INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-521 **Course Title**: Mathematics for Data Science
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Autumn **7. Subject Area:** PCC
8. **Pre-requisite:** Nil
9. **Objective:** To introduce students to the various Mathematical concepts to be used in ML and DS.
10. **Details of the Course**

S.No.	Contents	Contact hours
1.	**Basics of Linear Algebra**: Representation of vectors; Linear dependence and independence; vector space and subspaces (definition, examples and concepts of basis); linear transformations; range and null space; matrices associated with linear transformations; special matrices; eigenvalues and eigenvectors with applications to data problems; Least square and minimum normed solutions	10
2.	**Matrices in Machine Learning Algorithms**: projection transformation; orthogonal decomposition; singular value decomposition; principal component analysis and linear discriminant analysis	4
3.	**Gradient Calculus:** Basic concepts of calculus: partial derivatives, gradient, directional derivatives, Jacobian, Hessian.	4
4.	**Optimization:** Convex sets, Convex function and their properties, Unconstrained and Constrained Optimization, Numerical Optimization Techniques for Unconstrained Optimization, Derivative-Free methods (Golden Section, Fibonacci Search Method, Bisecting Method), Methods using Derivatives (Newton's Method, Steepest Descent Method), Penalty Function Methods for Constrained Optimization.	9
5.	**Probability:** Basic concepts of probability, conditional probability, total probability, independent events, Bayes' theorem, random variable, Moments, moment generating functions, some useful distributions, Joint distribution, conditional distribution, transformation of random variables, covariance, correlation.	8
6.	**Statistics:** Random sample, sampling techniques, statistics, sampling distributions, mixture models.	7
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1	M. P. Deisenroth, A. A. Faisal, C. S. Ong, Mathematics for Machine Learning, Cambridge University Press (1st edition)	2020
2	S. Axler, Linear Algebra Done Right. Springer International Publishing (3rd edition)	2015
3	J. Nocedal and S. J. Wright, Numerical Optimization. New York: Springer Science+Business Media	2006
4	E. Kreyszig, Advanced Engineering Mathematics, John Wiley and Sons, Inc., U.K. (10th Edition)	2015
5	R. A. Johnson, I. Miller, and J. E.Freund, "Miller & Freund's Probability and Statistics for Engineers", Prentice Hall PTR, (8th edition)	2011
6	C. Mohan and K. Deep: "Optimization Techniques", New Age Publishers, New Delhi.	2009

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-523 **Course Title**: Principles of Database Systems
2. **Contact Hours:** **L:** 3 **T:** 0 **P:** 2
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 10-25 **PRS:** 25 **MTE:** 15-25 **ETE:** 30-40 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Autumn **7. Subject Area:** PCC
8. **Pre-requisite:** Nil
9. **Objective:** To impart the knowledge of basic Data Base Management Systems.
10. **Details of the Course**

S.No.	Contents	Contact hours
1.	Purpose of Database System, Views of data, Data Models, Database Languages-Database System Architecture, Database users and Administrator, Entity Relationship model (E-R model) – E-R Diagrams, Introduction to relational databases.	8
2.	The relational Model – The catalog types, Keys, Relational Algebra, Domain Relational Calculus, Tuple Relational Calculus, Fundamental operations, Additional Operations, SQL fundamentals - Integrity, Triggers, Security, Advanced SQL features, Embedded SQL, Dynamic SQL, Missing Information, Views, Introduction to Distributed Databases and Client/Server Databases.	10
3.	PL/SQL- Basic and Advanced Concepts.	8
4.	Functional Dependencies – Non-loss Decomposition, Functional Dependencies, First, Second, Third Normal Forms, Dependency Preservation, Boyce-Codd Normal Form, Multi-valued Dependencies and Fourth Normal Form – Join Dependencies and Fifth Normal Form.	8
5.	Transaction Concepts - Transaction Recovery, ACID Properties, System Recovery, Media Recovery, Two Phase Commit, Save Points – SQL Facilities for recovery, Concurrency, Need for Concurrency, Locking Protocols, Two Phase Locking, Intent Locking, Deadlock, Serializability – Recovery Isolation Levels – SQL Facilities, for Concurrency.	8
Total		**42**

11. List of Practical:

1	DDL and DML commands in SQL-I
2	DDL and DML commands in SQL-II
3	Query designing in SQL
4	Aggregate functions and sorting concepts on created tables
5	Single row operation functions
6	View and displaying data from multiple tables
7	Aggregating data using group functions
8	Designing query with concepts of sub-queries
9	Pl-SQL
10	To implement the concepts of security and privileged
11	Implementing transaction control commands

12. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1	M. P. Deisenroth, A. A. Faisal, C. S. Ong, Mathematics for Machine Learning, Cambridge University Press (1st edition)	2020
2	S. Axler, Linear Algebra Done Right. Springer International Publishing (3rd edition)	2015
3	J. Nocedal and S. J. Wright, Numerical Optimization. New York: Springer Science+Business Media	2006
4	E. Kreyszig, Advanced Engineering Mathematics, John Wiley and Sons, Inc., U.K. (10th Edition)	2015
5	R. A. Johnson, I. Miller, and J. E.Freund, "Miller & Freund's Probability and Statistics for Engineers", Prentice Hall PTR, (8th edition)	2011
6	C. Mohan and K. Deep: "Optimization Techniques", New Age Publishers, New Delhi.	2009

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

Subject Code: AID-505 **Course Title**: Machine Learning

1. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Autumn **7. Subject Area:** PCC
8. **Pre-requisite:** Nil
9. **Objective:** To provide an understanding of the theoretical concepts of machine learning and prepare students for research or industry application of machine learning techniques.

10. **Details of the Course**

S.No.	Contents	Contact hours
1.	**Introduction:** Well-posed learning problems, examples of machine learning applications, model selection and generalization, concept learning, inductive learning hypothesis, inductive bias. Information theory: entropy, mutual information, KL divergence	4
2.	**Performance Optimization:** Directional Derivatives, Minima, Necessary Conditions for Optimality, Convex function, Gradient Descent, Stable learning rates, Newtons Method, Conjugate gradient method, The Levenberg-Marquardt algorithm.	4
3.	**Linear Classification:** Linear classifier, Logistic Regression, Decision Boundary, Cost Function Optimization, Multi-class Classification, Bias and Variance, L1 and L2 Regularization, feature reduction, Principal Component Analysis, Singular Value Decomposition	4
4.	**Artificial Neural Networks:** Perceptron, Linear Networks, Multi-layer Networks, Forward propagation, Backward propagation, Alternative activation functions, variations on backpropagation, Deep neural networks.	5
5.	**Decision tree learning:** Decision tree representation, appropriate problems for decision tree learning, hypothesis space search in decision tree learning, inductive bias in tree learning, avoiding overfitting the data, alternative measures for selecting attribute values, ensemble methods, bagging, boosting, random forest	5
6.	**Support Vector Machines:** Computational learning theory, probably approximately correct (PAC) learning, sample complexity and VC dimension, linear SVM, soft margin SVM, kernel functions, nonlinear SVM, Multiclassclassification using SVM, Support vector regression.	5
7.	**Instance based learning:** K-nearest neighbor learning, distance weighted neighbor learning, locally weighted regression, adaptive nearest neighbor methods, The Concept of Unsupervised Learning, Competition networks, K-means clustering algorithm.	3

8.	**Bayesian Learning:** Bayes theorem, maximum likelihood and least squared error hypotheses, Naive Bayes classifier, Bayesian belief networks, gradient ascent training of Bayesian networks, learning the structure of Bayesian networks, the EM algorithm, mixture of models, Markov models, hidden Markov models.	7
9.	**Reinforcement learning:** the learning task, Q learning, convergence, temporal difference learning, nondeterministic rewards and actions, generalization, relationship to dynamic programming.	5
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1.	T. Mitchell, Machine Learning, McGraw Hill	1997
2.	Christopher Bishop, Pattern Recognition and Machine Learning, Springer	2006
3.	K. Murphy. Machine Learning: A probabilistic perspective, MIT Press	2012
4.	Hastie, Tibshirani, Friedman, Elements of statistical learning, Springer	2011
5.	I. Goodfellow, Y. Bengio and A. Courville. Deep Learning. MIT Press	2016
6.	Richard S. Sutton and Andrew G. Barto, Reinforcement Learning: An Introduction, MIT Press	2018

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-525 **Course Title**: Data Structures and Algorithms
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Autumn **7. Subject Area:** PCC
8. **Pre-requisite:** Nil
9. **Objective:** To introduce advanced concepts in data structures and algorithms
10. **Details of the Course**

