भारतीय सूचना प्रौद्योगिकी संस्थान भागलपुर

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY BHAGALPUR

An Institute of National Importance Under Act of Parliament

2nd Meeting of Board of Academic Programs

for

M.Tech in Artificial Intelligence and Data Science

Dept. of Computer Science and Engineering (CSE)

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY BHAGALPUR

Dept. of Computer Science and Engineering (CSE)

M. Tech.

in

Artificial Intelligence and Data Science

Course Curriculum

Code	Course Name	L	T	P	C
1st Semester					
MA501	Statistical Foundation for Data Science	3	1	0	4
CS501	Advanced Data Structures and Algorithms	3	0	0	3
CS502	Artificial Intelligence	3	0	0	3
CS503	Data Mining	3	0	0	3
	Elective - I	3	0	0	3
CS531	Advanced Algorithms Lab	0	0	3	2
CS532	Artificial Intelligence Lab	0	0	3	2
CS581	Capstone Project – I	0	0	0	1
	Total Credits				21
2nd Semester					
MA502	Mathematical Foundation for Data Science	3	1	0	4
CS504	Machine Learning	3	0	0	3
EC503	Computational Intelligence	3	0	0	3
	Elective - II	3	0	0	3
	Elective - III	3	0	0	3
CS533	Machine Learning Lab	0	0	3	2
CS534	R Programming Lab	0	0	3	2
CS582	Capstone Project – II	0	0	0	1
	Total Credits				21
3rd Semester					
CS591	Major Project-I	0	0	0	10
4th Semester					
CS592	Major Project-II	0	0	0	14
Total Credits					66

Elective Courses (I, II & III)

Code	Course Name	L-T-P-C
CS551	Data Visualization	3-0-0-3
CS552	Advanced Database Systems	3-0-0-3
CS553	Big Data and Cloud Computing	3-0-0-3
CS554	Natural Language Processing	3-0-0-3
CS555	Fundamentals of Information Retrieval	3-0-0-3
CS556	Introduction to Reinforcement Learning	3-0-0-3
CS557	Introduction to Blockchain Technology	3-0-0-3
CS558	Deep Learning and its Applications	3-0-0-3
EC558	Introduction to Pattern Recognition	3-0-0-3

Syllabus

Course Code	Course name	L	T	P	C	Year	Semester
MA501	Statistical Foundations for Data Science	3	1	0	4	1st	1st
Topic	**Contents**						**No. of Lectures**
Module-I	Theory of Probability: Bayes theorem, Random variables, functions of random variables and distribution functions, probability distributions (Binomial, Poisson, Normal), Beta and Gamma Distribution expectations and moments, moment generating functions. Joint, marginal, and conditional distribution's function. Conditional expectations. Covariance, correlation and regression, standard multivariate distributions, Sequence of random variables						8
Module-II	Statistics: convergences in probability and in distribution, law of large numbers, linear Regression; Central limit theorem. Application of Central Limit Theorem.						10
Module-III	Sampling distributions of the sample mean and the sample variance for a normal population; Characteristics of Estimators ,Point and interval estimation; Sampling distributions (Chi-square, t, F,Z). Application of t distribution						10
Module-IV	Basics of hypothesis testing, The Wald test, Type I and Type II errors, t-test Kolmogorov-Smirnov test (KS test), p-values, Permutation test, Pearson correlation coefficient. Neyman Pearson Lemma Theorem. Chi-square test for independence						10
Module-V	Bayesian inference: Bayesian reasoning, Conjugate priors Regression: Simple Linear Regression, Multiple Linear Regression						10
	Total No. of Lectures						48
Text	1. Walepole, Myers, Myers, Ye; *Probability and Statistics for Engineers and Scientists*, Pearson, 9th Edition, 2013. 2. S.C.Gupta and V.K.Kapoor, *Fundamental of Mathematical Statistics*, Sultan Chand & Sons, 12th Edition, 2020. 3. H.C. Saxena and P.U. Surendran, *Statistical Inference*, S Chand & Company Pvt Ltd., 1994.						
Reference	1. R. V. Hogg, J. W. McKean and A. Craig, *Introduction to Mathematical Statistics*, Pearson, 8th Edition, 2019. 2. Larry Wasserman, *All of Statistics: A Concise Course in Statistical Inference*, Springer, Springer Texts in Statistics, 2010. 3. Peter Bruce, Andrew Bruce, Peter Gedeck, *Practical Statistics for Data Scientists*, 2nd Edition, O'Reilly, 2020.						

