School of Computer & Systems Sciences

M. Tech Programme in Statistical Computing

Course Structure

A student shall have to earn a minimum of 50 credits at the end of II year in order to be eligible for the award of M. Tech degree

### Semester I

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Course Name</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>*Data Structure &amp; Algorithms</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Optimization Techniques</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Probability &amp; Stochastic Processes</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Big Data Systems (Compulsory for Data Science Stream)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Data Communication &amp; Computer Networks (Compulsory for Data Communication Stream)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Statistical Computing Lab</td>
<td>3</td>
</tr>
</tbody>
</table>

*For students from non-computer science background, Data Structure & Algorithms is compulsory course. Students from Computer science background may opt for Linear algebra instead of Data Structure & Algorithms

### Semester II

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Course Name</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Statistical Inference &amp; Multivariate Techniques</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Big Data Algorithms (Compulsory for Data Science Stream)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Information Theory (Compulsory for Data Communication Stream)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Elective I</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Elective II</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Elective III</td>
<td>3</td>
</tr>
</tbody>
</table>

### Semester III

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Course Name</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Elective IV</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Elective V</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Seminar Course</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>Dissertation</td>
<td></td>
</tr>
</tbody>
</table>

### Semester IV

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Course Name</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dissertation (in continuation from semester III)</td>
<td>12</td>
</tr>
</tbody>
</table>
Elective courses Elective I, Elective II, Elective III, Elective IV and Elective V would be chosen from the stream in which the student has taken admission to the M.Tech. programme. The students from a given stream can opt at most two courses from Elective courses which are common elective courses to both streams. The elective courses in Data Science stream and Data Communication stream are given below:

<table>
<thead>
<tr>
<th>Data Science Stream</th>
<th>Data Communication Stream</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design Of Experiments</td>
<td>Design Of Experiments</td>
</tr>
<tr>
<td>Stochastic Modeling And Applications</td>
<td>Stochastic Modeling And Applications</td>
</tr>
<tr>
<td>Regression &amp; Time Series Analysis</td>
<td>Regression &amp; Time Series Analysis</td>
</tr>
<tr>
<td>Computational Intelligence &amp; Applications</td>
<td>Computational Intelligence &amp; Applications</td>
</tr>
<tr>
<td>Speech &amp; Natural Language Processing</td>
<td>Speech &amp; Natural Language Processing</td>
</tr>
<tr>
<td>Cloud Computing</td>
<td>Cloud Computing</td>
</tr>
<tr>
<td>Internet of Things</td>
<td>Internet of Things</td>
</tr>
<tr>
<td>Nature Inspired Algorithms</td>
<td>Nature Inspired Algorithms</td>
</tr>
<tr>
<td>Social Network Analytics</td>
<td>Social Network Analytics</td>
</tr>
<tr>
<td>Machine Learning</td>
<td>Statistical Signal Processing</td>
</tr>
<tr>
<td>Computational Finance</td>
<td>Wireless Sensor Networks</td>
</tr>
<tr>
<td>Bioinformatics</td>
<td>Multimedia Communication</td>
</tr>
<tr>
<td>Data Mining &amp; Pattern Recognition</td>
<td>Parallel &amp; Distributed Systems</td>
</tr>
<tr>
<td>Econometrics</td>
<td>Wireless Communication &amp; Mobile Computing</td>
</tr>
<tr>
<td>Spatial Data Analysis &amp;GIS</td>
<td>Game Theory &amp; Mechanism Design</td>
</tr>
<tr>
<td>Data Visualization Techniques</td>
<td>Vehicular Communication Networks</td>
</tr>
<tr>
<td>Big Data Analytics</td>
<td>Adhoc Networks</td>
</tr>
<tr>
<td>Data Stream Management</td>
<td>Data Compression</td>
</tr>
<tr>
<td>Multimedia and Video Analytics</td>
<td>Network Simulation</td>
</tr>
<tr>
<td>Web Mining</td>
<td>Network Security</td>
</tr>
<tr>
<td>Probabilistic Graphical Models</td>
<td>Multicast Communication</td>
</tr>
<tr>
<td>Probabilistic Risk Assessment</td>
<td>Performance Modeling of Computer Communication Networks</td>
</tr>
<tr>
<td>Data Security</td>
<td>Brain Theory &amp; Neural Networks</td>
</tr>
<tr>
<td>Smart Camera &amp; Visual Sensor Networks</td>
<td></td>
</tr>
</tbody>
</table>
Semester: I

Data Structures and Algorithms


Suggested Readings:

