Deep Learning Foundations and Applications (AI61002)

Announcements

Course Logistics

Prerequisites:

Linear Algebra and probability theory

Python and NumPy

Lecture Schedule:

Day	Time	Room
Wednesday	12 - 1 PM	NR 412
Thursday	11 AM - 12 PM	
Friday	9 - 11 AM	

Course Instructors:

Somdyuti Paul, Mahesh Mohan M R, Jiaul Hoque Paik

Teaching Assistants:

Anjali Raj, Sista Raviteja, Ashraf Haroon Rashid, Sumanta Chandra Mishra Sharma, Saurabh Mishra, Leo Lorence, Yashasvi Rathore

Grading Policy:

Midterm - 25%

Final - 35%
Continuous Evaluation - 40%
coding assignments
in-class surprise quizzes

Plagiarism Policy:

Strict anti-plagiarism policy would be enforced. Any detected attempt to plagiarize assignments would be subjected to negative marking.

Attendance Policy:

Attendance is mandatory and students failing to meet minimum attendance requirement would be subjected to de-registration from the course.

Recommended Books:

Deep Learning by Ian Goodfellow, Joshua Bengio and Aaron Courville, MIT Press (text) [online version]
Neural Networks and Deep Learning by Charu C. Aggarwal, Springer (reference) [download]
Dive into Deep Learning by Aston Zhang, Zachary C. Lipton, Mu Li and Alexander J. Smola, Cambridge University Press (coding reference) [online version]

Introduction to Deep Learning

Lecture 1: 03/01/2024

Introduction, course logistics and syllabus
Deep Learning (DL) in relation to Machine Learning (ML) and Artificial Intelligence (AI)
Computational models of artificial neurons
McCullouch-Pitts neuron
Perceptron
Sigmoid neuron

Lecture 2: 05/01/2024

Model of the primary visual cortex as an early motivation for convolutional neural networks (CNNs)
Neocognitron
LeNet
Emergence of modern CNNs
Timeline of the evolution of deep learning
Deep learning successes

Lecture slides: Introduction to Deep Learning
MP Neurons paper
Neocognitron Paper
LeNet-5 paper

Tutorial 1: Introduction to Python, NumPy and Matplotlib

Tutorial 1 - Part I: 04/01/2024

Introduction to Python

Keywords, identifiers, data types and operators
Control flow
Lists, tuples, sets and dictionaries
Strings and string manipulations
Functions, built-in functions, lambda functions, generators and decorators
Modules and Packages
Files
Debugging and Exception Handling
Classes and Objects

Tutorial 1 - Part II: 11/01/2024

Introduction to NumPy and Matplotlib

NumPy arrays - creating, indexing, slicing, reshaping, data types, joining splitting
Searching and sorting NumPy arrays
Mathematical operations on arrays and broadcasting
Visualizing data with Matplotlib - barplot, line chart, scatter plot, histogram, box plot
Adjusting plot properties - color, axes, title, labels etc.
Plotting common distributions.

Jupyter notebook on Python basics

Jupyter notebook on introduction to NumPy

Jupyter notebook on data visualization with matplotlib

Review of Linear Algebra and Probability Concepts

Lecture 3 - 10/01/2024

Review of Linear Algebra

Vectors - inner product, outer product and Hadamard product, vector norms, linear independence and span
Matrices - Identity and inverse matrices, existence of inverse, column space and null space, orthogonal and unitary matrices,
Eigen value decomposition and Singular value decomposition
System of linear equations and existence of solutions, Moore Penrose pseudoinverse

Lecture 4 - 12/01/2024

Review of Probablilty and Random Variables

Probablility space
Random variables (RVs) and their distributions - joint and marginal distributions, conditional distributions, independence and conditional independence.
Expectation, variance and covariance
Information content of a RV, entropy and KL divergence

Principal Component Analysis (PCA)

Goals of PCA
SVD of covariance matrix and projection onto principal components
Reconstruction error of PCA

Mathematics for Machine Learning

Programming Assignment 1

Programming Assignment 1

Machine Learning Basics

Lecture 5 - 17/01/2024

Types of machine learning - supervised, unsupervised, self-supervised and reinforcment learning, parametric and non-parametric learning, discriminative and generative learning
Components of a general parametric ML framework - task, data, model/hypotheis, cost, optimizer, performance metrics
ML performance metrics for regression and classification
Casting linear regression and logistic regression in the parametric ML framework
Overfitting and underfitting

Lecture 6 - 18/01/2024

Bias Variance trade-off,
Mitigating overfitting and underfitting
Gradient Descent - intuition, impact of learning rate,

Lecture 7 - 19/01/2024

Batch, minibatch and stochastic GD.

