Neural Network and Deep Learning

Lecture 1

Yanwei Fu
School of Data Science, Fudan University

Logistics

Neural Networks are taking over!

Optional subtitle

- Neural networks have become one of the major thrust areas recently in various pattern recognition, prediction, and analysis problems
- In many problems they have established the state of the art

  Often exceeding previous benchmarks by large margins

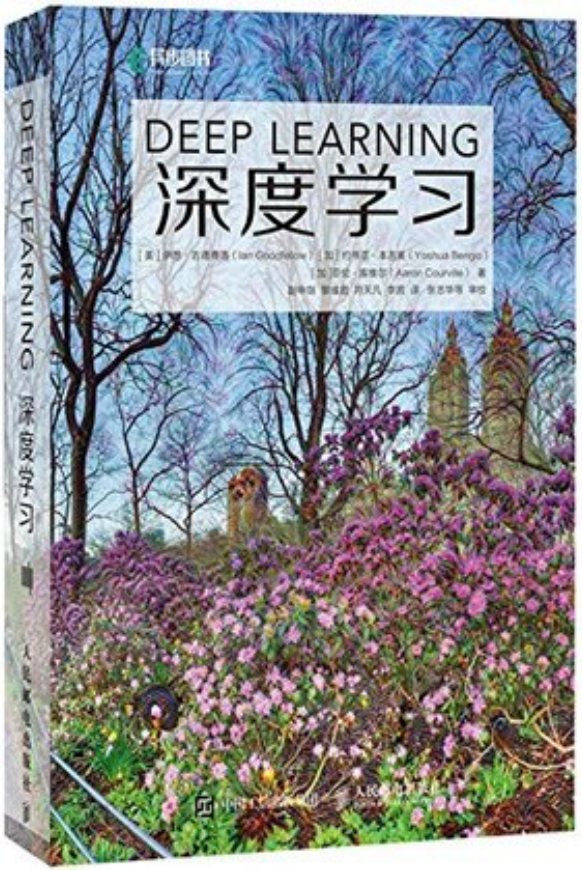
Course Information

Optional subtitle

- Instructor: 付彦伟
- Email: yanweifu@fudan.edu.cn
- Course Websites: http://yanweifu.github.io/courses/Neural_network/index.html
- Times&Venue:
  - Classroom:
  - H6406 (三 6-8 [1-2],[4-5],[7-8],[10-11],[13-14],16)
  - Hxinjinbo-1309 (三 6-8 3,6,9,12,15)
- Office Hours: Wed. 4:00-5:30pm, Xin jinbo building, 1506
- TAs: Chengming Xu (email: dlcourse.xcm@gmail); Yikai Wang (email: yikaiwang19@fudan.edu.cn); Boyan Jiang (email:)

Textbooks

Optional subtitle



Ian Goodfellow, Yoshua Bengio, Aaron Courville

《神经网络与深度学习》

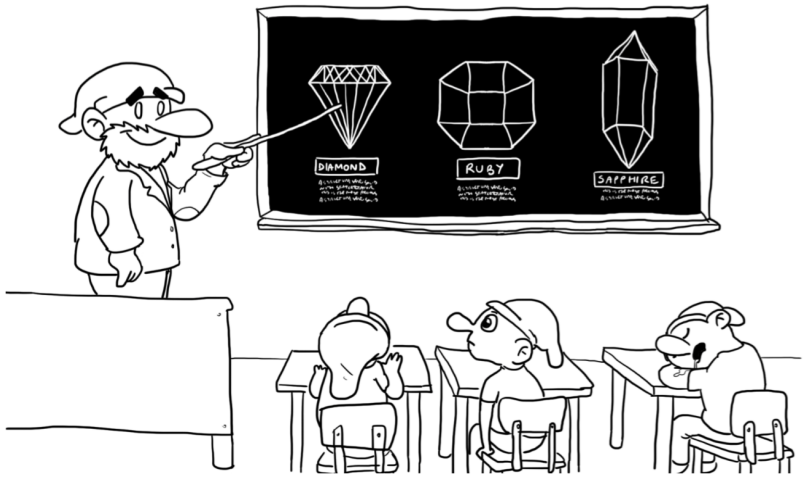
Neural Networks and Deep Learning

邱锡鹏
xpqiu@fudan.edu.cn

2019 年 11 月 21 日

Course Requirement

- Prerequisites on Math:
  - Basic linear Algebra/Calculus: vectors, matrices, eigenvalues;
  - Probability: conditional probability, expectations;
  - Multivariate calculus: gradients, optima;
- Prerequisites on programming:
  - Data structures: pointers, trees, heaps, hash maps, graphs;
  - Scientific computing: matrix factorisation .



Course Work

- Final scores=
  - +Class attendance/discussion (10%);
  - +Homework Assignment (10%)
  - +Projects (40%); 3-4; 每个人独立完成
  - +Final Project (40%): hit NeurPIS submissions; 每个队最多两个人。
- Courses:
  - We will give some tutorial.
- Pain and Happiness
  - Huge efforts to code, debug, read and think;
  - Worth doing it!!

Academic Integrity（学术诚信)

- **Academic integrity** is the moral code or ethical policy of academia. This includes values such as avoidance of cheating or plagiarism; maintenance of academic standards; honesty and rigor in research and academic publishing. (https://en.wikipedia.org/wiki/Academic_integrity)
- No cheating and plagiarism,
  - How to define *Plagiarism*? We follow ACM Policy on Plagiarism.
  - 抄袭和被抄袭双方的成绩都将被取消.
  - 作业、报告、期末论文的署名原则：署你名字的工作必须由自己完成；允许讨论，但作业必须独立完成，并在作业中列出所有参与讨论的人。不允许其他任何形式的合作——尤其是与已经完成作业的同学“讨论”。
  - 这是学术底线。

Questions?

