M.Tech. (CSE with Specialization in DA)

Programme Structure July 2022 onwards

FIRST SEMESTER

SNo	Course		Contact Hours				Credits
	No.	Title	L Total	T	P		
1.	17M11CS111	Data Structures and Algorithms for Big Data	3	-	-	03	3
2.	17M11CS112	Machine Learning and Data Mining	3	-	-	03	3
3.	21xxxxx	Elective – I	3	-	-	03	3
4.	21xxxxx	Elective – II	3	-	-	03	3
5.	21xxxxx	Elective – III	3	-	-	03	3
6.	18M11GE111	Research Methodology and Intellectual Property Right	2	-	-	02	2
7.	17M15CS111	Advanced Algorithms Lab	-	-	2	02	1
8.	17M15CS112	Machine Learning and Data Mining Lab			2	02	1
9.	17M15CS113	Cloud Technology Lab			2	02	1
		TOTAL	17	-	6	23	20

Elective-I
- A. Essential Statistics for Data Science
- B. Data Science Programming in R and Python

Elective-II
- A. Single Objective Optimization Algorithm
- B. E-Commerce and Social Web

Elective –III
- A. Large Scale Graph Algorithm and Analytics
- B. Big Data Engineering
- C. Computer Vision for Image Data Analysis

SECOND SEMESTER

SNo	Course		Contact Hours				Credits
	No.	Title	L	T	P	Total	
1.	17M11CS121	Cloud and Web Services Software Engineering	3	-	-	03	3
2.	17M11CS122	Performance Evaluation of Computing Systems	3	-	-	03	3
3.	21xxxxx	Elective – IV	3	-	-	03	3
4.	21xxxxx	Elective – V	3	-	-	03	3
5		Audit-I (HSS)	2	-	-	02	Qualifying
6.	17M17CS111	Project Based Learning-I (Open Source Software Development)		-	4	04	2
7.	17M15CS121	Cloud and Web Services Lab	-	-	2	02	1
8.	17M15CS122	Performance Engineering Lab			2	02	1
9.	17M15CS123	IOT Systems Development Lab			2	02	1
		TOTAL	15	-	10	24	17

Elective-IV
    A. Artificial Neural Network and Deep Learning
    B. IoT Data Analytics

Elective-V
    A. Multimedia Design and Analysis
    B. Technologies for Smart Cities

THIRD SEMESTER

SNo	Course		Contact Hours				Credits
	No.	Title	L	T	P	Total	
1.		OPEN ELECTIVE	3			3	3
2.	17M17CS212	Seminar & Term Paper/ Supervised Study or credits transfer from another university through normal course or MOOC				4	4
3.	17M17CS211	Project Based Learning-II (Software Development Automation)				8	4
4.	17M17CS213/ 17M17CS214/ 17M17CS215	Dissertation /Industrial Project/Entrepreneurial Project				8	4
5.		Audit-II (HSS)	2			2	Qualifying
		TOTAL					15

FOURTH SEMESTER

SNo	Course		Contact Hours				Credits
	No.	Title	L	T	P	Total	
1.	17M17CS223/ 17M17CS224/ 17M17CS225	Dissertation /Industrial Project/ Entrepreneurial Project				32	16
		TOTAL					16

TOTAL CREDITS : 68

Course outline for Electives

Elective-I

Essential Statistics for Data Science
Probability and Statistical inference: Modern Statistics, Statistics and Engineering, Probability, Conditional Probability, The Axioms of Probability, Sampling, Basic Statistics, Types of Convergence. Statistical Methods in Data Science: Data Distribution, Mathematical Statistics, Inferential Statistics, Descriptive Statistics, Gauss-Markov theorem, F-distribution. Hypothesis Testing, Z-test, ANOVA, T-Test, Redundancy Test, Chi-Square & F-test, Type-I and Type-II errors. Data Modeling- Cross validation, Monte Carlo methods, Cluster analysis, Time Series Modeling, Bayesian Modeling, Correlation Analysis, & Regression Models (Lasso, Ridge, Logistic, Linear, etc.), Ensemble Learning of Regression Models.

