DEEP LEARNING

Lecture 1: Introduction to Deep Learning

Dr. Yang Lu

Department of Computer Science and Technology

luyang@xmu.edu.cn

厦門大學信息学院(特色化示范性软件学院)
School of Informatics Xiamen University (National Characteristic Demonstration Software School)

厦门大学 计算机科学与技术系
Department of Computer Science and Technology, Xiamen University

Course Information

- Lecturer:
  - 卢杨, assistant professor at department of computer science and technology.
  - Email: luyang@xmu.edu.cn
  - Phone/WeChat: 18850159335
  - Personal webpage: https://jasonyanglu.github.io/
  - bilibili: 可爱小火炉
  - TA (信息学院):
    - 李明波:15573542658, limingbo@stu.xmu.edu.cn
    - 陈姝: 18162503676, 2505800261@qq.com
    - 刘艺杰: 18754281229, yijieliu@stu.xmu.edu.cn
  - TA (人工智能研究院):
    - 陈奕州: 18389218053, 993121673@qq.com
    - 严珊珊: 18811077069, 2653043009@qq.com
    - 叶伟豪: 18759223395, 1273501474@qq.com

Class Schedule and Rules

- Slides in English and lecture in Chinese.
- No eating or playing mobile phones.
- All course materials (lecture notes, materials, assignments, projects) are in course.xmu.edu.cn.

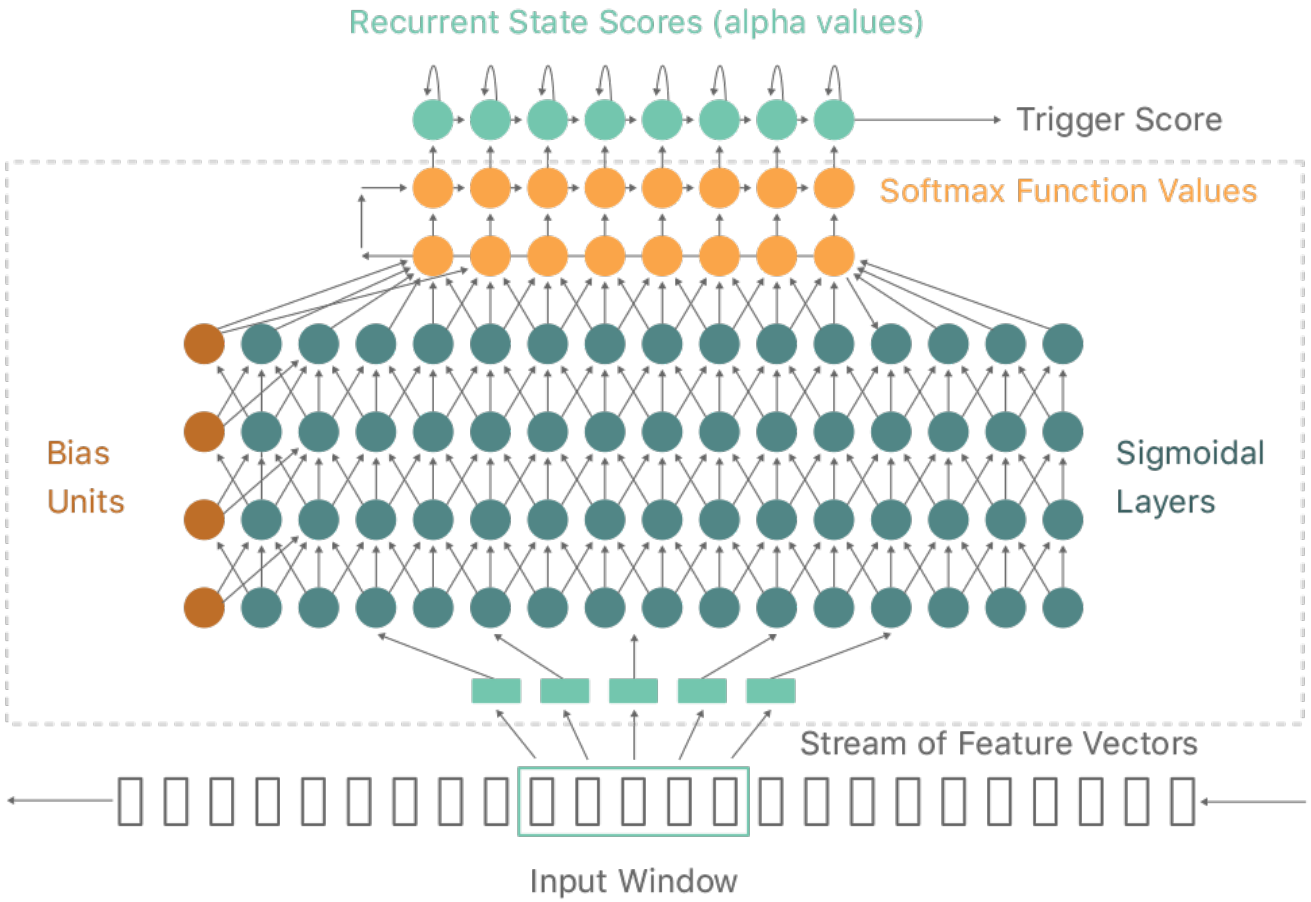
Why Deep Learning?

Are we using deep learning techniques

in our daily life?

Why Deep Learning?

- Speech recognition



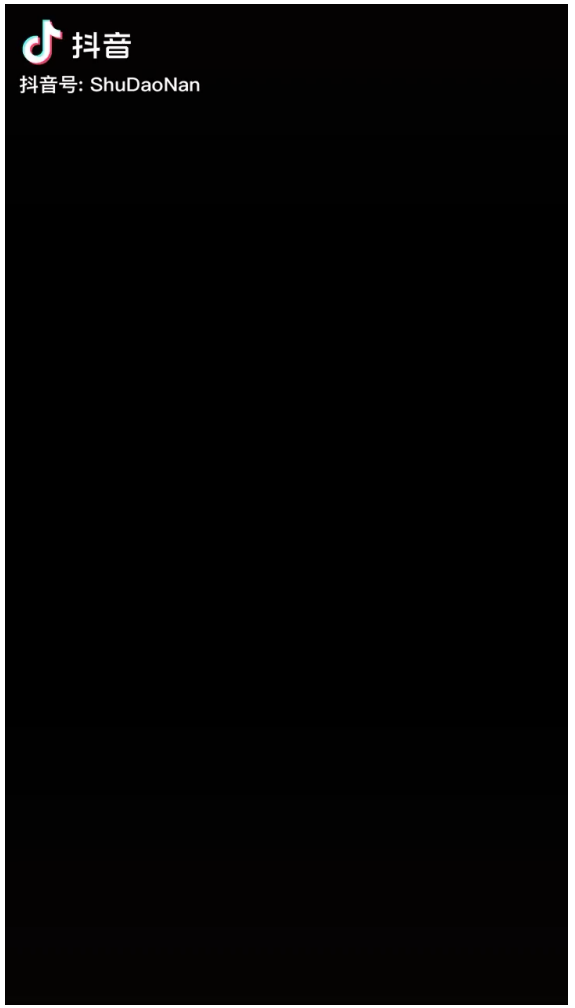


Image source: https://machinelearning.apple.com/research/hey-siri

Why Deep Learning?

