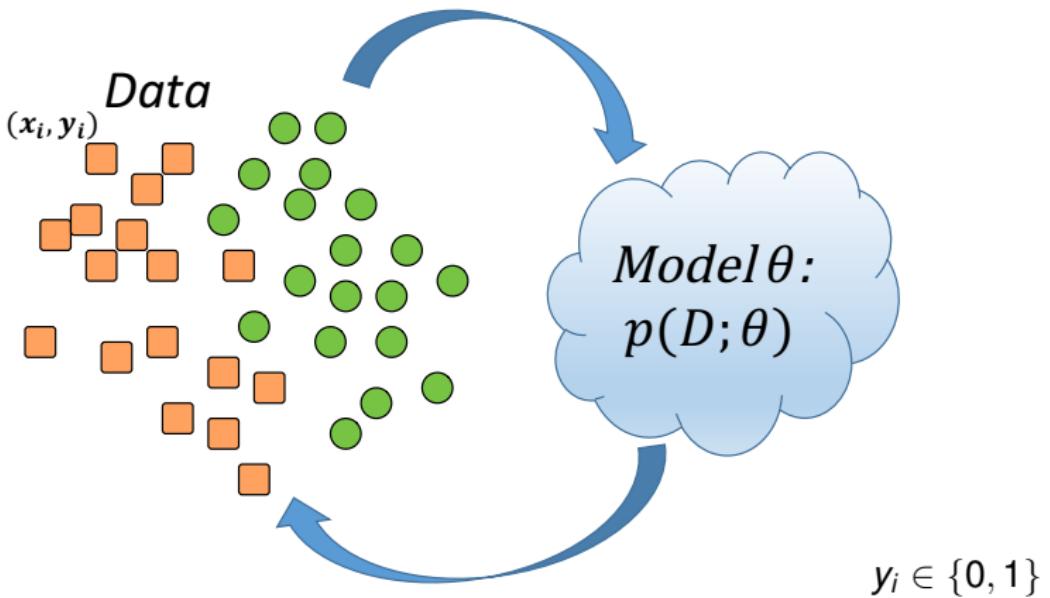
Generative AI Revolution



History of Deep Learning

- 1943: Neural Networks
- 1957-62: Perceptron
- 1970-86: Backpropagation, RBM, RNN
- 1979-98: CNN, MNIST, LSTM, Bidirectional RNN
- 2006: “Deep Learning”, DBN
- 2012: ImageNet + AlexNet
- 2014: GANs
- 2016-17: AlphaGo, AlphaZero
- 2017-19: Transformers
- 2020: GPT-3 (175 billion parameters)
- 2021: CLIP: Connecting text and images
- 2021: DALL·E: Creating images from text
- 2022: DALL·E 2, ChatGPT: Creating images from text, ChatBot
- More ...

