AI Engineer Training: I
In the Era of Deep Learning

IT21 Learning
Alvin Jin

Instructor Introduction

- Master of Computer Science @UWaterloo
- Director of Data Science
- 8 yrs industry experience with 4 paper published
- Research on AI and Big Data

Course Contents

- Total 10 sessions, 20 hours.

- AI Preliminaries
- Machine Learning Principles
- Machine Learning Practices
- Intro to Artificial Neural Network
- Deep Learning for Images I
- Deep Learning for Images II
- Deep Learning for Images III
- Deep Learning for Text I
- Deep Learning for Text II
- Advanced Topics

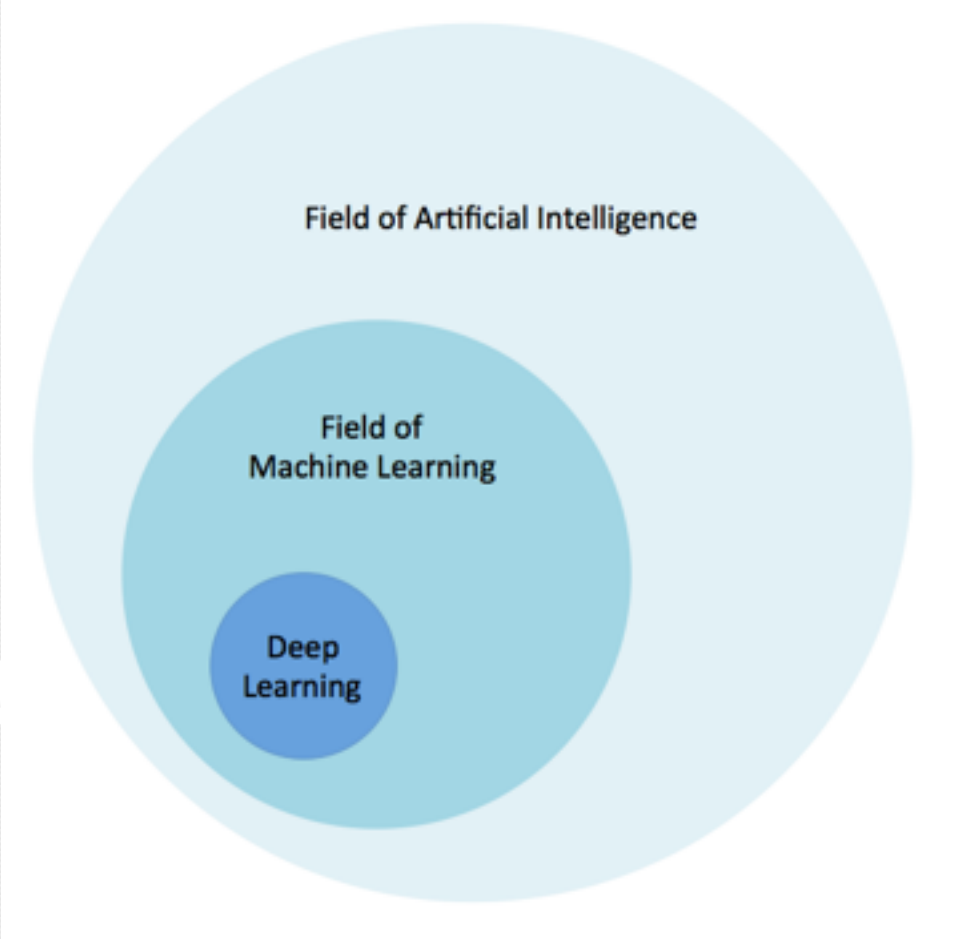
Course Objects

- Read AI books, articles and tutorials without obstacles
- Freely discuss AI relevant topics with your friends, colleagues and interviewers
- Work on your own AI side-projects to solve real-life problems
- Confidently apply AI opportunities inside/outside your organizations

Agenda

- Introduction to Artificial Intelligence
- Essential Mathematics for AI
- Introduction to Machine Learning

AI vs. ML vs. DL


Field of Artificial Intelligence
Field of
Machine Learning
Deep
Learning

AI vs. ML vs. DL

- Artificial intelligence
  - The effort to automate intellectual tasks normally performed by humans.
- Machine Learning
  - Provides systems the ability to automatically improve from data, without being explicitly programmed.
- Deep Learning
  - A specific subfield of machine learning
  - Emphasizes on learning successive layers of increasingly meaningful representations.