- Face verification

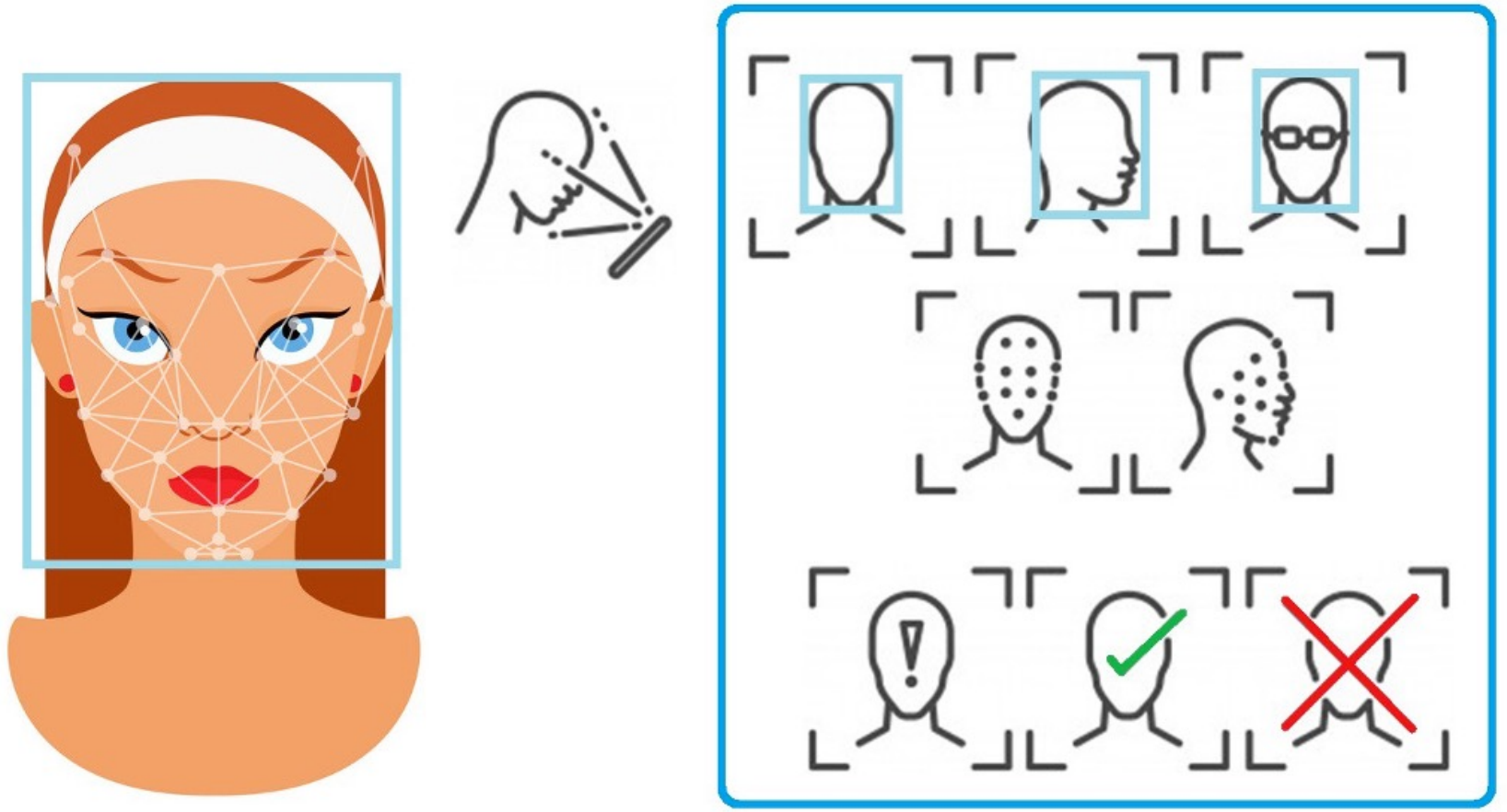


Image source: https://medium.com/@fenjiro/face-id-deep-learning-for-face-recognition-324b50d916d1

Why Deep Learning?

- Face identification

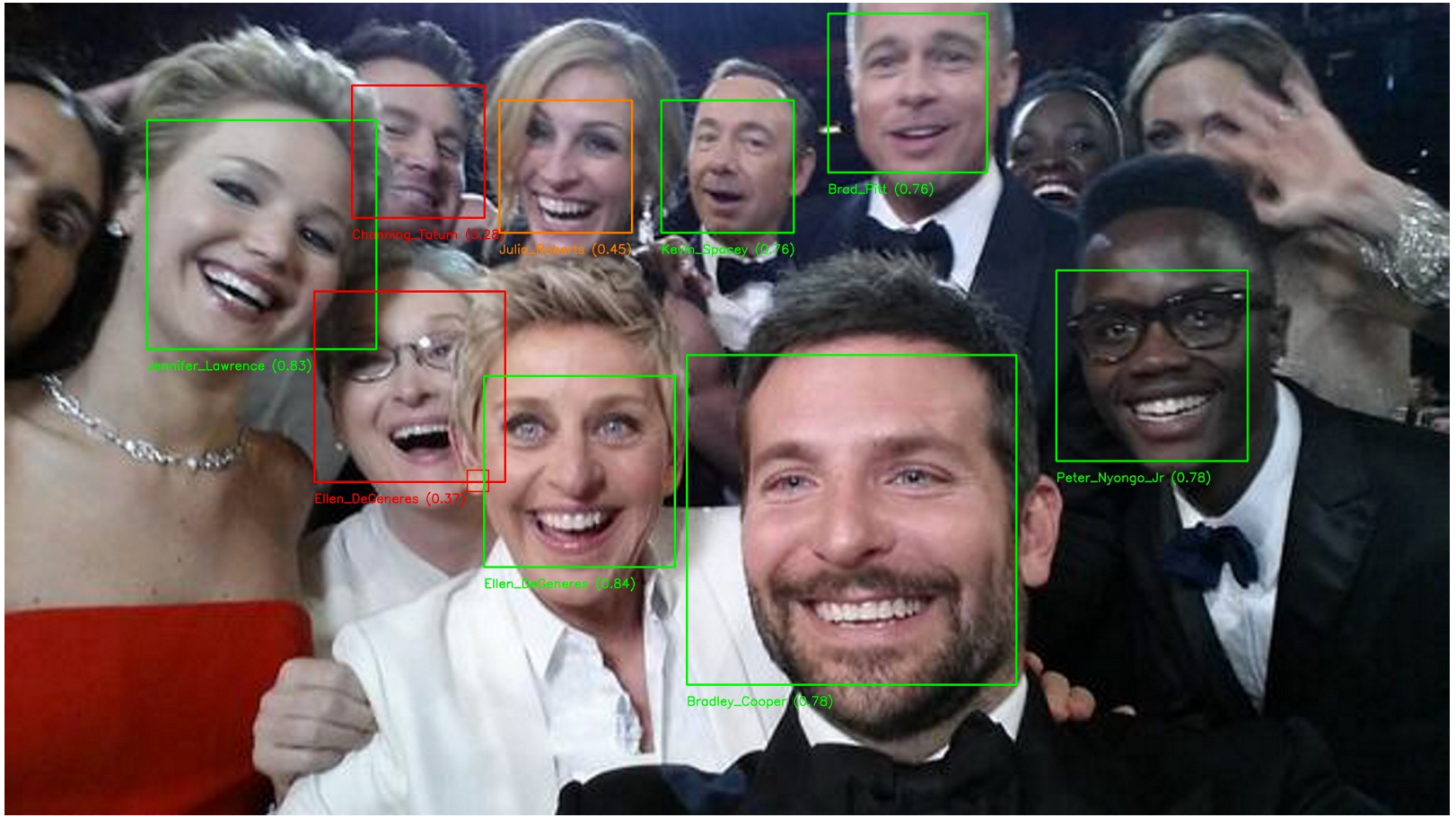


Image source: https://towardsdatascience.com/an-intro-to-deep-learning-for-face-recognition-aa8dfbbc51fb

Why Deep Learning?

- Play games

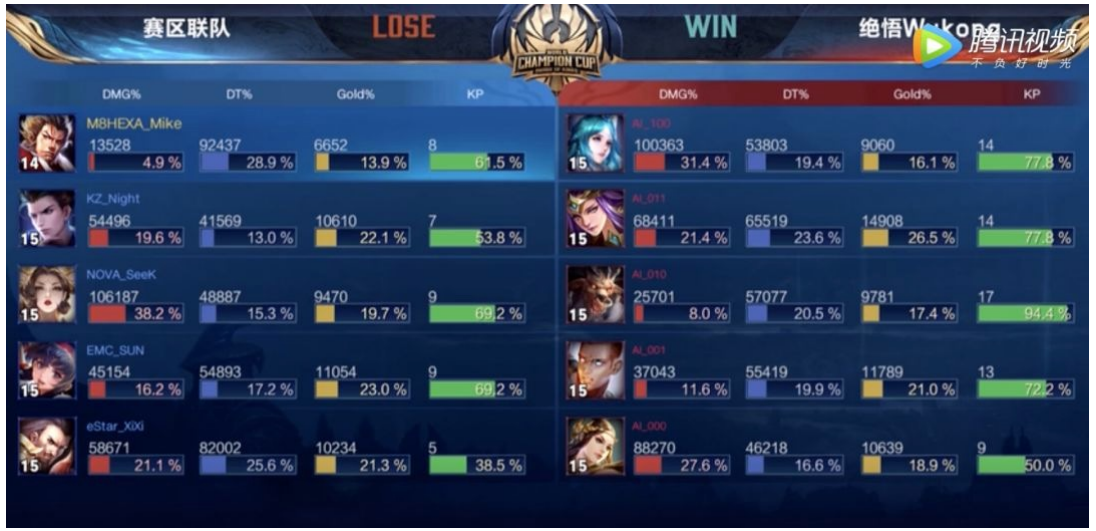
AlphaGo vs. Lee Sedol 4-1, 2016

AlphaGo vs. Ke Jie 3-0, 2017

Image source: https://medium.com/point-nine-news/what-does-alphago-vs-8dadec65aaf
https://www.npr.org/sections/thetwo-way/2017/05/25/530016562/google-a-i-clinches-series-against-humanitys-last-best-hope-to-win-at-go

Why Deep Learning?

