Minutes of the 16th Meeting of

Faculty Board of Studies (FBoS) Engineering Sciences

Contents

Minutes of the 16th Meeting of Faculty Board of Studies Engineering Sciences held on 30th May 2018 by VLC

Attendance:

BUIC

1. Prof. Dr. Muhammad Najam ul Islam	Dean ES	Chair
2. Prof. Dr. Tahseen Ullah Khan	HOD(E&ES)	Member
3. Associate Prof. Dr. Atif Raza Jafri	HOD(EE)	Member
4. Associate Prof. Dr. Faisal Bashir	HOD(CS)	Member
5. Associate Prof. Dr. Awais Majeed	HOD(SE)	Member
6. Associate Prof. Dr. Amina Jameel	HOD(CE)	Member

BUKC

7. Prof. Dr. Haroon Rasheed	HOD(EE)	Member
8. Associate Prof. Dr. Humera Farooq	HOD(CS)	Member
9. Associate Prof. Dr. Sohaib Ahmed	HOD(SE)	Member
10. Prof. Dr. Nargis Yasmeen	HOD(EES)	Member
11. Dr. Rizwan Iqbal	HOD(CE)	Member

BULC

12. Asstt Prof. Mr Farhan Saeed Sherazi	HOD(CS&IT)	Member

Proceedings

Preliminaries

1. The 16th Faculty Board of Studies, Engineering Sciences meeting took place on 30th May, 2018, with the quorum complete; the proceedings commenced at 12:00 hrs, with recitation from the Holy Quran.
2. The special FBoS meeting was called by the chair to deliberate on the reviews received from HEC on MS Data Science program. MS Data Science program was approved in 30th ACM held on 3rd October 2017 and case for obtaining NOC was forwarded to HEC. On 26th April 2018 reviewer comments were received and Dean (ES) on 2nd May 2018 constituted a committee comprising of following members to review the observations of the HEC on the proposed curriculum of MS Data Science. The committee presented their suggestions in Computer Sciences Departmental Board of Studies and approved minutes of DBoS were presented in 16th FBoS-ES.

    a. Assoc. Prof. Dr Faisal Bashir — HoD CS – Chair
    b. Prof. Dr. Muhammad Ramzan — Member
    c. Assoc. Prof.Dr. Imran Siddiqi — Member
    d. Asst. Prof. Dr. Asfand Yar — Member
    e. Asst. Prof. Dr. Awais Ahmad — Member

New Items:

Item 1601: Revision of MS Data Science Eligibility Criteria

Sponsor: HOD CS BUIC

Referral Authority: DBOS CS BUIC

Summary of the Case

The house was informed that HEC's eligibility criteria for MS Data Science recommends that Mathematicians and Statisticians are eligible for admission in MS Data Science. Basic computing knowledge is essential for Data Science students therefore as per HEC's eligibility criteria deficiency courses can be taught to students having non computing background. The following revised eligibility criteria was presented by the committee.

HEC recognized 4 years Bachelor degree in CS/SE/CE/EE/IT/Statistics/Mathematics or equivalent with CGPA 2.5/4.0 (Semester System) or 50% marks (Annual System). NTS-GAT (General)/ GRE/ University entry test passed with 50% marks.

The following core courses are recommended to be completed before entering the MS Data Science program.

a. Programming Fundamentals
b. Data Structures & Algorithms OR Design & Analysis of Algorithms
c. Database Systems

A student selected for admission having deficiency in the above stated courses shall be required to study the courses, which must be passed in the first two semesters. Deficiency courses shall be determined by the Departmental Selection Committee, before admitting the student. A student cannot register in MS courses, unless all specified deficiency courses have been passed.

Discussion:

The procedure for offering deficiency courses was discussed and it was agreed by the house that deficiency courses are to be enrolled with BS students and fee structure (per credit hour fee) of BS program shall be applicable to MS students in deficiency courses.

Decision 1601

The revised eligibility criteria for MS Data Science program is approved.