AI Evolutionary

- 1950s: Alan Turing shapes AI concepts
  - Published "Computing Machinery and Intelligence"
- 1980s: Symbolic AI is the dominant paradigm
  - Handcraft a large set of explicit rules for manipulating knowledge.
- 1990s: Machine learning starts to flourish.
  - Driven by faster hardware and larger datasets
- 2010s: Deep learning wows the world
  - Alex Krizhevsky won LSVRC; AlphaGo beat Master Lee Sedol

AI in Canada

- Tech giants like Google, Microsoft, Facebook, Uber, etc. built AI labs in Toronto, Montreal, Edmonton
- Government invested $125 million on AI research in March 2017. e.g. Vector, MILA, etc.
- Hundreds of local AI start-up companies was built in last 3 years.
- Large Canadian companies also start to invest in AI. e.g. RBC's Borealis AI, etc.

AI in China

- 2017 July, the State Council published white paper to make China the global AI leader by 2030 with industry worth US$150 billion.
- Investing $7 billion through 2030, including $2 billion for a research park in Beijing.
- Dominated global AI funding 48%, comparing US 38% in 2017
- Total AI companies, China hosts 23% comparing US 42% in 2017

AI in Educations

- USTC, NJU built School of Artificial Intelligence 2017.
- CMU offers undergraduate degree in AI in 09, 2018.
- UWaterloo opened AI research institute in April.

Canadian AI Pioneers



- Yann LeCun Geoffrey Hinton Yoshua Bengio

- 见“AI三巨头”，李开复的北美AI见闻录
- https://36kr.com/p/5107475.html

Artificial Neural Network

- Initially investigated in the 1950s. Started in 1980s.
- **Not** real models of the brain
- **Loosely inspired** by the neurobiology research

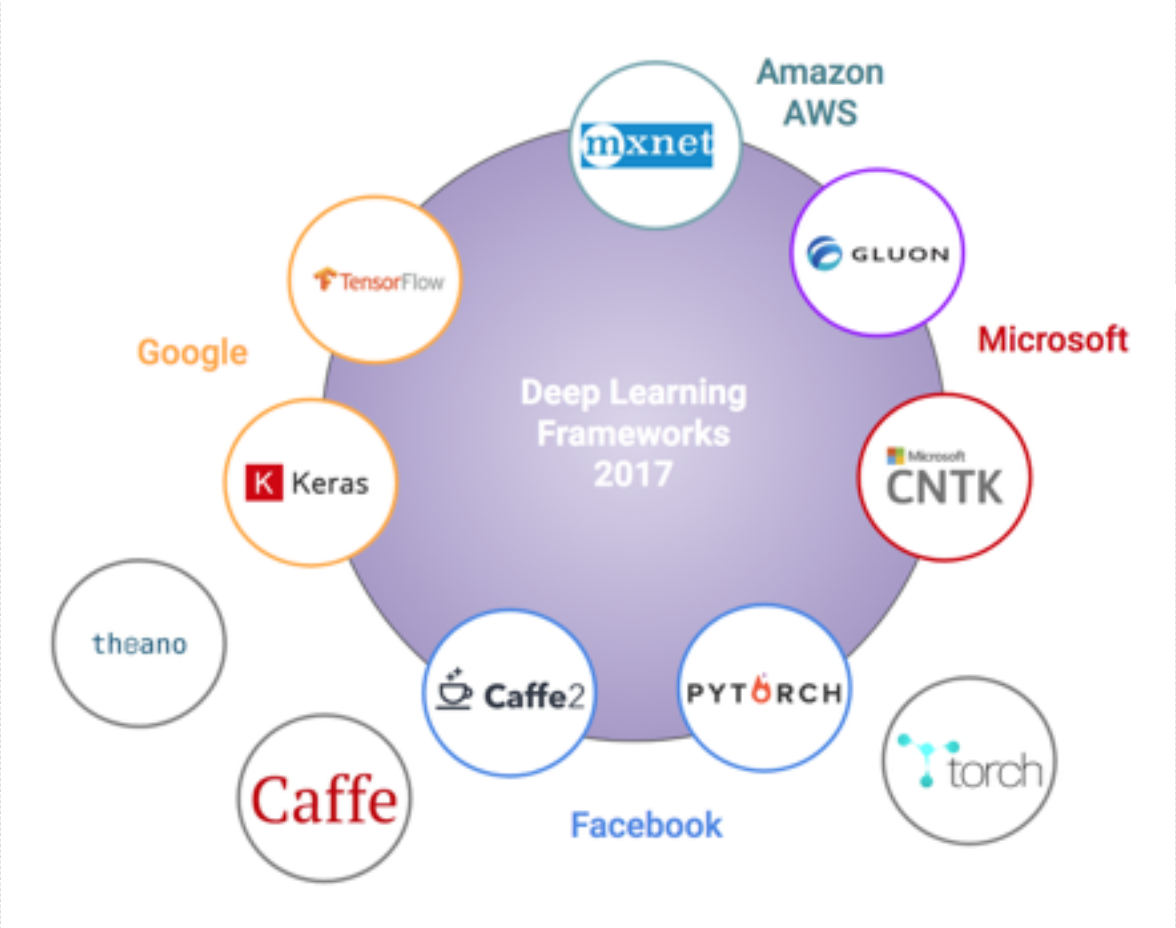
How ‘Deep’ is Deep Learning?