Data Science using R and Python
Python & R Fundamentals, Working with Data, Lists and Loops, Numeric, Date, String functions, Functions, and Classes, Basic Math and statistics, Exceptions and data structures, Object oriented design, Exploratory Data Analysis in R and Python: Analysing univariate and multivariate data, statistics for Categorical and numerical data

Data Visualization in R and Python: Line plot, Pie plot, Box plot, Heat Map, and Whisker plot, Scatter plot, Histograms, Bubble charts, , Data formatting, annotations, Time Series Data Visualization, Data Visualization of web, Choropleth Maps, Point Maps, Outlier Visualization.

Analysis in R and Python: Clusting using K-means, Predictive analysis, Association Analysis- Market Basket Analysis, Outlier analysis.

Elective-II

Single Objective Optimization Algorithms:
Introduction to Optimization, Linear Programming, Mixed Integer Linear Programming: Branch and Bound Method of NILP, Constraint Optimization, Constraint handling using Correction Approach. Optimization Algorithms- Binary Coded GA, Real Coded GA, Differential Evolution, Teacher-Learner Based Optimization, Particle Swarm Optimization, Artificial Bee Colony, Firefly Algorithm, Ant Colony Algorithm, Grey Wolf Optimization, MatLab Optimization Tool Box, GAMS/ IBM ILOG Optimization Studio, Optimization Real Context Case studies.

E-Commerce and Social Web
About E-Commerce, Selection of an E-commerce type and business model. Social Media , Social Media Analytics, Social Media Text analytics, Twitter as Marketing Tool, Social Web overview, data-types, format, Text cleaning, tagging and storage, Social media techniques, tools and platforms, data visualization of data, research, applications and challenges in social Web. Introduction to Social Commerce, Supporting Theories and Concepts for Social Commerce, Tools and Platforms for Social Commerce. Analyzing Social web, Nodes, Edges and Network measures, Centrality, Power and Bottlenecks, Concept of Cliques, Clusters and Components, Viral marketing, Graph data in real world, Business use of Social web, Privacy in Social web, Influencer Outreach. Social Media Marketing, Social Shopping: Concepts, Benefits, and Models, Customer Engagement and Metrics, Basic Social Marketing Strategies- Physical goods, Digital goods, Services, Affiliate Marketing, Guerrilla Marketing. Programming using API and RSS feeds, Twitter and Face book Data Analytics for Viral Marketing, Social Search Engine Optimization, Recommendation

Elective-III
Large Scale Graph Algorithm and Analytics
Introduction & Application of Large-scale Graph, Characteristics, Challenges, Hyper Graphs, Multi Graphs, Graph Duals, Complex Data Sources , Social graphs, Endorsement graphs, Technological graphs, Recommendation Graphs. Basic Large-scale Graph Analysis (Efficient Search – Graph Traversal and Search Algorithms; Pattern Discovery -Matching Algorithms, Centrality Computing Algorithms, List Ranking Algorithms; Partitioning – Connected Component Algorithms, Graph-Cut Algorithm. Graph indexing and ranking – Link Analysis Algorithms, Web Crawling, Page Ranking Personalized Page Rank, Page Rank Axioms, HITS. Adjacency Matrix Representation, Adjacency List Representation, V-Graph Representation (segmented vectors, storing graph topology), Graph Implementation Strategies & Software (RStudio, Python, Gephi, Pajek, SNAP, NetLogo, etc.). Power Law Distribution in Social Networks, Models of Power Law Random Graphs, Game-Theoretic Approach to Modeling Network Creation, Rank Aggregation and Voting Theory, Recommendation Systems

Big Data Engineering
Introduction to Big Data and Hadoop, Introduction to cloud and MapReduce Programming, Hadoop streaming using Python, Hadoop Eco System: PIG, Pig Latin, User defined functions, Hive, HIVE SQL over Haddop, Hive Shell, Hive services, HIVE QL, Querying data and user defined functions. Data Management and Relational Database modeling, NoSQL DataBases, Apache HBase, NoSQL Databases, and MongoDb, Real Time Data streaming. Apache Spark, Spark Programming, Spark RDD, Spark SQL, SparkMLLib, Analytics using PySpark. Machine Learning with Spark and Scale, Deep Learning Applications on Apache Hadoop and Apache Spark with GPU. Application Examples: Big AI chatbot, big spatial-temporal application i.e. geographical databases, big financial data, big multimedia data(Computer vision, Audio, Video), big medical/health data, big social media data, big scientific data i.e. bioinformatics data.