S.No.	Contents	Contact hours
1.	**Complexity Analysis:** Time and Space complexity of algorithms, asymptotic analysis, big O and other notations.	8
2.	**Linear Lists, Stacks and Queues, Hashing and Trees: A**bstract data types, sequential and linked implementations, equivalence problem, hash table, collision avoidance, linear open addressing, chains, uses of hash tables, insertion, deletion and search operations for sequential and linked lists, doubly linked lists, circular lists, skip lists, applications of lists in bin sort, radix sort, sparse tables, Binary trees and their properties, tree traversal methods and algorithms.	12
3.	**Algorithmic Techniques**: Algorithm design strategies, divide and conquer, merge sort, quick sort and its performance analysis, randomized quick sort, Strassen's matrix multiplication; Greedy method and its applications, knapsack problem; Dynamic programming and its performance analysis, optimal binary search trees, 0/1 knapsack problem; Traveling salesman problem; Back-tracking, n-queens problem; Branch and bound examples, 15-puzzle problem, 0/1 knapsack, traveling salesman problem.	12
4.	**Graph Algorithms**: DFS and BFS, spanning trees, bi-connectivity; Minimum cost spanning trees: Kruskal's, Prim's and Sollin's algorithms; Path finding and shortest path algorithms; Topological sorting; Bipartite graphs. P and NP-classes, NP-hard problems, reduction.	10
Total		**42**

11. **Suggested Books:**

S.No.	Name of Authors/Book/Publisher	Year of Publication / Reprint
1.	Wirth, N., "Algorithms and Data Structures", Prentice-Hall of India.	2017
2.	Brad Miller and David Ranum, Luther College, "Problem Solving with Algorithms and Data Structures Using Python," Franklin, Beedle &Associates.	2013
3.	Cormen T, Introduction to Algorithms, MIT Press, 3rd Edition.	2009

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code**: AID-527 **Course Title**: Programming for DS
2. **Contact Hours**: **L:** 0 **T:** 0 **P:** 4
3. **Examination Duration (Hrs.):** **Theory:** 0 **Practical:** 2
4. **Relative Weightage:** **CWS:** 0 **PRS:** 50 **MTE**: 0 **ETE:** 0 **PRE:** 50
5. **Credits:** 2 **6. Semester:** Autumn **7. Subject Area:** PCC
8. **Pre-requisite:** Nil
9. **Objective:** This course's objective is to provide hands-on experience on the various programming components for Data Science.
10. **Details of the Course:**

S.No.	Contents	Contact hours
1	Python: Basics, Numpy, Pandas, and Matplotlib	16
2	Scikit-Learn and NLTK	12
3	Tensor Flow and Keras	12
4	Tensor Flow Lite: Deploy machine learning systems on mobile (Android application) (Android Studio, Kotlin/Java)	16
Total		56

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication / Reprint
1	Jake VanderPlas “Python Data Science Handbook,” First Edition, O'Reilly Media, Inc.	2016
2	Wes McKinney “Python for Data Analysis: Data Wrangling with Pandas, NumPy, and I Python,” Second Edition, O'Reilly Media, Inc.	2017
3	Pramod Singh and Avinash Manure “Learn TensorFlow 2.0: Implement Machine Learning and Deep Learning Models with Python,” First Edition, Apress	2020
4	Aurélien Géron “Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow,” Second Edition, O'Reilly Media, Inc.	2019
5	Bill Phillips “Android Programming: The Big Nerd Ranch Guide,” Third Edition,Big Nerd Ranch Guide	2017

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-571 **Course Title**: Big Data Analytics
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.): Theory:** 3 **Practical:** 0
4. **Relative Weightage: CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** The purpose of this course is to introduce the students with Big Data Storage Systems and important algorithms that form the basis of Big Data Processing. The course also introduces the students with major application areas of Big Data Analytics.
10. **Details of the Course**

S.No.	Contents	Contact hours
1.	**Introduction to Big Data:** Introduction to Big Data, The four dimensions of Big Data: volume, velocity, variety, veracity, Drivers for Big Data, Introducing the Storage, Query Stack, Revisit useful technologies and concepts, Real-time Big Data Analytics	6
2.	**Distributed File Systems:** Hadoop Distributed File System, Google File System, DataConsistency	6
3.	**Big Data Storage Models:** Distributed Hash-table, Key-Value Storage Model (Amazon's Dynamo), Document Storage Model (Facebook's Cassandra), Graph storage models	10
4.	**Scalable Algorithms:** Mining large graphs, with focus on social networks and web graphs. Centrality, similarity, a 11-distances sketches, community detection,1 ink analysis, spectral techniques. Map-reduce, Pig Latin, and NoSQL, Algorithms for detecting similar items, Recommendation systems, Data stream analysis algorithms,Clustering algorithms, Detecting frequent items	10
5.	**Big Data Applications:** Advertising on the Web, Web Page Quality Ranking, Mining Social-Networking Group, Human Interaction with Big-Data. Recommendation systems with case studies of Amazon's, Item-to-item recommendations and Netflix Prize, Link Analysis with case studies of the PageRankalgorithm and the spam farm analysis, Crowdsourcing	6
6.	**Big Data Issues:** Privacy, Visualization, Compliance and Security, Structured vs Unstructured Data	4
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication / Reprint
1.	Ohlhorst, Frank J. Big data analytics: turning big data into big money. Vol. 65. John Wiley & Sons, 2012.	2012
2.	Russom, Philip. "Big data analytics." TDWI best practices report, fourthquarter 19, no. 4 (2011): 1-34.	2011
3.	Marr, Bernard. Big Data: Using SMART big data, analytics and metrics to make better decisions and improve performance. John Wiley & Sons, 2015.	2015
4.	LaValle, Steve, Eric Lesser, Rebecca Shockley, Michael S. Hopkins, and Nina Kruschwitz. "Big data, analytics and the path from insights to value." MIT sloan management review 52, no. 2 (2011): 21-32.	2011
5	Leskovec, Jure, Anand Rajaraman, and Jeffrey David Ullman. Mining of massive data sets. Cambridge university press, 2020.	2020

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Department of Computer Science and Engineering

1. **Subject Code:** CSN-515 **Course Title**: Data Mining and Warehousing
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Spring **7. Subject Area:** PEC
8. **Pre-requisite:** CS-102
9. **Objective:** To educate students to the various concepts, algorithms and techniques in data mining and warehousing and their applications.

10. **Details of the Course**

S.No.	Contents	Contact hours
1.	**Introduction to data mining:** Motivation and significance of data mining, data mining functionalities, interestingness measures, classification of data mining system, major issues in data mining.	3
2.	**Data pre-processing:** Need, data summarization, data cleaning, data integration and transformation, data reduction techniques —Singular Value Decomposition (SVD), Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT), data discretization and concept hierarchy generalization.	6
3.	**Data warehouse and OLAP technology:** Data warehouse definition, multidimensional data model(s), data warehouse architecture, OLAP server types, data warehouse implementation, on-line analytical processing and mining,	4
4.	**Data cube computation and data generalization:** Efficient methods for data cube computation, discovery driven exploration of data cubes, complex aggregation, attribute oriented induction for data generalization.	4
5.	**Mining frequent patterns, associations and correlations:** Basic concepts, efficient and scalable frequent item set mining algorithms, mining various kinds of association rules —multilevel and multidimensional, association rule mining versus correlation analysis, constraint based association mining.	6
6.	**Classification and prediction:** Definition, decision tree induction, Bayesian classification, rule based classification, classification by backpropagation and support vector machines, associative classification, lazy learners, prediction, accuracy and error measures.	6
7.	**Cluster Analysis:** Definition, Clustering Algorithms - partitioning, hierarchical, density based, grid based and model based; Clustering high dimensional data, constraint based cluster analysis, outlier analysis - density based and distance based.	6
8.	**Data mining on complex data and applications:** Algorithms for mining of spatial data, multimedia data, text data: data mining applications, social impacts of data mining, trends in data mining.	7
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication/ Reprint
1.	Marakas, George M. Modern data warehousing, mining, and visualization: core concepts. Upper Saddle River, NJ: Prentice Hall, 2003.	2003
2.	Pujari, Arun K. Data mining techniques. Universities press, 2001.	2001
3.	Lee, Mong Li, Hongjun Lu, Tok Wang Ling, and Yee Teng Ko. "Cleansing data for mining and warehousing." In International Conference on Database and Expert Systems Applications, pp. 751-760. Springer, Berlin, Heidelberg, 1999.	1999
4.	Wang, John, ed. Encyclopedia of data warehousing and mining. iGi Global, 2005.	2005
5.	Gupta, Gopal K. Introduction to data mining with case studies. PHI Learning Pvt. Ltd., 2014.	2014
6.	Tan, Pang-Ning, Michael Steinbach, and Vipin Kumar. Introduction to data mining. Pearson Education India, 2016.	2016

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code**: AID-552 **Course Title**: Deep Learning
2. **Contact Hours**: **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits**: 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Machine Learning
9. **Objective:** The objective of this course is to learn deep learning algorithms, concepts, experiments, research along with their application on generic use cases.
10. **Details of the Course:**