Course Code	Course name	L	T	P	C	Year	Semester
CS501	Advanced Data Structures and Algorithms	3	0	0	3	1st	1st
Topic	**Contents**						**No. of Lectures**
Module-I	Introduction to Data Structures and Algorithms. Random-access-machine model, concept of problem size, and asymptotic behavior of time/space complexity. Estimation of time/space complexity with order notations. Elementary data-structures: arrays, lists, queues, stacks and their applications. Concept of Randomized Algorithms.						6

Module-II	Sorting algorithms, including the average case analysis of quick-sort. Randomized Quicksort. Binary search algorithm. Analysis of Hashing algorithms. Hashing for insert, search, delete. Heap data structures. Binary trees, binary-search-tree data-structure. Balanced binary-search-tree: Red-Black trees.	8
Module-III	Amortized Analysis. Application to Splay Trees. External Memory ADT - B-Trees. Priority Queues and Their Extensions: Binomial heaps, Fibonacci heaps, applications to Shortest Path Algorithms. Concepts of graphs, paths, trees, cycles. Data structures for graphs: adjacency lists, adjacency matrix. Graph algorithms: Depth First Search, Breadth First Search, Minimum Spanning Tree.	9
Module-IV	Greedy paradigm with examples. Divide and conquer paradigm with examples. Dynamic-programming paradigm with examples. Single-source shortest path computation, topological sorting of a partially ordered set, convex- hull computation, string matching algorithms.	9
Module-V	Intractable Problems: Polynomial Time, class P, Polynomial Time Verifiable Algorithms, class NP, NP completeness and reducibility, NP Hard Problems, NP completeness proofs, Approximation Algorithms.	8
	Total No. of Lectures	40
Text	1. E. Horowitz, S. Sahni and S. Rajasekaran, "*Fundamentals of Computer Algorithms"* Univ. Press, 2nd Ed. 2018 (Reprint). 2. T H Cormen, C E Leiserson, R L Rivest and C Stein, "*Introduction to Algorithms*", PHI Learning, 3rd Ed. 2009.	
Reference	1. E. Horowitz, S. Sahni and S. A-Freed, "*Fundamentals of Data Structures in C*", 2nd Ed., Univ. Press, 2017 (Reprint) . 2. A. M. Tanenbaum, Y. Langsam and M. J. Augenstin, *"Data Structure using C"* by Tanenbaum, Pearson, 1st Ed. 2019 (Reprint).	

Course Code	**Course name**	**L**	**T**	**P**	**C**	**Year**	**Semester**
CS502	Artificial Intelligence	3	0	0	3	1st	1st
Topic	**Contents**						**No. of Lectures**
Module-I	Fundamental issues in intelligent systems: History of artificial intelligence; philosophical questions; fundamental definitions; modeling the world; the role of heuristics.						5
Module-II	Search and constraint satisfaction: Problem spaces; brute-force search; best-first search; two-player games; constraint satisfaction.						10
Module-III	Knowledge representation and reasoning: Formal methods (propositional, predicate logic, first order logic), resolution and unification; Informal methods (frames, scripts), answer extraction; knowledge based systems; logic programming, User interface: Human Computer Interaction, User Interface Components, modules of user interface.						8
Module-IV	AI planning systems: Definition and examples of planning systems; planning as search; operator-based planning; propositional planning; planning algorithms.						8

Module-V	Reasoning under Uncertainty and Learning: probabilistic reasoning; Bayes theorem; Introduction to neural networks and reinforcement learning; Case based reasoning, analytical reasoning, model based reasoning,	9
	Total	40
Text	1. Stuart Russell and Peter Norvig, "*Artifical Intelligence: A Modern Approach*", Pearson; 4th Edition , 2020. 2. Elaine Rich, Kevin Knight and Shivashankar B Nair, "*Artificial Intelligence*", Tata McGraw Hill, 3rd Edition 2017. 3. R.B. Mishra, "*Artificial Intelligence*", PHI Learning Pvt. Ltd., 1st edition, 2010.	
Reference	1. N. J. Nilsson, *"Principles of Artificial Intelligence"*, Narosa Publishing House, 2002. 2. Clocksin & Mellish, "*Programming in PROLOG*", Narosa Publ. House, 2002	