Item 1602: Revision of Course Contents and Course Outlines - MS Data Science

Sponsor: HOD CS BUIC Referral Authority: DBOS CS BUIC

Summary of the Case

Dr Faisal Bashir (HoD CS) informed that the course outlines and contents have been thoroughly reviewed. The core courses and their contents are now in accordance with the recommendations of the HEC. Following core courses are offered in MS Data Science program

a. Statistical and Mathematical Methods for Data Science

b. Tools and Techniques in Data Science

c. Machine Learning

The aforementioned courses not only address the concerns of the reviewer (both Mathematical and Statistical aspects of Data Science are covered) but also ensures that the curriculum is aligned with the latest guidelines of the HEC.

Dr Faisal Bashir presented the revised course titles, outlines and course contents that were approved by DBoS and are placed on Annex MSDS – Roadmap.

Decision 1602

The agenda item is approved, revised roadmaps and course outlines attached at **Annex MSDS - Roadmap**.

Item 1603: Revision of Elective Courses - MS Data Science

Sponsor: HOD CS BUIC Referral Authority: DBOS CS BUIC

Summary of the Case

The elective courses of MS Data Science program were presented and few courses that were not directly related with the program have been removed.

Decision 1603

The agenda item is approved, the revised list of electives for MS Data Science are placed on

Item 1604: Research Thesis - MS Data Science

Sponsor: HOD CS BUIC Referral Authority: DBOS CS BUIC

Annex MSDS - Electives and forwarded to ACM for consideration.

Summary of the Case

HoD CS highlighted the HEC's reviewer concern over inclusion of statistical aspects in the MS Thesis of Data Science Program. It was deliberated that research thesis at MS level include detailed design aspects were required and special emphasis is given to mathematical modelling of research problems.

Decision 1604

The agenda item is approved, mathematical and statistical aspects shall be covered in the Thesis of MS Data Science program where applicable.

Closing the Meeting

There being no further points, the Chair brought the meeting to close at about 14:00 hrs, thanking the participants for their wholehearted participation.

Prof. Dr. M. Najam-ul-Islam

Dean (ES), Head FBoS

04 June, 2018

Distribution:

BUHQ:	Rector, Pro-Rector, Registrar
	DAA
BUIC:	DG BUIC, DIC
	HOD(EES), HOD(EE), HOD(CS), HOD(SE), HOD(CE)
BUKC:	DG BUKC, DKC
	HOD(EES), HOD(EE), HOD(CS), HOD(SE), HOD(CE)
BULC:	DLC,
	HOD(CS)

Annex MSDS – Roadmap

Semester-wise Breakdown: MS Data Science

Semester 1

Course Code	Course Title	Credit Hours
DSC-500	Tools and Techniques in Data Science	03 (2+1)
DSC-501	Statistical and Mathematical Methods for Data Analysis	03
ESC-701	Research Methodology	03
	Total	09

Semester 2

Course Code	Course Title	Credit Hours
DSC-700	Big Data Analytics	03
CSC-719	Machine Learning	03
	Elective-I	03
	Total	09

Semester 3

Course Code	Course Title	Credit Hours
DSC-704	Deep Learning	03
	Elective-II	03
	Thesis-I / Elective – III	03
	Total	09

Semester 4

Course Code	Course Title	Credit Hours

	Thesis-II / Elective – IV	03
	Total	03
	TOTAL CREDIT HOURS	**30**

Elective Courses – MS Data Science			
Sr. No	**Course Code**	**Course Title**	**Credit Hours**
5	DSC-702	Data Visualization	3
6	DSC-703	Distributed Data Engineering	3
2	DSC-705	Unstructured Data Processing	3
3	CSC-518	Decision Support Systems	3
4	CSC-715	Intelligent Agents	3
5	CSC-741	Advanced Natural Language Processing	3
6	CEN-745	Advanced Digital Image Processing	3
7	CSC-749	Advanced Neural Networks and Fuzzy Logic	3
8	CSC-751	Pattern Recognition	3
9	CSC-764	Computer Vision	3
10	CSC-781	Cloud Computing	3
11	CSC-554	Advanced Information Theory	3
12	CSC-747	Text Mining	3
13	CSC-752	Advanced DBMS	3
14	CSC-760	Advanced Data Warehousing	3
15	SEN-764	Ontology Engineering	3
16	ESC-500	Thesis	6

MS Data Science Curriculum

Course Title: Tools and Techniques in Data Science

Course Code: DSC-500

Pre-Requisite: None

Objectives:

This course is aimed at introducing the students to the foundations of data science with hands-on exercises using the latest data science tools.

