Tribhuvan University

Institute of Science and Technology

SCHOOL OF MATHEMATICAL SCIENCES

Syllabus

Master's in Data Sciences (MDS)- THIRD SEMESTER

Compulsory Courses

Course Code	Course Titles	Credits	Nature
MDS 601	Research Methodology	3	Th.
MDS 602	Advanced Data Mining	3	Th.+ Pr.
MDS 603	Techniques for Big Data	3	Th.+ Pr.

Elective Courses (Any Two Available on School)

Course Code	Course Titles	Credits	Nature
MDS 604	Cloud Computing	3	Th.+ Pr.
MDS 605	Regression Analysis	3	Th.+ Pr.
MDS 606	Decision Analysis	3	Th.+ Pr.
MDS 607	Monte Carlo Methods	3	Th.

Code No. : MDS 601
Course Title: Research Methodology **Full Marks:** 75
Nature: Theory (Compulsory) **Credit: 3**

Course Description

The paper deals with different essential components of research methodology mostly focusing on quantitative research work. It explains the fundamental notions of research methodology, formulation of research problems and hypothesis, different aspects of review of literature and study designs. It also deals with the practical aspects of sampling, choice of appropriate data analysis technique(s) and interpretations of statistical findings with reference to the data problem. Finally, how a research report, journal article is prepared, tools and techniques used in preparing journal articles, ethical issues and the issues of plagiarism are included. The tutor should involve the students practically in preparing the project proposals and thesis proposals, which they will be going to conduct in the next semester.

Learning Objectives

After completion of the course, students will be able to

- Understand the need and importance of research methodology
- Formulate research problem, research hypothesis
- Know the skills of reviewing literature, citation and referencing
- Understand the different types of research design
- Perform online searching and prepare research/ project proposal
- Choose appropriate statistical tools and interpret the results
- Know how to prepare research reports and journal articles

Course Contents:

Unit 1: Introduction to Research Methodology **[5 Hrs.]**

Meaning, objectives, need, utility, research methods and methodology, Deductive and inductive theory, Characteristic of scientific method, Research language, concept, constructs, research process

Unit 2: Research Problem Identification and Formulation **[5 Hrs.]**

Understanding research problem and its formulation, Statement of research problem, Research question, Research hypothesis, Statistical hypothesis, Linkage between research hypothesis and statistical hypothesis

Unit 3: Review of Literature and Research Design **[12 Hrs.]**

Importance of review of literature in research work, Review of research journals, Use of encyclopedia, reference books, research guides, handbooks, academia database in the relevant field, Citations and referencing, American Psychological Association(APA), IEEE style, Bibliography; Need and importance of research design, Exploratory, descriptive, analytical, and experimental research design; Online searching: Different database, SCIFinder, Scopus, Science direct, Searching research articles, Citation index, Impact factor, H-index; Preparation of project proposals and research proposals

Unit 4: Data Analysis and Interpretations **[15 Hrs.]**

Concept of measurement and scaling, Validity and reliability, Levels of measurement- nominal, ordinal, interval and ratio scales; Sampling: Concepts of statistical population, sample survey, sampling frame, target population, sampling and non-sampling errors, Issues of choosing appropriate sampling technique(s) while selecting samples in the study, sample size, issues of practical considerations in sampling and sample size; Use of appropriate statistical tests and models, interpretations of statistical findings in terms of problem specific sense

Unit 5: Preparation of Research Report **[11 Hrs.]**

Preparation of academic and other reports, Formats of research reports (both academic and others), Research paper writing, layout of research paper, Journals in relevant field, Impact factor journals, Communications in the journals for publishing research article, Plagiarism, ethical issues in research and publishing the articles, Techniques and tools used for research: Effective way of searching research materials, Handling reference management software such as Mendeley, research paper formatting software such as LaTeX, MS Office.