S.No	Contents	Contact Hours
1	Introduction to deep learning, logical computations with neurons, perceptron, backpropagation, historical trends, applications, and use-cases for industry	6
2	Deep Networks: Training a deep neural network (DNN), hidden layers, activation functions, fine-tuning neural network hyper-parameters	7
3	Custom Deep Neural Networks: vanishing/exploding gradient issues, reusing pre trained layers, optimizers, l1 and l2 regularization, dropout	8
4	Convolutional neural networks (CNNs): convolutional layer, filters, stacking, pooling layer, CNN architectures	7
5	Recurrent neural networks (RNNs): recurrent neurons, unrolling, input and output sequences, training RNNs, deep RNNs, LSTM cell, GRU cell	7
6	Representation Learning and Generative Learning: Autoencoders: data representations, linear autoencoder, stacked autoencoders, variational autoencoders	7
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1	Aurélien Géron, "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems,"Second Edition, O'Reilly Media	2019
2	Ian Goodfellow, Yoshua Bengio, and Aaron Courville, "Deep Learning," First Edition, MIT Press	2017
3	François Chollet "Deep Learning with Python," First Edition, Manning Publication	2018
4	Rowel Atienza "Advanced Deep Learning with Keras," First Edition, Packt Publishing	2018
5	Sudharsan Ravichandran "Hands-On Deep Learning Algorithms with Python," First Edition, Packt Publishing	2019

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-572 **Course Title**: Ethics in Data Science
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration(Hrs):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** To introduce the concepts of ethics in data science
10. **Details of the Course:**

S.No.	Contents	Contact Hours
1	**Introduction and Philosophical frameworks for assessing fairness**: Foundations of ethics, early theories of fairness (Utilitarianism etc.); contemporary theories of fairness; significance of ethics in data science; ethics vs. law/compliance/public relations; cultural relativism; "professional" ethics in data science; individuals vs. collectives.	6
2	**Research Ethics:** Data driven research, methods of collection of data; different types of data: qualitative and quantitative; overview of ethical issues in data-driven organizations; doing ethical data analysis; responsible use of research data; plagiarism; fake data and fabrication of data; creation of data base.	8
3	**Data ownership, privacy and anonymity:** Understanding the difference between data ownership; data privacy and data anonymity; under- standing the idea behind data surveillance; data privacy vs. data security.	8
4	**Algorithmic fairness:** Discrimination and algorithms; obscure and un-intentional bias displayed by the algorithms; ethics of data scraping and storage; Mosaic data; found data; and designed data.	8
5	**Policies on data protection**: EU's general data protection rules (GDPR); digital India policy; personal data protection bill; 2019 ("PDP Bill"); ethical issues on data privacy in context with India, case studies.	8
6	**Case Studies**	4
Total		**42**

11. **Suggested Books:**

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1.	Michael J. Quinn "Ethics for the Information Age", Seventh Edition, Pearson.	2016
2.	DJ Patil, Hilary Mason, and Mike Loukides "Ethics and Data Science", O'Reilly Media Inc.	2018
3.	Bill Franks, "97 Things About Ethics Everyone in Data Science Should Know", O'Reilly Media Inc.	2020
4.	Kord Davis, "Ethics of Big Data: Balancing Risk and Innovation", O'Reilly Media Inc.	2012

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Department of Mathematics

1. **Subject Code:** MAN-628 **Course Title:** Evolutionary Algorithms
2. **Contact Hours:** **L:** 3 **T:** 0 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 3 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective**: To provide knowledge about basic concepts of Evolutionary Algorithms
10. **Details of the Course:**

S.No.	Contents	Contact Hours
1	**Genetic Algorithms:** Historical development, GA concepts – encoding, fitness function, population size, selection, crossover and mutation operators, along with the methodologies of applying these operators. Binary GA and their operators, Real Coded GA and their operators	12
2	**Particle Swarm Optimization:** PSO Model, global best, Local best, velocity update equations, position update equations, velocity clamping, inertia weight, constriction coefficients, synchronous and asynchronous updates, Binary PSO.	10
3	**Memetic Algorithms:** Concepts of memes, Incorporating local search as memes, single and multi-memes, hybridization with GA and PSO, Generation Gaps, Performance metrics.	5
4	**Differential Evolution:** DE as modified GA, generation of population, operators and their implementation.	5
5	**Artificial Bee Colony:** Historical development, types of bees and their role in the optimization process.	5
6	**Multi-Objective Optimization:** Linear and nonlinear multi-objective problems, convex and non – convex problems, dominance – concepts and properties, Pareto – optimality, Use of Evolutionary Computations to solve multi objective optimization, bi level optimization, Theoretical Foundations	5
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1.	Coello, C. A., Van Veldhuizen, D.A. and Lamont, G.B.: "Evolutionary Algorithms for solving Multi Objective Problems", Kluwer.	2002
2	Deb, K.: "Multi-Objective Optimization using Evolutionary Algorithms", John Wiley and Sons.	2002
3	Deb, K.: "Optimization for Engineering Design Algorithms and Examples", Prentice Hall of India.	1998
4	Gen, M. and Cheng, R.: "Genetic Algorithms and Engineering Design", Wiley, New York.	1997
5	Hart, W.E., Krasnogor, N. and Smith, J.E. : "Recent Advances in Memetic Algorithms", Springer Berlin Heidelberg, New York.	2005
6	Michalewicz, Z.: "Genetic Algorithms+Data structures=Evolution Programs", Springer-Verlag, 3rd edition, London, UK.	1992

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-573 **Course Title:** Intrusion Detection Systems
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** To introduce concepts in intrusion detection systems
10. **Details of the Course**

S.No.	Contents	Contact hours
1.	**Introduction to IDS:** Intruder types, intrusion methods, processes and detection, message integrity and authentication, honey pots	8
2.	**IDS Models:** General IDS model and taxonomy, data mining based IDS, Denning model, Framework for constructing features, and different models for intrusion detection systems, SVM, probabilistic, and statistical modelling, evaluation of IDS, cost sensitive IDS	8
3.	**Network Security Threat Detection:** NBAD, specification based and rate based DDOS, scans/probes, predicting attacks, network based anomaly detection, stealthy surveillance detection; defending against DOS attacks in scout, signature-based solutions, snort rules	9
4.	**Host based Threat Detection:** Host-based anomaly detection, taxonomy of security flaws in software, self-modelling system calls for intrusion detection with dynamic window size	9
5.	**Secure Intrusion Detection Systems:** Network security, secure intrusion detection environment, secure policy manager, secure IDS sensor, alarm management, intrusion detection system signatures, sensor configuration, signature and intrusion detection configuration, IP blocking configuration, intrusion detection system architecture.	8
Total		**42**

11. **Suggested Books:**

S.No.	Name of Authors/Book/Publisher	Year of Publication / Reprint
1.	J. Paul Guyer, "An Introduction to Intrusion Detection Systems," Create space Independent Publishers	2017
2.	Gerard Blokdyk, "Intrusion-detection System: How-to," Create space Independent Publishers.	2017
3.	Rash, M., Orebaugh, A. and Clark, G., "Intrusion Prevention and Active Response: Deploying Network and Host IPS," Syngress.	2005
4.	Endorf, C., Schultz E. and Mellander J., "Intrusion Detection and Prevention," McGraw-Hill.	2003

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Department of Computer Science and Engineering

1. **Subject Code:** CSN-528 **Course Title**: Natural Language Processing
2. **Contact Hours:** **L:** 3 **T:** 0 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 3 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Basic knowledge of Artificial Intelligence
9. **Objective:** To provide an understanding of the theoretical concepts of Natural Language Processing and prepare students for research or industry application of Natural Language Processing.

10. Details of the Course

S.No.	Contents	Contact hours
1.	Introduction to NLP, Corpus, Representation of Words, Preprocessing, Linguistic and Statistical Properties of Words, POS Tagging, Parsing, Performance Measures, Error Analysis, Confusion Matrix	6
2.	Probability and NLP, n-Gram, Language Model, Join and Conditional Probability, Chain Rule, Markov Assumption, Data Sparsity, Smoothing Techniques, Generative Models, Naive Bayes	6
3.	Distributed representation of words for NLP, Co-occurrence Matrix, Collocations, Dimensionality Reduction, Singular Value Decomposition	6
4.	Document Similarity, Inverted Index, Word2Vec, C-BoW, Skip-Gram Model, Sampling, Hierarchical Soft-max, Sequence Learning	6
5.	Neural Networks for NLP, Multi-Layer Perceptron, Activation Function, Gradient Descent, Sequence Modeling, Recurrent Neural Networks	6
6.	Gated Recurrent Unit, Long-Short Term Memory Networks, 1-D Convolutional Layer, Language Model using RNN, Forward Pass, Backward Pass	6
7.	Applications of NLP, Topic Modeling, Sentiment Analysis, Query Processing, ChatBoat, Machine Translation, Statistical Machine Translation, Neural Machine Translation, Spell Checker, Summarization	6
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication / Reprint
1.	Manning, Christopher, and Hinrich Schutze. Foundations of statistical natural language processing. MIT press, 1999.	1999
2.	Jurafsky, Dan. Speech & language processing. Pearson Education India, 2000.	2000
3.	Smith, Noah A. Linguistic structure prediction. Morgan and Claypool, 2011.	2011
4.	Kennedy, Graeme. An introduction to corpus linguistics. Routledge, 2014.	2014

