BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI

WORK INTEGRATED LEARNING PROGRAMMES

Digital

Part A: Content Design

Course Title	Deep Neural Network
Course No(s)	
Credit Units	4
Content Authors	Ms. Seetha Parameswaran
Version	1.0
Date	September 24th 2022

Course description

AIML* ZG511 Deep Neural Networks

Introduction to neural networks, approximation properties, back propagation, deep network training, regularization and optimization, convolution neural networks, recurrent neural networks, attention models, transformers, neural architecture search, federated learning, meta learning, applications in time series modelling and forecasting, online (incremental) learning.

Pre-requisistes:

AIML* ZC416 Mathematical Foundations for Machine Learning
AIML* ZC418 Introduction to Statistical Methods
AIML* ZG565 Machine Learning

Course Objectives

No	Course Objective
CO1	Introduce students to the basic concepts and techniques of Deep Learning.
CO2	Students will be able apply deep learning models to applications.
CO3	Students will be able to evaluate deep learning algorithms.

Text Book(s)

T1	Dive into Deep Learning by Aston Zhang, Zack C. Lipton, Mu Li, Alex J. Smola. https://d2l.ai/chapter_introduction/index.html

Reference Book(s) & other resources

R1	Deep Learning by Ian Goodfellow, Yoshua Bengio, Aaron Courville. MIT Press 2016.
R2	Introduction to Deep Learning by Eugene Charniak. The MIT Press 2019
R3	Deep Learning with Python by Francois Chollet. 1st Edition. Manning Publications https://livebook.manning.com/book/deep-learning-with-python/part-1/
R4	Deep Learning for Time Series Forecasting by Jason Brownlee
R5	Neural Architecture Search: A Survey by Thomas Elsken, Jan Hendrik Metzen, Frank Hutter https://arxiv.org/pdf/1808.05377.pdf

Content Structure

1 Fundamentals of Neural Network (4 hrs)
- 1.1 Objective of the course
- 1.2 Supervised, unsupervised, semi-supervised and reinforcement learning problems.
- 1.3 Why Deep Learning?
- 1.4 Applications of Deep Learning
- 1.5 Perceptron and Perceptron learning algorithm
- 1.6 Multilayer Perceptron (MLP)
- 1.7 MLP as classifiers and Universal approximators
- 1.8 Issue of Depth and Width

2 Deep Feedforward Neural Networks (4 hrs)
- 2.1 Forward and backward propagation
- 2.2 Computation graph
- 2.3 Gradient Descent algorithm
- 2.4 Impact of depth in DNN

3 Optimization of Deep models (2 hrs)
- 3.1 Challenges in Neural Network Optimization – saddle points and plateau
- 3.2 Non-convex optimization intuition
- 3.3 Overview of optimization algorithms
- 3.4 Mometum based algorithms
- 3.5 Algorithms with Adaptive Learning Rates

4 Regularization for Deep models (2 hrs)
- 4.1 Model Selection

- 4.2 Underfitting, and Overfitting
- 4.3 L1 and L2 Regularization
- 4.4 Dropout
- 4.5 Challenges - Vanishing and Exploding Gradients, Covariance shift
- 4.6 Parameter Initialization
- 4.7 Batch Normalization

5 Convolutional Networks (4 hrs)
- 5.1 Convolutions for Images
- 5.2 Learning a Kernel
- 5.3 Padding and stride, Channels, Pooling
- 5.4 Design of CNN
- 5.5 Popular CNN architectures
- 5.6 Transfer Learning
- 5.7 Applications of CNN

6 Sequence Models (6 hrs)
- 6.1 Recurrent Neural Networks
- 6.2 Back-propagation through time
- 6.3 Challenge - Exploding - Vanishing gradient and Gates
- 6.4 Popular RNN architectures
- 6.5 Applications of RNNs

7 Attention Mechanism (4 hrs)
- 7.1 Attention Pooling
- 7.2 Attention Scoring Functions
- 7.3 Multi-Head Attention, Self-Attention and Positional Encoding
- 7.4 Transformer architecture
- 7.5 Applications of Transformers

8 Neural Network search (2 hrs)
- 8.1 Search Space
- 8.2 Search algorithms
- 8.3 Evaluation Strategy

9 Time series Modelling and Forecasting (2 hrs)
- 9.1 Univariate, Multivariate and Multi-step CNN Models
- 9.2 Univariate, Multivariate and Multi-step LSTM Models

10 Other Learning Techniques (4 hrs)
- 10.1 Federated learning
- 10.2 Meta learning
- 10.3 Online (incremental) learning

Learning Outcomes:

No	Learning Outcomes
LO1	Able to understand the basics of Deep Learning.
LO2	Able to understand and apply techniques related to Deep Learning to applications.
LO3	Able to identify appropriate tools to implement the solutions to problems related to Deep Learning and implement solutions.