Course Code	**Course Name**	**L**	**T**	**P**	**C**	**Year**	**Semester**
CS503	Data Mining	3	0	0	3	1st	1st
Topic	**Contents**						**No. of Lectures**
Module-I	Overview of the Data Mining and Knowledge Discovery from Databases Process, Basic Data Mining Tasks, Problem Identification, Data Mining Metrics, Data Cleaning (pre-processing, feature selection, data reduction, feature encoding, noise and missing values, etc.).						06
Module-II	What is a data warehouse, A multidimensional data model, Data warehouse architecture, Data warehouse implementation, From data warehousing to data mining.						06
Module-III	Association rule mining, Mining single-dimensional Boolean association rules from transactional databases, Mining multilevel association rules from transaction databases, Association mining to correlation analysis, Constraint-based association mining.						09
Module-IV	Overview of Clustering Methods, Partitioning methods for Clustering, Hierarchical methods for Clustering, Density-based methods for Clustering, Grid-based methods for Clustering.						09
Module-V	Various topics in Classification and Prediction, Sequence mining, Mining complex types of data, Case studies of various Data mining applications, and trends in Data Mining.						10
	Total No. of Lectures						40
Text	1. J. Han and M. Kamber, "*Data Mining: Concepts and Techniques*", Morgan Kaufmann; 3rd edition, 2011.						

Reference	1. Hand David, Mannila Heikki, and Smyth Padhraic. "*Principles of Data Mining"*, Prentice Hall India Learning Private Limited, 2004 2. Pujari, Arun K. "*Data mining techniques.*" Paperback edition, The Orient Blackswan, 2016

Course Code	**Course Name**	**L-T-P-C**	**Year**	**Semester**
MA502	Mathematical Foundation for Data Science	3-1-0-4	1st	2nd
Module	**Description**			**No. of Lectures**
Module-I	Basic properties of matrix and vectors, idea about sparse and dense matrix, Hermitian, skew-Hermitian and unitary matrices, Matrix factorization concept/LU decomposition, Gaussian/Gauss-Jordan elimination, solving Ax=b linear system of equation. Vector spaces, basis, span, orthogonality, orthonormality, linear least square method.			10
Module-II	Eigenvalues, eigenvectors, and diagonalization, singular value decomposition (SVD), Applications. QR-Decomposition and its Applications. Vector Differential Calculus (Gradient, Divergence and Curl. Vector Integral Calculus.			8
Module-III	Convex sets and functions, Extreme points, Hyper planes and Half spaces, Convex cones, Polyhedral sets and cones, hyper planes, Operations preserving convexity. Linear Programming Problem Formulation, solution by Graphical Method, Theory of Simplex Method, Simplex Algorithm			10
Module-IV	Convex optimization: First-order optimality condition, quadratic minimization, Semi definite program, Conic program. Gradient descent: Fixed step size, Backtracking line search, Exact line search, Convergence analysis; Gradient boosting; Subgradient.			10
Module-V	Recurrence relations and equations, Integer programming, Constraint programming, knapsack problem. Graph properties: connected components, degree, maximum flow/minimum cut concepts, graph coloring.			10
	Total No. of Lectures			48

Text/Reference

1. Gilbert Strang, *Linear Algebra and Its Applications*, Cengage Learning, 4th Edition, 2007.
2. David S. Watkins, *Fundamentals of Matrix Computations*, Wiley, 3rd Edition 2010.
3. Dimitri P. Bertsekas, *Convex Optimization Theory*, Universities Press, 2010.
4. Kenneth Rosen, *Discrete Mathematics and Its Applications*, McGraw Hill Education, 7th Edition, 2011.f