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Department of Mathematics

1. **Subject Code:** MAN-613 **Course Title:** Operations Research
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester: Both** **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective**: To acquaint the students with the basic techniques of Operations Research.
10. **Details of the Course:**

S.No.	Contents	Contact Hours
1	Basics of LPP: Different Types of OR Models, Convex Sets, Graphical Method, Simplex Method, Big –M Method, Two Phase Method, Revised Simplex Method.	11
2	Duality Theory: Dual Simplex Method, Sensitivity Analysis, Parametric Linear Programming.	9
3	Integer Program: Cutting Plane and Branch and Bound Techniques for all Integer and Mixed Integer Programming Problems	5
4	Transportation Problems: Transportation Problems and Assignment Problems.	5
5	Game Theory: Graphical Method and Linear Programming Method for Rectangular Games, Saddle point, notion of dominance.	5
6	Queuing Theory: Steady -state solutions of Markovian Queuing Models: M/M/1, M/M/1 with limited waiting space, M/M/C, M/M/C with limited space, M/G/1, Inventory Models.	7
Total		**42**

11. **Suggested Books:**

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1.	Mohan, C. and Deep, K.: “Optimization Techniques”, New Age India Pvt. Ltd, New Delhi.	2009
2.	Mittal, K.V. and Mohan, C.: “Optimization Methods in System Analysis and Operations Research”, New Age India Pvt. Ltd, New Delhi.	1996
3.	Taha, H.A.: “Operations Research: An Introduction”, MacMillan Pub Co., NY, Ninth Edition (Reprint).	2013
4	Ravindran, A., Phillips, D.T. and Solberg, J.J.: “Operations Research: Principles and Practice”, John Wiley and Sons, NY, Second Edition (Reprint).	2012
5	Pant, J.C.: “Introduction to Optimization/ Operations Research”, Jain Brothers, New Delhi, Second Edition.	2012

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code**: AID-554 **Course Title**: Reinforcement Learning
2. **Contact Hours**: **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** This course aims to understand several reinforcement learning algorithms and their applications, along with emerging research trends.

10. Details of the Course:

S.No.	Contents	Contact Hours
1	Basics of probability and linear algebra, Definition of a stochastic multi-armed bandit, Definition of regret, Achieving sublinear regret, UCB algorithm, KL- UCB, Thompson Sampling.	6
2	Markov Decision Problem, policy, and value function, Reward models (infinite discounted, total, finite horizon, and average), Episodic & continuing tasks, Bellman's optimality operator, and Value iteration & policy iteration	8
3	The Reinforcement Learning problem, prediction and control problems, Model-based algorithm, Monte Carlo methods for prediction, and Online implementation of Monte Carlo policy evaluation	8
4	Bootstrapping; TD(0) algorithm; Convergence of Monte Carlo and batch TD(0) algorithms; Model-free control: Q-learning, Sarsa, Expected Sarsa.	6
5	n-step returns; TD(λ) algorithm; Need for generalization in practice; Linear function approximation and geometric view; Linear TD(λ).	6
6	Tile coding; Control with function approximation; Policy search; Policy gradient methods; Experience replay; Fitted Q Iteration; Case studies.	8
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication/ Reprint
1	Sutton, Richard S., and Andrew G. Barto. "Reinforcement learning: An introduction," First Edition, MIT press	2020
2	Sugiyama, Masashi. "Statistical reinforcement learning: modern machine learning approaches," First Edition, CRC Press	2015

3	Lattimore, T. and C. Szepesvári. “Bandit algorithms,” First Edition, Cambridge University Press.	2020
4	Boris Belousov, Hany Abdulsamad, Pascal Klink, Simone Parisi, and Jan Peters “Reinforcement Learning Algorithms: Analysis and Applications,”First Edition, Springer	2021
5	Alexander Zai and Brandon Brown “Deep Reinforcement Learning in Action,” First Edition, Manning Publications	2020

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-574 **Course Title:** Spreadsheet Modeling and Simulation
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** To introduce the concept of spreadsheet modelling and simulations to the students.
10. **Details of the Course:**

S.No.	Contents	Contact Hours
1	**Introduction:** Introduction to spreadsheets; historical development; basic capabilities of spreadsheets and their usage for creating models; types of data used in spreadsheets; spreadsheet notations for mathematical operations; common built-in formulas and functions; conditional expressions; relative and absolute references.	6
2	**Model building**: Designing spreadsheets reflecting assumptions; decision variables; and outcomes, creating basic cash-flow models; revaluating small business opportunities; incorporating what-if analysis; identifying key variables using sensitivity analysis; linear programming models and deterministic models.	10
3	**Optimization with Spreadsheets using Solver**: Linear programming, sensitivity analysis, transportation and assignment problems, network optimization problems, integer and nonlinear programming, multi-objective optimization, applications of optimization in different areas.	12
4	**Simulation and Optimization**: Use of spreadsheets to implement Monte Carlo simulations and linear programs for optimization; model uncertainty and risk in spreadsheets; and use of Excel's solver.	10
5	**Case Studies**	4
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1.	Hillier and Hillier "Introduction to Management Science: A Modeling and Case Studies Approach with Spreadsheets", McGraw-Hill/Irwin.	2013
2.	Cliff Ragsdale "Spreadsheet Modeling and Decision Analysis: A Practical Introduction to Business Analytics", Cengage India.	2018
3.	Barry Render, Nagraj Balakrishnan, and Ralph Stair, "Managerial Decision Modelling with Spreadsheets", Pearson.	2004
4.	S. Christian Albright and Wayne Winston "Spreadsheet Modeling and Applications: Essentials of Practical Management Science",Cengage.	2004

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Department of Mathematics

1. **Subject Code:** MAN-526 **Course Title**: Soft Computing
2. **Contact Hours:** **L:** 3 **T:** 0 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 3 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective**: To acquaint the students with the basic techniques of Soft Computing.
10. **Details of the Course:**

S.No.	Contents	Contact Hours
1	Introduction to Soft Computing, Historical Development, Definitions, advantages and disadvantages, solution of complex real life problems	2
2	Neural Networks: Fundamentals, Neural Network Architectures, Feedforward Networks, Backpropagation Networks.	10
3	Fuzzy Logic: Fuzzy Sets, Fuzzy numbers, Fuzzy Systems, membership functions, fuzzification, defuzzification.	8
4	Genetic Algorithms: Generation of population, Encoding, Fitness Function, Reproduction, Crossover, Mutation, probability of crossover and probability of mutation, convergence.	10
5	Hybrid Systems: Genetic Algorithm based Backpropagation Network, Fuzzy–Backpropagation, Fuzzy Logic Controlled Genetic Algorithms. Case studies.	7
6	Case studies in Engineering	5
Total		**42**

11. **Suggested Books:**

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1.	Jang, J-S. R., Sun,C-T, Mizutani, E.: "Neuro–Fuzzy and Soft Computing", Prentice Hall of India.	2002
2.	Klir, G. J. and Yuan, B.: "Fuzzy Sets and Fuzzy Logic: Theory and Applications", Prentice Hall.	1995
3.	Rajasekaran, S. and Vijayalakshmi Pai, G.A.: "Neural Networks, Fuzzy Logic and Genetic Algorithms: Synthesis and Applications", Prentice Hall of India.	2003
4	Sinha, N.K. and Gupta, M. M. : "Soft Computing and Intelligent Systems - Theory and Applications", Academic Press.	2000
5	Tettamanzi, A., Tomassini, M.: "Soft Computing: Integrating Evolutionary, Neural, and Fuzzy Systems", Springer.	2001

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Department of Mathematics

1. **Subject Code:** MAN-507 **Course Title**: Statistical Inference
2. **Contact Hours:** **L:** 3 **T:** 0 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 3 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** To introduce the concepts of statistical inference.
10. **Details of the Course**

S.No.	Contents	Contact hours
1.	**Principle of Data Reduction:** Sufficiency principle, Factorization criterion, minimal sufficiency, Completeness and bounded completeness, Likelihood principle, Equivariance principle.	8
2.	**Theory of Estimation:** Basic concepts of estimation, Point estimation, methods of estimation; method of moments, method of maximum likelihood; Unbiasedness, Minimum variance estimation, Cramer – Rao bound and its generalization, Rao Blackwell theorem, Existence of UMVUE estimators. Interval Estimation, Some results for normal population case.	12
3.	**Testing of Hypothesis:** Null and alternative hypothesis, Type I and II errors error probability and power function, Method of finding tests, Neyman – Pearson lemma, Uniformly most powerful tests, Likelihood ratio principle, Likelihood ratio test, Sequential probability ratio test, Some results based on normal population.	18
4.	**Analysis of Variance:** one-way classification; simple linear regression analysis with normal distribution.	4
Total		**42**

11. **Suggested Books:**

S.No.	Name of Authors/Book/Publisher	Year of Publication/ Reprint
1.	Miller, I. and Miller, M.,"Freund's Mathematical Statistics with Applications", Prentice Hall PTR, 7th edition	2006
2.	Lehman, E.L., "Testing of Statistical Hypothesis", Wiley Eastern Ltd	1959
3.	G. Casella, R. L. Berger, "Statistical Inference", Duxbury Press	2002
4.	Lehman, E.L., "Point Estimation", John Wiley & sons	1984
5.	Rohatgi, V.K., "Statistical Inference", Dover Publications	2011