References:

1. Kothari C. R. (2014). *Research Methodology, Methods and Techniques*, 3rd Edition, New Age international Publishers, New Delhi, India.
2. Kumar Ranjit(2011). *Research Methodology: A Step by Step Guide for Beginners*, SAGE Publications PVt. Ltd, India.
3. Creswell J.W.(2014). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*, 4th Edition, SAGE Publications, India.
4. Day Robert A. (1996). *How to Write and Publish Scientific paper*, Cambridge University Press
5. Relevant updated References from the internet

●●●●●

Code No.: MDS 602
Course Title: Advanced Data Mining **Full Marks: 75**
Nature: Theory + Practical (Compulsory) **Credit: 3**

***Course Description*:**

This course covers some basic and advanced concepts of data mining techniques including techniques for association analysis, classification, clustering, and outlier detection.

***Learning Objectives*:**

After completion of this course, students should be able to

- Understand concepts like pre-processing, measures of similarity, summary statistics and visualization of data
- Understand and implement different algorithms for association, classification, clustering, and anomaly detection

***Course Contents*:**

Unit 1: Introduction **[6 Hrs.]**

Introduction of data mining, Origins of data mining, Data mining tasks, Types of data, Data quality, Data pre-processing, Measures of similarity and dissimilarity, Summary statistics and visualization.

Unit 2: Association Analysis **[7 Hrs.]**

Basic concepts, Apriori algorithm, Pattern-growth approach, Handling categorical and continuous attributes, Handling concept hierarchy, Sequential and subgraph patterns

Unit 3: Classification **[14 Hrs.]**

Decision trees induction, Rule-based classifier, Nearest-neighbor classifier, Bayesian classifier, Artificial neural network (ANN), Support vector machines, Ensemble methods, Model evaluation and selection

Unit 4: Clustering **[12 Hrs.]**

Overview; K-Means, Agglomerative hierarchical clustering, DBSCAN, Fuzzy clustering, Graph based clustering, Clustering evaluation

Unit 5: Outlier Detection **[9 Hrs.]**

Preliminaries, Outliers and types, Statistical approaches, Proximity-based approaches, Clustering-based approaches, Classification based approaches

Laboratory Works:

Laboratory work includes writing computer programs using high level programming language like Python to implement all the data mining techniques and algorithms studied in the course.

References:

1. Pang-Ning Tan, Michael Steinbach, and Vipin Kumar (2014). *Introduction to Data Mining*, Pearson.
2. Jiawei Han, Micheline Kamber, and Jian Pei, Data Mining (2012). *Concepts and Techniques*, Third Edition, MK.
3. David Forsyth (2019). *Applied Machine Learning*, Springer.

●●●●●

Code No.: MDS 603

Course Title: Techniques for Big Data **Full Marks: 75**

Nature: Theory + Practical (Compulsory) **Credit: 3**

***Course Description*:**

In this course students will learn fundamental knowledge to handle the challenges of Big Data. This course shall first introduce the overview of Big Data, its scope, applications, challenges and current trends. Then, it will introduce the fundamental platforms, such as Hadoop, MongoDB and HBase as NoSQL databases, Spark, Hive and Pig. Students will then have fundamental knowledge on Big Data Analytics to solve various real-world problem posed by Big Data.

Learning Objectives:

Upon the conclusion of the course, students should be able to:

- Understand the fundamentals of big data, its characteristics, scopes and challenges
- Setup Hadoop cluster and perform basic file operations in HDFS
- Write MapReduce Programs to solve big data problems in real life run the job in Hadoop Environment
- Understand different types of NoSQL database and analyze Structured, Unstructured and Semi Structured Data
- Perform Real Time Analytics with Spark Streaming
- Preform SQL Like queries on top of Hadoop using Spark SQL, Pig and Hive.

***Course Contents*:**

Unit 1: Introduction **[4 Hrs.]**

Introduction to big data, Characteristic of big data, Current trend and real-life applications of big data, Scope and challenges of big data, Tools and technologies used in big data.