Course Code	Course Name	L	T	P	C	Year	Semester
CS504	Machine Learning	3	0	0	3	1st	2nd
Topic	**Content**						**No. of Lectures**
Module I	Introduction: History of machine learning, Basic concepts						5
Module II	Supervised learning: Supervised learning setup, LMS, Logistic regression, Perceptron, Backpropagation, neural networks, Exponential family, Generative learning algorithms, Gaussian discriminant analysis, Naive Bayes, Support vector machines, Model selection and feature selection, Ensemble methods: Bagging, boosting.						10
Module III	Learning theory: Bias/variance trade-off, Union and Chernoff/Hoeffding bounds, VC dimension, Worst case (online) learning.						8
Module IV	Unsupervised learning: Clustering K-means, EM. Mixture of Gaussians, Factor analysis, PCA (Principal components analysis), ICA (Independent components analysis).						9
Module V	Miscellaneous topics: Hypothesis testing, cross-validation, quadratic discriminant Analysis, adaptive hierarchical clustering, gradient boosting.						8
						Total No. of Lectures	40
Text	1. Ethem Alpaydin, "*Introduction to Machine Learning"*, PHI, Third Edition, 2015. 2. Marsland, Stephen. "*Machine learning: an algorithmic perspective"*, Chapman and Hall/CRC, 2nd edition, 2014. 3. Tom Mitchell, "*Machine Learning*", McGraw Hill, First edition 2017.						
Reference	1. Murphy, Kevin, "*Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series)"* , The MIT Press; Illustrated edition, 2012. 2. Müller, Andreas C., and Sarah Guido, "*Introduction to machine learning with Python: a guide for data scientists"*, O'Reilly, 1st edition, 2016.						

Course Code	Course name	L	T	P	C	Semester
EC503	Computational Intelligence	3	0	0	3	2nd
Topic	**Contents**					**No. of Lectures**
Module-I	**Introduction to Computational Intelligence:** Intelligence machines, Computational intelligence paradigms, Soft computing constituents and conventional Artificial intelligence, Neuro-Fuzzy and soft computing characteristics					7
Module-II	**Rule-Based Expert Systems and Fuzzy Expert Systems:** Rule-based expert systems, Uncertainty management, Fuzzy sets and operations of fuzzy sets, Fuzzy rules and fuzzy inference, Fuzzy expert systems, Case study: fuzzy logic controller for various applications					9
Module-III	**Artificial Neural Networks:** Fundamental neuro-computing concepts: artificial neurons, activation functions, Neural network architectures, learning rules, Supervised learning neural networks: multi-layer feed forward neural networks, simple recurrent neural networks, time delay neural networks, supervised learning algorithms, Back propagation algorithm, Radial basis function networks Unsupervised learning neural networks, self-organizing feature maps, Deep neural networks and learning algorithms					9

Module-IV	**Evolutionary techniques**: Genetic Algorithm, Evolutionary computation**:** Chromosomes, fitness functions, and selection mechanisms, Genetic algorithms: crossover and mutation, Genetic programming, Evolution strategies, PSO, ACO, BFO	9
Module-V	**Hybrid Intelligent Systems:** Neural expert system, neuro-fuzzy systems, Evolutionary neural network, case study of Neuro-fuzzy based systems.	7
	Total No. of Lectures	**40**
Text	1. S. Rajasekaran, G. A. Vijayalaksmi Pai, *Neutral Networks, Fuzzy Logic and Genetic Algorithm: Synthesis and Applications*, PHI Learning, 2nd edition, 2017. 2. J. S .R. Jng, C. T. Sun, E. Mizutani, *Neuro-Fuzzy and Soft Computing*, Pearson Education, 1st edition, 2015. 3. S. N. Deepa, S. N. Sivanandam, *Principles of Soft Computing*, John Wiley, 3rd edition, 2018.	
Reference	1. Timothy J. Ross, *Fuzzy logic with Engineering Applications*, McGraw-Hill, 3rd edition,2011. 2. Simon Haykin, *Neural Networks: A Comprehensive Foundation*, Pearson, 3rd edition, 2009.	

Elective (I, II & III) Syllabus

Course Code	**Course name**	**L**	**T**	**P**	**C**	**Year**	**Semester**
CS551	Data Visualization	3	0	0	3	1st	
Topic	**Content**						**No. of Lectures**
Module-I	**Introduction to Data Visualization and Basic Statistics** Overview of data visualization - Data Abstraction -Analysis: Four Levels for Validation- Task Abstraction - Analysis: Four Levels for Validation, Basic Graphics: line, bar, box, histogram plots, Trellis, Scatter plots, Basic Statistics: mean median, mode, percentile, quantile, Frequency Distribution, Histogram Analysis, Data: Distribution, Types of Data Distribution, Hypothesis Testing						12
Module-II	**Visualization Techniques and Tools** Scalar and point techniques Color maps Contouring Height Plots - Vector visualization techniques Vector properties Vector Glyphs Vector Color Coding Stream Objects, Introduction to data visualization tools- Tableau - Visualization using R.						10
Module-III	**Visual Analytics** Visual Variables- Networks and Trees - Map Color and Other Channels- Manipulate View, Arrange Tables Geo Spatial data Reduce Items and Attributes.						07
Module-IV	**Diverse Types Of Visual Analysis** Time-Series data visualization Text data visualization Multivariate data visualization and case studies						05
Module-V	**Visualization Dashboard Creations** Dashboard creation using visualization tools for the use cases: Finance-marketing-insurance healthcare etc., and Recent Trends						06
	Total No. of Lectures						40