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code**: AID-555 **Course Title**: Time Series Data Analysis
2. **Contact Hours**: **L:**3 **T:**1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits**: 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** The objective of this course is to understand and analyze time-series data facilitated by R programming

10. Details of the Course:

S.No.	Contents	Contact Hours
1	Basic Properties of time-series data: Distribution and moments, Stationarity, Autocorrelation, Heteroscedasticity, Normality	4
2	Autoregressive models and forecasting: AR, ARMA, ARIMA models	4
3	Random walk model: Non-stationarity and unit-root process, Drift and Trend models	4
4	Regression analysis with time-series data using R programming	5
5	Principal Component Analysis (PCA) and Factor Analysis	5
6	Conditional Heteroscedastic Models: ARCH, GARCH. T-GARCH, BEKK-GARCH	6
7	Introduction to Non-linear and regime-switching models: Markov regime-switching models, Quantile regression, Contagion models	5
8	Introduction to Vector Auto-regressive (VAR) models: Impulse Response Function (IRF), Error Correction Models, Co-integration	5
9	Introduction to Panel data models: Fixed-Effect and Random-Effect models	4
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication /Reprint
1	Chris Brooks "Introductory Econometrics for Finance," Fourth Edition, Cambridge University Press	2019
2	Ruey S. Tsay "Analysis of Time-series data," Third Edition, Wiley	2014
3	John Fox and Sanford Weisberg "An R Companion to Applied Regression," Third Edition, SAGE	2018
4	Yves Croissant and Giovanni Millo "Panel Data Econometrics with R," First Edition, Wiley	2018

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-575 **Course Title**: Blockchain Technology
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** To provide knowledge about cryptographic and cybersecurity concepts of blockchain technology with some applications.
10. **Details of the Course**

S.No.	Contents	Contact hours
1.	**Basics of Blockchain:** Distributed Database, Two General Problem, Byzantine General problem and Fault Tolerance, Hadoop Distributed File System, Distributed Hash Table, ASIC resistance, Turing Complete	5
2.	**Crypto Primitives:** Hash functions, security aspects of hash function, Collison resistant hash, digital signatures, public key cryptography, verifiable random functions, NIST standards	8
3.	**Blockchain Theory:** Advantage over conventional distributed database, Blockchain Network, Mining Mechanism, Distributed Consensus, Merkle Patricia Tree, Gas Limit, Transactions and Fee, Anonymity, Reward, Chain Policy, Life of Blockchain application, Soft & Hard Fork, Private and Public blockchain.	8
3.	**Distributed Consensus:** Nakamoto consensus, Proof of Work, Proof of Stake, Proof of Burn, Difficulty Level, Sybil Attack, Energy utilization and alternate	6
4.	**Cryptocurrency and regulations:** Distributed Ledger, Bitcoin protocols - Mining strategy and rewards, Ethereum - Construction, DAO, Smart Contract, GHOST, Vulnerability, Attacks, Sidechain, Namecoin, IBM hyper ledger, Stakeholders, Roots of Bitcoin, Legal Aspects - Cryptocurrency Exchange, Black Market and Global Economy	10
5.	**Blockchain Applications:** Good blockchain examples and how to identify potential use-cases, Design Thinking, Internet of Things, Medical Record Management System, Domain Name Service	5
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication / Reprint
1.	Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller and Steven Goldfeder, Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction, Princeton University Press	2016
2.	Bettina Warburg, Bill Wanger, Tom Serres, "Basics of Blockchain " Independently published	2019
3.	Andreas M. Antonopoulos, "Mastering Bitcoin: unlocking digital cryptocurrencies", O'Reilly Media Inc.,	2014
4.	Wattenhofer, Roger, "Blockchain Science", Inverted Forest Publishing, 3rd Edition	2019

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code**: AID-568 **Course Title**: ML and AI Applications in Earth Sciences
2. **Contact Hours**: L: 3 T: 1 P: 0
3. **Examination Duration (Hrs): Theory:** 3 **Practical**: 0
4. **Relative Weightage: CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Good foundation in Mathematics and Physics with specific exposure in Numerical Methods. Understanding of fundamental principles of Geology and Geophysics would be preferable.
9. **Objective:** To make the participants familiar with tools and techniques in Earth Sciences and the use of Machine Learning and Artificial Intelligence for optimizing the workflows for more accurate prediction of events and properties of the subsurface.
10. **Details of the Course**:

S.No.	Contents	Contact Hours
1	**Familiarization with Major Domains and Data Types in Earth Sciences:** Earthquake Seismology, Engineering Geology and Rock Mechanics, Reservoir Characterization, Paleontology	4
2	**General Introduction to ML and AI in Earth Sciences:** **ML and statistical pattern recognition:** Supervised learning (generative/ descriptive learning, parametric/ non-parametric learning, neural networks, Support vector machines), Unsupervised learning (clustering, dimensionality reduction, kernel methods); time series modelling, linear regression, regularization, linear classifiers, ensemble methods, neural networks, model selection and evaluation, scalable algorithms for big data, and data ethics. **Data science:** Extreme value statistics, multi-variate analysis, factor analysis, compositional data analysis, spatial information aggregation models, spatial estimation, geo-statistical simulation, treating data of different scales of observation, spatio-temporal modelling (geo-statistics).	6
3	**Automating Data Mining and Analysis in Seismology:** Basics of earthquake detection and phase picking using short-term average (STA)/long-term average (LTA); detection using waveform similarity: Network Matched Filtering/template matching, Fingerprint And Similarity Thresholding (FAST). Associating seismic phases across all stations using deep-learning techniques and combining the ones have the same origin source (PhaseLink). Generic workflow of data collection, preprocessing, model training, model evaluation, and production. Applications of ML in ground motion synthesis, and future directions.	6
4	**Classification of Earthquake Sources:** Using supervised learning for classifying earthquakes and finding their occurrence mechanism. Training dataset (waveforms) on different kinds of sources: earthquake, glacial, volcanic, landslide, explosion, etc. A brief discussion on seismic sources and radiation pattern of emerging waves.	4
5	**Deep learning (DL) based Seismic Inversion:** Theory of Seismic Inversion, Convolutional neural network (CNN) and fully connected network (FCN) architectures, Performance evaluation, Geophysical inversion versus ML, their applications to reflectivity inversion in seismic, Numerical examples.	4

6	**Automation in 3D Reservoir Property Prediction:** Data Mining, Automated Petrophysics, Statistical and Regression Methods for Elastic Property Prediction, ML and AI application in Geostatistics, Convoluted Neural Networks for Seismic Interpretation, Deep Learning for Impedance Inversion and Porosity Prediction.	4
7	**Data-Driven Analytics in Shale Resources:** Concepts of shale as source-reservoir-seal, Modeling Production from Shale, Shale Analytics, Decline Curve Analysis, Shale Production Optimization Technology (SPOT), Numerical Simulation and Smart Proxy	4
8	**Machine learning Applications in Engineering Geology and Rock Mechanics:** ML in rock mass characterization, Rock Mass Rating, Slope Mass Rating, Q-System, Engineering properties of rock and various rock engineering applications, AI in Landslides study.	6
9	**Separation and Taxonomic Identification of Microfossil:** 3D object recognition and segmentation applied to X-ray MicroCT images; Testing different algorithms for identifying and localizing individual microfossils in rock samples: Automated Computer Vision, Deep learning-based CNN semantic, and other segmentation architectures.	4
	Total	**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publications/ Reprint
1	Patrick Wong, Fred Aminzadeh, and Masoud Nikravesh, Soft Computing for Reservoir Characterization and Modeling, Springer-Verlag Berlin Heidelberg GmbH	2002
2	William Sandham & Miles Leggett, Geophysical Applications of Artificial Neural Network and Fuzzy Logic, Springer	2003
3	C. Cranganu, H. Luchian, M. E. Breaban, Artificial Intelligent Approached in Petroleum Geosciences, Springer	2015
4	Shahab D. Mohaghegh , Data-Driven Analytics in Unconventional Resources, Springer	2017

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-576 **Course Title**: Data Science in Bioinformatics
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** The course provides exposure to the Data Science within the context of its importance in biology. The course discusses various methodologies and techniques as well as relevant problems in biology addressed using Data Science.