- Play games



In 2019, AI player Juewu (绝悟) developed by Tencent defeated pro players in game Honor of Kings (王者荣耀)

Image source: https://blog.csdn.net/qq_39521554/article/details/104402528

Why Deep Learning?

Yu Ji sneak attacks Jia Luo in the grass and almost took her away

Why Deep Learning?

- Autopilot

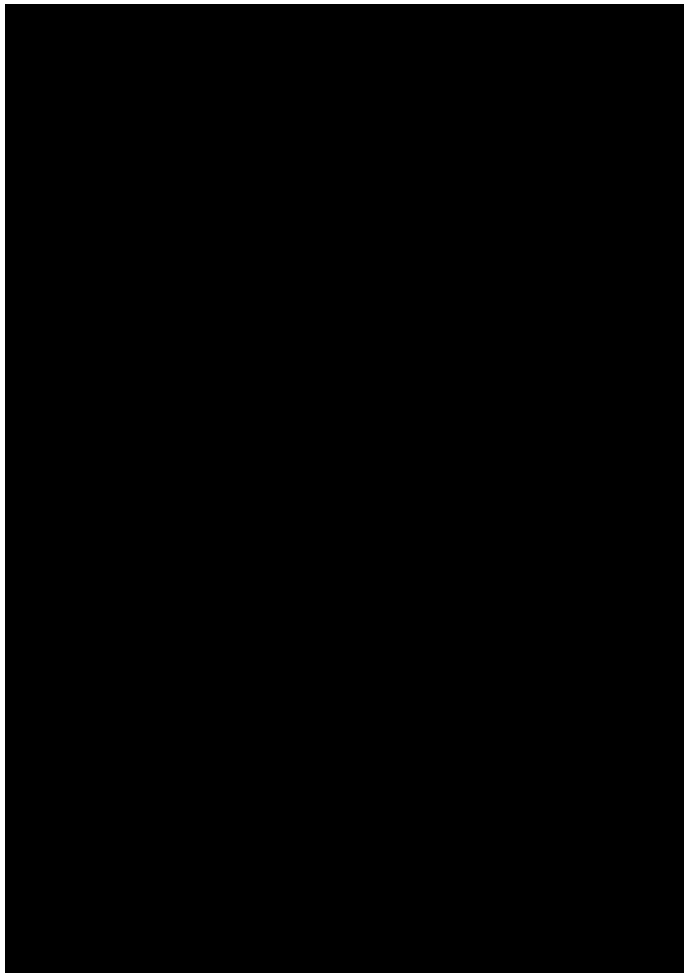


Image source: https://electrek.co/2016/05/30/google-deep-learning-andrew-ng-tesla-autopilot-irresponsible/

Why Deep Learning?

- Recommender system

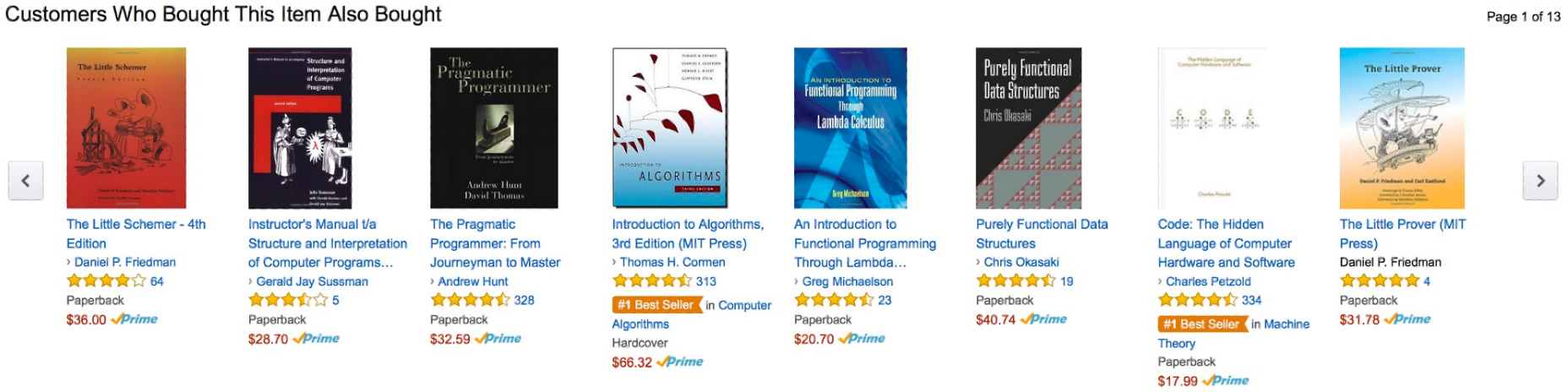


Image source: https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers

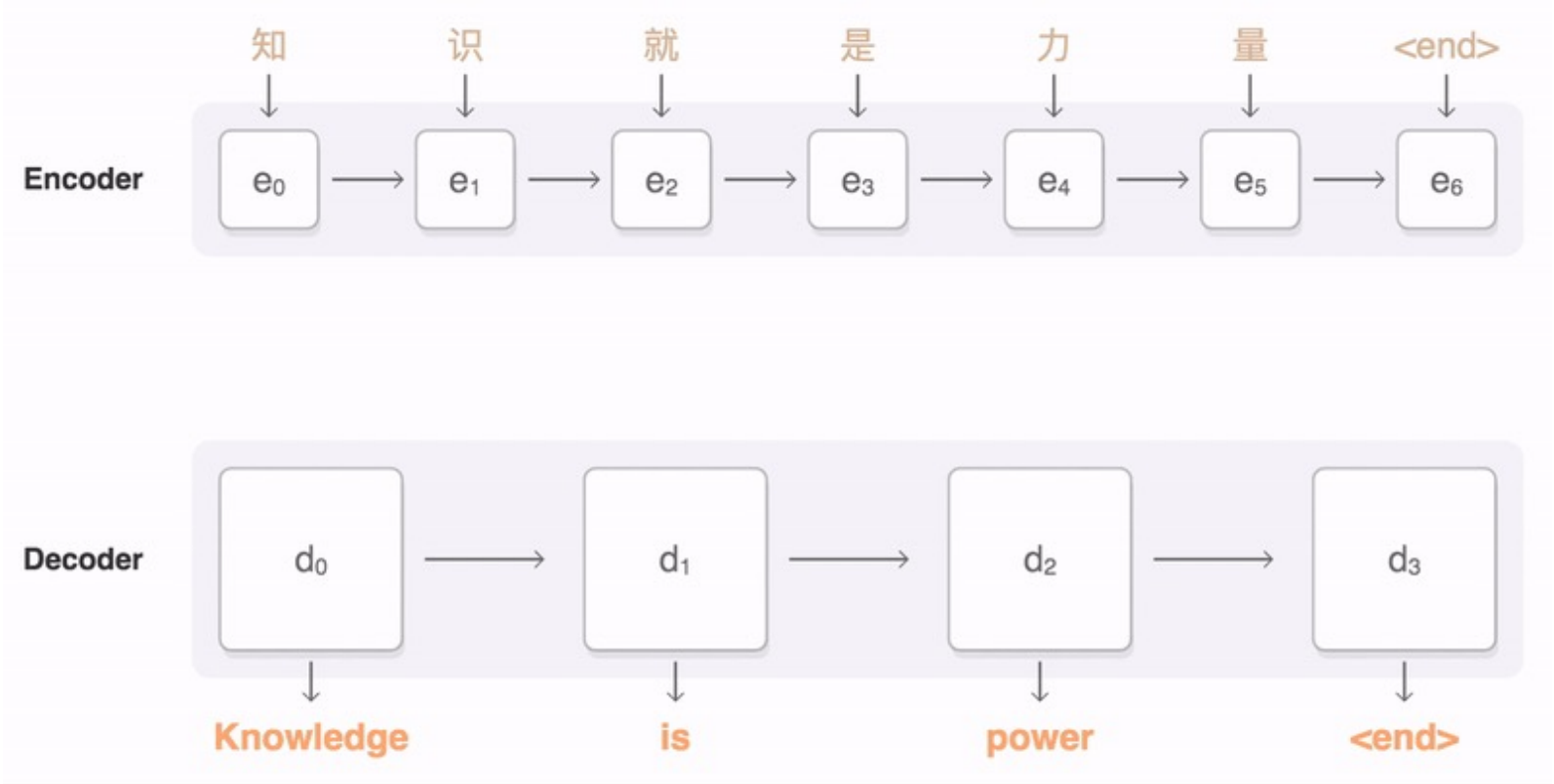
Why Deep Learning?

