Birla Institute of Technology and Science Pilani

K.K. Birla Goa Campus

AY 2019–20, Semester I

Course Handout

Course metadata

Course: Neural Networks and Fuzzy Logic (BITS F312)
Instructor-in-charge/Instructor: Tirtharaj Dash
Chamber No. and contact: A-420, tirtharaj@goa.bits-pilani.ac.in
Lecture timing: M W F 9 (A 605)
Lab timing: Will be discussed
Mode of lectures: Blackboard (sometimes, PPT)

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 and 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.

Text Book(s)

1. Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning*, MIT Press.

Reference Book(s)

1. Graves, A. (2012). Supervised sequence labelling with recurrent neural networks.
2. Aggarwal, C. C. (2018). Neural networks and deep learning, Springer.
3. Haykin, S. S. (2009). Neural networks and learning machines/Simon Haykin, Pentice Hall.

Conferences

1. NeuroIPS (Formerly, NIPS), ICML, ECML, ACML, ICONIP, ICANN, IJCNN.

Course plan

- INTRODUCTION
  - 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
  - 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
  - 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
- EARCH AS OPTIMIZATION
  - 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 (MLP)
  - 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
- CONVOLUTIONAL NEURAL NETWORKS (CNN)
  - Convolution operation, Pooling
  - Variants of convolution functions
  - Learning in CNN
  - Backpropagation in CNN
  - Application demonstration
- RECURRENT NEURAL NETWORKS (RNN)
  - Recurrence in input and recurrence in hidden features
  - Back propagation through time (BPTT)
  - RNN with Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU)
  - Backpropagation in LSTM and GRU
- AUTOENCODERS
  - Is the representation good?
  - Loss function and its gradient
  - Over- and under-complete autoencoders
  - De-noising autoencoders
- GENERATIVE MODELS
  - 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 ADVANCED APPLICATIONS
  - Adversarial Machine Learning
  - Generative Adversarial Networks (GAN)
  - One-shot learning
  - Capsule Network
  - Some talks by experts on related areas of Deep Learning* (Depending on availability)

Evaluation Scheme

Component	Mark	Type
Mid-sem test	30	Closed Book
Lab	10	Open Book
Project	20	Open Book
Comprehensive test	40	Closed Book

Course Notices

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

Attendence, Make-up and Malpractice policies

- Either come to all the lectures or don't come at all!
- 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

E-mail me, if you want to meet. Don't email about unofficial or other projects, where I am not involved.

Tirtharaj Dash
Instructor-in-charge
BITS F312