10. **Details of the Course**

S.No.	Contents	Contact hours
1.	Relevance of Data Science in Bioinformatics. Why Data Science in Biology and Healthcare? Visualization tools for biological and bioinformatics datasets; data handling; transformations of data	7
2.	Data Science in genomics. From genetics to genomes. Alignment and phylogenetic trees.	7
3.	Structural bioinformatics, Proteomics, Protein structure prediction, integrative structural modeling, and structure-based drug design.	7
4.	AI algorithms, statistical tools, graph algorithms for bioinformatics data analytics	7
5.	Deep learning algorithms in perspective of bioinformatics applications; GANs for biological applications	7
6.	Whole-cell modeling approaches, Big Data Consortiums; Hands-on experience ofapplying Data Science in Biology	7
Total		**42**

11. **Suggested Books:**

S.No.	Name of Authors/Book/Publisher	Year of Publication / Reprint
1.	Arthur M. Lesk, “Introduction to Bioinformatics”, Oxford University Press) (Fifth Edition)	2019
2.	Jeil Grus, “Data Science from Scratch: First Principles with Python”, O’Reilly Media Inc. (Second Edition,)	2019
3.	Vince Buffalo, “Bioinformatics Data skills”, O’Reilly Media Inc.	2015
4.	Neil C. Jones and Pavel A. Pevzner, “An introduction to Bioinformatics Algorithms”, The MIT Press	2004

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-577 **Course Title:** Data Science for Decision Making
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** To introduce the concept of data driven decision making systems to the students.
10. **Details of the Course:**

S.No.	Contents	Contact Hours
1	**Fundamentals of Analytics**: Introduction to data-driven decision making; general introduction to data driven strategy and its importance; use of examples and mini-case studies to illustrate the role of statistical analysis in decision making.	8
2	**Basic Data Analysis**: Various types of data that are commonly collected by firms; methods to be used and inferences/insights that can be obtained depending on the type of data that are available (stated versus revealed preference, level of aggregation, cross- sectional, time series, panel data and so forth); use of frequency distributions, mean comparisons, and cross tabulation; statistical inferences using chi-square; t-test and ANOVA.	10
3	**Experimental Design and Natural Experiments**: Issues of design of experiments and internal and external validity; case studies in marketing; economics; and medicine etc.; A-B testing; and circumstances that provide us with "natural" experiments.	10
4	**Decision making tools**: Regression analysis and its applications; use of regression output in forecasting; promotional planning and optimal pricing; multivariate analysis (unsupervised learning) cluster analysis; factor analysis decision trees; elastic nets and random forests.	10
5	**Case Studies:** To understand the problem at an intuitive level; use of simple data analysis and visualization to verify (or falsify) the intuition; use of appropriate statistical analysis to present your arguments.	4
	Total	**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1.	F.S. Hillier and G.J. Liberman "Introduction to Operations Research*"* Tata McGraw Hill Education Private Limited.	2001
2.	Gregory S. Parnel, Terry A. Bresnick, Steven N. Tani, Eric R.Johnson "Handbook of Decision Analysis", Wiley.	2013
3.	Emily Moberg and Igor Linkov "Multi-Criteria Decision Analysis: Environmental Applications and Case Studies", CRCPress, Taylor and Francis group.	2011
4.	Adiel Teixeira de Almeida, Emel Aktas, Sarah Ben Amor, João Luis de Miranda "Advanced Studies in Multi-Criteria Decision Making", CRC Press.	2020

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code**: AID-562 **Course Title**: AI for Investment
2. **Contact Hours**: **L:** 3 **T:** 0 **P:** 2
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 10-25 **PRS:** 25 **MTE:** 15-25 **ETE:** 30-40 **PRE:** 0
5. **Credits**: 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** The objective of this course is to understand the application of Artificial Intelligence and Machine Learning techniques in financial markets, trading, and asset management.
10. **Details of the Course:**

S.No.	Contents	Contact Hours
1	Introduction to financial markets and market microstructure	4
2	Introduction to risk-return framework	4
3	Introduction to asset management and portfolio optimization	4
4	Market efficiency and behavioral finance	4
5	Prediction in Financial markets using AI and machine learning models, AI and machine learning in Trading execution and portfolio management	6
6	Credit scoring and credit modeling with non-linear machine learning models and deep learning	4
7	Model risk management and stress testing	4
8	Robo advisory, social and quantitative investing	5
9	Machine learning for asset management	4
10	AI and machine learning in regulatory compliance and supervision	3
Total		**42**

11. **Suggested Books:**

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1	M. Dixon, I Halperin, and P. Bilokon "Machine Learning in Finance," First Edition, Springer	2020
2	Marcos Lopez "Advances in Financial Machine Learning," First Edition, Wiley	2018
3	Marcos Lopez "Machine Learning for Asset Managers," First Edition,Cambridge University Press	2020
4	Stefan Jansen "Machine Learning for Algorithmic Trading," Second Edition, Packt	2020
5	Elton and Gruber, "Modern Portfolio Theory," Ninth Edition, Wiley	2014

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-553 **Course Title:** Digital Image Processing
2. **Contact Hours: L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.): Theory:** 3 **Practical:** 0
4. **Relative Weight: CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
9. **Pre-requisite:** Nil
10. **Objective:** The objective of this course is to introduce the fundamental techniques and algorithms used for acquiring, processing and extracting useful information from digital images.

11. **Details of the Course**

S.No.	Contents	Contact Hours
1.	**Introduction:** Signal processing overview; Image processing basics; Fundamental signals (1-D and 2-D); Classification of systems; Characteristics of LTI/LSI systems. Introduction to the DIP areas and applications.	4
2.	D**igital Image Fundamentals:** Human visual system and visual perception; Image sensing and acquisition Image file types; Pixel representation and spatial relationship	4
3.	Image Digitization: Sampling and quantization. Image Transforms: 2- D DSFT and 2-D DFT, 2-D discrete cosine transform (DCT), 1-D and 2-D Karhonen Loeve (KL) or principal component analysis (PCA) and 1-D and 2-D discrete wavelet transforms and relation to filter banks.	8
4.	Image Enhancement: Point and algebraic operations, edge detection and sharpening, filtering in the spatial and transformed domains. Rotation, interpolation, image filtering, spatial operators, morphologicaloperators.	6
5.	**Image Segmentation:** Thresholding; Edge based segmentation; Region growing; Watershed transform. Image Restoration: Degradation models, inverse and pseudo-inverse filtering, 2-D Wiener filtering and implementation	6
6.	Image Compression and Encoding: Entropy-based schemes, Transform- based encoding, Predictive encoding and DPCM, Vector quantization, Huffman coding.	4
7.	Feature Extraction and Segmentation: Contour and shape dependent feature extraction, textural features, region-based and feature-based segmentation.	5
8.	Pattern Classification: Standard linear and Bayesian classifiers, supervised Vs unsupervised classification, classification performance index. Applications in satellite, sonar, radar and medical areas.	5
Total		**42**

12. Suggested Books:

S.No.	Name of Authors/Book/ Publisher	Year of Publication/Reprint
1.	Gonzalez R. C. and Woods R. E., "Digital image processing," FourthEdition, Prentice Hall.	2017
2.	Lim J. S., "Two-dimensional signal and image processing," Prentice Hall.	1990
3.	Dudgeon D.E. and Merserau R. M., "Multidimensional digital signalprocessing," Prentice Hall Signal Processing Series.	1984
4.	Bose T., "Digital Signal and Image Processing", Wiley India.	2010
5.	Sonaka M., Hlavac V. and Boyle R., "Image Processing, Analysis and Machine Vision," Fourth edition, Cengage India Private Limited.	2017
6.	W.K. Pratt. "Digital Image Processing," Fourth Edition, John Wiley &Sons, New York.	2007

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-578 **Course Title**: Graphs Algorithms in Data Science
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective**: To acquaint the students with the knowledge of graph algorithms in ML and Data Science.
10. **Details of the Course:**

S.No.	Contents	Contact Hours
1	Fundamentals: graph models, Isomorphic graphs, Spanning tree, connectivity in graphs, Eulerian and Hamiltonian Graphs, matching, vertex colouring and domination, random graphs.	5
2	Graph Modelling with Neo4j: Graph Databases, directed vs undirected, weighted vs unweighted, cyclic vs acyclic, dense vs sparse, connected vs disconnected, graph traversal, Cypher Query Language, nodes and relationships, managing databases with Neo4j, creating, selecting a node, filtering, creating a relationship, selecting relationship, updating and deleting nodes and relationships, pattern matching and data retrieval, aggregation functions, importing data from CSV to JSON, Empowering business with pure Cypher, knowledge graphs, graph-based search, recommendation engines.	14
3	Graph Algorithms: The Graph Data Science Library and Path finding, Dijkstra's shortest path algorithm, A-star algorithm, k-shortest path, optimizing processes using graphs, travelling salesman problem, spanning tress, prims algorithm, minimum spanning tree in a Neo4j graph.	10
4	Spatial data: Node importance, representation spatial attributes, creating a geometry layer with Neo4j, spatial queries, visualization spatial data with Neo4j, Community detection and similarity measures.	8
5	Machine Learning on Graphs: Using graph-based features in machine Learning, predicting relationships, graph embedding from graphs to matrices, Applications of Neo4j in web applications.	5
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/ Publisher	Year of Publication/Reprint
1	Jonathan Gross and Jay Yellen, Graph Theory and its Applications, SecondEdition, CRC Press.	2018
2	Estelle Scifo, Hands-On Graph Analytics with Neo4j, Kindle Edition.	2020
3	Bondy J.A. and Murty U.S.R., Graph Theory Ⅱ, Springer.	2013
4	Bela Bollobas, Random Graphs, Cambridge University Press.	2008
5	Douglas B. West ―Graph Theory, Prentice Hall.	2014

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Department of Computer Science and Engineering

1. **Subject Code:** CSN-527 **Course Title**: Internet of Things
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Knowledge of computer networks
9. **Objective:** To impart the know-how of Internet of Things and their applications, architectures and protocols, building IoT applications/systems, securing the IoT systems, and their recent advances.