Unit 2: Hadoop Ecosystem **[6 Hrs.]**

Introduction to Hadoop, History of Hadoop, Hadoop ecosystem,
Core components of the Hadoop ecosystem: Hadoop common, Hadoop distributed file system (HDFS), Map reduce framework and YARN (Yet another resource negotiator),
Hadoop master/slave architecture, Hadoop daemons, Hadoop configuration modes, Hadoop cluster setup, Hadoop streaming.

Unit 3: Distributed Storage and Processing of Big Data **[12 Hrs.]**

HDFS: The Design of HDFS, HDFS Concepts, Command line interface, Hadoop file system interfaces, Data flow, Basic HDFS commands,
MapReduce: Functional programing, MapReduce fundamentals, Execution overview of MapReduce, Basic MapReduce API Concepts, Setting up the development environment, Writing unit test with MRUnit, Running locally on test data, Running on a cluster, Anatomy of a MapReduce job run, Failures, Shuffle and sort, Task execution, Map reduce types and formats.

Unit 4: NoSQL Databases **[10 Hrs.]**

Types of data, Introduction to NoSQL, Need of NoSQL, Types of NoSQL Databases, NoSQL v's Relational databases, The CAP theorem,

MongoDB (Collections, Documents, Object Ids, Queries on MongoDB, Aggregation pipeline, Nested documents),

HBase (Overview, HBase vs. RDBMS, HBase vs. HDFS, HBase Architecture, HBase Data Model, Concept of Row Keys and Column Families, HBase Regions, Creating a Table, Writing Queries to insert and retrieve data to and from HBase.

Unit 5: Data Analytics with Spark **[6 Hrs.]**

Introduction to spark, Need of spark, Evolution of spark, Spark shell, Spark context, Resilient distributed dataset (RDD), Transformations, Programming with RDD, Spark core, Spark SQL, MLib, Spark streaming, GraphX.

Unit 6: Querying Big Data with Pig and Hive **[10 Hrs.]**

Pig: Introduction to pig, Execution modes of pig, Comparison of pig with databases, Grunt, Pig Latin, User defined functions, Data processing operators.

Hive: Hive shell, Hive services, Hive metastore, Comparison with traditional databases, HiveQL, tables, Querying data and User defined functions

Practical Works:

1. Setup Hadoop cluster in pseudo distributed and fully distributed mode
2. Perform basic file system operation on HDFS
3. Write MapReduce programs in Java/Python and run the job in Hadoop cluster
4. Analyze Big Data (CSV and JSON) using Hadoop MapReduce
5. Install MongoDB and perform CRUD operations
6. Twitter data analysis using MongoDB
7. Install HBase in pseudo distributed and fully distributed mode and perform CRUD operations
8. Use SparkSQL and dataframe to analyze CSV and JSON data
9. Use spark streaming for real time analytics
10. Installing and running pig
11. Installing and running hive
12. Project work to analyze big data (CSV and JSON) using any 3 technologies among Hadoop MapReduce, MongoDB, HBase, Spark, Pig and Hive

References:

1. Tom White (2015). *Hadoop: The Definitive Guide Fourth Edition*, O'Reilly Media.
2. Jeffrey Dean and Sanjay Ghemawat (2004). *MapReduce: Simplified Data Processing on Large Clusters*, Google.
3. Judith Hurwitz (2013). Alan Nugent, Dr. Fern Halper, and Marcia Kaufman, *Big Data for Dummies*, Wiley
4. Chuk Lam (2011). *Hadoop in Action*, Manning
5. Adam Fowler (2015), *NoSQL for Dummies*, Wiley

●●●●●

Code No.: MDS 604
Course Title: Cloud Computing **Full Marks: 75**
Nature: Theory +Practical (Elective) **Credit: 3**

Course Description:

In this course students will learn the fundamental concept of cloud computing, the course contents the cloud computing infrastructure, specification of cloud system structure and implementation of specifications, the evolution of cloud computing methods and tools for development of cloud infrastructure in an economic and timely manner. The goal of the course is to provide the knowledge of as software as service, platform as service and infrastructure as service.

