

MEHRAN UNIVERSITY OF ENGINEERING & TECHNOLOGY,
JAMSHORO
DEPARTMENT OF COMPUTER SYSTEMS ENGINEERING
INSTITUTE OF INFORMATION AND COMMUNICATION
TECHNOLOGIES

MS Data Science (MS DS)

FRAMEWORK OF COURSES AND CURRICULUM
(FROM BATCH 2023 TO ONWARDS)

Curriculum for MS Data Science (MS-DS).

Deficiency Courses:

These non-credit courses are compulsory for the Non-IT field candidates.

1. Programming Fundamentals (Core Programming Course)
2. Data Structures & Algorithms OR Design & Analysis of Algorithms
3. Database Systems

Outline of the MS (DS) Program:

The program would be spread over 4 semesters, with a 6-credit hour thesis being offered in the second year.

Course Offering Plan:

Course Types	Cumulative Credits
Program Core Courses (3CH)	9
Specialization Requirement Courses (2CH)	6
Electives (3 CH)	9
Thesis (6 CH)	6
Total	30

Proposed Core Courses:

1. Statistical and mathematical models for data science. (3 CH)
2. Tools and Techniques for Data Science (2 CH +1 CH)
3. Machine Learning (3 CH)
4. Data Mining (2CH)
5. Research Methodology (2 CH)

Proposed Specialization Core Courses: (Choose any two)

1. Big Data Analytics (3CH)
2. Deep Learning (3 CH)
3. Natural Language Processing (3 CH)
4. Distributed Data Processing (3 CH)

Elective Courses: (Choose any three)

Following is a non-exhaustive list of elective courses. New elective courses may be added to this list. Students may be recommended to make their choice of electives, in the light of a soft specialization within the field of data science.

1. Advanced Computer Vision
2. Algorithmic trading
3. Bayesian Data Analysis
4. Bioinformatics
5. Cloud computing
6. Computational Genomics
7. Data Visualization
8. Deep Reinforcement Learning
9. Distributed Machine Learning in Apache Spark
10. High performance computing
11. Inference & Representation
12. Optimization Methods for Data Science and Machine Learning
13. Probabilistic Graphical Models
14. Scientific Computing in Finance
15. Social network analysis

Deficiency Courses for Non-It (Non-Credit)

Deficiency Courses for Non-IT		
S. No	Course Code	Course Title
1	MSDS-601	Programming Fundamentals
2	MSDS-604	Data Structures & Algorithms
3	MSDS-607	Database Systems

Semester-Wise Course Offering Plan:

S. No	Course Code	Course Title	C.H (TH)	C.H (PR)
First Semester				
1	MSDS-611	Data Science Tools & Techniques	2	1
2	MSDS-616	Statistical and Mathematical Models for Data Science	3	----
3		Elective-I	3	----
		Total Cr. Hrs for first semester		9
2nd Semester				
1	MSDS-631	Machine Learning	3	----
2	MSDS-636	Specialized Core: Big Data Analytics	3	----
3		Elective-II	3	----
		Total Cr. Hrs for 2 nd semester		9
3rd Semester				
1	MSDS-651	Specialized Core: Deep Learning	3	----
2		Elective-III	3	----
3	MSDS-655	Natural Language Processing	3	----
		Total Cr. Hrs for 3 rd semester		9

Mehran University of Engineering and Technology
Institute of Information and Communication Technologies
Department of Computer Systems Engineering

4th Semester				
1	MSDS-681	MS Thesis/MS Project	0	6

The course will be commenced from Batch 2023.

Mehran University of Engineering and Technology
 Institute of Information and Communication Technologies
 Department of Computer Systems Engineering

Title of Subject:	Statistics and Mathematical models for Data Science (MSDS-616)		
Disciplines:	MS in Data Science		
Pre-requisites:	None		
Assessment:	Sessional: 10%	Mid Semester: 30%	Final Exam: 60%
Term:	1 st		
Effective:	2023		
Marks:	100		
Credit Hours:	3+0	Min. Contact Hours	42 Hrs.