Part B: Learning Plan

Session No.	Topic Title	Resource Reference
1	Fundamentals of Neural Network • Objective of the course • Why Deep Learning? • Applications of Deep Learning • Biological neuron vs artificial neuron • Connectionism model • Perceptron • Perceptron learning algorithm • XOR Problem	T1 – Ch1 http://mlsp.cs.cmu.edu/people/rsingh/docs/Chapter1_Introduction.pdf
2	Fundamentals of Neural Network • Multilayer Perceptron (MLP), • MLP on Boolean, reals and continuous values • MLP as classifiers • MLP as Universal approximators • Issue of Depth and Width	http://mlsp.cs.cmu.edu/people/rsingh/docs/Chapter1_Introduction.pdf http://mlsp.cs.cmu.edu/people/rsingh/docs/Chapter2_UniversalApproximators.pdf
3	Deep Feedforward Neural Network • MLP with hidden Layers • Forward Propagation • Backward Propagation (review) • Training a DNN using Gradient Descent algorithm (review) • Computational Graphs (review as already	T1 – Ch4 and Ch3.4

	discussed in MFML)	
4	Deep Feedforward Neural Network • Activation Functions • Softmax Regression	T1 – Ch4 and Ch3.4
5	Optimization algorithms for Deep models • Challenges – Saddle points and plateau • Non-convex optimization intuition • Stochastic Gradient Descent (SGD), Minibatch SGD • Overview of Rprop, Quickprop • Momentum, Nastrov's Accelarated Momentum • Algorithms with Adaptive Learning Rates, Adagrad, RMSprop, ADAM	T1 – Ch11
6	Regularization for Deep models • Model Selection, Underfitting, and Overfitting • L1 and L2 Regularization • Dropout • Challenge - Vanishing and Exploding Gradients • Parameter Initialization • Challenge Covariance Shift • Batch Normalization	T1 – Ch4 , 7.5
7	Convolutional Neural Network • Basics of Computer Vision and Invariance • Convolutions for Images • Learning a Kernel • Padding and stride • Channels • Pooling • Designing a CNN	T1 – Ch6
8	Popular CNN architectures • LeNet • AlexNet • VGG16 • RCNN and Fast RCNN	T1 – Ch7

	• Network in Network (NiN) • Inception Net • ResNet • DenseNet • Transfer Learning • Applications of CNN	
9	Sequence Models • Recurrent Neural Networks • Types of Sequences and RNNs • Back-propagation Through Time (discuss the paper) • Gates and Exploding / Vanishing gradient	T1 – Ch8
10	Popular RNN architectures • Gated Recurrent Units (GRU) • Long Short-Term Memory (LSTM) Networks (discuss why and how LSTM solves) • Bidirectional models • ~~Sequence to sequence learning with an RNN encoder and an RNN decoder~~	T1 – Ch9
11	Attention Mechanism • Attention Pooling • Attention Scoring Functions • Multi-Head Attention	T1 – Ch10
12	Attention Mechanism • Self-Attention • Positional Encoding • Transformer architecture • Applications of Transformers	T1 – Ch10
13	Neural Network search overview • Search Space • Search algorithms • Evaluation Strategy	R5
14	Time series Modelling and Forecasting	R4

	• Using CNN • Using LSTM (Trends and Seasonality will be discussed. No overview of statistical techniques)	Ch-8, Ch-9
15	Federated learning • Federated Learning Of Out-Of-Vocabulary Words Meta learning	https://arxiv.org/pdf/1902.04885.pdf https://arxiv.org/pdf/1903.10635.pdf https://arxiv.org/pdf/2004.05439.pdf
16	Online (incremental) learning • Continual Lifelong Learning with Neural Networks • Three scenarios for continual learning	https://arxiv.org/pdf/1802.07569.pdf https://arxiv.org/pdf/1904.07734v1.pdf

Detailed Plan for Lab work

Lab No.	Lab Objective	Lab Sheet Access URL	Session Reference
1	Introduction to Tensorflow, Keras		2
2	Computational graph in Pytorch		3
3	Deep Neural Network with Back-propagation and optimization		4
4	CNN		6
5	RNN		9
6	LSTM		10
7	Transformers		12
8	Time series forecasting		15

Evaluation Scheme:
Legend: EC = Evaluation Component; AN = After Noon Session; FN = Fore Noon Session

No	Name	Type	Duration	Weight	Day, Date, Session, Time
EC-1(a)	Quizzes	Online		10%	
EC-1(b)	Assignments	Online		20%	
EC-2	Mid-Semester Test	Closed Book		30%	
EC-3	Comprehensive Exam	Open Book		40%	

Note:
Syllabus for Mid-Semester Test (Closed Book): Topics in Session Nos. 1 to 8
Syllabus for Comprehensive Exam (Open Book): All topics (Session Nos. 1 to 16)

Important links and information:

Elearn portal: https://elearn.bits-pilani.ac.in or Canvas
Students are expected to visit the Elearn portal on a regular basis and stay up to date with the latest announcements and deadlines.

Contact sessions: Students should attend the online lectures as per the schedule provided on the Elearn portal.

Evaluation Guidelines:

1. EC-1a consists of two Quizzes. Students will attempt them through the course pages on the Elearn portal. Announcements will be made on the portal, in a timely manner.
2. EC-1b consists of two Assignments. Students will attempt them through the course pages on the Elearn portal. Announcements will be made on the portal, in a timely manner.
3. For Closed Book tests: No books or reference material of any kind will be permitted. Use of calculators is permitted in all exams.
4. For Open Book exams: Use of books and any printed / written reference material (filed or bound) is permitted. However, loose sheets of paper will not be allowed. Use of calculators is permitted in all exams. Laptops/Mobiles of any kind are not allowed. Exchange of any material is not allowed.
5. If a student is unable to appear for the Regular Test/Exam due to genuine exigencies, the student should follow the procedure to apply for the Make-Up Test/Exam which will be made available on the Elearn portal. The Make-Up Test/Exam will be conducted only at selected exam centers on the dates to be announced later.

It shall be the responsibility of the individual student to be regular in maintaining the self-study schedule as given in the course hand-out, attend the online lectures, and take all the prescribed evaluation components such as Assignment/Quiz, Mid-Semester Test and Comprehensive Exam according to the evaluation scheme provided in the hand-out.