Learning Objectives:

After completion of the course, students will be able to

- Understand the concepts and principles of cloud computing its role
- Describe about Cloud reference model, Cloud deployment models, Cluster and Grid computing: Grid computing Versus Cloud computing.
- Explain the concept of virtualization, MapReduce and benefit of virtualization.
- Describe the concept, principles and practice of security use in cloud computing and Security Architecture Design
- Describe the cloud platform, application and some case studies related to cloud computing

***Course Contents*:**

Unit 1: Introduction to Cloud Computing [6 Hrs.]

History and development of Cloud computing , Characteristics of Cloud computing, Types of cloud, Cloud services: Benefits and challenges of cloud computing, Evolution of Cloud computing , Applications cloud computing, Business models around Cloud, Cloud Architecture, Cloud storage, Cloud services requirements, Cloud and dynamic infrastructure, Cloud adoption.

Unit 2: Architecture of Cloud Computing [6 Hrs.]

Cloud computing Characteristics, Cloud reference model -platform as service, software as a service, infrastructure as service, Cloud deployment models -Public clouds, private clouds, Community cloud, hybrid clouds, Cloud design and implementation using SOA, security, trust and privacy

Unit 3: Data in the Cloud [8 Hrs.]

Map-Reduce and extensions: Parallel computing, The map-Reduce model, Parallel efficiency of Map-Reduce, Relational operations using Map-Reduce, Enterprise batch processing using Map-Reduce, Introduction to cloud development, Map-Reduce model

Unit 4: Cloud Virtualization Technology [10 Hrs.]

Virtualization defined, Types of Virtualization, Implementation Levels of Virtualization Structures, virtualization benefits, server virtualization, hypervisor management software, virtual infrastructure requirements.

Unit 5: Cloud Security **[6 Hrs.]**

Introduction to security, Cloud security challenges and risks, Software as-a-service security, Security monitoring, Security architecture design, Data security, Application security, Virtual machine security, Identity management and access control, Autonomic security**.**

Unit 6: Cloud Platforms, Applications and Case Studies **[12 Hrs.]**

Web services, appEngine, azure Platform, Aneka, open challenges, scientific applications, business and consumer applications.

Practical Works:

The practical work consists of Cloud programming models such as Thread programming, Task programming and Map-reduce programming.

References:

1. Dr. Kumar Saurabh (2012). *Cloud Computing, 4th ed .,* John Wiley and Sons
2. Raj Kumar Buyya, Christian Vecchiola, S. Thamarai Selvi (2013). *Mastering Cloud Computing* , Tata McGraw-Hill Education
3. David S. Linthicum (2010). *Cloud Computing and SOA Convergence in your enterprise,* Pearson Education, Inc., India.
4. Barrie Sosinsky (2011). *Cloud Computing Bible* John Wiley and Sons.
5. Saurabh, K. (2011). *Cloud Computing – Insights into New -Era Infrastructure,* Wiley India.

●●●●●

Code No. : MDS 605

Course Title: Regression Analysis **Full Marks:** 75

Nature: Theory and Practical (Elective) **Credit: 3**

Course Description:

The paper covers important intermediate to advanced level topics on regression analysis useful for statistical data analysis and modeling. It deals with multiple regression analysis, its uses and associated model adequacy tests. Additionally, it covers issues on modeling techniques and methods to deal with quantitative and qualitative variables. Finally, techniques of choosing the best regression is included.

Learning Objectives:

The general objective of the course is to make students understand and enable them to apply regression analysis methods effectively so that they are able to associate and establish relationship between variables statistically. Specifically, the objectives are to make them capable with handling multiple linear regression, regression with qualitative variables, and selection of best regression appropriately during data analysis with focus in data science.