Machine Learning at a Glance



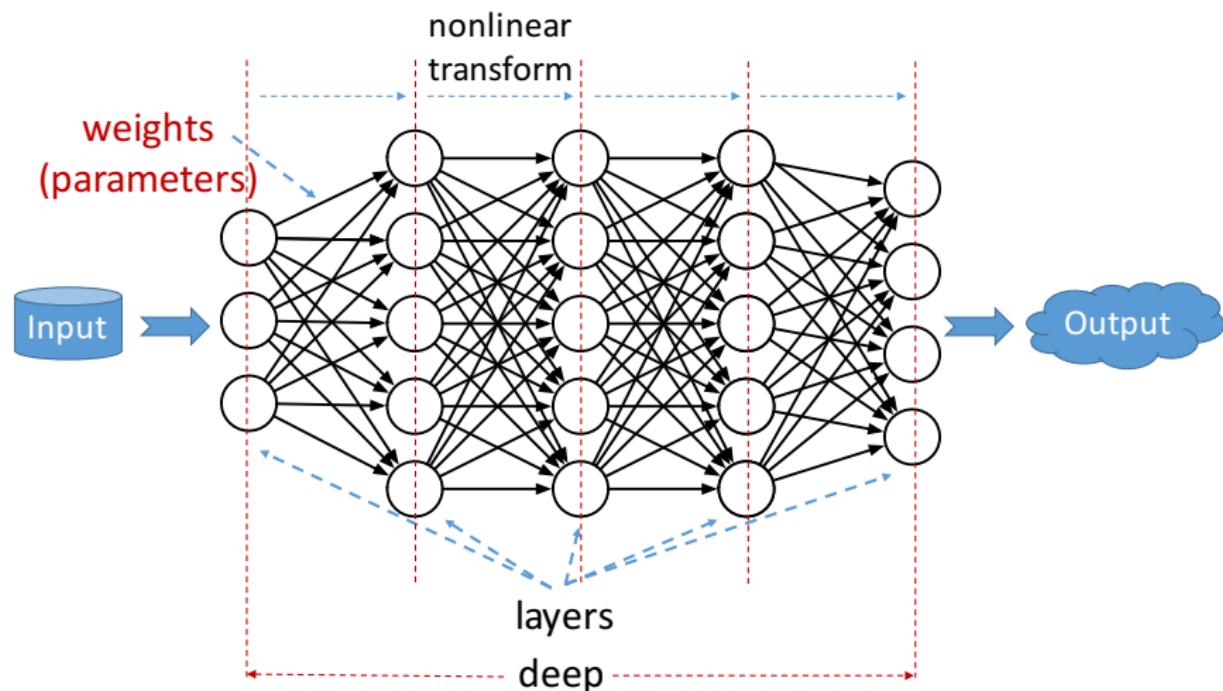
Logistic regression: $p(y_i; \theta, x_i) = \frac{e^{y_i \theta^T x_i}}{1 + e^{\theta^T x_i}}$, $\mathcal{D} = \{\mathbf{d}_1, \dots, \mathbf{d}_N\}$, $\mathbf{d}_i = \{\mathbf{x}_i, y_i\}$

Machine Learning

Learning model θ from the data \mathcal{D}

Deep Neural Networks

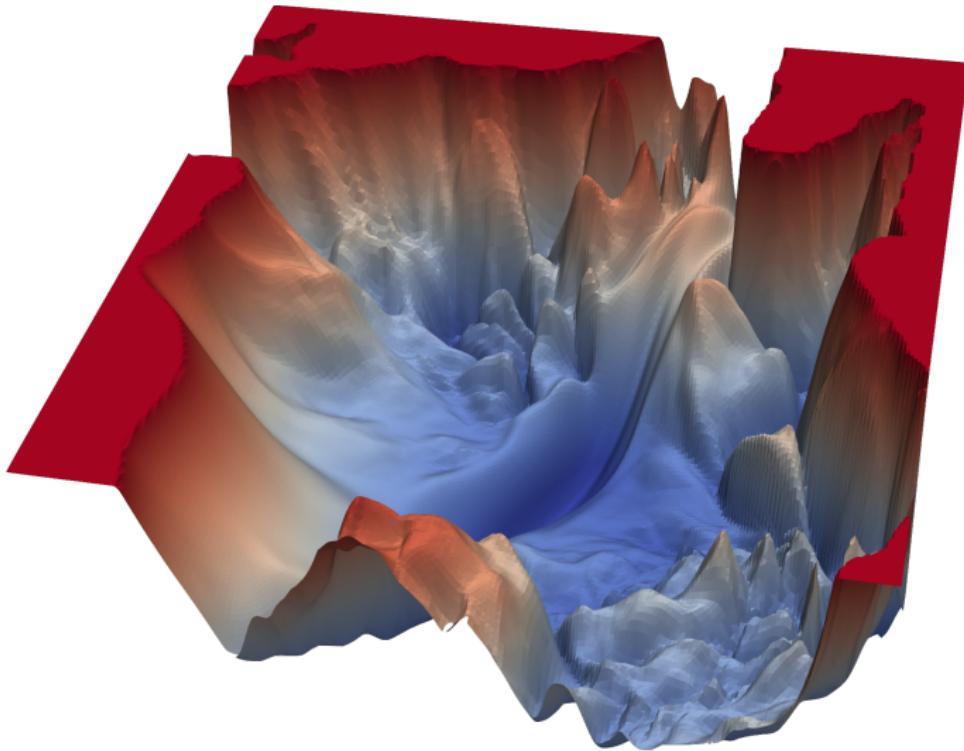
A sequence of **nonlinear transformations** from input to output



Learning network weights from data.