- Deep Learning is a rebrand of Artificial Neural Network with more than two layers.
- “Deep” isn’t a reference to deeper understanding achieved by the approach
- It stands for the idea of successive layers of representations.

Teaching “the whole game”

- David Perkins, Harvard education professor with PhD from MIT in AI
- Good education starts with “the whole game”
- Good education equips you to work on the questions you care about
- Good education motivates the study of underlying technical concepts.

Deep Learning Frameworks



Environment & Tools

- AWS cloud service for Deep Learning
- Python 3, Jupyter
- Numpy: the fundamental library for scientific computing
- Tensorflow: library for numerical computation.
- Keras: high-level neural networks API, capable of running on top of TensorFlow, CNTK, or MXNet.

Essential Mathematics for AI

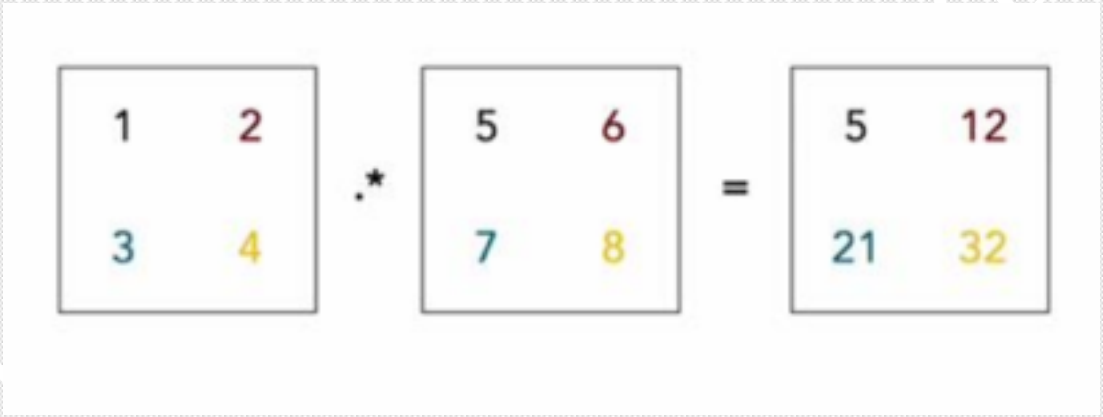
- Linear Algebra: Matrices operations
- Calculus: Differentiation, gradient descent
- Statistics: Probability