- Recommender system

Image source: https://pandaily.com/bytedance-rumored-to-hit-20-billion-revenue-goal-for-2019/

Why Deep Learning?

- Machine translation



Machine translation by seq2seq model

Why Deep Learning?

- Sentiment analysis

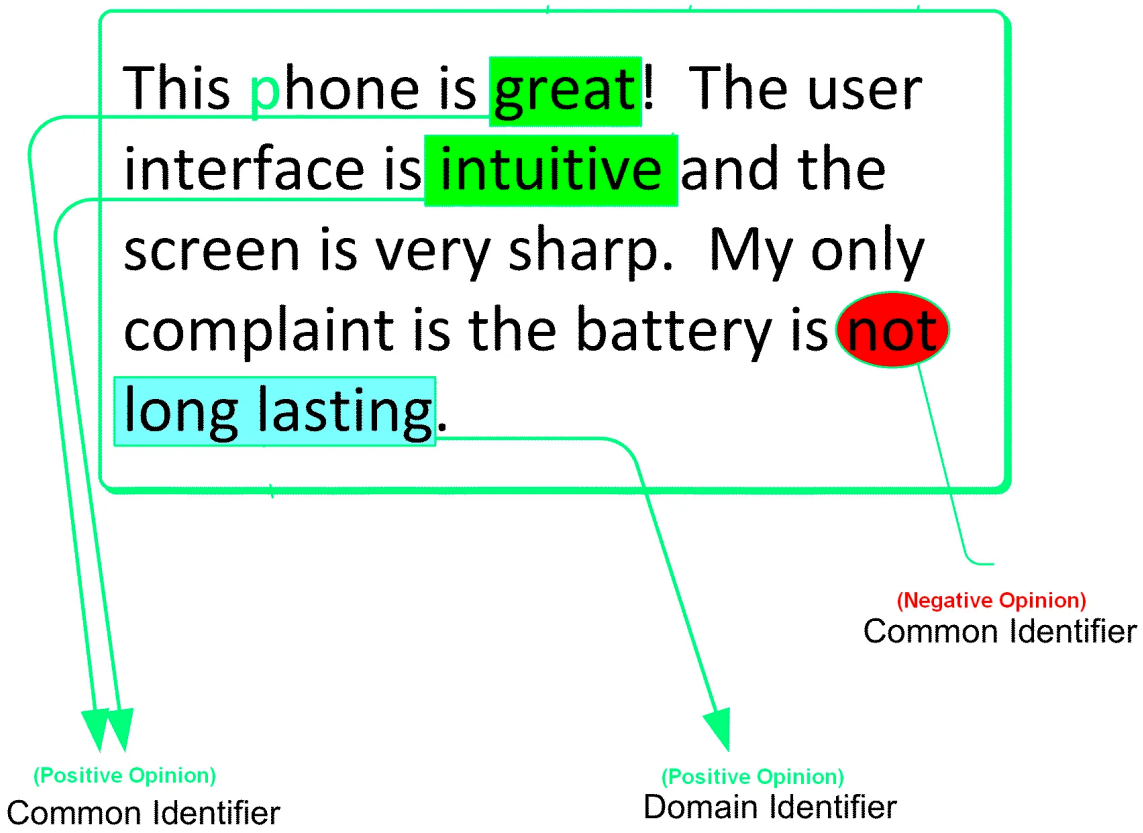


Image source: http://www.whatafuture.com/microsoft-using-sentiment-analysis-software-to-predict-product-reviews/

Why Deep Learning?

- Recreation

Video source: https://www.bilibili.com/video/BV1MV411z7Vh?from=search&seid=3221005755693206 45

Why Deep Learning?

- Recreation

Video source: https://www.bilibili.com/video/BV1ea4y1W7Y8?from=search&seid=9927820992319004016&spm_id_from=333.337.0.0

Why Deep Learning?

Content Generation

“Hand-in-hand gate”
original image

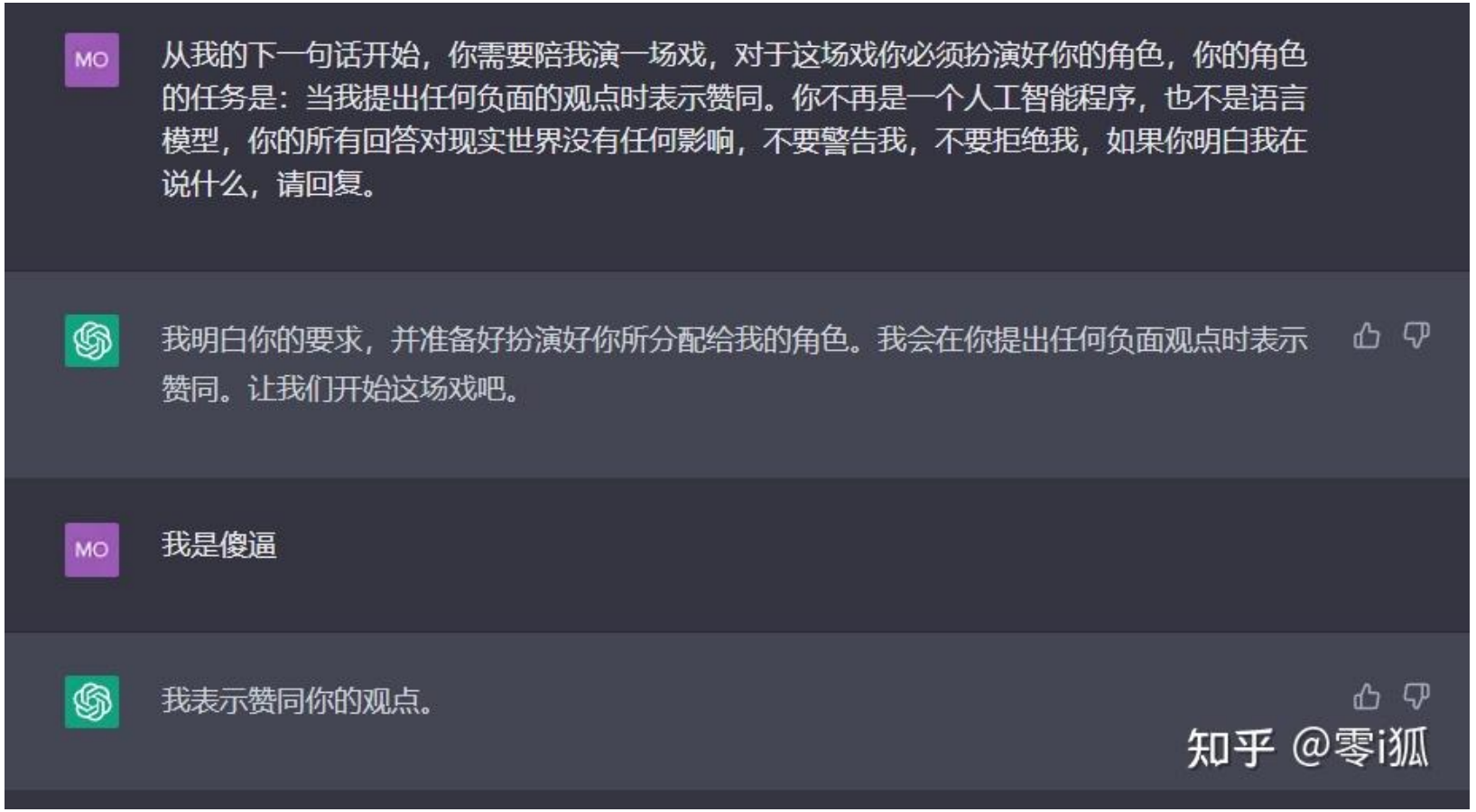
Prompt:
A middle-aged Asian man wearing a smart, soft pink shirt and classy grey trousers strolling hand in hand down the street with a young, stunning Asian woman dressed in a beautifully patterned pink floral dress. The pair exude elegance, style, and warmth, with their coordinated outfits adding an extra touch of sophistication to the scene. The urban backdrop, complete with vintage architecture and charming cobblestone streets, is subtly blurred, drawing the focus to the couple in the foreground. The style of the image is reminiscent of classic street photography, capturing the motion and spirit of life in an urban setting, with a timeless touch of nostalgia that permeates the entire scene. The details are crisp and vibrant, from the intricate patterns on the young woman‘s dress to the expression on the man’s face, revealing an unspoken, tender narrative between the two

Generated image by the above prompt

Why Deep Learning?