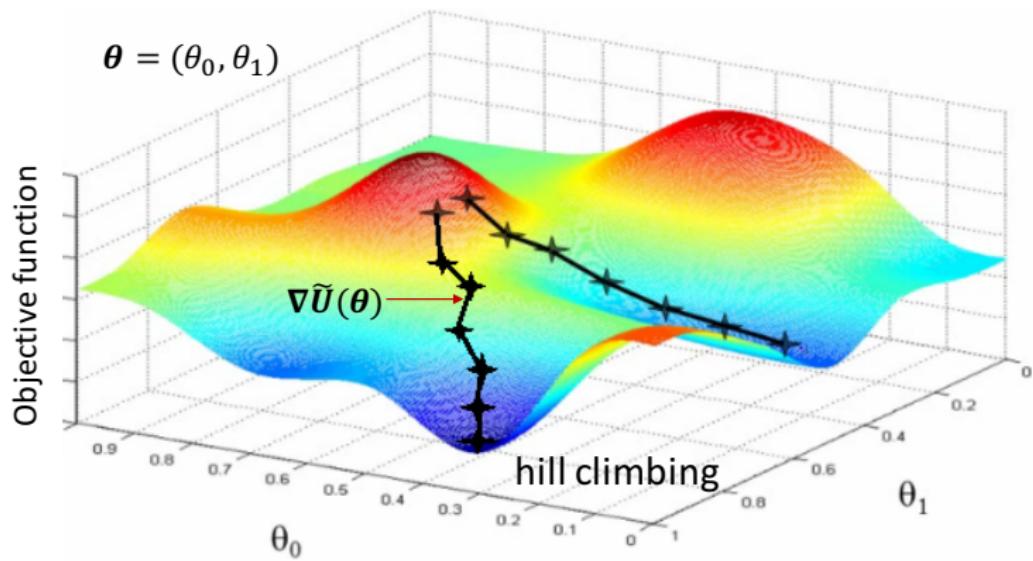
Learning from Data: Loss Function

- Extremely complex geometrics structure.



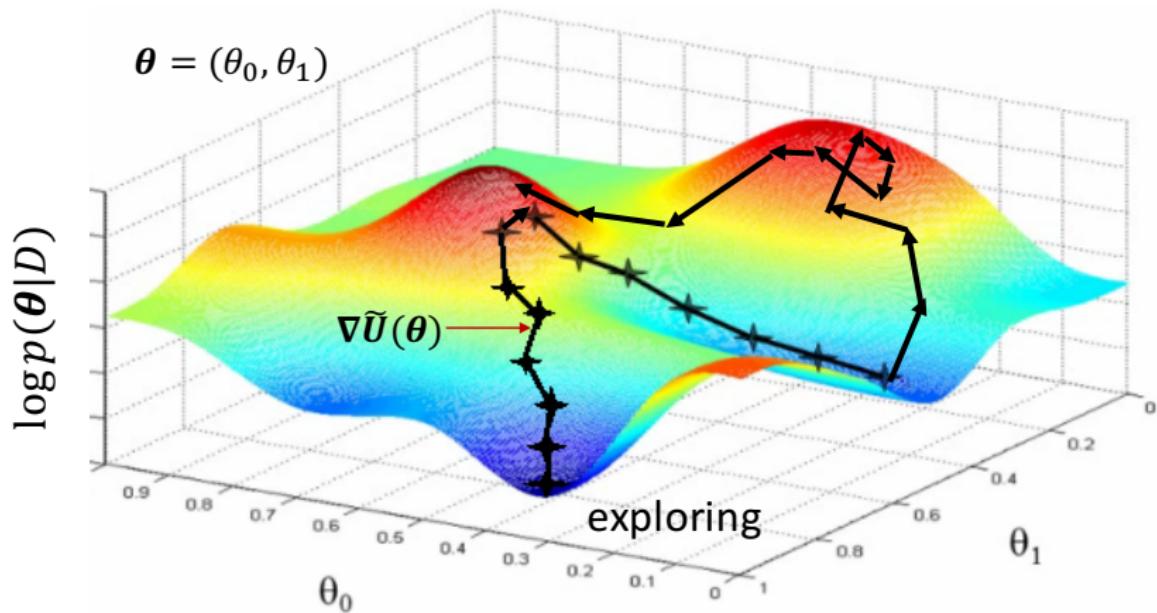
Learning with Stochastic Gradient Descent