Optional subtitle

Post on eLearning

Machine Learning & AI

人工智能研究的主要方法：

1，**符号主义方法**：认知是一种符号处理过程，人类思维过程可用符号来描述，思维就是计算，这种思想一度构成了人工智能的基础理论。

代表人物：司马贺（西蒙，Herbert Alexander Simon）和纽厄尔（Allen Newell），物理符号系统，1975 年图灵奖获得者。

2，**联结主义方法**：模拟人的智能要依靠仿生学，特别是需要模拟人脑，建立脑模型。人类思维的基本单元是神经元，而不是符号，智能是相互联结的神经元竞争与协作结果。

代表人物：麦卡洛克（Warren McCulloch), 皮茨（Walter Pitts)提出的神经元的数理模型。

3，**行为主义方法**：模拟人在控制过程中的智能行为和作用，研制所谓的控制论动物。

代表人物：博德（H.W.Bode) 和埃文斯（W.R.Evans)等。

Resources—Conferences

Machine Learning:

- Neural Information Processing Systems (NeurIPS)
- International Conference on Machine Learning (ICML)
- European Conference on Machine Learning (ECML)
- Uncertainty in Artificial Intelligence (UAI)
- Computational Learning Theory (COLT)
- International Conference on AI & Statistics (AISTATS)

AI In general

- AAAI Conference on Artificial Intelligence (AAAI) (AAAI: Association for the Advancement of Artificial Intelligence);
- International Joint Conference on Artificial Intelligence (IJCAI);

Pattern Recognition&Computer Vision:

- European Conference on Computer Vision (ECCV)
- IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
- IEEE International Conference on Computer Vision (ICCV)
- ICPR/ACCV/BMVC;

- Data Mining:
  - ACM SIGKDD (Knowledge discovery and Data Mining);
  - ACM SIGIR/ICDM;

Natural Language Processing:

- ACL/EMNLP/COLING

Resources—Journals

Machine Learning&AI:

- Journal of Machine Learning Research (JMLR)
- IEEE Trans on Pattern Analysis and Machine Intelligence (TPAMI);
- Artificial Intelligence;
- International Journal of Computer Vision (IJCV);

Statistics:

- The Annals of Statistics;

Pattern Recognition &NN:

- Neural Computation
- Neural Networks;
- IEEE Transactions on Neural Networks and Learning System

Data Mining:

- IEEE Transactions on Knowledge and Data Engineering (TKDE)

What is Machine Learning?

- **Definition of ML (Mitchell, 1997): WELL-POSED LEARNING PROBLEMS.**
  - A computer program is said to learn from experience ***E*** with respect to some class of tasks ***T*** and performance measure ***P***, if its performance at tasks in ***T***, as measured by ***P***, improves with experience ***E***.
- **Example: A computer program that learns to play checkers**
  - **T**ask: playing checkers games;
  - **E**xperience: obtained by playing games against itself;
  - **P**erformance Measure: percent of games won against opponents

A handwriting recognition learning problem:

- Task ***T***: recognizing and classifying handwritten words within images
- Performance measure ***P***: percent of words correctly classified
- Training experience ***E***: a database of handwritten words with given classifications

A robot driving learning problem: an example from (Mitchell, 1997)

- Task ***T:*** driving on public four-lane highways using vision sensors;
- Performance measure ***P***: average distance traveled before an error (as judged by human overseer)
- Training experience ***E***: a sequence of images and steering commands recorded while observing a human driver;

Example: Spam classification

- Task ***T***: determine if emails are Spam or non-Spam.
- Experience ***E:*** Incoming emails with human classification
- Performance Measure ***P***: percentage of correct decisions

Notations, formally

Task:

$\mathcal{X}$ input variables (from input set), a.k.a., features, predictors, independent variables.

$\mathcal{Y}$ output variables (from output set), a.k.a., response or dependent variable.

$f : \mathcal{X} \to \mathcal{Y}$ Prediction function,

Performance:

$l : \mathcal{X} \to \mathcal{Y}$ Loss function,

$l\left(y, y'\right)$ is the cost of predicting y' if y is correct.

Experience: task-dependent, many different scenarios

- Supervised Learning, Unsupervised Learning, Reinforcement Learning,
- Semi－supervised Learning, Multiple Instance Learning, Active Learning.

Supervised Learning

- A labeled training set examples with outputs provided by an expert,

$$\mathcal{D} = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\} \subset \mathcal{X} \times \mathcal{Y}$$

- Regression Vs. Classification problems,
  - ***Regression***: *Y* is is **quantitative** (e.g price, blood pressure);
  - ***Classification***: *Y* takes values in a finite, unordered set (survived/died, digit 0-9, cancer class of tissue sample), **qualitative**.

Other problems such as ranking is often formulated as either problem.

Definition,

- A supervised learning system (or learner), *L* is a (computable) function from the set of (finite) training sets to the set of prediction functions:

$$L : \mathbb{P}^{<\infty}(\mathcal{X} \times \mathcal{Y}) \to \mathcal{Y}^{\mathcal{X}}$$
$$L : \mathcal{D} \mapsto f$$

So if presented with a training set $\mathcal{D}$, it provides a decision rule/function $f : \mathcal{X} \to \mathcal{Y}$

Let *L* be a learning system.

- Process of computing is $f = L(\mathcal{D})$ called training (phase).

• Applying f to new data is called prediction, or testing. (phase).