Tensors

- Tensor:
  - a container for numerical data
  - a multidimensional array
  - basic data structure for machine learning

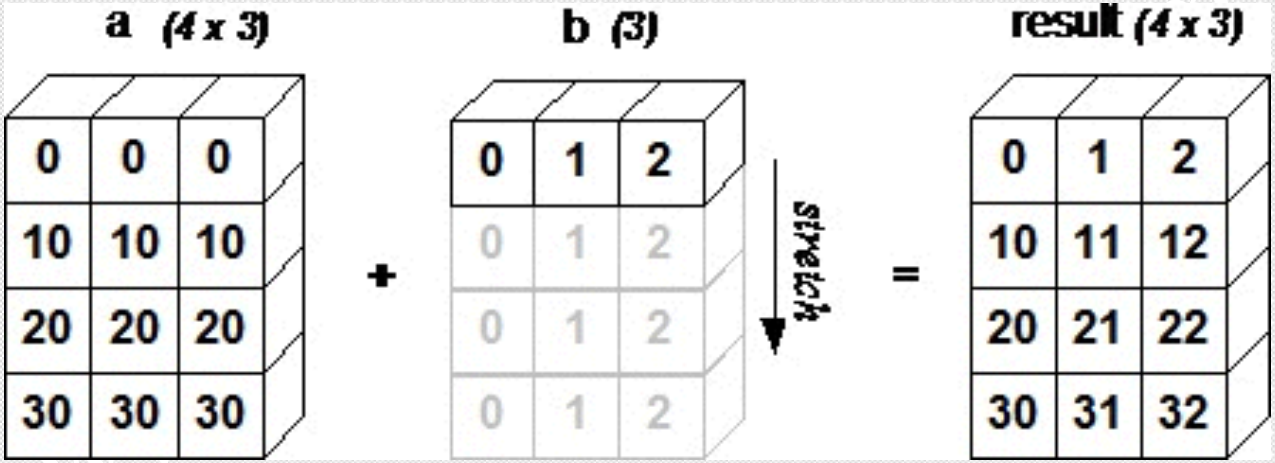
Tensor Operations

- Element-wise product
  - multiply each element of one with each element of the other