10. **Details of the Course**

S.No.	Contents	Contact hours
1.	**Basic concepts revisited:** Introduction to sensing & actuating, Basic networking, Wireless networks, Wireless sensor networks (WSN), Communication protocols, and other enabling technologies, IoT standards, Data storage & management issues and approaches, Cloud computing, Key challenges, research, and future directions of IoT, and security & privacy issues.	7
2.	**Embedded Systems:** Hardware and software of IoT, Microcontrollers, Understanding and programming Arduino, Raspberry Pi, NodeMCU, Lora, etc. Integrating microcontrollers with sensors and actuators, Building the IoT applications with any microcontroller.	6
3.	**IoT Architectures and Protocols:** Layers of communication, Architectures: State-of-the-art, IoT architecture reference models, Different views of IoT architectures and frameworks design, Protocols: Application protocols, Service discovery protocols, Infrastructure protocols, and other protocols. Understanding various types of protocols like HTTP, MQTT, CoAP, AMQP, 6LoWPAN, etc. Cross-layer implementations, and Data dissemination.	9
4.	**Support Technologies for IoT:** Big Data, Data Analytics, Artificial Intelligence, Mobile, Cloud, Software defined networks, 5G, and Fog/Edge computing. IoT integration with recent technologies. State-of-the-art. Design goals, challenges, and components.	8
5.	**Cyber Physical Systems:** Industry 4.0, Society 5.0, Design & use cases, Development, and implementation insights some examples like smart cities, smart homes, smart grids, smart agriculture, smart healthcare, smart transportation, smart manufacturing, and other smart systems. State-of-the-art. Conceptualizing the new IoT-based smart systems using a case study.	6
6.	**IoT Security & Privacy:** –, IoT Security and Privacy issues and challenges, Risks involved with IoT infrastructures, Trust in IoT platforms and other integrating technologies, Data aggregation, storage, retrieval, and other management issues including fault tolerance, interoperability, security, and privacy, Cyber-physical-systems and their security and privacy, Mitigation approaches.	6
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication/ Reprint
1.	Edited by: Buyya, Rajkumar, and Amir Vahid Dastjerdi, Internet of Things: Principles and paradigms. Elsevier/Morgan Kaufmann	2016
2.	Bahga, Arshdeep; Madisetti, Vijay, Internet of Things (A Hands-on-Approach), AbeBooks.com	2014
3.	Sohraby, Kazem, Daniel Minoli, and Taieb Znati. Wireless sensor networks: technology, protocols, and applications. John Wiley & Sons	2007
4.	Marinescu, Dan C., Cloud computing: theory and practice. Elsevier/ Morgan Kaufmann	2017

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Department of Computer Science and Engineering

1. **Subject Code**: AID-579 **Course Title**: Leveraging Data Science for Finance
2. **Contact Hours**: **L:** 3 **T:** 0 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits**: 3 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** The objective of this course is to understand and apply the knowledge of data science related applications in the domain of finance.

10. **Details of the Course:**

S.No.	Contents	Contact Hours
1	Data Science basics: Preparation, organizing, and visualization of financial market data and examination of basic properties of security prices	4
2	Quantiative models of risk-return framework in financial markets	4
3	Linear and non-linear price dynamics and modelling of security prices	4
4	Stock market prediction modelling, portfolio optimization, and wealth market maximization	5
5	Application of latent factor and commonality models in financial markets	5
6	Modelling of financial market volatility using Conditional Heteroscedastic Models	6
7	Introduction to Crisis/Non-crisis models (Non-linearity, extreme-value modelling, Markov regime-switching models, Quantile regression, Contagion models)	5
8	Introduction to data modelling for high-frequency algorithmic trading	5
9	Use cases for application of data science in Finance: Investment Management, Sharpe ratio analysis, Capital Asset Pricing Model, etc. (using R programming)	4
Total		**42**

11. **Suggested Books:**

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1	Chris Brooks "Introductory Econometrics," Fourth Edition, Cambridge University Press	2019
2	Ruey S. Tsay "Analysis of Time-series data," Third Edition, Wiley	2014
3	John Fox and Sanford Weisberg "An R Companion to Applied Regression," Third Edition, SAGE	2018
4	Yves Croissant and Giovanni Millo "Panel Data Econometrics with R," First Edition, Wiley	2018

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-580 **Course Title**: Multi-Objective and Multi-Criteria Decision Making
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.): Theory:** 3 **Practical:** 0
4. **Relative Weightage: CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** To introduce the concept of multi-objective and multi criteria decision making systems to the students.

10. **Details of the Course**:

S.No.	Contents	Contact Hours
1	**Introduction:** Review of decision making process in optimization and operations research models; overview of machine learning algorithms; ranking methods.	8
2	**Multi Objective Optimization (MOO)**: Introduction to multi objective optimization, classical and recent methods for multi objective optimization like genetic algorithms and particle swarm optimization.	10
3	**Multi Criteria Decision Making (MCDM):** Introduction to MCDM methods; group decision making, weighing methods and ranking methods.	10
4	**Data Manipulation:** Data wrangling and data management for large sized multi objective and multi criteria problems.	10
5	**Implementation:** Implementation of the models developed in 2, 3 and 4 in Python	4
Total		**42**

11. **Suggested Books:**

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1.	G.H. Tzeng, J.J. Huang, "Multiple Attribute Decision Making:Methods and Applications", CRC Press	2011
2.	M. Köksalan. J. Wallenius, S. Zionts, "Multiple Criteria Decision Making. From Early History to the 21st Century", World Scientific	2011
3.	J. Branke, K. Deb, K. Miettinen, R. Slowinski (Eds.), "Multiobjective Optimization: Interactive and Evolutionary Approaches", Springer-Verlag, Berlin, Heidelberg	2008
4.	A. Ishizaka, P. Nemery, "Multicriteria Decision Aid: Methods and software", Wiley.	2013
5.	K Deb, "Multi-Objective Optimization Using Evolutionary Algorithms", Wiley.	2011
6.	Michael Carter, Camille C. Price and Ghaith Rabadi "Operations Research, A Practical Introduction", CRC Press.	2018

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Department of Mathematics

1. **Subject Code:** MAN-634 **Course Title**: Parallel Computing
2. **Contact Hours:** **L:** 3 **T:** 0 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:**0
5. **Credits:** 3 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective**: To provide knowledge about parallel computing.
10. **Details of the Course:**

S.No.	Contents	Contact Hours
1	Introduction, history, temporal parallelism, data parallelism, combined temporal and data parallelism, data parallelism with dynamic and quasi-dynamic assignment, specialist data parallelism, coarse-grained specialized temporal parallelism, agenda parallelism, task dependencies and task graphs.	7
2	Structures of parallel computers: classification of parallel computers based on data / instruction flow, coupling, mode of accessing memory, grain size, vector supercomputers, systolic processors.	8
3	Shared memory parallel computers based on shared bus& intercommunication networks, direct and indirect networks.	5
4	Message Passing Systems, MPI Programming, point-to-point communications, collective communications	6
5	CUDA Programming, host, device, threads, blocks, indexing, synchronization, performance optimization.	6
6	Performance evaluation, parallel balance point, concurrency, scalability, speedup, Amdahl's law, Gustafson's law, Sun and Ni's law.	5
7	Parallel algorithms, matrix multiplication, system of linear equations, sorting, discrete Fourier transforms, numerical integration.	5
Total		**42**

11. **Suggested Books:**

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1	Aki, Selim G.: "The Design and Analysis of Parallel Algorithms", Prentice Hall, Englewood Cliffs, New Jersey.	1989
2	Krik, David B. and Hwu, W.W.: "Programming Massively Parallel Processors - A Hands on Approach: Applications of GPU ComputingSeries", Elsevier Inc.	2010
3	Pacheco, Peter S.: "Parallel Programming with MPI", Morgan Kaufmann Publishers, Inc., California.	1997
4	Quinn, M. J.: "Parallel Computing: Theory and Practice", Tata McGraw Hill.	1994
5	Rajaraman, V and Murthy, C. Siva Ram: "Parallel Computers Architecture and Programming", Prentice Hall of India.	2000

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-581 **Course Title:** Pattern Recognition
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** To introduce various pattern recognition algorithms.