Linear regression update rule using gradient descent
Maximum Likelihood estimation

Lecture slides: Machine Learning Basics

Feedforward Neural Networks

Lecture 7 - 19/01/2024

The perceptron algorithm
Convergence of the perceptron algorithm
Limitation of single layer perceptrons
Multi-layer perceptrons (MLPs) / feedforward neural networks
Solving the XOR problem with multi-layer perceptron

Lecture 8 - 24/01/2024

Representation power of MLPs
Need for deeper network architectures
Common Activation Functions
Universal approximation theorem
Computational Graphs of Neural Networks

Lecture slides: Feedforward Neural Networks

Tutorial 2: Basics of PyTorch and Feedforward Neural Networks

Tutorial 2: 25/02/2024

Useful modules and libraries
Defining and using tensors
Dataloaders and custom dataloaders
Designing feedforward neural networks
Specifying loss function and optimization algorithm
Training feedforward neural network and validation

Jupyter notebook on PyTorch basics and feedforward neural networks

Programming Assignment 2

Programming Assignment 2

Backpropagation and Gradient Descent Algorithms

Lecture 9 - 30/01/2024

Learning neural network parameters through backpropagation
Recursive re-use of gradients
Gradient flow
Computing gradients
Backpropagation through successive layers
Vanishing and exploding gradients

Lecture 10 - 07/02/2024

Overcoming vanishing and exploding gradients

Choice of activation function
Appropriate weight initialization
Batch normalization
Gradient clipping
Residual connections
Gradient descent with momentum
Gradient descent with adaptive learning rates
Adagrad
RMSProp
Adadelta
Adam

Lecture Slides: Backpropagation and Gradient Descent Algorithms

Adagrad paper

Adam paper

Adadelta paper

Batch normalization paper

Tutorial 3: Gradient Descent Algorithms

Tutorial 3: 08/02/2024

Learning rate scheduling and early stopping
Batch, minibatch and stochastic gradient descents
Implementation of gradient descent with momentum
Implementation of adagrad optimizer
Implementation of RMSProp optimizer
Using built-in PyTorch optimizers

Jupyter Notebook on Gradient Descent Algorithms

Regularization Techniques for Deep Neural Networks

Lecture 11: 09/02/2024

Norm Penalties
Data Augmentation
Early Stopping
Ensemble Methods
Dropout
Paramater Sharing

Lecture slides: Regularization techniques for deep neural networks

Dropout paper

Questions: class tests and last year's midterm

Class test 1 solutions

DLFA 2023 midterm question paper

Class test 2 Solns

Programming Assignment 3

Programming Assignment 3

Class Test Scores

Class test scores

Convolutional Neural Network

Convolutional Neural Network

1. Basics of CNN and Understanding Different Layers
2. CNN for Regression
3. Understanding and Visualizing CNN
4. Standard CNN Architectures

Further Reading: Deep Learning Textbook, Chapter 9

https://www.deeplearningbook.org/contents/convnets.html

CNN Part 1: Basics of CNN

CNN Part 2: Regression, Visualization and Models

Reading Exercise: Visualizing and Understanding CNN Paper

Reading Exercise: CNN Architectures: Alexnet, VGGnet, Googlenet, Resnet,

CNN Lab Assignment

CNN Tutorial

Tutorial 4: Regularization Techniques for Deep Neural Networks

Tutorial 4: 07/03/2024

Implementation of L1 and L2 regularizations
Data Augmentation
Implementation of dropout

Jupyter notebook on regularization techniques

Recurrent Neural Nets

RNN-1

In this lecture we cover the following things:

1. Why we need RNN
2. Structure and computation over RNN

LSTM

Please note that I did not use any slide for this.

The chapter named "Sequence Models" from the above link is the main reference I have used in my lectures.

Attention

Attention

This lecture covers encoder-decoder model and attention net.

Transformer

transformer

These lectures cover transformer network.

BERT and Fine Tuning