AIM:

Statistical and probabilistic methods are used for analysis of different datasets for forecasting the values, predicting the unknowns, relating the variables for getting deeper insights and relating data differences with real world complexities. Data Science extracts knowledge from data on the basis of hidden patterns which can be made explicit by incorporating the statistical algorithms in it. The student will not only gain an understanding of the theoretical foundation of statistical learning, but the practical skills necessary for their successful application to new problems in science and industry.

Objectives:

After completion of this course, the students should be able:

- To understand the significance and core concepts behind statistical and probabilistic methods and their implications
- To analyze, correlate, apply preprocessing on data by using different statistical and probabilistic techniques.
- To apply fundamental prediction, classification and forecasting algorithms in various applications.

Contents:

Probability: Probability basics (axioms of probability, conditional probability, random variables, expectation, independence, etc.), multivariate distributions, Maximum a posteriori and maximum likelihood estimation; Statistics: introduction to concentration bounds, laws of large numbers, central limit theorem, minimum mean-squared error estimation, confidence intervals; Linear algebra: Vector spaces, Projections (will also cover the least regression), linear transformations, singular value decomposition (this substitute for PCA), eigen decomposition, power method; Optimization: Matrix calculus with Lagrange Multipliers, gradient descent, coordinate descent, introduction to convex optimization.

Books Recommended:

Mehran University of Engineering and Technology
Institute of Information and Communication Technologies
Department of Computer Systems Engineering

- Michael Baron, Probability and Statistics for Computer Scientists, latest edition.
- David C. Lay and Steven R. Lay, Linear Algebra and Its Applications, latest edition.
- Gilbert Strang, Introduction to Linear Algebra, latest edition.
- David Forsyth, Probability for Computer Scientists, online latest edition.

Approval:

Board of Studies	Resolution No. 1.2	Dated: 04-03-2022
Board of Faculty of EEC		
Engg: Advanced Studies and	Resolution No.19.4	Dated: 03-06-2022
Research Board	Resolution No.184.08(a)	Dated: 06-07-2022
Academic Council	Resolution No. 104.10(ii)	Dated: 29-07-2022

Mehran University of Engineering and Technology
Institute of Information and Communication Technologies
Department of Computer Systems Engineering

Title of Subject:	Data Science Tools and Techniques (MSDS-611)		
Disciplines:	MS in Data Science		
Pre-requisites:	None		
Assessment:	Sessional: 10%	Mid Semester: 30%	Final Exam: 60%
Term:	1 st		
Effective:	2023		
Marks:	50		
Credit Hours:	2+1	Min. Contact Hours	28 Hrs.+42Hrs.

AIM:

The aim of this course is to: Introduce students to this rapidly growing field of Data Science and equip them with some of its basic principles and tools as well as its general mindset. Also, to explain the significance of exploratory data analysis in data science. Identify common approaches used for Feature Generation as well as Feature Selection, and finally discuss the Ethical and Privacy issues. Python, R and few modern tools and libraries will be preferred for the practical work of this course.

Objectives:

After completion of this course, the students should be able:

- To Understand the core concepts of Data science, Gain understanding for data transformation
- To Identify common approaches used for Feature Generation as well as Feature Selection, and extraction. Also, to generate meaningful representations of data.
- To develop a practical understanding of the skills, advance tools and algorithms necessary for the effective application of data science and Write codes for the algorithm implementation.