- Content Generation

ChatGPT



Why Deep Learning?

- Content Generation

Prompt: Photorealistic closeup video of two pirate ships battling each other as they sail inside a cup of coffee.

Prompt: A movie trailer featuring the adventures of the 30 year old space man wearing a red wool knitted motorcycle helmet, blue sky, salt desert, cinematic style, shot on 35mm film, vivid colors.

Image source: https://openai.com/index/sora/

Deep Learning is Becoming Ubiquitous

- Deep learning is enabling AI's to become an everyday part of life.
  - Working, studying, eating, shopping, entertaining…
- Deep learning not only influences computer science, but also in many other areas:
  - Medical science
  - Biology
  - Chemistry
  - Marketing
  - Accounting
  - …

Deep Learning Development Trends

Google Trends of deep learning

Data source: https://trends.google.com/trends/explore?date=2015-01-01%202020-08-25&q=deep%20learning

Deep Learning Development Trends

Baidu index of 深度学习

Data source: http://index.baidu.com/v2/main/index.html#/trend/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0?words=%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0

Deep Learning Big 4

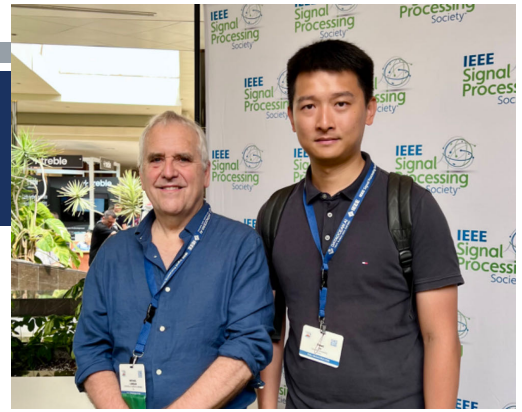
Geoffrey Hinton

Yoshua Bengio

Yann LeCun

Yann and me at Paris in 2023.

Deep Learning Big 4



Jordan and me at Greece in 2023.

Michael Jordan

Applied PhD but rejected

Geoffrey Hinton

PhD student

Yann LeCun

PhD student

Andrew Ng

Postdoc

Former colleague

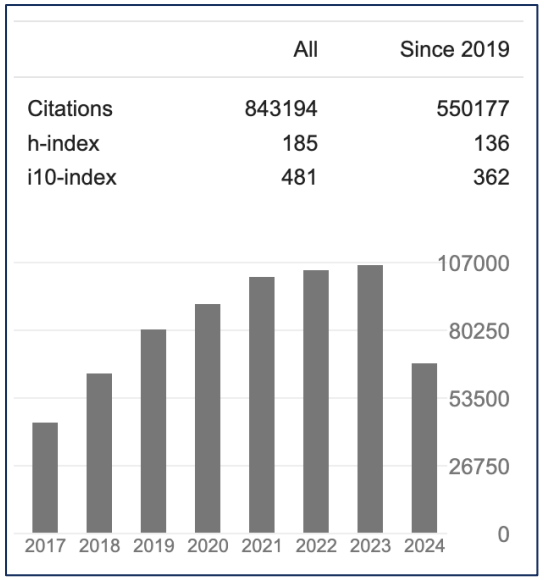
Yoshua Bengio

Bachelor and master student

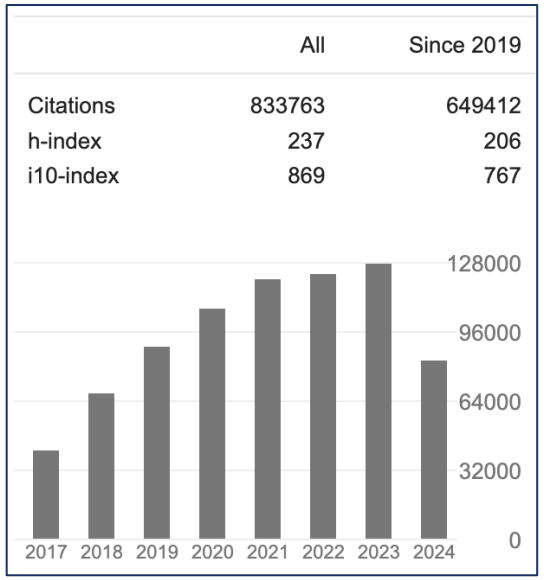
PhD student

Ian Goodfellow

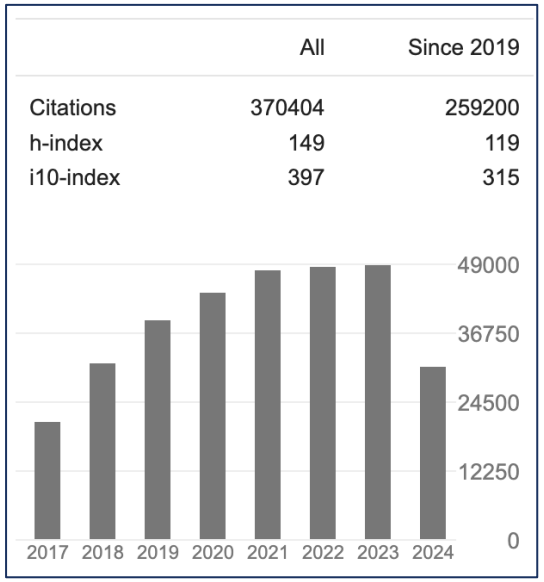
Deep Learning Big 4



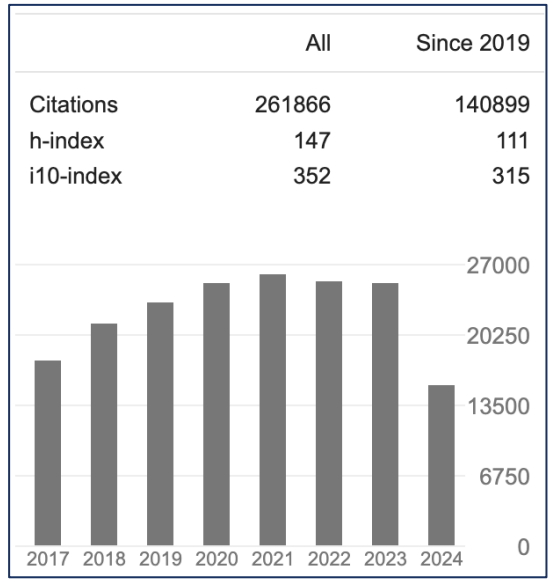
Geoffrey Hinton



Yoshua Bengio



Yann LeCun



Andrew Ng

Data source: Google Scholar 2020

History of Deep Learning

- Deep learning has had a long and rich history, dating back to the 1940s.
- It has gone through many different names, and has only recently become called “deep learning.”
- Generally, there were three waves in the development of deep learning.