***Course Contents*:**

Unit 1: The Multiple Linear Regression **[10 Hrs.]**

The *k* variable linear regression model, model specification including in matrix form, assumptions, OLS estimation of parameters and their interpretations, the hat matrix, properties of least square estimates, standard error and confidence interval of estimates and mean response, test of significance, analysis of variance for goodness of fit, prediction, unadjusted and adjusted multiple coefficient of determinations, standardized regression coefficients, lack of fit tests, use of computer packages in regression analysis including R package.
Problems and examples focusing data science

Unit 2: Model Adequacy Tests: Normality and Multicollinearity **[10 Hrs.]**

Normality: Kolmogorov-Smirnov and Anderson-Darling tests, Normal probability plots: P-P and Q-Q plots for assessing normality
Multicollinearity: Nature of multicollinearity, detection of multicollinearity, checking for correlations between explanatory variables, wrong regression coefficient signs, variance inflation factors, condition index, consequences and remedial measures

Problems and examples focusing data science

Unit 3: Model Adequacy Tests: Heteroscadasticity and Residual analysis **[10 Hrs.]**

Heteroscadasticity: Nature of heteroscedasticity and its detection through residual plots, Goldfeld-Quandt test, Breusch-Pagan-Godfrey test, Park test, treatment of heteroscedasticity: method of weighted least squares.

Autocorrelation and residual analysis: Definitions of residuals, raw, standardized and studentized residuals, the PRESS residuals, the nature of autocorrelation, detection and consequences of auto-correlation, residual plots for detecting autocorrelation including partial residual plots, Durbin Watson test, remedial measures: changing the functional form, Cochrane-Orcutt iterative procedure, detection and treatments of outliers.

Problems and examples focusing data science

Unit 4: Regression with Categorical Variables [12 Hrs.]

Regression with explanatory categorical or indicator variables: Nature of categorical or dummy variables and their effects and uses in regression models, regression with categorical independent variables: regression with one quantitative and one qualitative variables with two classes and more than two classes, use of multiple quantitative and qualitative independent variables, Use of measuring interaction effects using dummy variables.

Regression with qualitative dependent variables: The nature of categorical dependent variables in regression modeling, The binary logistic model, its importance and uses in data modeling, assumptions, estimation of parameters and their interpretations, odds ratio and its interpretation, model adequacy tests: Hosmer-Lemeshow test, ROC curve.
Problems and examples focusing data science

Unit 5: Methods of Selection of Best Regression [6 Hrs.]

All possible regressions and best subset regression, Stepwise regression: Forward and backward methods, Selecting significance levels in stepwise regression, The use of R^2 Statistic, Residual mean square and the Mellows C_p Statistic, Akaike information criterion
Problems and examples focusing data science.

Practical works:

The practical works include analysis and modeling for multiple regression, models for categorical data and selection of best possible regression including their statistical model adequacy tests using any statistical software.

References:

1. Douglass C. Montgomery, Elizabeth A. Peck and G. Geoffrey Vining (2012). *Introduction to linear regression analysis*, Fifth edition, Wiley.
2. Drapper, N. R. and Smith, H. (1998). *Applied Regression Analysis*, Third edition, Wiley.
3. Maddala, G. S. (2002). *Introduction to Econometrics*, John Wiley and Sons.
4. Ramanathan, B. (2002). *Introductory Econometrics with Applications*, South-Western Thomson Learning, Singapore.
5. Brian Caffo (2015). *Regression models for data science in R*, Lean publishing.

●●●●●

Code No. : MDS 606
Course Title: Decision Analysis **Full Marks: 75**
Nature: Theory + Practical (Elective) **Credit: 3**

***Course Description*:**

This course provides an overview on formal, structured, systematic and visual approach to evaluate problems that leads to decisions and action. It covers both philosophical aspects of decision theory and practical aspects of decision making environments under probabilistic and non-probabilistic situations. Moreover, it covers an interactive decision approach: theory of games which is indispensible while assessing the enterprise risks. It also covers the multi-objective functions related to optimization problem under goal programing.