Optimization Techniques

Linear programming–Formulation, Simplex Method, Interior Point Methods, Duality and Sensitivity Analysis, Constrained Non-linear programming – Lagrange Multiplier Methods, Karush-Kuhn-Tucker Optimality Conditions, Quadratic and Separable Programming Methods, Unconstrained Non-linear Programming–Gradient Search, Newton Method

Suggested Readings:

Probability and Stochastic Processes

Introduction to Probability Theory, Bayes Formulae, Random Variables, Expectation, Moment Generating Function, Characteristic Function, Jointly Distributed Random Variables, Weak Law and Strong Law of Large Numbers, Modes of Convergence, Limit Theorems. Sample Moments and Their Distributions; Introduction to Stochastic
Processes, Markov Chains, Chapman-Kolmogorov Equations, Classification of States, Time Reversible Markov Chains, Random Walk and Gambler's ruin Problem; Poisson Processes, Continuous-time Markov Chains, Birth-Death Processes, Uniformization, Renewal Theory, Limit Theorem, Semi-Markov Processes, Queuing Theory, M/M/1, M/G/1 Systems, Network of Queues, Reliability Theory, Martingales, Stopping Times, Brownian Motion, Using Martingales to Analyze Brownian Motion Process

**Suggested Readings:**

**Big Data Systems (for Data Science Stream)**

Introduction to Big Data: Big Data characteristics, types of Big Data, Traditional vs. Big Data business approach; Big Data Architecture: Tradition Information Architecture, Big Data Architecture Capabilities: Storage, Management, Database, Processing, Data Integration; Introduction to MapReduce and Hadoop: Core Hadoop Components; Hadoop Ecosystem; Physical Architecture; Hadoop limitations; Distributed File System, HDFS; Data Management Techniques to Store Data Locally and in Cloud Infrastructures; Data-Intensive Computations on Cluster and Cloud Infrastructures using MapReduce; NoSQL: introduction, architectural patterns; Challenges and Opportunities in Big Data Management

**Suggested Readings:**
1. Rajaraman, A., Ullman, J. D., Mining of Massive Datasets, Cambridge University Press, United Kingdom, 2012  

**Data Communication and Computer Networks (for Data Communications Stream)**

Data Communication – Analog and digital communications, Channel characteristics, modulation, encoding schemes; Error Detection and correction, Flow control, multiplexing switching, Multiple access techniques, Routing– shortest path algorithms, routing protocols, virtual path routing, Network Protocols – IP, TCP, UDP, FTP, SMTP,
Suggested Readings:

Semester II

Statistical Inference and Multivariate Techniques


Suggested Readings:

**Big Data Algorithms (for Data Science Stream)**


**Suggested Readings:**

**Information Theory (for Data Communications Stream)**


**Suggested Readings:**

**Design of Experiments**

Introduction: Guidelines for Designing Experiments, Experiments with a Single Factor, Analysis of Variance, Randomized Blocks, Latin Squares and Related Designs, Factorial Designs: $2^k$ factorial design, Fractional Factorial Designs- Two level, Three level and Mixed Level Designs, Response Surface Methods: Process Optimization, Response Surface, Mixture Experiments, Robust Designs, Nested and Split plot Designs, Non-normal Responses and Transformations, Analysis of Covariance, Computer Solutions, Factorial Experiments with Covariates (Emphasis will be placed on computational methods in the design of experiments)

**Suggested Readings:**

**Stochastic Modeling and Applications**


**Suggested Readings:**

**Regression and Time Series Analysis**


**Suggested Readings:**

**Computational Intelligence and Applications**


**Suggested Readings:**
1. C. M Bishop, “Neural Network for Pattern Recognition”, Oxford University Press, 2014

Speech and Natural Language Processing

Biology of Speech Processing; Place and Manner of Articulation; Word Boundary Detection; Argmax based computations; HMM and Speech Recognition. Morphological Diversity of Languages; Morphology Paradigms; Finite State Machine Based Morphology; Automatic Morphology Learning; Shallow Parsing; Named Entities; Maximum Entropy Models; Random Fields. Theories of Parsing, Parsing Algorithms; Robust and Scalable Parsing on Noisy Text as in Web documents; Hybrid of Rule Based and Probabilistic Parsing; Scope Ambiguity and Attachment Ambiguity resolution. Lexical Knowledge Networks, Wordnet Theory; Indian Language Wordnets and Multilingual Dictionaries; Semantic Roles; Word Sense Disambiguation; WSD and Multilinguality; Metaphors; Coreferences. Sentiment Analysis; Text Entailment; Robust and Scalable Machine Translation; Question Answering in Multilingual Setting; Cross Lingual Information Retrieval (CLIR).