Text	1. Tamara Munzer, "*Visualization Analysis and Design"*, CRC Press, 2nd edition, 2014 2. AlexandruTelea, "*Data Visualization Principles and Practice"*, CRC Press, 2nd edition, 2014. 3. Paul J. Deitel, Harvey Deitel, "*Java SE8 for Programmers (Deitel Developer Series)"*, 3rd Edition, 2014. 4. Y. Daniel Liang, "*Introduction to Java programming-comprehensive version"*, Pearson ltd., 10th Edition, 2015
Reference	1. Paul Deitel Harvey Deitel ; "*Java, How to Program"*, Prentice Hall; 9th edition , 2011. 2. Cay Horstmann, "*BIG JAVA"*, John Wiley Sons, 4th edition, 2009 3. Nicholas S. Williams, "*Professional Java for Web Applications"*, Wrox Press, 1st edition, 2014.

Course Code	**Course name**	**L**	**T**	**P**	**C**	**Year**	**Semester**
CS552	Advanced Database Systems	3	0	0	3	1st	
Topic	**Contents**						**No. of Lectures**
Module-I	**Overview of Databases and Relational Databases** Introduction to database management, data abstraction and system Structure. Entity relational model, entity set, relationship sets, Mapping cardinalities, keys, E-R diagrams, Relational model, database schema, relational algebra, outer join and manipulation of databases.						09
Module-II	**Normalization, Query Processing and Optimization** Functional dependencies, Rules about functional dependencies, Muiltivalued, Join Dependency, Normal Forms, 1NF, 2NF, 3NF, BCNF, 4NF, Translating SQL queries into relational algebra, Algorithms for External Sorting, Algorithm for SELECT, JOIN, and etc. Operations, Using Heuristics in Query Optimization, Using Selectivity and Cost Estimates in Query Optimization, Overview of Query Optimization in Oracle.						08
Module-III	**Advanced Transaction Processing and Recovery** Introduction, ACID property, State of transaction, Serializability and recoverability, lock based protocols, Deadlock handling, Recovery in DBMS. multiple level granularity, B and B+ trees, Introduction of NoSQL, Need of NoSQL, CAP Theorem, different NoSQL data models: Key-value stores, Column families, Document databases, Graph databases.						08
Module-IV	**Objected Oriented and Object Relational Databases** Modeling Complex Data Semantics, Specialization, Generalization, Aggregation and Association, Objects, Object Identity, Equality and Object Reference, Architecture of Object Oriented and Object Relational Databases, Object Definition Language, Object Query Language.						05
Module-V	**Parallel, Distributed, and Spatial Databases** Design of Parallel Databases, Parallel Query Evaluation, Distributed Data Storage – Fragmentation & Replication, Location and Fragment Transparency, Distributed Query Processing and Optimization, Distributed Transaction Modeling and Concurrency Control, Distributed Deadlock, Distributed Commit Protocols. Definition & key terms, Data Standards Coordinate Reference Systems Spatial Metadata, Non Spatial and Spatial Indexes, Spatial query language.						10
						Total No. of Lectures	40

Text	1. Elmasri, Navathe, Somayajulu, and Gupta, "*Fundamental of Database Systems*", Pearson Education, 7th edition, 2016. 2. A.Silberschatz, H. F. Korth, & S. Sudhatshan, "*Database System Concepts*", McGraw Hill, 7th edition, 2019.
Reference	1. Thomas Connolly and Carolyn Begg, "*Database Systems*", Pearson Education, 6th edition, 2015 2. C.J. Date, A. Kanana, S. Swamynathan, "*An Introduction to database system*", Pearson Education, 8th edition, 2020. 3. Rani Chakrabarti and Shilbhadra Dasgupta, "*Advanced Database Management System*", DreamTech, 2016. 4. Rajesh Narang, "*Database management System*", PHI, 3rd edition, 2019. 5. Introduction to Database Management system by ISRD Group, Tata McGraw Hill, 2005.