Contents:

Introduction to Data Science, Data Science Life cycle & Process (Asking Right Questions, Obtaining Data, Understanding Data, Building Predictive Models, Generating Visualizations) For Building Data Products, Introduction to Data (Types of Data and Datasets), Data Quality (Measurement and Data Collection Issues), Data pre-processing Stages (Aggregation, Sampling, Dimensionality Reduction, Feature subset selection, Feature creation etc.), Algebraic & Probabilistic View of Data, Introduction to Python Data Science Stack (Python, Numpy, Pandas, Matplotlib), Relational Algebra & SQL, Scraping & Data Wrangling (assessing, structuring, cleaning & munging of data), Basic Descriptive & Exploratory Data Analysis, Introduction to Text Analysis (Stemming, Lemmatization, Bag of Words, TF-IDF), Introduction to Prediction and Inference (Supervised & Unsupervised) Algorithms, Introduction to Scikit Learn, Bias-Variance Tradeoff, Model Evaluation & Performance Metrics (Accuracy, Contingency Matrix, Precision-Recall, F-1 Score, Lift, etc.), Introduction to Map-Reduce paradigm.

Note: Practical will be based upon theory.

Books Recommended:

Mehran University of Engineering and Technology
Institute of Information and Communication Technologies
Department of Computer Systems Engineering

- William McKinney, Python for Data Analysis, latest edition.
- G. James, D. Witten, T. Hastie and R. Tibshirani, An Introduction to Statistical Learning with Applications in R, latest edition.
- A. Adhikari and J. DeNero, Computational and Inferential Thinking: The Foundations of Data Science, latest edition.
- M. Zaki & W. Meira, Data Mining and Analysis: Fundamental Concepts and Algorithms, latest edition.
- Joel Grus, Data Science from Scratch, latest edition.
- Cathy O'Neil and Rachel Schutt, Doing Data Science, latest edition.
- Laura Igual, Introduction to Data Science. A Python Approach to Concepts, Techniques and Applications, latest edition.

Approval:

Board of Studies	Resolution No. 1.2	Dated: 04-03-2022
Board of Faculty of EEC		
Engg. Advanced Studies and	Resolution No. 19.4	Dated: 03-06-2022
Research Board	Resolution No. 184.08(a)	Dated: 06-07-2022
Academic Council	Resolution No. 104.10(ii)	Dated: 29-07-2022

Mehran University of Engineering and Technology
Institute of Information and Communication Technologies
Department of Computer Systems Engineering

Title of Subject:	Data Visualization (MSDS-621)		
Disciplines:	MS in Data Science		
Pre-requisites:	Database Systems		
Assessment:	Sessional: 10%	Mid Semester: 30%	Final Exam: 60%
Term:	1 st		
Effective:	2023		
Marks:	50		
Credit Hours:	3+0	Min. Contact Hours	28 Hrs.

AIM:

Data visualization is an interdisciplinary field that deals with the graphic representation of data. In this course we will study techniques and algorithms for creating effective visualizations based on principles from graphic design, visual art, perceptual psychology, and cognitive science.

Objectives:

After completion of this course, the students should be able:

- To explore how to design and create data visualizations based on data available and tasks to be achieved.
- To evaluate the effectiveness of visualization designs, and think critically about each design decision, such as choice of color and choice of visual encoding.

Contents:

Introduction (Ugly, Bad, and Wrong Figures), Visualizing Data: Mapping Data onto Aesthetics. Coordinate Systems and Axes. Color Scales.
Directory of Visualizations (Amounts, Distributions, Proportions, x–y relationships, Geospatial Data, Uncertainty, Visualizing Amounts(Bar Plots, Grouped and Stacked Bars, Dot Plots and Heatmaps)Visualizing Distributions: (Histograms and Density Plots, Empirical Cumulative Distribution Functions and Q-Q Plots, Distribution Functions, Highly Skewed DistributionsVisualizing Many Distributions at Once (Visualizing DistributionsAlong the Vertical Axis, Visualizing Distributions Along the Horizontal Axis)
Visualizing Proportions. Visualizing Time Series and Other Functions of an IndependentVisualizing Trends. Visualizing Geospatial Data (Projections, Layers, Choropleth Mapping, Cartograms. Visualizing Uncertainty.