Suggested Readings:
2. A. James, “Introduction to Natural Language Understanding”, Addison Wesley, 1991

Cloud Computing


Suggested Readings:

**Internet of Things**

Introduction to the Internet of Things (IoT), Technology and Business drivers for IoT, IoT Architectures and design considerations; IoT paradigms and frameworks; semantics, security, privacy, network and standardisation issues; IoT Integration with Cloud technologies, big data, cyber-physical systems, components, network technologies; IoT application and device programming; Data analytics for IoT; Typical IoT applications, Trends and implications; Challenges and Opportunities with IoT.

**Suggested Readings:**

5. Adrian McEwen, Hakim Cassimally, Designing the Internet of Things, John Wiley and Sons, 2014

**Nature Inspired Algorithms**


**Suggested Readings:**

Social Network Analytics

Background: Review of graph theory, graph isomorphism, types of graph, connectivity, paths and cycles, trees, graph representation, spectral properties of graphs; Introduction: Real world complex networks – technological networks, information networks and social networks, topological properties of networks, network data sets; Social Networks: Strong and weak ties, strength and network structure in large scale data, homophily, tracking link formation in online data; Structure of the web: Hypertext and associative memory, web as a directed graph, link analysis and web search; Network effects: Information cascades, stability, instability and tipping points, power laws and rich – get richer phenomenon; Network dynamics: Modeling diffusion through a network small world phenomenon, epidemics; Case studies: Facebook, LinkedIn, Google Plus and Twitter

Suggested Readings
1. David Easley and Jon Kleinberg, “Networks, Crowds and Markets”, Cambridge University Press, 2010

Machine Learning (Data Science Stream)

Overview of Machine Learning; Concept Learning and the General - to - Specific Ordering, Decision Tree Learning, Neural Networks, Evaluating Hypothesis, Bayesian Learning, Computational Learning Theory; Instance Based Learning, Genetic Algorithms, Learning Sets of Rules, Analytical Learning, Combining Inductive and Analytical Learning, Reinforcement Learning, Markov and Hidden Markov Model

Suggested Readings:

**Computational Finance (Data Science Stream)**

Basic Functional Mathematics- Time Value of Money, Annuities, Yields, Bonds, Bond Price Volatility, Term Structure of Interest rates; Option Basics, Exchange Traded Options, Arbitrage in Option Pricing, Relative Option Prices, Put-Call Parity and its Consequences; Option Pricing Models- Binomial Option Pricing Model, Black- Scholes Formula; Forwards, Futures, Future Options, Forward Contracts; Continuous Time Functional Mathematics- Stochastic Integrals, Black- Scholes Differential Equation, Hedging and Futures, Hedging and Options.

**Suggested Readings:**


**Bioinformatics (Data Science Stream)**

Suggested Readings:

Data Mining & Pattern Recognition (Data Science Stream)


Suggested Readings:

**Econometrics (Data Science Stream)**


**Suggested Readings:**

**Spatial Data Analysis and GIS (Data Science Stream)**

Introduction review of non-spatial statistics, overview of different types of spatial data
Geostatistics: Variograms and covariance functions, fitting variogram functions, kriging, spatial regression
Areal data: Neighborhoods, testing for spatial association, Global and local tests of association, CAR and SAR models, inference, phenomena mapping
Point process data: types of spatial pattern, spatial clustering
GIS conceptual framework, Database, Visualization, Modelling and Analysis
Special topics: Non-stationary Covariance, Bayesian methods, Spatio-temporal modelling, Current Topics

**Suggested Readings:**
2. T. C. Bailey and A. Gatrell, "Interactive Data Analysis", Longman, 1995
3. S. Fotheringham, A. Stewart, C. Brunsdon and C. Martin, "Quantitative Geography: Perspectives on Spatial Data Analysis", SAGE publication, 2000

**Data Visualization Techniques (Data Science Stream)**

Interactive and Dynamic Graphics: Scatter Plots; Brushing and Linked Brushing; Focussing, Zooming; Rotations and Projections; Parallel Coordinate Plots, Andrews Plots; Density Plots; Categorical Data; Virtual Reality, Interactive 3D Graphics, 3d Representation of Statistical Data; Exploratory Spatial Data Analysis, Application of Interactive and Dynamic Graphics. Applications from Geography, Medicine and Environmental Sciences, Interactive Micromaps, Choropleth maps; Graphical Software; Limitations of Graphics

