# BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI

## WORK INTEGRATED LEARNING PROGRAMMES

### Digital

Part A: Content Design

| Course Title | Deep Learning |
|---|---|
| Course No(s) | |
| Credit Units | 4 |
| Credit Model | 2 - 0.5 - 1.5. <br> 1 unit for class room hours, 0.5 unit for Tutorial, 1.5 units for Student preparation. 1 unit = 32 hours |
| Content Authors | Ms. Seetha Parameswaran |
| Version | 3.0 |
| Date | August 12th 2022 |

## 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<br>https://livebook.manning.com/book/deep-learning-with-python/part-1/ |

**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    Representation Learning   ( 4 hrs)
   8.1    Regularized Autoencoders
   8.2    Denoising Autoencoders
   8.3    Variational Autoencoders
   8.4    Applications of Autoencoders

9    Generative Adversarial Networks  ( 2 hrs)
   9.1    GAN architecture
   9.2    Applications of GAN

**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

| Academic Term | |
|---|---|
| **Course Title** | Deep Learning |
| **Course No** | |
| **Lead Instructor** | |

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

| | | |
|---|---|---|
| 5 | Deep Feedforward Neural Network<br>&bull; Activation Functions<br>&bull; Softmax Regression | T1 – Ch4 and Ch3.4 |
| 6 | Optimization algorithms for Deep models<br>&bull; Challenges – Saddle points and plateau<br>&bull; Non-convex optimization intuition<br>&bull; Stochastic Gradient Descent (SGD), Minibatch SGD<br>&bull; Overiew of Rprop, Quickprop<br>&bull; Momentum, Nastrov's Accelarated Momentum<br>&bull; Algorithms with Adaptive Learning Rates, Adagrad, RMSprop, ADAM | T1 – Ch11 |
| 7 | Regularization for Deep models<br>&bull; Model Selection, Underfitting, and Overfitting<br>&bull; L1 and L2 Regularization<br>&bull; Dropout<br>&bull; Challenge - Vanishing and Exploding Gradients<br>&bull; Parameter Initialization<br>&bull; Challenge Covariance Shift<br>&bull; Batch Normalization | T1 – Ch4 , 7.5 |
| 8 | Convolutional Neural Network<br>&bull; Basics of Computer Vision and Invariance<br>&bull; Convolutions for Images<br>&bull; Learning a Kernel<br>&bull; Padding and stride<br>&bull; Channels<br>&bull; Pooling<br>&bull; Designing a CNN | T1 – Ch6 |
| 9 | Popular CNN architectures<br>&bull; LeNet<br>&bull; AlexNet<br>&bull; VGG16<br>&bull; Network in Network (NiN)<br>&bull; Inception Net<br>&bull; ResNet<br>&bull; DenseNet | T1 – Ch7 |

| | | |
|---|---|---|
| | • Transfer Learning<br>• Applications of CNN | |
| 10 | Sequence Models<br>• Recurrent Neural Networks<br>• Types of Sequences and RNNs<br>• Back-propagation Through Time<br>• Gates and Exploding / Vanishing gradient | T1 – Ch8 |
| 11 | Popular RNN architectures<br>• Gated Recurrent Units (GRU)<br>• Long Short-Term Memory (LSTM) Networks<br>• Bidirectional models<br>• Sequence to sequence learning with an RNN encoder and an RNN decoder | T1 – Ch9 |
| 12 | Attention Mechanism<br>• Attention Pooling<br>• Attention Scoring Functions<br>• Multi-Head Attention | T1 – Ch10 |
| 13 | Attention Mechanism<br>• Self-Attention<br>• Positional Encoding<br>• Transformer architecture<br>• Applications of Transformers | T1 – Ch10 |
| 14 | Representation Learning<br>• Review of PCA<br>• Autoencoder<br>• Denoising Autoencoders | R1 – Ch14 |
| 15 | Representation Learning<br>• Variational Autoencoders<br>• Applications of Autoencoders | R1 – Ch14 |
| 16 | Generative Adversarial Networks - An overview, | T1 – Ch19 |

| | applications of GAN | | |
|---|---|---|---|

**Detailed Plan for Lab work**

| Lab No. | Lab Objective | Lab Sheet Access URL | Session Reference |
|---|---|---|---|
| 1 | Introduction to Tensorflow and Keras | | 2 |
| 2 | Deep Neural Network with Back-propagation and optimization | | 4 |
| 3 | CNN | | 6 |
| 4 | RNN | | 9 |
| 5 | LSTM | | 10 |
| 6 | Autoencoders | | 11 |

**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.
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.