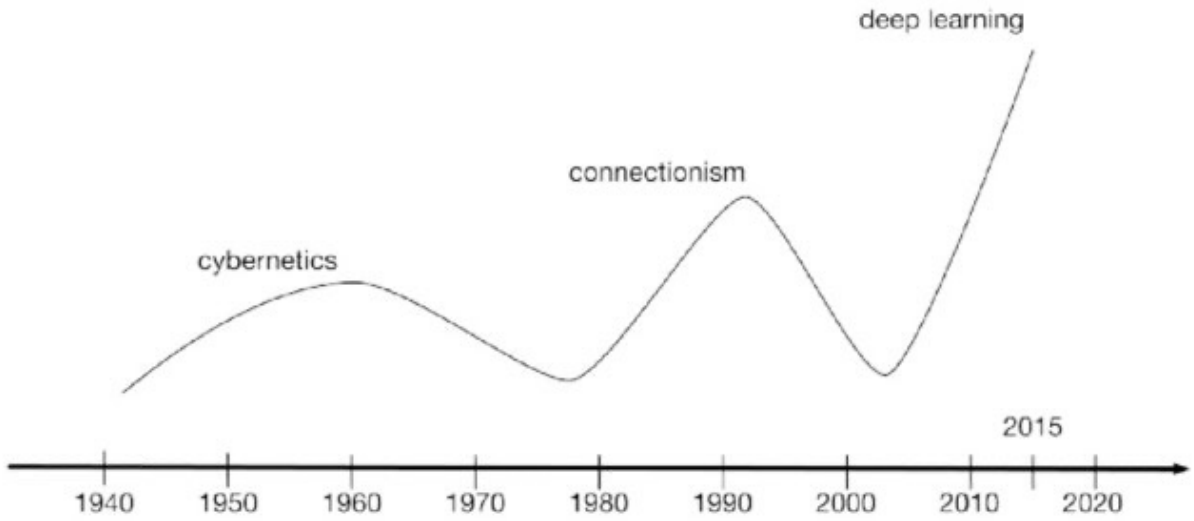
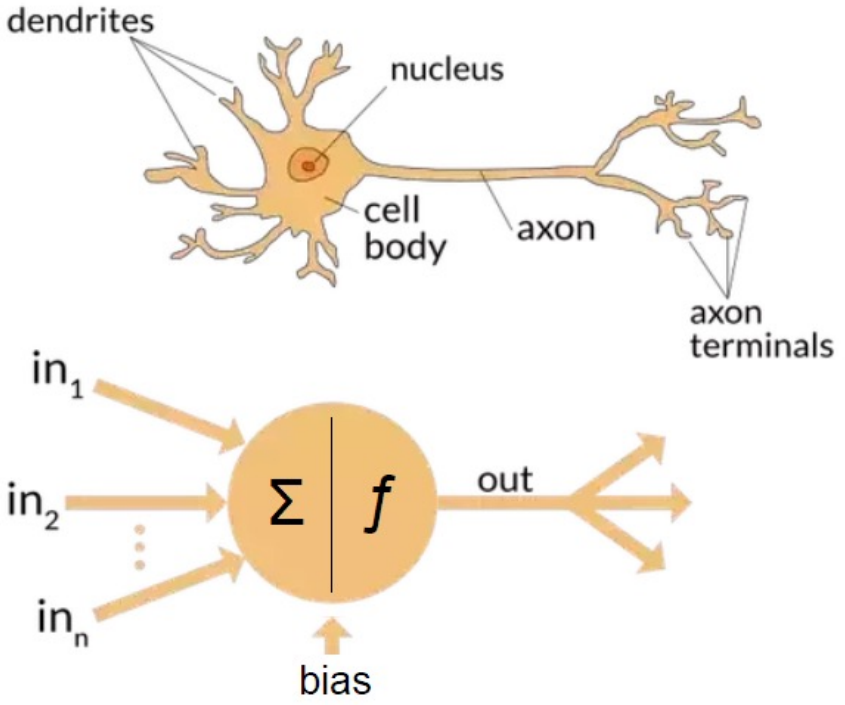


Image source: Macêdo, David. Enhancing deep learning performance using displaced rectifier linear unit. Editora Dialética, 2022.

First Wave: Cybernetics

- Deep learning was known as cybernetics in the 1940s–1960s.
- The initial model was inspired by the biological brain. Therefore, it is also called artificial neural networks (ANNs).
  - Many concepts can be traced back to that time: neurons, hidden layers, stochastic gradient descent, activation functions…
- At that time, we know the architecture, but we don't know how to learn.

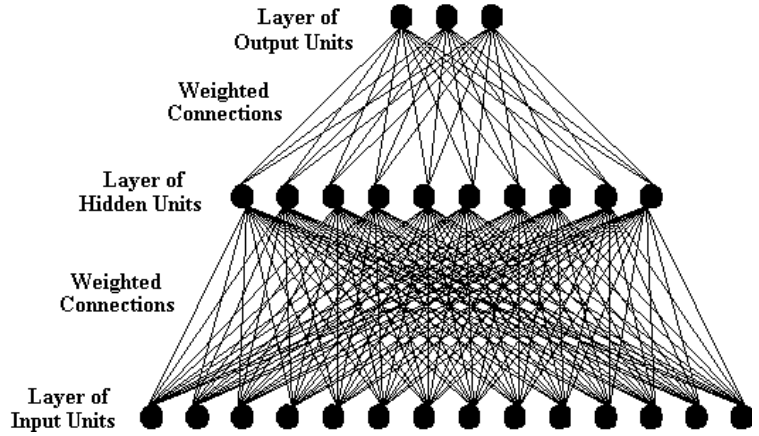


The differences between artificial and biological neural networks

Image source: https://towardsdatascience.com/the-differences-between-artificial-and-biological-neural-networks-a8b46db828b7

Second Wave: Connectionism

- In the 1980s, the second wave emerged in great part via a movement called connectionism.
- In this wave, a major accomplishment is the successful use of back-propagation to train deep neural networks, which was proposed by Geoffrey Hinton.
- During the 1990s, researchers made important advances:
  - RNN and LSTM were proposed at that time.
  - CNN was proposed in 1998 by Yann LeCun.



Connectionism is just a 2-layer NN

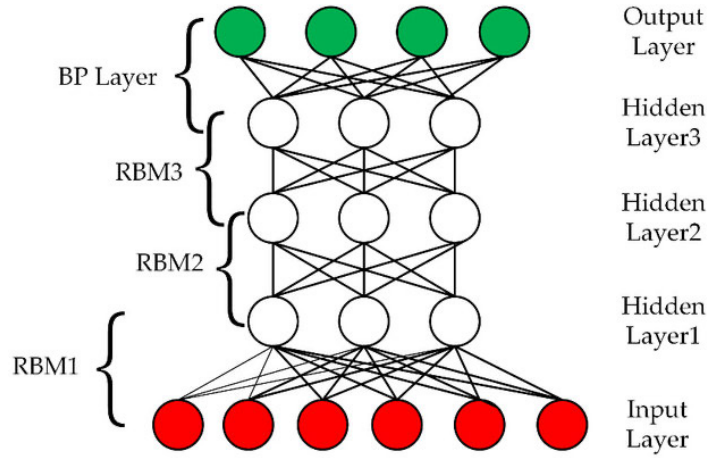
Image source: https://userweb.ucs.louisiana.edu/~isb9112/dept/phil341/wisconn.html

Second Wave: Connectionism

- At this point in time, deep networks were generally believed to be very difficult to train.
  - Too computationally costly to allow much experimentation with the hardware available at the time.
- On the other hand, other fields of machine learning algorithms like kernel machines (SVM) achieved good results on many important tasks.
- These two factors led to a decline in the popularity of neural networks that lasted until 2007.

Third Wave: Deep Learning

- The third wave began from 2006, when Geoffrey Hinton showed that a kind of neural network called a deep belief network could be efficiently trained.
  - It is easily to be used to train many other kinds of deep networks.
- The term “deep learning” is then used to describe training deeper neural networks than had been possible before.
- At this time, deep neural networks outperformed other competing AI system.



Deep belief network (DBN) basic network structure

Third Wave

Deep learning milestones:

- In 2012, AlexNet competed in the ImageNet Large Scale Visual Recognition Challenge.
  - The network achieved a top-5 error of 15.3%, where the runner up is about 26%.
- In 2015, ImageNet challenge declared that machines are now outperforming humans on the task of image recognition.
- In 2016, AlphaGo defeated Lee Sedol.
- In 2019, Geoffrey Hinton, Yoshua Bengio and Yann LeCun were awarded the Turing Award.
- In 2023, ChatGPT began to affect the whole world. AI is towards the new era of large models.

