Birla Institute of Technology and Science Pilani

K.K. Birla Goa Campus

AY 2017–18, Semester II

Course Handout

Course metadata

Course: Neural Networks and Fuzzy Logic (BITS F312)
Instructor-in-charge/Instructor: Tirtharaj Dash
Chamber No. and contact: C. Ex-2, tirtharaj@goa.bits-pilani.ac.in
Lecture timing: M W F 5 (A-602)
Lab timing: will be decided

Scope and Objective of the course

Deep learning has had a long and rich history, but has gone by many names like cybernetics (1940s–1960s), connectionisms (1980s–1990s), presently named so since 2006, reflecting different philosophical viewpoints. The same viewpoints has now been shifted towards the term "Deep Learning" since the extensive use of powerful graphics processors.

This course on Neural Networks would focus on the conceptual and mathematical foundation of Deep Learning along with computational investigations of various models as a part of series of laboratory experiments or projects.

At the end of this course, students should be able to deal with various real-world problems and can model such deep learning machines which, of course, depends on whether the problem is a machine learnable (or neuro-computable) problem.

Books

1. Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning*, MIT Press.
2. Simon Haykin, *Neural Networks and Learning Machines*, Third Edition, Pearson.

Other references and venues

1. [LECTURES] Lectures by Prof. Ashwin Srinivasan, Machine Learning course, BITS F464. (runs in parallel with Neural Net course)
2. [LECTURES] Lectures by Prof. Ashwin Srinivasan and Tirtharaj Dash, Artificial Intelligence, CS F407 (runs every alternative semester by the same lecturers)
3. [BOOK] Christos H. Papadimitriou, Kenneth Steiglitz, Combinatorial Optimization: Algorithms and Complexity
4. [CONFERENCS] NIPS, IJCNN, ICML, ECML, ACML, ICONIP ... (such like)

Course plan

- [INTRODUCTION] What do we intend to do?
  - Course Introduction
  - Basic building block of learning machines (all ML kind) — [(0) Data – (0.1) Partition – (1) Model – (2) Learning algo – (3) Parameters – (3.1) Go to (2) – (4) Analysis – (5) Go to (0.1)]
  - Various design challenges in modeling learning problems such as (Is it learnable?, Preprocessing, Saturation in activation, Feature extraction, Feature encoding, Missing data, Imbalanced data, Background knowledge)
  - Neuron architecture — Defining a 'neural network'
  - Limitations of deep learning — Why Deep Learning is not just Artificial Brain? Or, why biological brain can't be imitated as a deep learning machine?
- [DATA AND PARAMETERS] What is given to us?
  - Explanation of a dataset $\mathcal{D} = \{\mathbf{X}, \mathbf{Y}\}$ — more on $\mathbf{x} \in \mathbf{X}$ and $y \in \mathbf{Y}$
  - **Search** for unknowns using $\mathcal{D}$
  - A discussion on random search in learning neural network parameters
- [PERCEPTRON] How do we do? – some simple ways and a series of impossibilities
  - Introducing Hebbian learning, McCulloh and Pitt's model of neuron, activation functions, neural network weights as parameters, neural network as a representation machine
  - Search in a continuous space (restricted version) and optimum
  - Generalization vs. specialization
  - Perceptron learning algorithm and proving convergence within finite time
  - Perceptron failure — Relation between learnability and separability of patterns
  - Learnability in kernelized space — impact on separability issue
  - More on kernel trick in neural networks
- [SEARCH AS OPTIMIZATION] Who can help us?
  - Objective functions
  - Introduction of more unknowns, search space
  - Derivative based optimization, (a small discussion on derivative free optimization)
  - Stochastic Gradient Descent and Batch Gradient Descent
  - Effects of various parameters on search
- [MULTILAYER PERCEPTRON] Opening possibilities on 'How do we do?'
  - Cost functions for MLP, Study of search as optimization in MLP
  - Parameter initialization strategies, parameters tuning using cross-validation, Adaptive learning rate and momentum factor
  - Regularization
- [CNN] How do we do on multidimensional input data?

  - Convolution operation, Pooling
  - Variants of convolution functions
  - Learning in CNN
- [RNN] How do we do on sequential input data?
  - Recurrence in input and recurrence in hidden features
  - Back propagation through time (BPTT)
  - RNN with Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) and such like
- [AUTOENCODERS] How do we represent efficiently?
  - Loss function and its gradient
  - Over- and under-complete autoencoders
  - De-noising autoencoders
- [GENERATIVE MODELS] How do we generate distributions?
  - Restricted Boltzmann Machine (RBM) – Energy function as joint distribution of input and its representation
  - Relatedness to Markov network
  - Generating conditional distribution from network
  - Synthesize input back
  - Deep Belief Network (DBN)
- [SOME SERIOUS STUFF] Let's do it.
  - Adversarial Machine Learning
  - Generative Adversarial Networks (GAN)
  - One-shot learning
  - Capsule Network
  - Some talks by experts on related areas of Deep Learning

Table 1: **Evaluation Scheme**

Component	**Mark**	**Type**
Mid-sem test	30	CB
Lab	10	OB
Project	20	OB
Comprehensive test	40	CB

Course Notices

All the announcement will be made in LMS, in class, or via e-mail.

Make-up and malpractice policies

- Make-up shall be granted only in genuine cases based on individual's need and circumstances.
- No marks will be awarded without make-up for that component
- Malpractice policies are as per institute regulations.

Chamber Consultation Hour

No specific time. Please let me know the topic of discussion and time. I will be sharing my present semester time schedule with you.

Tirtharaj Dash
Instructor-in-charge
BITS F312