Simply follow the noisy gradient directions to find an optimal parameter



Learning with Markov Chain Monte Carlo

Draw samples from the posterior distribution of the parameter



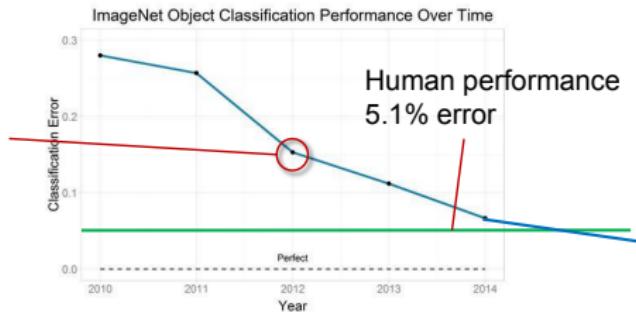
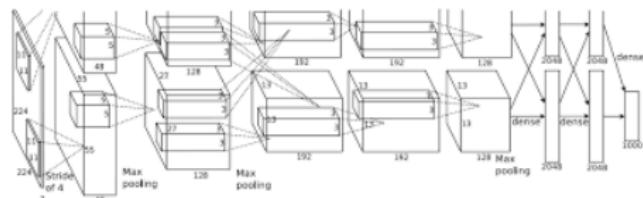
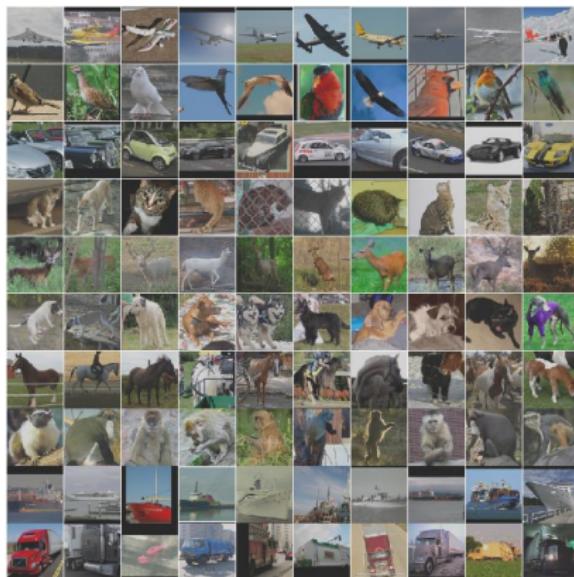
Machine Learning Tasks

- ① Supervised learning:
 - ▶ models have both input (features) and output (labels) data
- ② Unsupervised learning:
 - ▶ models only have input data
- ③ Semi-supervised learning:
 - ▶ models have input data and partial output data
- ④ Reinforcement learning:
 - ▶ models in general do not have data (interact with an environment to get the data)

Neural networks are used to build models under these settings.

Supervised Learning: Large-Scalable Classification²

Convolutional NNs: AlexNet (2012): trained on 200 GB of ImageNet Data

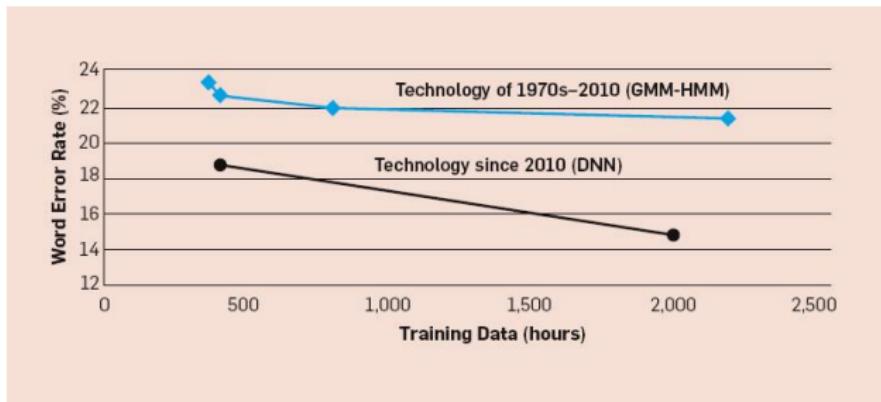
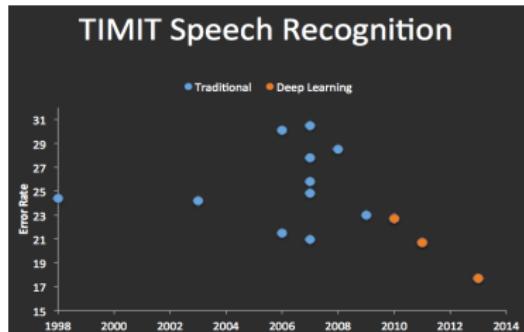
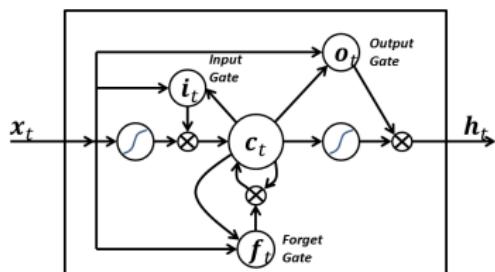