**Suggested Readings:**

**Big Data Analytics (Data Science Stream)**

Introduction to Data Mining, Data Analytics, Predictive Analysis and Business Intelligence, Large Scale File System, Mining Big Data, Advanced Data Analytics and Machine Learning, Big Data Streams and Real Time Predictive Analysis, Tools and Visualization, Link Analysis, Web Analytics, Collaborative Filtering, Social Network Analysis, Issues, Challenges and Opportunities with Big Data and its Analytics
Suggested Readings:
1. Rajaraman, A., Ullman, J. D., Mining of Massive Datasets, Cambridge University Press, United Kingdom, 2012

Data Stream Management (Data Science Stream)

Introduction to data streams, overview of streaming applications, architecture of Data Stream Management Systems, issues in data stream management, data models and query semantics for streams, streaming operators and languages, query processing and optimization, algorithmic issues, resource management, multiple and distributed streams, mining and analysis of data streams, streaming applications and systems, security and privacy, stream reduction, data stream management in mobile environments

Suggested Readings:
2. Joao Gama, “Knowledge Discovery from Data Streams”, CRC Press, Taylor and Francis Group, 2010

Multimedia and Video Analytics (Data Science Stream)

Introduction to multimedia systems. Multimedia compression including fundamentals of compression, text compression, image compression, audio and speech compression, video compression. Multimedia information storage and retrieval including text. Audio, image and video storage and retrieval methods multimedia programming multimedia security. Usability of multimedia.

Basics of image processing, computer vision and machine learning (assuming background in these areas). Video content analysis including: moving object detection and tracking algorithms. Image/video features for human activity detection and recognition in images and video image/video event classification and recognition objects in video counting approaches. Anomaly detection in images and videos multi-camera video analysis. Analyzing videos for video analytics applications such as retail analytics, healthcare, traffic analytics.

Suggested Readings:
Web Mining (Data Science Stream)


Suggested Readings:
1. Bing Liu, Web Data Mining, Springer Publication
2. Somen Chakrabarti, Web mining, Elsevier Publication
4. Witton Frank, Data Mining, Morgan Kauffan Publishers

Probabilistic Graphical Models (Data Science Stream)


Suggested Readings:
2. Adnan Darwiche, “Modeling and Reasoning with Bayesian networks”.

References:

**Probabilistic Risk Assessment (Data Science Stream)**

Concept of risk, objective and scope of risk assessment, probabilistic risk, risk perception and acceptability, Quantitative aspects of risk. Three levels of risk quantification, PRA management, preliminary hazard analysis, HAZOP and HAZAN, FMEA and FMECA analysis, Fault tree Analysis, Digraph and other approaches. Computation of Hazard probability, unavailability and other parameters using fault tree methodology. Monte Carlo Simulation technique, Event tree analysis, identification of initiating events, sequence and scenario development, system analysis, external events and dependent failures and quantification, Accident-consequence Analysis, uncertainty analysis, sensitivity analysis and importance measures, Bayesian approaches. Human Reliability analysis.

**Suggested Readings:**

1. Mohammad Modarres, “Probabilistic Risk Assessment”, Springer

**Data Security (Data Science Stream)**


**Suggested Readings:**

2. Dorothy Elizabeth Robling Denning, Cryptography and Data Security, publisher Addison-Wesley, 1982
3. Brent Mullins, Data Security Complete Certification Kit, Publisher Emereo Publishing company, May 2016

**Smart Camera & Visual Sensor Networks (Data Science Stream)**
Basics of image sensors and processing, computer vision and visual sensor networks, Calibration of smart camera networks, Camera network tracking and re-identification, Visual analytics in a smart camera network, Data association in visual sensor network, Distributed camera networks, Collaborative sensing and analysis in visual sensor networks.

**Suggested Readings:**
4. B. Bhanu, C. Ravishankar, A. Roy-Chowdhury, H., Distributed Video Sensor Networks

**Statistical Signal Processing (Data Communication Stream)**


**Suggested Readings:**

**Wireless Sensor Networks (Data Communication Stream)**


**Suggested Readings:**

**Multimedia Communications (Data Communication Stream)**


**Suggested Readings:**

**Parallel and Distributed Systems (Data Communication Stream)**

Parallel processing concept, Parallelism in conventional machine, Pipelining, Flynn’s classification, Feng’s classification, Array processor, Amdahl’s law, Minsky’s conjecture. Static and dynamic networks, Single stage and multistage interconnection network, Blocking and nonblocking network, Star, Ring, Mesh, Torus, Pyramid etc. topology, Elementary permutations used in Interconnection network, Crossbar, Clos, Benes network, Shuffle exchange, Hypercube, PM21 network. Simple addition on various network topologies, Recurrence computation, Matrix multiplication, Sorting networks 0-1 Principle, Bitonic sorter, Merger, SorterPRAM Model, EREW, ERCW, CREW, CRCW algorithms. Distributed computation, characteristics of distributed systems, overview of related networking, operating systems and programming language concepts. Interprocess communication, message passing communication, remote procedure call (RPC), atomic transactions. Distributed coordination, physical and logical clocks, synchronization, mutual exclusion, leader election.