Learning Objectives:

After the completion of the course, students will be able to

- Utilize a range of methods and tools to aid in the capture, analysis and synthesis of information for the effective decision making
- Identify the values, objectives, attributes, decisions, uncertainties, consequences, and trade-offs in a real decision problem
- Assess the enterprise risk management and adapt the interactive decision approach in real life
- Deal with multiple objectives taking into consideration of their priority
- Use Microsoft Excel / LINGO or latest software in decision analysis

Course Contents:

Unit 1: Introduction to Decision Theory [6 Hrs.]

History and introduction of decision analysis, Decision making philosophy, Elements of decision problems, Framework of decision making, Decision making processes, Problem solving and creative thinking, , Problem or opportunity findings, Types of decision , Decision making conditions, Cognitive biases, Decision making styles, Decision theories: Classical decision theory, Behavioral theory, Normative and descriptive decision theory; Group decision making , Techniques of group decision making

Unit 2: Decision Making Under Uncertainty and Risk [16 Hrs.]

Structuring a decision problem, Decision making under uncertainty (Criterion of optimism, Criterion of pessimism, Laplace criterion, Criterion of realism, Criterion of regret), Conditional probability and use of Bayes theorem, Posterior probabilities and Bayesian analysis ,Value functions, Decision tree analysis, Decision making with utilities: Utility function and loss function, Expected utility, Visualizing uncertainty, Utility curves ; Decision making under risk (EMV, EOL, EVPI), Decision making under ignorance ; Risk analysis and Sensitivity analysis.

Unit 3: Enterprise Risk Management [6 Hrs.]

Introduction, History, Trends in ERM, Risk management standard and guidance, Enterprise risk management integrated framework, Risk profile, Risk Appetite and Risk tolerance

Unit 4: Theory of Games [8 Hrs.]

Introduction, Game models: Two-Person Zero-Sum games, pure strategies: games with saddle point, mixed strategies: games without saddle point, Principles of dominance, solution methods of games without saddle points.

Unit 5: Goal Programming **[12 Hrs.]**

Philosophy of goal programming: Satisficing, Optimizing, Ordering, Balancing; Goal programming variants: Generic goal program, Distance metric based variants, Decision variables, Goal-based variants; Formulation of goal programming: Formulating goals and setting targets, Variant choice, Normalization; solving and analyzing goal programming, application of goal programming in various areas.

Practical Works:

The practical work consist of real-world problems with application of various techniques and methods including multi-attribute utility models, decision trees, and Bayesian models using Microsoft Excel and LINGO/ or latest Software. It also includes Project work with applications to real-world data.

References:

1. Charles Yoe. (2019). *Principle of Risk Analysis Decision Making Under Uncertainty*. New York: Taylor and Francis Group . Retrieved from https://b-ok.asia/book/3711279/e97f5b

2. Anderson D.R., Sweeney D.J., Williams T.A. & Martin K. (2011). *An introduction to Management Science Quantitative Approaches to Decision Making*. Delhi: Cengage Learning India Private Limited.

3. Dylan Jones & Mehrdad Tamiz. (2010). *Practical Goal Programming*. New York: Springer. Retrieved from https://b-ok.asia/book/1056846/f876da

4. Howard, Ronald A. and Abbas, Ali E. (2015). *Foundations of Decision Analysis. England: Pearson Education Ltd.*
5. Sharma J.K. (2004). *Quantitative Techniques for Managerial Decisions*. New Delhi: Rajiv Beri for Macmillan India Ltd.
6. Martin Peterson. (2009). *An introduction to Decision Theory*. New York: Cambridge University Press. Retrieved from https://b-ok.asia/book/2479794/2849b4
7. Vohra N.D. (2004). Quantitative *Techniques in Management*. New Delhi: Tata McGraw-Hill Publishing Company Limited.
8. Sven Ove Hansson, (2015). *Decision Theory: A Brief Introduction.* Stockholm: Royal institute of Technology