- ResNet (2015): 3.57% error – surpass human performance.

²Some of the following slides are credited to online resources.

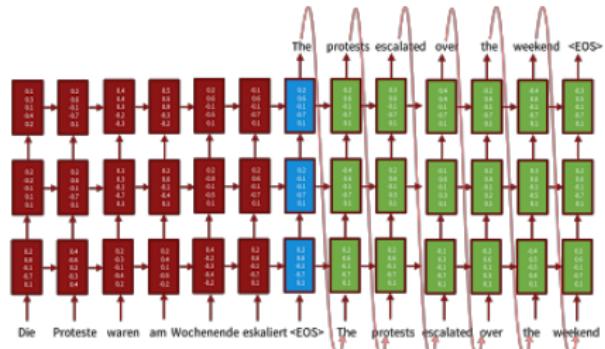
Supervised Learning: Speech Recognition

Recurrent Nets: LSTMs (1997):



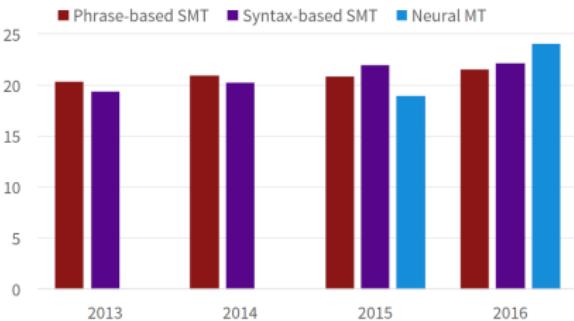
Supervised Learning: Machine Translation

Sequence-to-sequence models with LSTMs and attention:



Progress in Machine Translation

[Edinburgh En-De WMT newstest2013 Cased BLEU; NMT 2015 from U. Montréal]



From [Sennrich 2016, http://www.meta-net.eu/events/meta-forum-2016/slides/09_sennrich.pdf]

In 2016, Google launched a deep-learning based translation system on real product, which significantly improves traditional methods.

Unsupervised Learning: Generative Models

How to deal with data without labels?

- Density estimation



- Sample generation



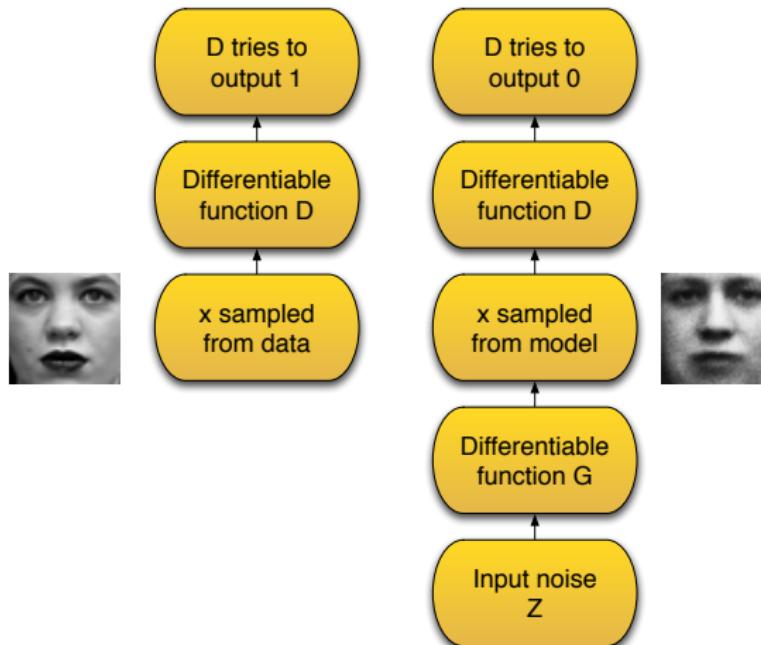
Training examples

Model samples

(Goodfellow 2016)