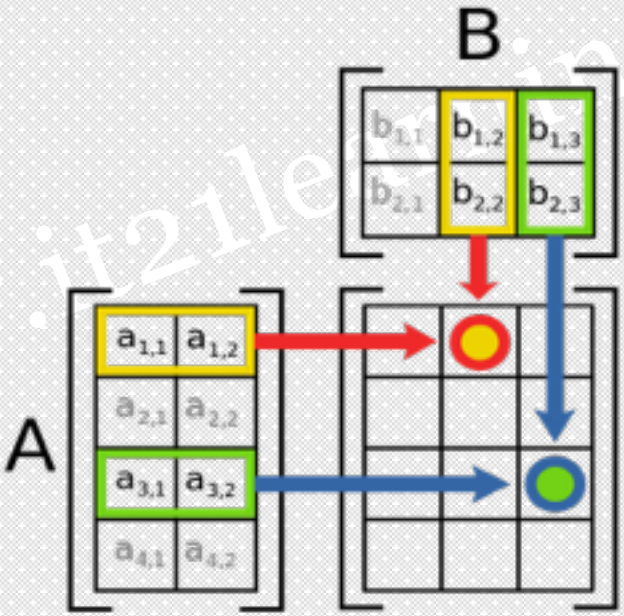
Broadcasting

- Smaller tensor is broadcasted to match the shape of larger tensor



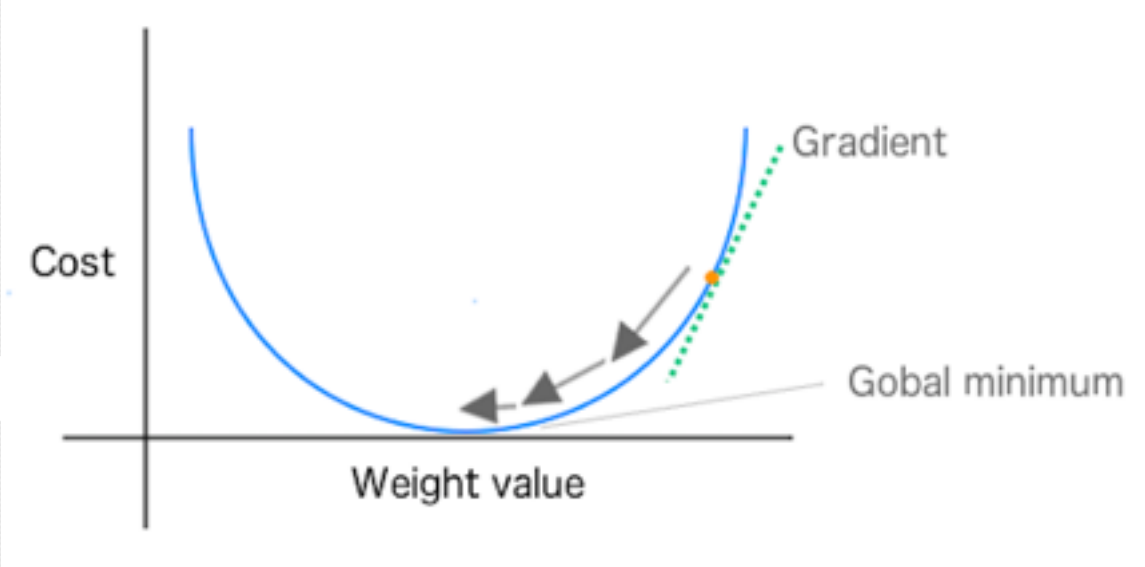
Dot Product Example

- The number of columns of the first matrix must be equal to the number of rows of the second matrix.
- A(m, n) x B(n, k) = C(m, k)



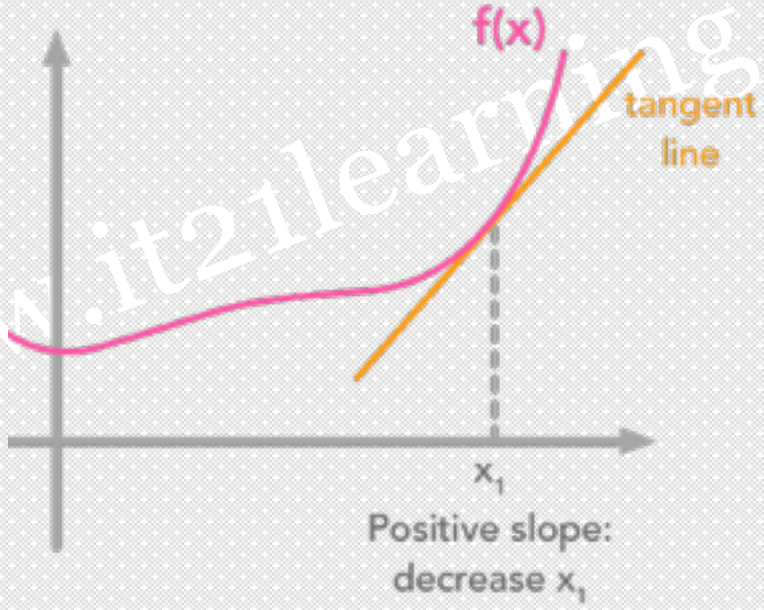
Derivative

- Derivative measures “rate of change” of a function
- Differentiable: Smooth, continuous functions can be derived