•••••

Code No.: MDS 607

Course Title: Monte Carlo Methods **Full Marks: 75**

Nature: Theory +Practical (Elective) **Credit: 3**

Course Description:

The goal of this course is to prepare the student to get an idea of Monte Carlo methods and able to apply to solve problems of data science. In this course, students will learn the fundamental concept of Bayesian statistics and the difference between frequentist and Bayesian inference. The students will study the concept of simple and Markov chain Monte Carlo methods and will be able to apply these methods to solve some common problems in Mathematics, Statistics and to handle big data.

Learning Objectives:

After completion of the course, students will be able to

- Differentiate between Frequentist and Bayes statistics
- Differentiate between simple and Markov Chain Monte Carlo methods
- Write appropriate programs language to apply Monte Carlo methods to solve some problems of data science
- Understand some of algorithms of Marko Chain Monte Carlo methods

Course Contents:

Unit 1: Approaches for Statistical Inference **[6 Hrs.]**

Introduction, Motivating vignettes, Defining the approaches, Bayes vs frequentist approach , Some basic Bayesian models

Unit 2: The Bayes Approach **[6 Hrs.]**

Introduction, Prior distributions, Bayesian inference, Hierarchical modeling, Model assessment

Unit 3: Monte Carlo Methods **[13 Hrs.]**

Introduction, Motivating examples, Random numbers, Pseudorandom number generators, Random walks, Markov processes; Simple, Importance and Rejection sampling

Fundamental concepts of transformation, reweighting; Monte Carlo integration, estimation of pie by using programing code in R /Python/Fortran /C (Note: Students are supposed to design algorithm and code for such simple problems in practical classes)

Unit 4: Markov Chain Monte Carlo **[6 Hrs.]**

Introduction, Definition and transition probabilities, Decomposition of the state space, Stationary distributions, Limiting theorems, Reversible chains, Continuous state space

Unit 5: Gibbs Sampling (Algorithms) **[10 Hrs.]**

Introduction, Definition and properties, implementation and optimization, forming the sample, scanning strategies, Using the sample, Reparametrization, convergence diagnostics: Rate of convergence, Informal convergence monitaors, convergence prescription, formal convergence methods, Applications: Hierarchical models, Dynamics models, spatial models, MCMC Code (Students should be able to write simple MCMC code in python/R/Fortran/C)

Unit 6: Metropolis-Hastings Algorithms **[7 Hrs.]**

Introduction, Definition and properties, Special cases, Hybrid Algorithms, Applications – case studies and examples

Practical Works:

The practical work consist of Applying Monte Carlo and Markov Chain Monte Carlo methods to solve different problems from Text books writing codes in Python/R/Frotran/C (any one).

References:

1. Dani G., & Lopes H.F. (2006). *Markov Chain Monte Carlo Stochastic Simulation for Bayesian Inference*, 2nd edition, CRC Boca Raton FL: Chapman & Hall
2. Carlin B.P., & Louis T.A. (2009). *Bayesian Methods for Data Analysis* 3rd edition, CRC Boca Raton FL: Chapman & Hall
3. Gilks W.R., Richardson S. & Spiegelhalter D.J. (1996). *Markov Chain Monte Carlo in practice*, CRC Boca Raton FL: Chapman & Hall
4. Kendall W.S., Liang F. & Wang J.-S. (2005). *MARHOV CHAIN MONTE CARLO Innovations and Applications*, Singapore World Scientific
5. Anagnostopoulos K. N. (2014). *Computational Physics*, National Technical University of Athens, Greece

●●●●●