Unsupervised Learning: Adversarial Nets Framework

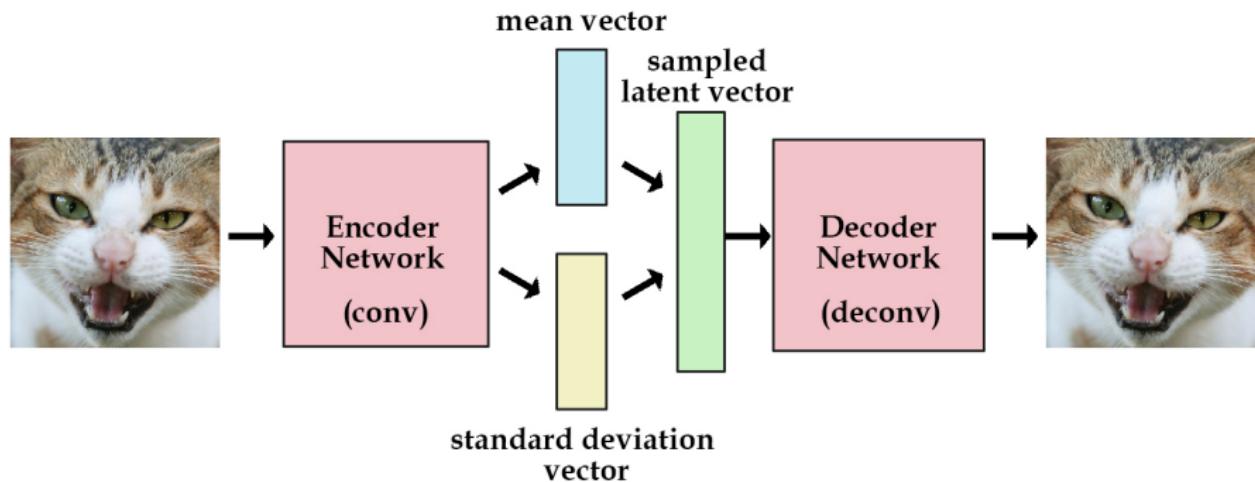
- Two players (represented by deep neural networks) compete with each other in the training to make samples generated from a model look real.



Goodfellow, 2016

Unsupervised Learning: Variational Autoencoder

- Model data with a **probabilistic model** (represented by deep neural networks) so that it can reconstruct itself well.



Vector Space Arithmetic

For a good latent representation, we hope:

Arithmetic on latent space → Arithmetic on image space



-

+

=

Man
with glasses

Man

Woman

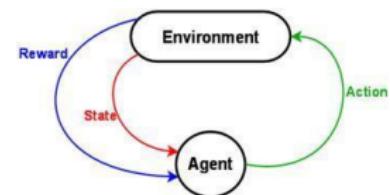


Woman with Glasses

Reinforcement Learning

Markov Decision Process (MDP)

- Finite set of states \mathcal{S} .
- Finite set of actions \mathcal{A} .
- Immediate reward function $\mathcal{R} : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$.
- Transition (next-state) function:
 $\mathcal{T} : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{S}$.
- A discount factor, $\gamma \in (0, 1]$, for long-term reward.



$$s(0) \rightarrow a(0) \rightarrow r(0) \rightarrow s(1) \rightarrow a(1) \rightarrow r(1) \rightarrow \dots$$

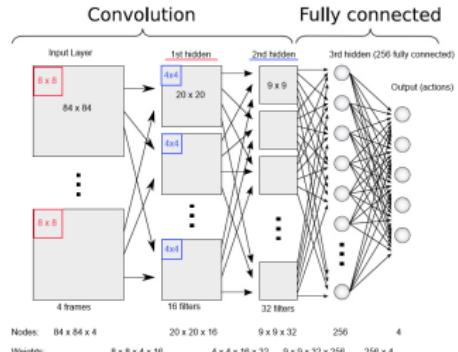
Goal: Learn to choose actions that maximize the cumulative reward:
 $r(0) + \gamma r(1) + \gamma^2 r(2) + \dots$

Deep Reinforcement Learning

In 2013, DeepMind's arcade player beats human expert on six Atari Games. Acquired by Google in 2014.