Books Recommended:

Mehran University of Engineering and Technology
Institute of Information and Communication Technologies
Department of Computer Systems Engineering

- Roger D. Peng, “Exploratory Data Analysis with R”, latest edition.
- Claus O. Wilke, Fundamentals of Data Visualization, latest edition, O'Reilly Media, Inc.
- Laura Igual, Introduction to Data Science. A Python Approach to Concepts, Techniques and Applications, latest edition.

Approval:

Board of Studies	Resolution No. 1.2	Dated: 04-03-2022
Board of Faculty of EEC		
Engg. Advanced Studies and	Resolution No. 19.4	Dated: 03-06-2022
Research Board	Resolution No. 184.08(a)	Dated: 06-07-2022
Academic Council	Resolution No. 104.10(ii)	Dated: 29-07-2022

Mehran University of Engineering and Technology
Institute of Information and Communication Technologies
Department of Computer Systems Engineering

Title of Subject:	Machine Learning (MSDS-631)		
Disciplines:	MS in Data Science		
Pre-requisites:	None		
Assessment:	Sessional: 10%	Mid Semester: 30%	Final Exam: 60%
Term:	2 nd		
Effective:	2023		
Marks:	100		
Credit Hours:	3+0	Min. Contact Hours	42 Hrs.

AIM:

Today, machine learning is one of the most active areas of engineering and is enjoying unprecedented levels of success.

The primary aim of the course is to enable the student to think coherently and confidently about machine learning problems and present the student with a set of practical tools that can be applied to solve real-world problems in machine learning, coupled with an appropriate, principled approach to formulating a solution.

Objectives:

After successful completion of this unit, students will be able to:

- Understand, describe, and critique advanced machine learning techniques.
- Identify and select suitable modelling, learning and prediction techniques to solve a complex data problem.
- Design and implement a refined machine learning solution; and appraise ethical and privacy issues of artificial intelligence techniques.

Contents:

Introduction to machine learning and statistical pattern recognition. Supervised learning:

Part I (Graphical models (full Bayes, Naïve Bayes), Decision trees for classification & regression for both categorical & numerical data, Ensemble methods, Random forests, Boosting (Adaboost and Xgboost), Stacking.

Part II (Four Components of Machine Learning Algorithm (Hypothesis, Loss Functions, Derivatives and Optimization Algorithms), Gradient Descent, Stochastic Gradient Descent, Linear Regression, Nonlinear Regression, Perceptron, Support vector machines, Kernel Methods, Logistic Regression, Softmax, Neural networks); Unsupervised learning: K-means, Density Based Clustering Methods (DBSCAN, etc.), Gaussian mixture models, EM algorithm, etc.; Reinforcement learning; Tuning model complexity; Bias-Variance Tradeoff; Grid Search, Random Search; Evaluation Metrics; Reporting predictive performance

Books Recommended:

Mehran University of Engineering and Technology
Institute of Information and Communication Technologies
Department of Computer Systems Engineering

- C.M. Bishop, Pattern Recognition and Machine Learning Springer, latest edition.
- Andreas C. Müller and Sarah Guido, Introduction to Machine Learning with Python, A Guide for Data Scientists, O'Reilly, latest edition.
- Kevin R Murphy, Machine Learning: A Probabilistic Perspective, latest edition.
- David Forsyth, Applied Machine Learning, online latest edition,
<http://luthuli.cs.uiuc.edu/~daf/courses/LearningCourse17/learning-book-6-April-nnrevision.pdf>

Approval:

Board of Studies	Resolution No. 1.2	Dated: 04-03-2022
Board of Faculty of EEC		
Engg: Advanced Studies and	Resolution No. 19.4	Dated: 03-06-2022
Research Board	Resolution No. 184.08(a)	Dated: 06-07-2022
Academic Council	Resolution No. 104.10(ii)	Dated: 29-07-2022

Mehran University of Engineering and Technology
Institute of Information and Communication Technologies
Department of Computer Systems Engineering

Title of Subject:	Big Data Analytics (MSDS-636)		
Disciplines:	MS in Data Science		
Pre-requisites:	Database Systems		
Assessment:	Sessional: 10%	Mid Semester: 30%	Final Exam: 60%
Term:	2 nd		
Effective:	2023		
Marks:	100		
Credit Hours:	3+0	Min. Contact Hours	42 Hrs.