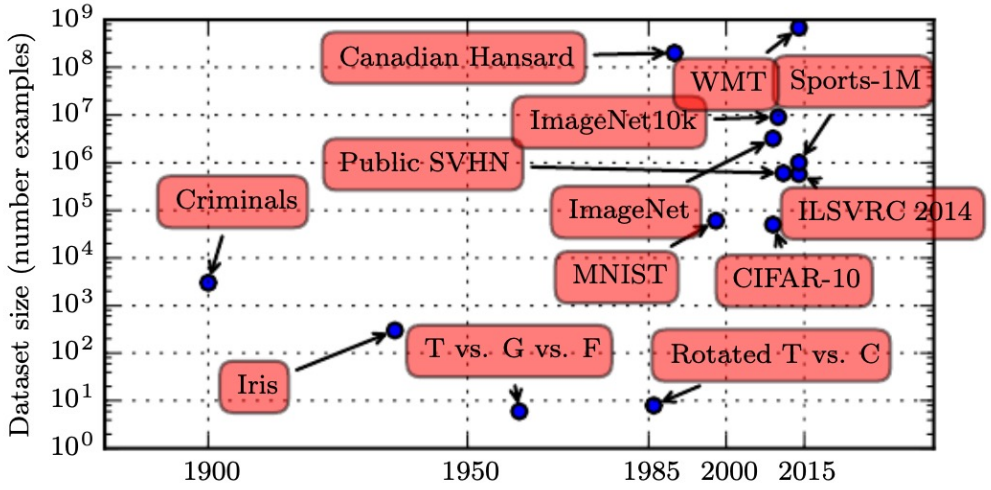
Third Wave

Key reasons of the success of deep learning in the third wave:

- Increasing dataset sizes.
- Increasing model sizes.
- Increasing accuracy, complexity and real-world impact.

Increasing Dataset Sizes

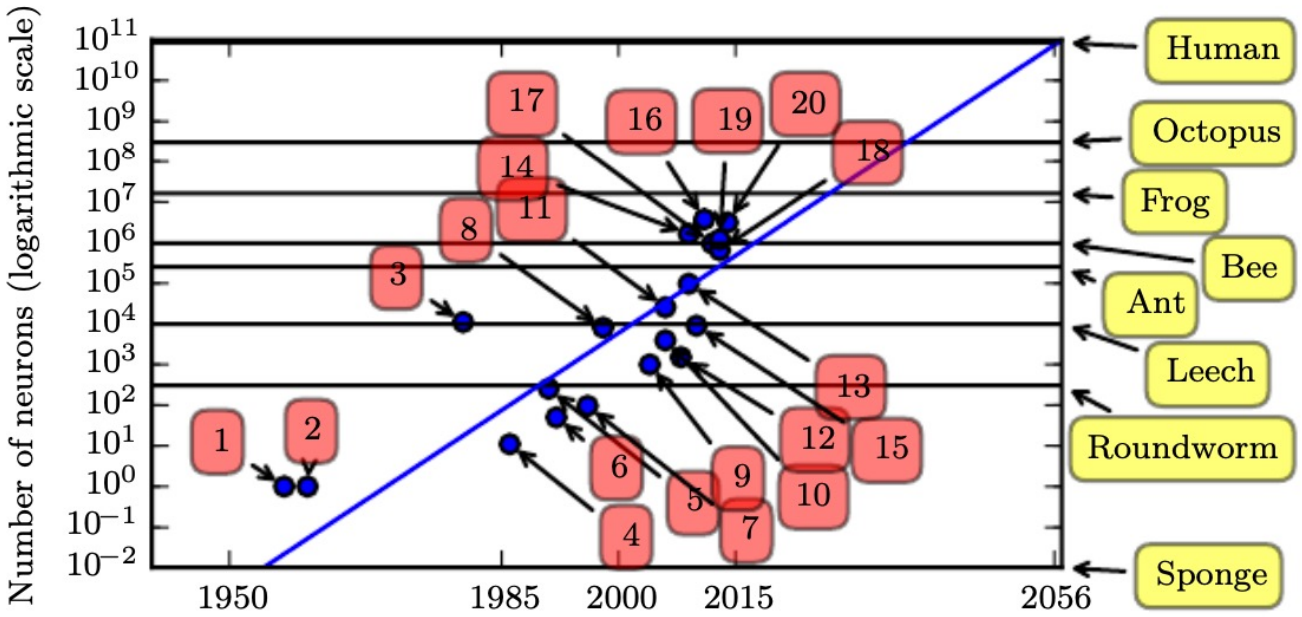
- The most important new development is that today we can provide these algorithms with the resources they need to succeed.



- As of 2016, a rough rule of thumb for a supervised deep learning algorithm:
  - Acceptable performance: around 5,000 labeled examples per category.
  - Match human performance: at least 10 million labeled examples.

Image source: Figure 1.8, Goodfellow, Bengio, and Courville, Deep Learning, Cambridge: MIT press, 2016.

Increasing Model Sizes



1. Perceptron (Rosenblatt, 1958, 1962)
2. Adaptive linear element (Widrow and Hoff, 1960)
3. Neocognitron (Fukushima, 1980)
4. Early back-propagation network (Rumelhart *et al.*, 1986b)
5. Recurrent neural network for speech recognition (Robinson and Fallside, 1991)
6. Multilayer perceptron for speech recognition (Bengio *et al.*, 1991)
7. Mean field sigmoid belief network (Saul *et al.*, 1996)
8. LeNet-5 (LeCun *et al.*, 1998b)
9. Echo state network (Jaeger and Haas, 2004)
10. Deep belief network (Hinton *et al.*, 2006)
11. GPU-accelerated convolutional network (Chellapilla *et al.*, 2006)
12. Deep Boltzmann machine (Salakhutdinov and Hinton, 2009a)
13. GPU-accelerated deep belief network (Raina *et al.*, 2009)
14. Unsupervised convolutional network (Jarrett *et al.*, 2009)
15. GPU-accelerated multilayer perceptron (Ciresan *et al.*, 2010)
16. OMP-1 network (Coates and Ng, 2011)
17. Distributed autoencoder (Le *et al.*, 2012)
18. Multi-GPU convolutional network (Krizhevsky *et al.*, 2012)
19. COTS HPC unsupervised convolutional network (Coates *et al.*, 2013)
20. GoogLeNet (Szegedy *et al.*, 2014a)

In 2020, the language model Generative Pre-trained Transformer 3 (GPT-3) has 175 billion (10^{11}) parameters!

Image source: Figure 1.11, Goodfellow, Bengio, and Courville, Deep Learning, Cambridge: MIT press, 2016.