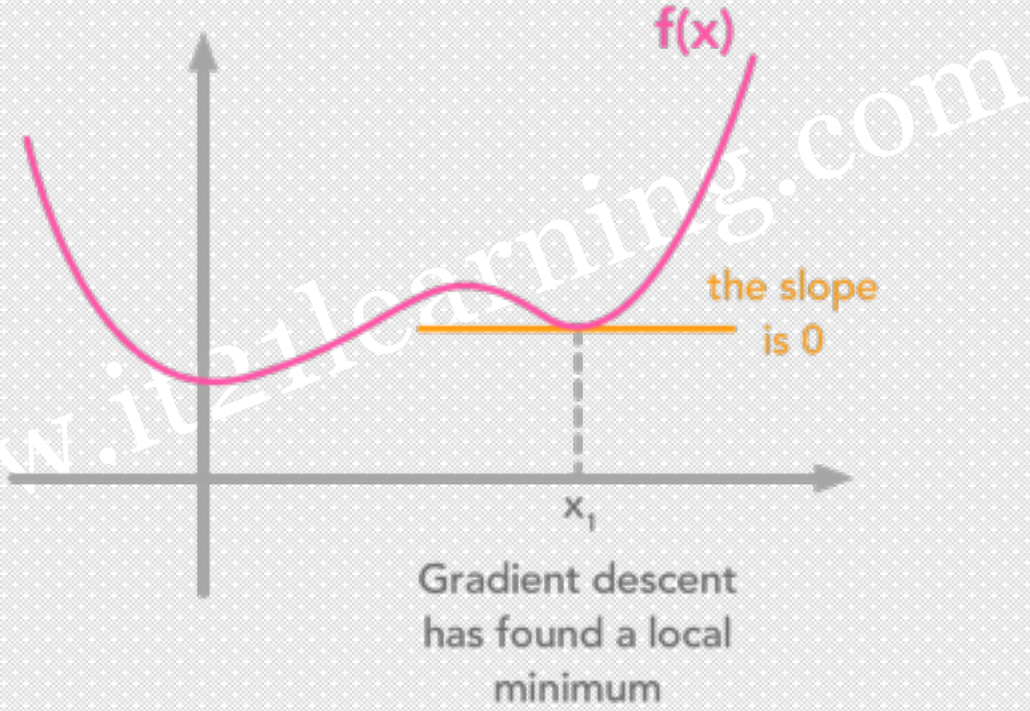
Gradient Descent

- Gradient: the derivative of a tensor operation in multi-dimensions.
- Calculate the slope of the loss function by taking a derivative.

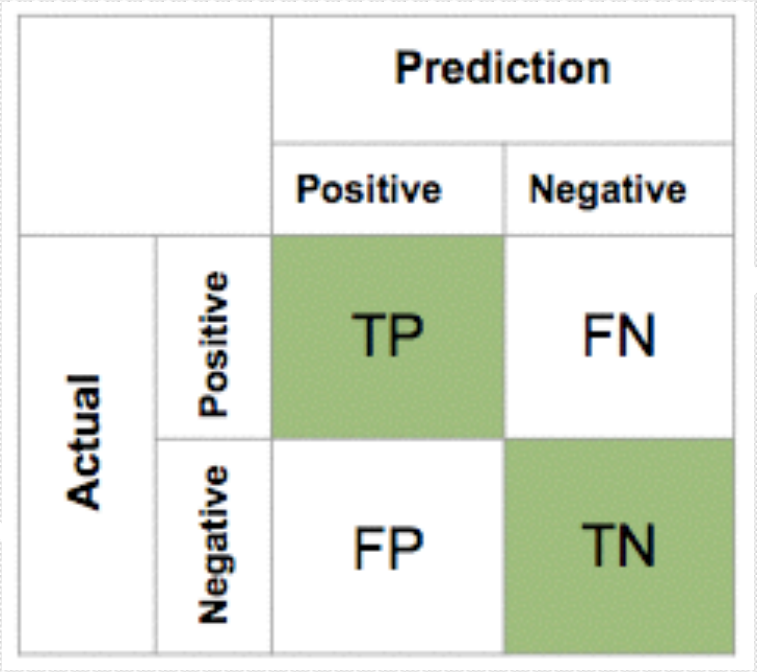


Local Minimum

- The cost function tells us how far we are from the global minimum.



Confusion Matrix



Accuracy vs. Precision

- Accuracy: the closeness to quantity's true value

$$acc = \frac{TP + TN}{TP + TN + FP + FN}$$

- Precision: positive prediction value

$$p = \frac{TP}{TP + FP}$$

- A measurement is valid if it is both accurate and precise.