**Suggested Readings:**

**Wireless Communication and Mobile Computing (Data Communication Stream)**

Mobile radio systems, Paging systems, cordless telephone system, cellular telephone system, Cellular Concept: Frequency reuse, channel assignment, hand off, Interference and cell splitting, sectoring, Improving Coverage and capacity in Cellular systems. Propagation modeling: Outdoor/Indoor Propagation models, Small scale Multipath
propagation - Rayleigh fading, Ricean Fading, Nakagami fading, Shadowing, lognormal shadowing fading model, outage probability, coverage estimation under shadowing, and multipath fading. Wireless Networks 802.11, frequency-hopping, encoding and modulation, MAC Layer Protocol Architecture Multiple access with collision avoidance protocol, Virtual Carrier-Sensing, DCF Protocol, PCF Operation. Mobility: challenges, limits and connectivity, mobile TCP, mobile IP and cellular IP in mobile computing.

Suggested Readings:
2. Matthews. Gast, 802.11 wireless networks, o’reilly
4. Jochen Schiller, Mobile communications, phi/person edu., 2nd ed.

Game Theory and Mechanism Design (Data Communication Stream)


Learning: Theories of learning, reinforcement learning, belief learning, Bayesian learning.

Introduction to mechanism design: Social choice functions, compatibility and revelation theorem, Vickery-Clarke-Groves (VCG) mechanisms, and individual rationality.

Auctions: Types and desirable properties, combinatorial auctions, optimal mechanisms, and Myerson auction.

Cooperative Game theory: Two person bargaining problem, the Shapley value.

Suggested Readings:
1. Y Narahari, Game Theory and Mechanism Design, IISc Lecture Notes Series, Vol.4

Vehicular Communication Networks (Data Communication Stream)


Suggested Readings:

**Ad Hoc Networks (Data Communication Stream)**


**Suggested Readings:**

**Data Compression (Data Communication Stream)**


**Suggested Readings:**

**Network Simulation (Data Communication Stream)**


**Suggested Readings:**

**Network Security (Data Communication Stream)**

Introduction, Security goals, attacks, services and mechanisms, cryptography and steganography, Symmetric Key cipher-substitution ciphers, Transposition ciphers, stream and block ciphers, Modern block ciphers, Modern stream ciphers, DES and AES, Elliptic curve cryptosystems, RSA, Message integrity, Digital signature, Public key distribution, IPSec, SET, ESP, PGP, SSL, Security in wireless.

**Suggested Readings:**

**Multicast Communication (Data Communication Stream)**

Introduction, Application, Characteristics, Multicast Backbone Architecture, Multicast Routing, Basic Routing Algorithm, Group Dynamics, Multicast routing between domains, Ip multicast, Multicast in transport protocols, address allocation, Multicast LANs, Reliable Multicast, Congestion control, Security issues.

**Suggested Readings:**

**Performance Modeling of Computer Communication Networks (Data Communication Stream)**

Role of Modeling and Analysis, Examples of Performance Modeling, Analytic Models, Elements of Stochastic process, Poisson Process, Basic Queuing models, M/M/1; M/M/∞; M/G/∞; M/M/m; M/M/m/m Queues with Product formula. Cell and Burst scale Traffic Models: Round trip time distribution, PING data, Markov modulated Poisson Process, Long Range Dependence, Heavy Tail Distribution. Traffic Control: Admission Control, Effective Bandwidth, Statistical Multiplexing gain, Access Control: Leaky bucket System. Multi access Modelling: Slotted ALOHA Markov chain, Diffusion
Approximation Approach, CSMA, Congestion Control, Window Control, Modelling TCP, Window Size, TCP Window Dynamics.

**Suggested Readings:**
2. I. Kaj, 2002, Stochastic Modeling in Broadband Communications Systems, SIAM.

---

**Brain Theory and Neural Networks (Data Communication Stream)**

Introducing the Neuron, Basic properties of Neurons, receptors and effectors, Neural models

Biological neurons and networks, detection and classification of extracellular action-potential recordings, information theoretic analysis of neural data, identification of non-linear dynamics in neural population activity.


**Suggested Readings:**