Computer vision for Image Data Analysis

About Computer Vision, Concepts of image processing in the spatial and frequency domains, pattern recognition. Binary image analysis, Greyscale image manipulation, linear and nonlinear filtering, feature extraction, image enhancement, image segmentation and recognition. Camera calibration and projective geometry, reconstructed three-dimensional information from single images, stereo pairs of images and motion sequences. vision technology applied across a broad range of business and consumer products. Computer vision potential applications: satellite imaging, control and measurement, industrial inspection, surveillance, face recognition and medical applications, Business intelligence.

Elective-IV

Artificial Neural Network and Deep Learning

Introduction, Processing units, Hebbian Learning, Linear Models, LMS Algorithm, Types of neural network architectures, Simple Neuron Model, Perceptron, backpropagation algorithm, Activation Functions, error functions, loss functions, optimization, About Bias-Variance, Overfitting, Inductive Bias, Momentum Methods, RmsProp, Gradient Descent, Regularization, Bayesian Approach, drop-out, Deep Learning, Tensorflow, Convolutional Net, Object Recognition, Neural Nets for sequence- Recurrent, LSTM, GRU. Unsupervised Learning: Autoencoder, Variation autoencoder, Probabilistic Neural Network, Boltzmann Machine, Restricted Boltzmann machine (RBM), Adversial nets, Deep Reinforcement Learning.

IoT Data Analytics (Taken from M.Tech- IoT)

Introduction to data analytics for IoT – Structured Versus Unstructured Data – Data in Motion Versus Data at Rest – IoT analytics overview and challenges ;Machine learning and Deep learning for getting Intelligence from IoT Big Data – IoT Predictive Analytics; Geographical Concepts and Spatial Technology for IoT; Big data Platform for IoT Analytics Massively parallel processing databases - Hadoop Ecosystem, Lambda Architecture- NoSQL Databases – Cloud based Amazon web services, Azure Data Lake and IoT Hub, Node RED; Edge and Fog Computing –Architecture of Edge and Fog Computing - Edge Analytics Core Functions – Distributed Analytics Systems - Fog Computing -Big Data Metadata Management – Data lifecycle - Data analytics at different Fog Layers ; Flexible NetFlow Architecture – FNF components – Flexible NetFlow in Multiservice IoT Networks - IoT Network Analytics; Web enhanced IoT -Design layers, design complexity-Web Enhanced Building Automation Systems – Smart City Control and Monitoring – Smart Environment Monitoring.

Elective-V

Multimedia Design and Analysis

Sensory Perception, Technologies, Taxonomies, and Applications. MM Types - Text; Graphical – static, dynamic, medical, remote sensed imaging; Audio, Video, web pages, UI designs, Games including VR System, etc; MM Standards, datatypes and Compression; MM Authoring tools; Principles of multimedia design and production for creation, corrections, enhancement of new or existing content; Understanding and using multimedia features, video analysis and management, retrieval techniques, spatial indexing methods, long-term learning and Relevance Feedback, audio analysis and retrieval, semantic based retrieval techniques; MM databases and data mining – storage, searching, indexing, retrieval, etc; Visual Data Mapping; Large-scale image retrieval; Learning from 3D sensors; Learning Methods for Images and Audio data sets.

Technologies for Smart Cities

Smart Cities: Definition, technology (AR/ VR/ Iot, Drone, etc), Smart Cities Marco View & Case Studies, Smart City Planning & Performance, Smart cities open Data and analytics : Transportation data, Smart Government, Smart Building and Home devices, Smart Community, Smart Energy and Water, Financial Data, Health data, Environment data, Public safety data, tourism data. GIS and Geo Data, Cybersecurity and privacy, Data complexity, Multiagent systems, design thinking, Introduction of sensor technology and Instrumentation, Sustainable development, Modeling and simulation Frameworks and Tools, Data grids and information visualization, data organization, Smart city informatics, Urban Network, Community resource sensing, Database handling, Social Media & Digital Inclusion.