AIM:

Advances in our capability to generate and collect information coupled with decreasing disk space prices are pushing us towards a world centered around data management. Databases are at the heart of modern commercial application development. Their use extends beyond this to many other environments and domains where large amounts of data must be stored for efficient update and retrieval. The purpose of this course is to provide a comprehensive introduction to the use of database management systems for applications.

Objectives:

After completion of this course, the students should be able to:

- Understand the concepts underlying databases design.
- Analyze problems to identify data requirements, types and relations.
- Create Entity-Relationship and relational designs from problem statements.
- Design data structures and functions to store and process the information.
- Design and implement databases using SQL language, manipulate and interact SQL database.
- Describe the Big Data landscape and explain the V's of Big
- Get value out of Big Data by using a 5-step process to structure your analysis.
- Provide an explanation of the architectural components and programming models used for scalable big data analysis.

Contents:

Introduction to Big Data Analytics, Big Data Platforms, Data Store & Processing using Hadoop, Big Data Storage and Analytics, Big Data Analytics ML Algorithms, Recommendation, Clustering, and Classification, Linked Big Data: Graph Computing and Graph Analytics, Graphical Models and Bayesian Networks, Big Data Visualization, Cognitive Mobile Analytics.

Books Recommended:

- Elmasri&Navathe,Fundamentals of Database Systems, latest edition, Addison-Wesley.
 - R. Ramakrishan,Database Management Systems, latest edition, McGraw-Hill.
-

Mehran University of Engineering and Technology
Institute of Information and Communication Technologies
Department of Computer Systems Engineering

- D. Maier, The Theory of Relational Databases, latest edition, Computer Science Press.

Approval:

Board of Studies	Resolution No. 1.2	Dated: 04-03-2022
Board of Faculty of EEC		
Engg: Advanced Studies and	Resolution No. 19.4	Dated: 03-06-2022
Research Board	Resolution No. 184.08(a)	Dated: 06-07-2022
Academic Council	Resolution No. 104.10(ii)	Dated: 29-07-2022

Mehran University of Engineering and Technology
Institute of Information and Communication Technologies
Department of Computer Systems Engineering

Title of Subject:	Natural Language Processing (MSDS-655)		
Disciplines:	MS in Data Science		
Pre-requisites:	None		
Assessment:	Sessional: 10%	Mid Semester: 30%	Final Exam: 60%
Term:	2 nd		
Effective:	2023		
Marks:	100		
Credit Hours:	3+0	Min. Contact Hours	42 Hrs.

AIM:

Students who complete this course will gain a foundational understanding in natural language processing methods and strategies. They will also learn how to evaluate the strengths and weaknesses of various NLP technologies and frameworks as they gain practical experience in the NLP toolkits available. Students will also learn how to employ literary-historical NLP-based analytic techniques like stylometry, topic modeling, synsetting and named entity recognition in their personal research.

Objectives:

After completion of this course, the students should be able

- Identify techniques for information retrieval, language translation, and text classification.
- List the advantages of using standard corpora. Identify examples of current corpora for a variety of NLP tasks.
- Define and contrast deterministic and stochastic grammars, providing examples to show the adequacy of each. 4. Simulate, apply, or implement classic and stochastic algorithms for parsing natural language.

Contents:

Deterministic and stochastic grammars, Parsing algorithms, CFGs, Representing meaning / Semantics, Semantic roles, Temporal representations, Corpus-based methods, N-grams and HMMs, Smoothing and backoff, POS tagging and morphology, Dependency parsing, Information retrieval, Vector space model, Precision and recall, Information extraction, Language translation, Text classification (logreg, MLP and convolutional neural nets), categorization, Bag of words model, Vector semantics and static word embeddings, Machine translation, Social NLP.