In 2016, DeepMind's alphaGo defeats former world champion Lee Sedol.

In 2017, DeepMind's alphaGo defeats current world champion Jie Ke.

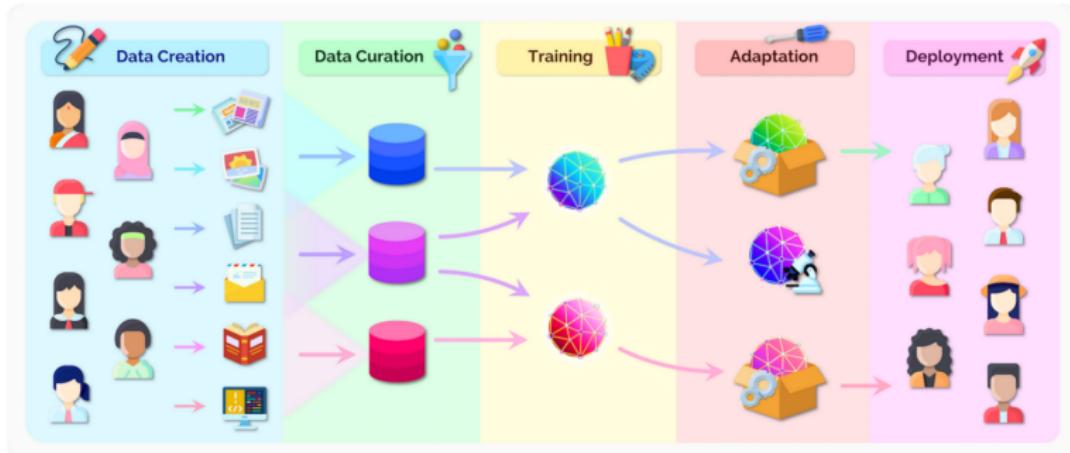


Deep Reinforcement Learning

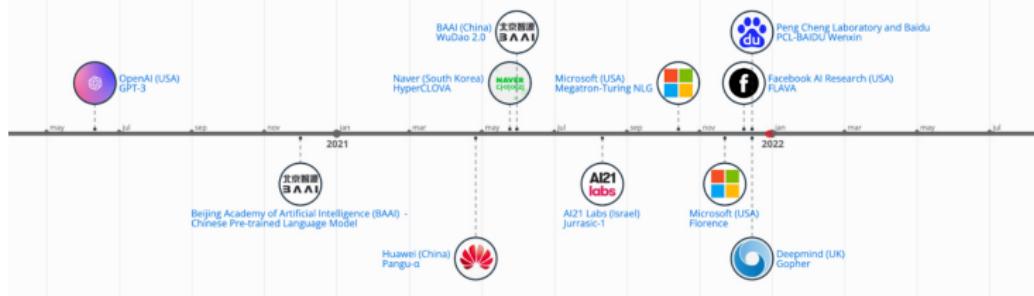


Figure: Top: OpenAI DOTA; Bottom: DeepMind Starcraft.

Foundation Models



Foundation Models Since GPT-3



Course Logistics

Course Information

- Time & Place:
 - ▶ 7:00 - 8:20 PM, Tuesdays/Fridays
 - ▶ Nsc 215
- Course website:
 - ▶ Piazza:
<https://piazza.com/buffalo/spring2025/cse676sectione/home>
- Slides typically uploaded before each lecture.

Instructor & TAs

Instructor

- Vishnu Lokhande

Homepage: <https://www.cse.buffalo.edu/~vishnulo/>

Email: vishnulo@buffalo.edu

Office: 332 Davis Hall

- Office hours:

6:20:-7:20PM Mondays and Wednesdays in Office 332 Davis Hall

TAs

- Sai Bharadwaj Sirigadi: sirigadi@buffalo.edu

- Office hours:

4:00-6:00PM, Fridays, Zoom

Meeting ID: 91337397538, Passcode: 997353

- Primary medium for course communication.
- Course information, announcements, slides, homeworks, etc, will be posted on Piazza.
- Homework are submitted to UBLearn:
 - ▶ Details will be posted on Piazza.
- Encourage to post questions via Piazza:
 - ▶ General post to all (name visible), choose appropriate folder.
 - ▶ Private post to instructor.

