

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**LIST OF COURSES OFFERED FOR HONOR PROGRAM (R20)****(DATA SCIENCE)**

Course code	Course Title	Contact hours/week				Credits
		L	T	P	Total	
20CSH1	Mathematical Foundation for Data Science	3	1	0	4	4
20CSH2	Computer vision	3	1	0	4	4
20CSH3	Deep Learning	3	1	0	4	4
20CSH4	Natural Language Processing	3	1	0	4	4

B.Tech – CSE

**20CSH1 – MATHEMATICAL FOUNDATION FOR
DATA SCIENCE**

L	T	P	Cr.
3	1	0	4

Pre-requisite : Good understanding of Fundamentals of mathematics (especially Algebra and Arithmetic).

Course Educational Objective: The Objective of the course is to introduce the basic notions of Probability Distributions, Inferences concerning mean and variance and the use of Matrices in applications to Data Science.

Course Outcomes: At the end of this course, the student will be able to

- CO1:** Describe the basic and intermediate concepts of Probability, Statistics, and Distributions. (**Understand - L2**)
- CO2:** Analyze hypothesis to accept/reject alternative hypothesis based on statistical evidence available. (**Apply- L3**)
- CO3:** Understand fundamentals of Matrices in Data Science (**Understand - L2**)
- CO4:** Demonstrate proficiency on the topics Eigenvalues, Eigenvectors, and Inner Product Spaces (**Apply- L3**)
- CO5:** Understand fundamentals Stochastic Process and Markov Chains. (**Understand - L2**)

UNIT-I**Probability Distributions**

Random Variables, The Binomial Distribution, The Hypergeometric Distribution, The Mean and the Variance of a Probability Distribution, Chebyshev's Theorem, The Poisson Distribution and Rare Events, Poisson Processes, The Geometric and Negative, Binomial Distribution, The Multinomial Distribution, Simulation.

UNIT-II**Inferences concerning mean and variance**

Statistical Approaches to Making, Generalizations, Point Estimation, Interval Estimation, Maximum Likelihood Estimation, Tests of Hypotheses, Null Hypotheses and Tests of Hypotheses, Hypotheses Concerning One Mean, The Relation between Tests and Confidence Intervals, Power, Sample Size, and Operating Characteristic Curve, The Estimation of Variances, Hypotheses Concerning One Variance, Hypotheses Concerning Two Variances.

UNIT-III**Matrices**

Types of Matrices, Symmetric; Hermitian; Skew-symmetric; Skew-Hermitian; orthogonal matrices; Unitary Matrices; rank of a matrix by Echelon form and Normal form, Inverse of Non-singular matrices by Gauss-Jordan method; System of linear equations; solving system of Homogeneous and Non-Homogeneous equations. Gauss elimination method; Gauss Seidel Iteration Method.

UNIT-IV**Eigenvalues, Eigenvectors, and inner product Spaces**

Eigenvalues and Eigenvectors - Eigenvectors and Upper Triangular matrices - Eigenspaces and Diagonal Matrices - Inner Products and Norms - Linear functionals on Inner Product spaces.

UNIT-IV**Stochastic Processes and Markov Chains**

Introduction to Stochastic processes- Markov process. Transition Probability, Transition Probability Matrix, First order and Higher order Markov process, n-step transition probabilities, Markov chain, Steady state condition, Markov analysis.

TEXTBOOKS:

1. Miller & Freund's, Richard a. Johnson, "Probability and statistics for engineers", Pearson, 9th edition, 2017
2. Devore. J.L., "Probability and Statistics for Engineering and the Sciences", Cengage Learning, New Delhi, 8th Edition, 2012.
3. B.S. Grewal, Higher Engineering Mathematics, Khanna Publishers, 36th Edition, 2010

4. Erwin kreyszig, Advanced Engineering Mathematics, 9th Edition, John Wiley & Sons, 2006.
5. Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, Probability & Statistics for Engineers & Scientists, 9th Ed. Pearson Publishers ,2013

B.Tech – CSE

20CSH2 – COMPUTER VISION

L	T	P	Cr.
3	1	0	4

Pre-requisite : **Concepts of Computer Graphics**

Course Educational Objective: The Objective of the course is to make learn the basic principles of digital image processing, Image data structures, different image encoding techniques, segmentation, and Morphological Image Processing.