Course Code	**Course name**	**L**	**T**	**P**	**C**	**Year**	**Semester**
CS553	Big Data and Cloud Computing	3	0	0	3	1st	
Topic	**Contents**						**No. of Lectures**
Module-I	Foundations of cloud computing: Distributed systems, grid, parallel and cloud computing systems, cloud computing service models, deployments models, characteristics of cloud computing.						8
Module-II	Virtualization and Cloud Programming Paradigms: Virtualization tools, server virtualization, storage virtualization, network virtualization, application virtualization. Cloud programming paradigms: Google, Microsoft, AWS.						8
Module-III	Cloud Security and Introduction to Big Data: Securing the cloud, data security, network security, host security, recovery in cloud. Introduction to big data, characteristics and applications of big data						8
Module-IV	Big Data Technologies: Big data storage and computation, HDFS, MapReduce programming paradigm, MapReduce algorithms, implementation and running of MapReduce Job.						8
Module-V	Big Data Analysis and Applications: Big data analysis: Spark, Hive, Pig, HBase. Data streaming, Programming big data applications.						8
Total No. of Lectures							40
Text	1. Kai Hwang, Geoffrey C. Fox and Jack J. Dongarra, "*Distributed and Cloud Computing from Parallel Processing to the Internet of Things*", Morgan Kaufmann, 1st Edition, 2012. 2. Barrie Sosinsky, "*Cloud Computing Bible*", Wiley, 1st Edition, 2011. 3. Jules J. Berman. "*Principles and Practice of Big Data*", Academic Press, 2nd Edition, 2018.						
Reference	1. Buyya R., Broberg J., Goscinski A. M., "*Cloud Computing – Principles and Paradigms*",Wiley, 1st Edition, 2011. 2. Tom White,"*Hadoop: The Definitive Guide*", O'Reilly, 4th Edition, 2015.						

	3. Edward Capriolo, Dean Wampler, Jason Rutherglen, "*Programming Hive: Data Warehouse and Query Language for Hadoop*", 1st edition, O'Reilly, 2012.

Course Code	**Course name**	**L**	**T**	**P**	**C**	**Year**	**Semester**
CS554	Natural Language Processing	3	0	0	3	1st	
Topic	**Contents**						**No. of Lectures**
Module-I	Introduction to Natural Language (NL): why is it hard to process NL, linguistics fundamentals, etc. why NLP is useful, classical problems.						**7**
Module-II	Words: Structure (spellcheck, morphology using FSTs). Words: Semantics (Basic ideas in Lexical Semantics, WordNet and WordNet based similarity measures, Distributional measures of similarity, Concept Mining using Latent Semantic Analysis). Semantics (Word Sense Disambiguation; supervised, unsupervised and semi-supervised approaches) Words: Parts of Speech (POST using Brill's Tagger and HMMs).						9
Module-III	Sentences: Basic ideas in compositional semantics, Classical Parsing (Bottom up, top down, Dynamic Programming: CYK parser) Sentences: Parsing using Probabilistic Context Free Grammars and EM based approaches for learning PCFG parameters. Language Models: n-grams, smoothing, class-based, brown clustering.						8
Module-IV	Machine Translation (rule-based techniques, Statistical Machine Translation (SMT), parameter learning in SMT (IBM models) using EM). Information Extraction: Introduction to Named Entity Recognition and Relation Extraction.						8
Module-V	Natural Language Generation (NLG): the potential of using ML for NLG Additional topics: Advanced Language Modelling (including LDA), other applications like summarization and text classifications.						8
Total No. of Lectures							40
Text	1. Daniel Jurafsky and James H. Martin, "*Speech and Language Processing: An Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics",* 2nd Ed. Pearson, 2020. 2. Jacob Eisenstein, *"Introduction to Natural Language Processing"*, MIT Press, New Ed, Oct. 2019.						
Reference	1. L. Rabiner, B-H Juang and B. Yegnanarayana, "*Fundamentals of Speech Recognition",* Pearson, 1st Ed. 2009. 2. Christopher D. Manning and Hinrich Schütze, "*Foundations of Statistical Natural Language Processing"*. MIT Press Ed. May 1999.						