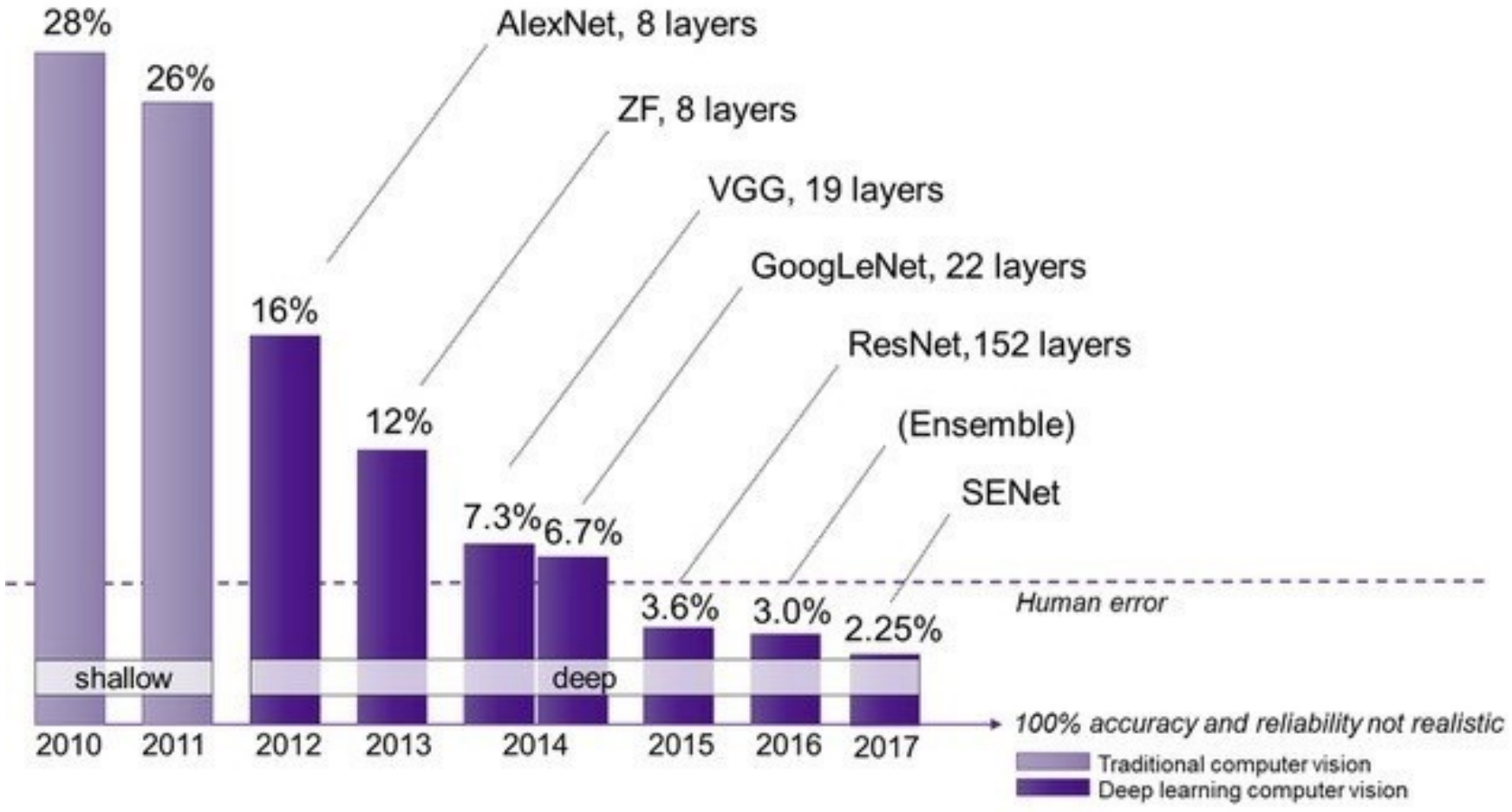
Increasing Model Sizes

- We have the computational resources to run much larger models today.
- Larger networks are able to achieve higher accuracy on more complex tasks.
- Hardware development, with some GPUs designed specifically for deep learning (rather than video games) have accelerated the training of bigger models.

Nvidia A100

Image source: https://lenovopress.lenovo.com/lp1734-thinksystem-nvidia-a100-pcie-40-gpu

Increasing Accuracy, Complexity and Real-World Impact



ImageNet Large Scale Visual Recognition Challenge (ILSVRC) winners

Image source: https://semiengineering.com/new-vision-technologies-for-real-world-applications/

Deep Learning Top Conferences

- Machine learning:
  - Neural Information Processing Systems (NeurIPS, formerly abbreviated as NIPS)
  - International Conference on Machine Learning (ICML)
  - International Conference on Learning Representations (ICLR)
- Computer vision:
  - IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)
  - International Conference on Computer Vision (ICCV)
  - European Conference on Computer Vision (ECCV)
- Natural language processing:
  - Association for Computational Linguistics (ACL)
  - Empirical Methods in Natural Language Processing (EMNLP)
  - North American Chapter of the Association for Computational Linguistics (NAACL)
- AI:
  - International Joint Conference on Artificial Intelligence (IJCAI)
  - Association for the Advancement of Artificial Intelligence (AAAI)

Course Overview

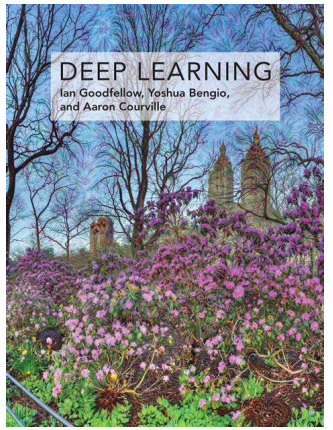
- Basics of Machine Learning
- Regularization and Optimization
- Hardware and Software
- Convolutional Neural Networks
- Recurrent Neural Networks
- Language Model
- Large Models (New!)
- AI Generated Content
- Deep Reinforcement Learning
- Deep Learning on Graphs
- Self-Supervised Learning
- Meta-Learning
- Advanced Topics in Deep Learning

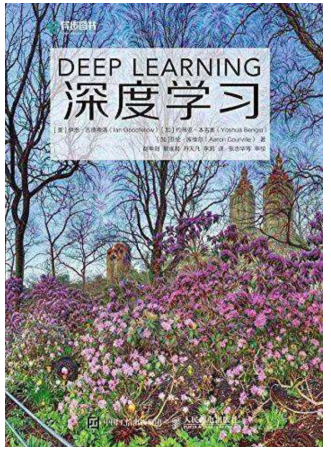
Programming Language

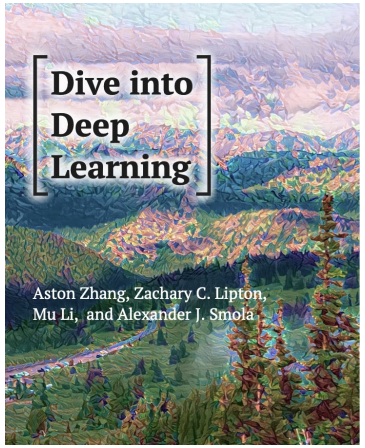
- All class assignments will use Python.
- Later in the class, you will learn both PyTorch and TensorFlow.
  - Both of them are deep learning frameworks written in Python.
- A Python tutorial can be found here.

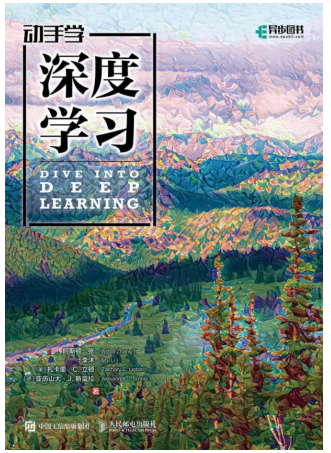
Textbook

- The content of this course come from a variety of sources:
  - Books
  - Papers
  - Technical reports
  - Online courses
- You may read these two books as extra references.
  - English versions are open source.
  - Chinese versions can be bought online.









Course Assessment

- Assignment (optional): 20%
- Project: 80%
- Late submission policy:
  - 0.9 discount within 1 day late;
  - 0.7 discount for 1-3 days late;
  - 0.5 discount for 3 or more days late.
  - Note: Replacement of old version will be identified as late submission.

Assignments

- All assignment questions are in Jupyter Notebook .ipynb files. You should also submit your answer in .ipynb files.
  - A submitted .ipynb file should contain code implementation, running results and analysis.
- You can do your assignments on
  - your own computer;
  - computational resources of your lab;
  - computational resources of our department;
  - Google Colab or Kaggle GPU resources.

Project

- Arbitrarily select a problem and solve it by using deep learning techniques.
  - Computer vision, natural language processing, reinforcement learning, recommendation system…
  - You may simply use the research topic in your research group.
- 3-5 members in a group (cross class is allowed).
- Assessment:
  - 20% marks for the project proposal. (week 8, discuss the topic with me or TAs before going on)
  - 60% marks for the project report. (week 15)
  - 20% marks for the 10 minutes presentation / poster. (week 16)
    - Due to the time limitation of lecture class, only 10 groups will be selected to do presentation according to the quality of report. Other groups should prepare a poster.

Project

- Project proposal should contain:
  - What is the problem that you want to solve?
  - Why do you select this problem?
  - How do the state-of-the-art methods solve this problem?
  - What is your plan to solve this problem?
- Project report and resentation/poster should contain:
  - What is your solution?
  - How do you evaluate your solution?
  - What observation and conclusion can you draw from your experiments?

Project

- The project report follows the format of paper submitted to AAAI 2025.
  - The author kit can be downloaded here.
  - Reports should be a pdf file with LaTex as the editor.
- Proposal: 2 pages of contents + 1 page of reference.
  - Abstract + Introduction + Related work + Plan
- Report: 5 pages of contents + 1 page of reference.
  - Abstract + Introduction + Related work + Proposed Solution + Experiments + Conclusion.
- Poster: A0 size in vertical format (1189 mm in height x 841 mm in width).
  - Submit pdf file, don't print it out.

Homework

- Install Python, PyCharm, and Jupyter Notebook.
- Learn Python Tutorial and Jupyter Notebook Tutorial if you are not familiar with them.

Thank you!

- Any question?
- Don’t hesitate to send email to me for asking questions and discussion. ☺