Course Outcomes: At the end of this course, the student will be able to

- CO1:** Summarize the image fundamentals and mathematical transforms necessary for image processing (**Understand - L2**)
- CO2:** Apply image enhancement techniques (**Apply- L3**)
- CO3:** Demonstrate smoothing and sharpening and segmentation of color images (**Understand - L2**)
- CO4:** Apply image compression procedures (**Apply- L3**)
- CO5:** Demonstrate the concept of segmentation and Morphological image processing. (**Understand - L2**)

UNIT – I

Introduction: Examples of fields that use digital image processing, fundamental steps in digital image processing, components of image processing system. **Digital Image Fundamentals:** A simple image formation model, image sampling and quantization, basic relationships between pixels

UNIT – II

Image enhancement in the spatial domain: Basic gray-level transformation, histogram processing, enhancement using arithmetic and logic operators, basic spatial filtering, smoothing and sharpening spatial filters, combining the spatial enhancement methods

UNIT – III

Color Image Processing: Color fundamentals, color models, pseudo color image processing, basics of full-color image processing, color transforms, smoothing and sharpening, color segmentation

UNIT – IV

Image Compression: Fundamentals, image compression models, error-free compression, lossy predictive coding, image compression standards

UNIT – V

Image Segmentation: Detection of discontinuous, edge linking and boundary detection, thresholding, region-based segmentation

Morphological Image Processing: Preliminaries, dilation, erosion, open and closing, hit or miss transformation, basic morphologic algorithms

TEXTBOOK:

1. Rafeal C. Gonzalez, Richard E.Woods, “Digital Image Processing”, Pearson Education/PHI, Second Edition,2011

REFERENCE BOOKS:

1. Milan Sonka, Vaclav Hlavac and Roger Boyle, “Image Processing, Analysis, and Machine Vision”, Thomson Learning, Second Edition,2014.
2. Alasdair McAndrew , “Introduction to Digital Image Processing with Matlab” , Thomson Course Technology,1st Edition,2015.
3. B. Chanda, D. DattaMajumder, “Digital Image Processing and Analysis”, Prentice Hall of India, 2nd Edition, 2003.
4. Adrian Low , “Computer Vision and Image Processing”, Second Edition, B.S.Publications.2010
5. Rafeal C. Gonzalez, Richard E.Woods Steven L. Eddins, ”Digital Image Processing using Matlab”, Pearson Education,2012.

B.Tech - CSE

20CSH3 - DEEP LEARNING

L	T	P	Cr.
3	1	0	4

Pre-requisite : Machine Learning

Course Educational Objective: The Objective of the course is to make learn the major deep learning algorithms, the problem settings, and their applications to solve real world problems

Course Outcomes: At the end of this course, the student will be able to

- CO1:** Understand basic concepts of neural networks and back propagation Algorithms. . **(Understand - L2)**
- CO2:** Analyze the layers in the architecture of convolution neural networks. **(Analyze – L4)**
- CO3:** Apply operations in convolution neural networks for encoding frameworks. **(Apply - L3)**
- CO4:** Explore deep learning models for sequence analysis. **(Apply- L3)**
- CO5:** Analyze performance of LSTM Networks. **(Analyze – L4)**

UNIT – I

The Neural Network : Building Intelligent Machines , The Limits of Traditional Computer Programs , The Mechanics of Machine Learning , The Neuron ,Expressing Linear Perceptrons as Neurons , Feed-Forward Neural Networks, Linear Neurons and Their Limitations , Sigmoid, Tanh, and ReLU

Training Feed-Forward Neural Network: Gradient Descent, The Delta Rule and Learning Rates ,Gradient Descent with Sigmoidal Neurons, The Back propagation Algorithm , Stochastic and minibatch Gradient Descent , Test Sets, Validation Sets, and Over fitting, Preventing Over fitting in Deep Neural Networks

UNIT – II

Convolutional Neural Networks: Neurons in Human Vision, The Shortcomings of Feature Selection, Vanilla Deep Neural Networks, Filters and Feature Maps, Full Description of the Convolutional Layer, Max Pooling

Full Architectural Description of Convolution Networks: Closing the Loop on MNIST with Convolutional Networks, Image Pre-processing, Pipelines Enable More Robust Models, Accelerating Training with Batch Normalization