Books Recommended:

- Python Machine Learning, Sebastian Raschka. Publisher: Packt Publishing, 2015.
- Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit Latest Edition, Steven Bird, Ewan Klein and Edward Loper

Mehran University of Engineering and Technology
Institute of Information and Communication Technologies
Department of Computer Systems Engineering

Publisher: O'Reilly Media, 2009.

- Speech and Language Processing, Latest Edition, Daniel Jurafsky and James H. Martin Publisher: Prentice Hall, 2000.

Approval:

Board of Studies	Resolution No. 1.2	Dated: 04-03-2022
Board of Faculty of EEC		
Engg. Advanced Studies and	Resolution No. 19.4	Dated: 03-06-2022
Research Board	Resolution No. 184.08(a)	Dated: 06-07-2022
Academic Council	Resolution No. 104.10(ii)	Dated: 29-07-2022

Mehran University of Engineering and Technology
Institute of Information and Communication Technologies
Department of Computer Systems Engineering

Title of Subject:	Deep Learning (MSDS- 651)		
Disciplines:	MS in Data Science		
Pre-requisites:	None		
Assessment:	Sessional: 10%	Mid Semester: 30%	Final Exam: 60%
Term:	3 rd		
Effective:	2023		
Marks:	100		
Credit Hours:	3+0	Min. Contact Hours	42 Hrs.

AIM:

Deep learning is removing the responsibility of humans to design features, instead Deep Learning is given a task to find the appropriate representation.

Objectives:

After completion of this course, the students should be able to:

- Understand key concepts related to Deep Learning. Apply DNN to real-life problems. Understand how to frame problems in the NN framework. Understand DNN architecture and parameter.

Contents:

Introduction to Deep learning, Review of Linear classification (Multi-class Support Vector Machines, Soft max) and Regularization, Gradient Descent & Stochastic Gradient Descent (SGD), Back propagation (Intuitions, back propagation as flow graph).

Introduction to Neural Networks (model of a biological neuron, activation functions, neural net architecture, representational power, etc.), Building Neural Networks (data pre-processing, loss functions, weight initialization, regularization, dropout, batch normalization), Learning Neural Networks (Learning and Evaluation gradient checks, sanity checks), Variants of SGD (momentum, Adagrad/RMSprop, ADAM).

Introduction to Convolutional Neural Networks (CNN) and its components (Convolution and Pooling Layers), Convolutional Neural Network case studies (AlexNet/ZFNet/VGGNet), Understanding and Visualizing Convolutional Neural Networks, Convolutional networks for other visual Recognition Tasks (Localization, Detection, Segmentation, etc.), Transfer Learning and Fine-tuning Convolutional Neural Networks.

Introduction to Natural Language Processing (NLP), Learning word and sentences embedding (wordvec, glove, sentvec),

Introduction to recurrent networks (RNNs, LSTMS, etc.), Applications of Recurrent neural networks to different NLP tasks (e.g. sentiment analysis, parsing, NER tagging, etc.), Introduction to Reinforcement Learning and QLearning, Deep Q-Networks (DQN) and Game playing using DQN, Introduction to Policy gradients and their applications.

Books Recommended:

- YoshuaBengio, Ian Goodfellow, Aaron Courville, Deep Learning, latest edition.
- Michael A. Nielsen, Deep learning, latest edition.
- AurélienGéron, Hands On Machine Learning with Scikit Learn and Tensor Flow, latest

Mehran University of Engineering and Technology
Institute of Information and Communication Technologies
Department of Computer Systems Engineering

edition.