10. **Details of the Course**

S.No.	Contents	Contact hours
1.	**Introduction to Pattern Recognition and Bayesian Theory:** Pattern recognition systems, The design cycle, Modeling using continuous and discrete features, Discriminant functions, The Gaussian density, Error estimation, Some basic examples	8
2.	**Parametric Models:** Maximum-likelihood estimation, Bayesian estimation, Expectation-Maximization and mixture density estimation, Hidden Markov Models, Bayesian Belief Networks	6
3.	**Non-parametric Methods and Feature Reduction:** Density estimation, Parzen windows estimation, Nearest neighbor estimation, Curse of dimensionality, Principal Component Analysis, Linear Discriminant Analysis, Feature selection	8
4.	**Non-Bayesian Classifiers and Clustering:** K-nearest neighbor classifier, Linear discriminant functions, Support vector machines, Neural networks, Decision trees, Random Forests, Criterion functions for clustering, k-means clustering, Hierarchical clustering, Graph-theoretic clustering, Cluster validity	8
5.	**Algorithm-Independent Learning Issues:** No Free Lunch Theorem, Resampling for classifier design, Comparing classifiers, Combining classifiers	6
6.	**Structural and Syntactic Pattern Recognition:** Recognition with strings, Grammatical methods, Graph-theoretic methods	6
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1.	R. O. Duda, P. E. Hart, D. G. Stork, Pattern Classification, 2nd edition, John Wiley & Sons, Inc	2000
2.	C. M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press	1995
3.	K. Fukunaga, Introduction to Statistical Pattern Recognition, Academic Press	1990
4.	R. Schalkoff, Pattern Recognition: Statistical, Structural and Neural Approaches, John Wiley & Sons, Inc.	1992
5.	A. K. Jain, R. C. Dubes, Algorithms for Clustering Data, Prentice Hall	1988

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

Subject Code: AID-582**Course Title**: Recommender Systems

1. **Contact Hours**: **L:** 3 **T:** 0 **P:** 2
2. **Examination Duration (Hrs.): Theory:** 3 **Practical:** 0
3. **Relative Weightage: CWS:** 10-25 **PRS:** 25 **MTE:** 15-25 **ETE:** 30-40 **PRE:** 0
4. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** The objective of this course is to learn and understand the algorithms, theories, and designs of recommender systems with relevant use cases.
10. **Details of the Course:**

S.No.	Contents	Contact Hours
1	Basic concepts for recommender systems, Detailed taxonomy of recommender systems, Evaluation of recommender systems	4
2	Content-based filtering algorithms, Collaborative filtering algorithms	6
3	Neighborhood-based collaborative filtering algorithms (Memory-Based Algorithms)	6
4	Model-Based Collaborative Filtering Algorithms and Dimensionality Reduction	8
5	Ensemble-Based and Hybrid Recommender Systems	6
6	Advanced Topics in Recommendation Systems: The Cold Start, Context-aware recommender systems, time-sensitive, location-sensitive, social, and multi-criteria	12
Total		**42**

11. Suggested Books:

S.No.	Name of Authors/Book/Publisher	Year of Publication/Reprint
1	Charu Aggarwal “Recommender Systems: The Textbook,” First Edition, Springer	2016
2	Francesco Ricci, Lior Rokach, and Bracha Shapira “Recommender SystemsHandbook,” First Edition, Springer	2015
3	Rounak Banik “Hands-On Recommendation Systems with Python,” First Edition, Packt Publishing	2018
4	Kim Falk “Practical Recommender Systems,” First Edition, Manning Publications	2019
5	Deepak Agarwal and Bee-Chung Chen “Statistical Methods for RecommenderSystems,” First Edition, Cambridge University Press	2016

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Department of Computer Science and Engineering

1. **Subject Code:** CSN-519 **Course Title**: Social Network Analysis
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Spring **7. Subject Area:** PEC
8. **Pre-requisite:** Knowledge of computer networks
9. **Objective:** To introduce the basic notions used for social network analysis.

10. **Details of the Course:**

S.No.	Contents	Contact hours
1.	Social Network Analysis: Preliminaries and definitions, Erdos Number Project, Centrality measures, Balance and Homophily.	4
2.	Random graph models: Random graphs and alternative models, Models of network growth, Navigation in social Networks	4
3.	Network topology anddiffusion, Contagion in Networks, Complex contagion, Percolation and information, Epidemics and information cascades	4
4.	Cohesive subgroups, Multidimensional Scaling, Structural equivalence, Roles and positions, Ego networks, Weak ties, Structural holes	6
5.	Small world experiments, Small world models, Origins of small world, Heavy tails, Small Diameter, Clustering of connectivity	6
6.	The Erdos Renyi Model, Clustering Models, Preferential Attachment	6
7.	Navigation in Networks Revisited, Important vertices and page rank algorithm, Towards rational dynamics in networks, Basics of game theory	6
8.	Coloring and consensus, biased voting, network formation games, network structure and equilibrium, behavioral experiments, Spatial and agent-based models	6
Total		**42**

11. **Suggested Books:**

S.No.	Name of Authors/Book/Publisher	Year of Publication / Reprint
1.	Wasserman, Stanley, and Joseph Galaskiewicz. Advances in social network analysis: Research in the social and behavioral sciences. Sage	1994
2.	Knoke, David, and Song Yang. Social network analysis. Sage Publications.	2019
3.	Carrington, Peter J., John Scott, and Stanley Wasserman, eds. Models and methods in social network analysis. Vol. 28. Cambridge university press.	2005
4.	Liu, Bing. "Social network analysis." In Web data mining, pp. 269-309. Springer, Berlin, Heidelberg.	2011

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

NAME OF DEPARTMENT/CENTRE: Mehta Family School of Data Science and Artificial Intelligence

1. **Subject Code:** AID-583 **Course Title**: Data-driven Analytics for Smart Transportation Systems
2. **Contact Hours:** **L:** 3 **T:** 1 **P:** 0
3. **Examination Duration (Hrs.):** **Theory:** 3 **Practical:** 0
4. **Relative Weightage:** **CWS:** 20-35 **PRS:** 0 **MTE:** 20-30 **ETE:** 40-50 **PRE:** 0
5. **Credits:** 4 **6. Semester:** Both **7. Subject Area:** PEC
8. **Pre-requisite:** Nil
9. **Objective:** To familiarize with the applications of data science in traffic and transportation engineering and to demonstrate the applications of the data science in smart transportation planning

10. **Details of the Course:**

S.No.	Contents	Contact hours
1.	**Data Science in Transportation** Overview and Practical Applications; Transportation Data Sources; Data Collection; Data Preparation and Visualization.	4
2.	**Sensing and Data Mining for Smart Transportation Systems** Intelligent Transportation Systems, Incident Management Program, Efficient Emergency Vehicle Movement (Pre-Emption), Crash Detection, Reporting, and Clearance; Traffic Surveillance, Identification of Hotspots, Violation Detection; Road Network Asset Management, Inventory of Potholes, other Deficiencies; Adaptive Traffic Signal.	8
3.	**Data Analytics in Urban Transportation Planning** Basics of Urban Transportation Planning, Data Collection and Advanced Data Sources, Household Surveys, Demand Modeling using WiFi/ Bluetooth/ Call Data Record, Data Extraction and Analysis using APIs, Trip Distribution Modelling Approaches, Route Choice Models, Choice Set Generation Methods, Genetic Algorithms, Transportation Planning Example using Data-Driven Modeling and Simulation.	10
4.	**Urban Mass Transit System** Basics of Urban Mass Transit System, Static and Dynamic GTFS, Real-Time Transit, Travel Time Variability, Transit Reliability, Transit Planning using Smart-Card Data, Real-Time Accessibility Analysis.	6
5.	**Applications in Environmental Impact of Transport System** IOT based Air pollution, Real-Time Air Pollution Monitoring and Data Analysis, Placement of Mobile Sensors, Pollution Prediction using ML, Noise Data, Analysis of Key Parameters, Development of Policy Framework.	6
6.	**Crash Data Analytics** Crash Data, Data Preparation, Model Estimation, Real-Time Data-Driven Analysis; Emergency Vehicle Data, Crash Prone Stretches, Ambulance Deployment; Near-misses/Traffic Conflict Data, Surrogate Approach, Proactive Assessment and Safety Interventions.	8
	Total	**42**

11. Suggested books

S.No.	Name of Authors/Book/Publisher	Year of Publication/ Reprint
1.	Fumitaka Kurauchi, Jan-Dirk Schmöcker "Public transport planning with smart card data" CRC Press	2021
2.	Juan de Dios Ortúzar, Luis G. Willumsen "Modelling Transport", Wiley	2011
3.	Vukan R. Vuchic "Urban Transit: Operations, Planning, and Economics" Wiley	2005
4.	Constantinos Antoniou, Loukas Dimitriou, Francisco Pereira "Mobility Patterns, Big Data and Transport Analytics" Elsevier	2018
5.	Sara Moridpour, Alireza Toran Pour, Tayebeh Saghapour "Big Data Analytics in Traffic and Transportation Engineering: Emerging Research and Opportunities" IGI Global	2019
6.	Khaled R. Ahmed, Aboul-Ella Hassanien "Deep Learning and Big Data for Intelligent Transportation" Springer	2021
7.	Davy Janssens, Ansar-Ul-Haque Yasar and Luk Knapen "Data Science and Simulation in Transportation Research" IGI Global	2013