Course Code	Course name	L	T	P	C	Year	Semester
CS555	Fundamentals of Information Retrieval	3	0	0	3	1st	
Topic	**Contents**						**No. of Lectures**
Module-I	Introduction to Information retrieval, Information Retrieval Vs Information Extraction, The nature of unstructured and semi-structured text. Inverted index and Boolean queries.						6
Module-II	Text encoding: tokenization, stemming, stop words, phrases, index optimization. Index Compression: lexicon compression and postings lists compression. Gap encoding, gamma codes, Zipf's Law. Index construction. Postings size estimation, merge sort, dynamic indexing, positional indexes, n-gram indexes, real-world issues.						8
Module-III	Retrieval Models; Boolean, vector space, TFIDF, Okapi, probabilistic, language modeling, latent semantic indexing. Vector space scoring. The cosine measure. Efficiency considerations. Document Length normalization. Relevance feedback and query expansion. Rocchio.						8
Module-IV	Performance Evaluation; Evaluating search engines. User happiness, precision, recall, F-measure. Creating test collections: kappa measure, interjudge agreement. Introduction to text classification. Naive Bayes models. Spam filtering.						9
Module-V	Clustering versus classification. Partitioning methods. k-means clustering. Mixture of gaussians model. Hierarchical agglomerative clustering. Clustering terms using documents. Introduction to advanced topics: Summarization, Topic detection and tracking, Personalization, Question answering, Cross Language information retrieval, Hypertext, web crawling, search engines, ranking, link analysis, XML retrieval.						9
	Total						40
Text	1. Christopher D. Manning, Raghavan Prabhakar, Schütze Hinrich, "*Introduction to Information Retrieva*l", Cambridge University Press, Illustrated Edition, 2008. 2. Ricardo Baeza-Yates and Berthier Ribeiro-Neto, "*Modern Information Retrieval*", Pearson Education India, 1st edition, 2003.						
Reference	1. William B. Frakes, Ricardo Baeza-Yates, "*Information Retrieval Data Structures and Algorithms*", Prentice Hall, 1st edition, 1992. 2. Soumen Chakrabarti, "*Mining the Web*", Morgan-Kaufmann Publishers, 1st edition, 2002. 3. Bing Liu, "*Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*", Springer, Corr. 2nd edition, 2011.						

Course Code	Course name	L	T	P	C	Year	Semester
CS556	Introduction to Reinforcement Learning	3	0	0	3	1st	
Topic	**Contents**						**No. of Lectures**
Module-I	Introduction: Recent Advances in Reinforcement Learning (RL), Atari Game Player, Alpha Go, and other case studies, Rewards and returns, Markov Decision Processes, Value functions, optimality and approximation.						7

Module-II	Dynamic programming: value iteration, policy iteration, asynchronous DP, generalized policy iteration.	7
Module-III	Monte-Carlo methods: policy evaluation, roll outs, on policy and off policy learning, importance sampling.	8
Module-IV	Temporal Difference learning : TD prediction, Optimality of TD(0), SARSA, Q-learning, R-learning, Games and after states, n-step TD prediction, TD (lambda), forward and backward views, Q (lambda), SARSA (lambda).	8
Module-V	Function Approximation: Value prediction, gradient descent methods, linear function approximation, ANN based function approximation, Policy Gradient methods.	10
	Total	40
Text	1. Richard S. Sutton and Andrew G. Barto. "*Introduction to Reinforcement Learning*", MIT Press. 2nd edition, 2017. 2. Dimitri Bertsikas and John G. Tsitsiklis, "*Neuro Dynamic Programming.*", Athena Scientific, 1st addition, 1996.	
Reference	Lapan, M. "*Deep Reinforcement Learning Hands-On*", Packt Publishing Ltd.2nd edition, 2020	

Course Code	**Course name**	**L**	**T**	**P**	**C**	**Year**	**Semester**
CS557	Introduction to Blockchain Technology	3	0	0	3	1st	
Topic	**Contents**						**No. of Lectures**
Module-I	Basics of Blockchain Technology: Cryptography, Hashing, MD5 message digest algorithm, secure hash algorithm (SHA-1), security of hash functions, digital signatures.						8
Module-II	Introduction to Blockchain Technology: Blockchain introduction, applications, opportunities and challenges in blockchain technology.						6
Module-III	Bitcoin and Cryptocurrency: Bitcoin introduction, bitcoin mining, bitcoin case studies, understanding cryptocurrency.						9
Module-IV	Blockchain Technology Applications: Ethereum blockchain, ethereum virtual machine and gas, smart contracts.						9
Module-V	Blockchain Case studies: ICO case study, banking case study, blockchain white papers, study of recent trends and features of blockchain technology.						8
	Total No. of Lectures						40