Approval:

Board of Studies	Resolution No. 1.2	Dated: 04-03-2022
Board of Faculty of EEC		
Engg: Advanced Studies and	Resolution No. 19.4	Dated: 03-06-2022
Research Board	Resolution No. 184.08(a)	Dated: 06-07-2022
Academic Council	Resolution No. 104.10(ii)	Dated: 29-07-2022

Mehran University of Engineering and Technology
 Institute of Information and Communication Technologies
 Department of Computer Systems Engineering

Title of Subject:	Cloud Computing (MSDS-656)		
Disciplines:	MS in Data Science		
Pre-requisites:	None		
Assessment:	Sessional: 10%	Mid Semester: 30%	Final Exam: 60%
Term:	3 rd		
Effective:	2023		
Marks:	100		
Credit Hours:	3+0	Min. Contact Hours	42 Hrs.

AIM:

Advances in our capability to generate and collect information coupled with decreasing disk space prices are pushing us towards a world centered around data management. Databases are at the heart of modern commercial application development. Their use extends beyond this to many other environments and domains where large amounts of data must be stored for efficient update and retrieval. The purpose of this course is to provide a comprehensive introduction to the use of database management systems for applications.

Objectives:

After completion of this course, the students should be able to:

- Explain the core concepts of the cloud computing paradigm: how and why this paradigm shift came about, the characteristics, advantages and challenges brought about by the various models and services in cloud computing.
- Apply fundamental concepts in cloud infrastructures to understand the tradeoffs in power, efficiency, and cost, and then study how to leverage and manage single and multiple datacenters to build and deploy cloud applications that are resilient, elastic and cost-efficient.
- Discuss system, network and storage virtualization and outline their role in enabling the cloud computing system model.
- Illustrate the fundamental concepts of cloud storage and demonstrate their use in storage systems such as Amazon S3 and HDFS.
- Analyze various cloud programming models and apply them to solve problems on the cloud.

Contents:

Definition and evolution of Cloud Computing, Enabling Technologies, Service and Deployment Models, Popular Cloud Stacks and Use Cases, Benefits, Risks, and Challenges of Cloud Computing, Economic Models and SLAs, Topics in Cloud Security.

Historical Perspective of Data Centers, Datacenter Components: IT Equipment and Facilities, Design Considerations: Requirements, Power, Efficiency, & Redundancy, Power Calculations, PUE and Challenges in Cloud Data Centers, Cloud Management and CloudSoftware Deployment ConsiderationsVirtualization (CPU, Memory, I/O), Case Study: Amazon EC2Software Defined Networks (SDN), Software Defined Storage (SDS).

Introduction to Storage Systems, Cloud Storage Concepts, Distributed File Systems (HDFS, Ceph FS), Cloud Databases (HBase, MongoDB, Cassandra, DynamoDB), Cloud Object Storage (Amazon S3, OpenStack Swift, Ceph).

Distributed Programming for the Cloud, Data-Parallel Analytics with Hadoop MapReduce (YARN), Iterative Data-Parallel Analytics with Apache Spark, Graph-Parallel Analytics with GraphLab 2.0 (PowerGraph).

Books Recommended:

- Ricardo Puttini, Thomas Erl, and Zaigham Mahmood, Cloud Computing: Concepts, Technology & Architecture, latest edition, The Prentice Hall Service Technology Series from Thomas Erl, ISBN-13: 978-0133387520, ISBN-10: 013338752.
- Kris Jamsa, Cloud Computing, latest edition, Jones & Bartlett Publishers, ISBN-10: 9380853777, ISBN-13: 978-9380853772.
- Fang Liu, Jin Tong, Jian Mao, Robert B. Bohn, John V. Messina, Mark L. Badger, Dawn M. Leaf, NIST Cloud Computing Reference Architecture, Special Publication NIST SP, ISBN: 500-292.

Approval:

Board of Studies	Resolution No. 1.2	Dated: 04-03-2022
Board of Faculty of EEC	Resolution No. 19.4	Dated: 03-06-2022
Engg: Advanced Studies and	Resolution No. 184.08(a)	Dated: 06-07-2022
Research Board	Resolution No. 104.10(ii)	Dated: 29-07-2022
Academic Council		