UNIT – III

Embedding and Representation Learning: Learning Lower-Dimensional Representations, Principal Component Analysis, Motivating the Auto encoder Architecture , Implementing an Auto encoder, Denoising to Force Robust Representations, Sparsity in Auto encoders, The Word2Vec Framework , Implementing the Skip-Gram Architecture

UNIT – IV

Sequence Modelling: Recurrent and Recursive nets: Unfolding Computational Graphs, Recurrent neural networks, Bidirectional RNNs, Encoder-Decoder sequence-to –sequence architectures, Deep Recurrent networks, Recursive neural networks

UNIT – V

The Challenge of Long-Term Dependencies: Echo State Networks, Leaky Units & Other strategies for multiple timescales, The Long Short-Term memory and other Gated RNNs, Optimization for Long-Term Dependencies.

TEXTBOOKS:

1. Nikhil Buduma, Nicholas Locascio, “Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms”, O Reilly Media, 2017.
2. Ian Goodfellow, Yoshua Benjio, Aaron Courville, “Deep Learning” MIT Press,2017

REFERENCE BOOKS:

1. Li Deng and Dong Yu, “Deep learning Methods and Applications”, Now publishers, 2013.
2. Michael Nielsen, “Neural Networks and Deep Learning”, Determination Press 2015.
3. François Duval, “Deep Learning for Beginners: Practical Guide with Python and Tensor flow”, Data Science Series, Create Space Independent Publishers.

B.Tech. - CSE

20CSH4- NATURAL LANGUAGE PROCESSING

L	T	P	Cr.
3	1	0	4

Pre-requisite : Nil

Course Educational Objective: The Objective of the course is to make learn the basic elements of C programming, control structures, derived data types, Modular programming, user defined structures, basics of files and its I/O operations.

Course Outcomes: At the end of this course, the student will be able to

- CO1:** Familiar with the basic components of NLP. (**Understand - L2**)
- CO2:** Applying N-gram models to predict a sequence of text. (**Apply - L3**)
- CO3:** Build a basic language understanding system using preliminary concepts of NLTK library. (**Apply - L3**)
- CO4:** Exposure on advanced techniques for understanding patterns in text (**Apply- L3**)
- CO5:** Understand the semantics of linguistic components in a natural dialogue (**Understand - L2**)

UNIT – I

Introduction

Knowledge in Speech and Language Processing; Ambiguity; Models and Algorithms; Language, Thought and Understanding; History

Regular Expressions

Regular Expression; Words; Corpora; Text Normalization; Minimum Edit Distance

UNIT – II

N-gram Language Models

N-Grams; Evaluating Language Models.

Generalization and Zeros.

Smoothing: Laplace Smoothing; Add-k Smoothing; Backoff and Interpolation; Kneser-Ney Smoothing

UNIT – III

Natural language processing tools in Python (NLTK Package)

Part-I: Introduction to NLTK; Tokenizing; Filtering Stop words; Stemming; Tagging parts of speech; Lemmatizing; Chunking; Chinking

Part-II:

Using Named Entity Recognition (NER); Getting Text to Analyze; Using a Concordance; Making a Dispersion Plot;

UNIT – IV

Information Extraction: Relation Extraction Algorithms; Using Patterns to extract relations; Relation extraction via supervised learning; Semi supervised relation extraction via bootstrapping; Distant Supervision for Relation Extraction; Evaluation of Relation Extraction; Extracting Times; Extracting Events and their Times; Template Filling

UNIT – V

Word Senses and WordNet

- Defining Word Senses; How many senses do words have?
- Relations between senses

WordNet: Sense relations in WordNet; Word Sense Disambiguation; Alternate WSD algorithms and Tasks

TEXTBOOKS:

1. Daniel Jurafsky, James H. Martin, "Speech and Language Processing", Third Edition, PHI, 2020.
2. <https://realpython.com/nltk-nlp-python/#getting-text-to-analyze>

REFERENCE BOOKS:

1. Natural Language Processing with Python: Analysing Text with the Natural Language Toolkit, Steven Bird, Ewan Klein, 2011
2. Applied Text Analysis with Python: Enabling Language-Aware Data Products with Machine Learning, Benjamin Bengfort, Rebecca Bilbro, 2018
3. Speech and Language Processing, 2nd Edition, Daniel Jurafsky, James H. Martin, 2009.