Text	1. Roger Wattenhofer, "*The Science of the Blockchain*", Createspace Independent Pub, 1st edition, 2016 2. Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller, and Steven Goldfeder, "*Bitcoin and cryptocurrency technologies: a comprehensive introduction*", Princeton University Press; Illustrated edition, 2016
Reference	1. Behrouz A Forouzan, Debdeep Mukhopadhyay, "*Cryptography and Network Security*", Mc-GrawHill, 3nd Edition, 2016. 2. Melanie Swan, "*Blockchain"*, O′Reilly, 1st Edition, 2015.

Course Code	**Course name**	**L**	**T**	**P**	**C**	**Year**	**Semester**
CS558	Deep Learning and its Applications	3	0	0	6	1st	
Topic	**Contents**						**No. of Lectures**
Module-I	**Introduction**: Review of Feedforward neural network, brief review of concepts from optimization, Multilayer Perceptron, Difficulty of training deep neural networks, Discussion on deep learning frameworks.						06
Module-II	**Convolutional Neural Networks**: Construction of foundational layers of CNNs (pooling, convolutions) and to stack them properly in a deep network to solve multi-class image classification problems. Discussions on various convnet architectures: LeNet, AlexNet.						09
Module-III	**Recurrent Neural Networks**: Back propagation through time, Long Short Term Memory, Gated Recurrent Units, Bidirectional LSTMs, Bidirectional RNNs						10
Module-IV	**Generative models:** Generative Adversarial Networks (GAN), Deep Convolutional GAN (generative adversarial network).						10
Module-V	Recent trends: Variational Autoencoders, Multi-task Deep Learning, Applications: Vision, NLP (just an overview of different applications in PyTorch)						5
	Total No. of Lectures						40
Text	1. Ian Goodfellow and Yoshua Bengio and Aaron Courville, "*Deep Learning*," MIT Press, 2016. 2. Christopher Bishop, "*Pattern Recognition and Machine Learning*" , Springer; 1st edition, 2006.						
Reference	1. Ian Pointer, "*Programming PyTorch for Deep Learning*", Shroff/O'Reilly; First edition 2019. **2.** Sherin Thomas & Sudhanshu Passi, "*PyTorch Deep Learning Hands-On*" , Packt Publishing, 2019						

Course Code	**Course name**	**L**	**T**	**P**	**C**	**Semester**
EC558	Introduction to Pattern Recognition	3	0	0	3	
Topic	**Contents**					**No. of Lectures**
Module-I	**Introduction:** Introduction fundamentals and definitions, Feature vectors, Classifiers, Supervised and Unsupervised learning, Bayesian decision theory					7

Module-II	**Features**: types and traits, scaling ordering, measurements, normalization, invariance, feature properties, dimensionality reduction of feature space, dimensionality reduction by feature selection, PCA, KPCA, ICA	8
Module-III	**Parameter estimation**: Maximum likelihood estimation (MLE), least squares estimation (LSE), Method of minimum variance & unbiased Estimation (MVUE); parameter free methods: KNN, Clustering; Special classifiers: linear regression, LDA, SVM, CNN.	8
Module-IV	**Classifiers and learning:** Fundamentals of classifiers, Linear Classifiers, Nonlinear Classifiers. Unsupervised and semi supervised learning: Learning from unclassified data, **Clustering:** Basic Concepts of clustering, Hierarchical agglomerative clustering, K-means partitional clustering, semi-supervised learning with expectation maximization using labelled and unlabelled data; Characteristics analysis of different classifiers	9
Module-V	**Classification with nominal features** : decision tree, random forest; classifier independent concepts, **Combinations of classifiers:** boosting, voting , stacking	8
	Total No. of Lectures	40
Text	1. Christopher Bishop, *Pattern Recognition and Machine Learning*, Springer, 2nd printing edition, 2011. 2. Richard O. Duda, Peter E. Hart, David G. Stork, *Pattern Classification*, John Wiley, 2nd edition, 2002.	
Reference	1. Kevin P. Murphy, *Machine Learning: A Probabilistic Perspective*, The MIT Press; Illustrated edition, 2012. 2. S. Theodoridis, K. Koutroumbas, "*Pattern Recognition"*, Academic Press, 4th edition, 2008.	