Topics

The following topics will be covered

- Machine learning/mathematics basic
- Feedforward neural networks and regularization
- Backpropagation
- Convolutional neural networks
- Recurrent neural networks
- Transformer networks
- Training algorithms
- Deep generative models
- Deep reinforcement learning
- Some advanced topics

Prerequisites

- Proficiency in Python
- Multivariate Calculus, Linear Algebra
- Basic Probability and Statistics
- Foundation of Machine Learning

Text Books & References³

- I. Goodfellow, Y. Bengio and A. Courville, Deep Learning, MIT Press, 2016.
- Prof. Srihari's Deep Learning course:
<http://www.cedar.buffalo.edu/~srihari/CSE676/index.html>
- Andrej Karpathy's Deep Learning course:
<http://cs231n.stanford.edu/slides/2017/>
- Tons of online materials and papers:
 - ▶ e.g., excellent tutorials for ML, DL, NLP and others:
<https://medium.com/machine-learning-in-practice/over-200-of-the-best-machine-learning-nlp-and-python-tutorials-2018-edition-dd8cf53cb7dc>

³The slides will follow closely with the references.

Goals of the Course

- ① To introduce students to the primary deep learning tasks and the techniques that have been developed for those tasks.
- ② To enable students to apply cutting edge deep learning techniques to solve machine learning problems, and effectively communicate their research results.
- ③ To enable students to gain a horizon on an up-to-date deep learning techniques, and equipped knowledge for their future research.

Grading Policy

- Two or Three assignments: 40%
- Midterm: 30%
- Final: 30%
- More details in the syllabus.

Grading Policy⁴

Final grade

A	$[92.5, 100]$	$A-$	$[87.5, 92.5)$
$B+$	$[82.5, 87.5)$	B	$[77.5, 82.5)$
$B-$	$[72.5, 77.5)$	$C+$	$[67.5, 72.5)$
C	$[62.5, 67.5)$	$C-$	$[57.5, 62.5)$
F	$[0, 57.5)$		

- Instructor reserves the right for adjustment:
 - ▶ Apply the same scaling for all students

⁴Reference to CSE 574

Exam Guideline

- Midterm and Final exam will be submitted in the form of hard papers.
 - ▶ Problem types can be choice, QA, calculation, and design problems.
 - ▶ Zero tolerance on cheating: no discussion, copy, reference with each other. If identified, you will fail the course.

Exam Guideline

- Midterm exam:
 - ▶ Location: Davis101
 - ▶ Time: 5:00-6:30PM, March 12, 2025
 - ▶ Only standard calculator without communication function is allowed; no other electronic devices
- Final exam:
 - ▶ Location: Nsc 216
 - ▶ Time: 7:15-10:15PM, May 9th, 2025
 - ▶ Only standard calculator without communication function is allowed; no other electronic devices
- If one is identified AI in the final exam, he/she will fail the course.

Python Resources⁵

- Install Python, ipython: <http://ipython.org/install.html>
- Python IDE: <https://pythonhosted.org/spyder/>
- More about ipython notebooks: <http://ipython.org/notebook.html>
- Python for Developers, a complete book on Python programming by Ricardo Duarte:
<http://ricardoduarte.github.io/python-for-developers/>
- CodeAmerica - Python:
<http://www.codecademy.com/en/tracks/python>
- An introduction to machine learning with Python and scikit-learn (repo and overview) by Hannes Schulz and Andreas Mueller:
http://nbviewer.ipython.org/github/temporaer/tutorial_ml_gkbionics/blob/master/2%20-%20KMeans.ipynb

⁵ Adapted from CSE 574

Academic Integrity and Honor Code



<https://engineering.buffalo.edu/computer-science-engineering/information-for-students/policies/academic-integrity.html>

Important Disclaimer

- This course requires background on machine learning, linear algebra and probability theory:
 - ▶ Only get enrolled if you have these background; otherwise, it is likely that you will under-perform other students.
- This course will focus on introducing the underlying principles of different deep learning techniques, which needs enough mathematics background.
- **This is not a course to teach how to code deep learning!**
 - ▶ You will learn coding by yourself – the most basic requirement.
- It is a course to teach why and how deep learning works.
- Be respectful to each other in class: no talking but please be interactive with the instructor.

A recommended resource to learn to code deep learning with Pytorch

<https://github.com/udacity/deep-learning-v2-pytorch>

- Need to do it by yourself.
- Will not go over these in class.

To Do

- Decide whether to enroll this course.
- Register on Piazza.
- Post questions if you have any.