# Artificial Neural Network: Orientation & Introduction

Tatpong Katanyukul


มหาวิทยาลัยขอนแก่น
KHON KAEN UNIVERSITY

20 กรกฎาคม พ.ศ. 2562

# Milestones

- Introduction & Sweet Treats
- Optimization & Machine Learning
- 2-Layer MLP & Backpropagation
- Selected Topic: Dimension Reduction
- Selected Topic: CNN
- Selected Topic: LSTM

} Tentative

# A Rought Intro to Artificial Neural Network

Artificial Neural Network (ANN)

- A computing (math) model
- Prediction: given $x$, predict (compute) $y$
- Simply, ANN $\equiv f : x \mapsto y$.
- $x$ is a question, $y$ is an answer, and $f$ is a model
  If an answer is an egg, a model is a hen.
  If you want good eggs, take good care of a hen.
- They can be "anything".
  E.g., categorical data, numerical data, text, image, audio, vdo, etc.
  $\Rightarrow$ (This is not completely true,
  but we learn this complicated fact as we go along.)

# This Course (I)

- We won't learn everything. (Nobody can!)
- We will cover the essensial.
- We may learn some related topics that are not "ANNs" in the conventional meaning.
- Hand-on
- It is about MATH and MATH PROGRAMMING.
- It requires TENACITY, curiosity, creativity, and skills, in this order.

# This Course (II)

The Essential

- Optimization. (NOT AN OPTION! IT'S THE ENGINE.)
- Model selection and Evaluation.
- Simple model, e.g., polynomial (basic modeling)
- 2-Layer ANN
- Other models and applications, e.g., PCA, CNN, LSTM, etc.
- Reading ~~an~~ academic papers (object recognition, etc.)

  Learn from the source and be able to stand on your own feet

Grading is based on:

Participation in class, assignments, ~~and~~ class project, Ex + Ex.

# Related Terms

- Machine Learning (ML): a broader term covers other families of models, e.g., SVM, PCA, LDA, LSTM, CRF, etc.
- Data Science: a fancy term combining mainly ML and techniques, e.g., MapReduce, to handle big data.
- Artificial Intelligence (AI): an even broader term covers various approaches including heuristic search and logic based approaches.
  $\Rightarrow$ 'A' = Artificial is in the sense of "not biological".
  It does not mean "fake" or "charade".
  AI $\neq$ I can be lazy and stupid and let the machine does the job.

# Common Theme (I)

- Predictive model $f : x \mapsto y$
- Answer $y$ for question $x$.
- We do not find answer $y$ directly.
- We find a good $f$. $\Rightarrow$ We do not lay eggs $y$ ourselves.
  $\Rightarrow$ Take good care of our hen $f$.
- $f$ is a tunable function: $f$ has parameters
  - Linear model: $f(x) = mx + c$
  - Cubic model: $f(x) = a_3x^3 + a_2x^2 + a_1x + a_0$
  - Polynomial: $f(x) = a_nx^n + \ldots + a_1x + a_0$
  - Perceptron: $f(x_1, x_2) = h(w_1x_1 + w_2x_2 + b)$, where $h(a) = 1$ for $a > 0$ and $h(a) = 0$ otherwise.

# Common Theme (II)

- Predictive model $f : x \mapsto y$
- Answer $y$ for question $x$.
- We do not find answer $y$ directly.
- We find a good $f$.
- $f$ is a tunable function, e.g., $f(x) = mx + c$.
- Find a good set of parameters $\Rightarrow\sim$ good $f$.
- "Two-Phase" Strategy
  - Find a good $f$ ($\sim$ find a good set of param's)
    $\Rightarrow$ "Training" / "Learning" (Take care of the hen)
  - Use good $f$ for prediction.
    $\Rightarrow$ "Testing" / "Predicting" / "Using" (Get the eggs)

# Numerical Programming I

- Equality comparison

  Naive match: x == y vs Tolerance: y - e < x < y + e
- Computation sequence, e.g., $y = \frac{a}{b} \cdot \log_{\exp c} \exp b$

  Naive vs Refined $\frac{a}{c}$
- $y = \sum_i w_i x_i$

  Conventional loop vs Vectorization: $y = \mathbf{w}^T \cdot \mathbf{x}$
- $[\mathbf{y} : y_n = \sum_i w_i x_{i,n}]$

  Loop vs Vectorization
- Probabilities

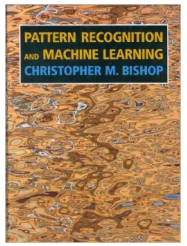
  Probability: $\prod_i p_i$ vs Log Probability: $\sum_i \log p_i$

# Numerical Programming II

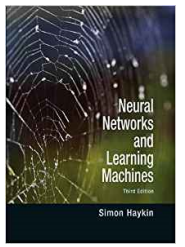
- Softmax $y_k = \frac{\exp(a_k)}{\sum_i \exp(a_i)}$
  Straightforward vs $y_k = \frac{\exp(a_k - max(\mathbf{x}))}{\sum_i \exp(a_i - max(\mathbf{x}))}$
- Cross-Entropy: $\hat{y} \log y + (1 - \hat{y}) \log(1 - y)$,
  where $\hat{y} \in \{0, 1\}$ and $y \in [0, 1]$
  $\Rightarrow$ Naive vs Selective
- Multi-Class Cross-Entropy: $\sum_k \hat{y}_k \log y_k$
  Naive vs Selective
- NaN : Not A Number

# Further Study

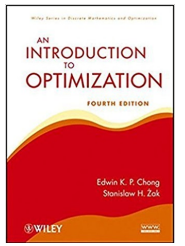
C. M. Bishop, Pattern Recognition and Machine Learning. Springer 2007



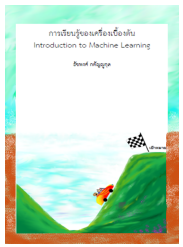

S. Haykin, Neural Networks and Learning Machines. Prentice Hall 2009




E. K.P. Chong and S. H. Żak, Introduction To Optimization. Wiley 2013




ธ. กตัญญูกุล, การเรียนรู้ของเครื่องเบื้องต้น. ม. ขอนแก่น 2017



# Words of Wisdom

Bruce Lee:
"Don't pray for an easy life.
Pray for strength to endure the hard one."