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

Text Books:

- Python for Data Analysis, 1st Edition, William McKinney, 2012
- Data Science from Scratch, 1st Edition, Joel Grus, 2015

Reference Books:

- An Introduction to Statistical Learning with Applications in R, 1st Edition, G. James, D. Witten, T. Hastie and R. Tibshirani, 2013
- Computational and Inferential Thinking: The Foundations of Data Science, 1st Edition, A. Adhikari and J. DeNero, 2017
- Doing Data Science, 1st Edition, Cathy O'Neil and Rachel Schutt, 2013
- Introduction to Data Science. A Python Approach to Concepts, Techniques and Applications, 1st Edition, Laura Igual, 2017

Course Title: Statistical and Mathematical Methods for Data Analysis

Course Code: DSC-501

Pre-Requisite: None

Objectives:

This course is designed to teach learners the basic math you will need in order to be successful in almost any data science stat and math course and was created for learners who have basic math skills but may not have taken Stat, algebra or pre-calculus. Also, it brings students up to speed with mathematical and statistical concepts from discrete mathematics, calculus and elementary linear algebra - all with a view of data science, statistics and machine learning applications that follow. Data Science Math and Stat skills introduces the core math that data science is built upon, with no extra complexity, introducing unfamiliar ideas and math symbols one-at-a-time.

Contents:

Students learn the skills of set theory, including Venn diagrams, properties of the real number line, interval notation and algebra with inequalities, uses for summation and sigma notation, math on the cartesian (x,y) plane, slope and distance formulas, graphing and describing functions and their inverses on the x-y plane, probability distributions (How to judge the probability of an event, based on certain conditions), statistical significance, hypothesis testing, and regression, Basics of Linear Regression . The concept of instantaneous rate of change and tangent lines to a curve, Exponents, logarithms, and the natural log function.

Text Books:

- Mathematical Problems in Data Science: Theoretical and Practical Methods, Book by Bo Jiang, Li Chen, and Zhixun Su, December 15, 2015, Springer

Reference Books

- Digital and Discrete Geometry: Theory and Algorithms, Book by Li. M. Chen. ISBN 978-3-319-12098-0, 2014, Springer
- Convexity and Discrete Geometry Including Graph Theory, Book by Karim Adiprasito, Imre Barany, Costin Vilcu, ISSN 2194-1009, Springer, 2016

Course Title: Research Methodology

Course Code: ESC-701

Pre-Requisite: None

Objectives:

This course is aimed at providing the students with an ability to undertake postgraduate level research and an appreciation of relevant ethical and professional issues. After completing this course, students will be able to: Formulate research questions and carry out research investigations, identify various sources of information and critically analyze the collected information, Identify and apply appropriate research methods in order to plan, conduct and evaluate their research, Effectively report/publish the results of research activities and Develop and deliver presentations to disseminate research findings.

Contents:

Introduction to research, Qualitative and Quantitative research, The scientific method of research, Choosing a research problem, Choosing a research advisor, Literature Review – Conducting and writing, Formulating the research question, Identifying variables and generating hypothesis, Research Design/Methodology, Information gathering and data collection, Data representation, analysis and interpretation, Writing a research proposal, Ethics of research – Plagiarism and Intellectual property rights, Organizing and managing conferences and workshops, Writing research papers/Reviewing research papers, Planning and delivering scientific presentations, Writing thesis/dissertations.

Text Books:

- How to Research, L. Blaxter, C. Hughes, M. Tight, 4th Edition, 2010.

Reference Books:

- Research Methodologies – A step by step guide for beginners, Ranjit Kumar, 2005.
- Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, John W. Creswell, 2008.

Course Title: Big Data Analytics

Course Code: DSC-700

Pre-Requisite: None

Objectives:

Analyzing data is a challenging task as it has to be figured out which type of data analytics are to be used, as well as defeat the challenges that come up when it comes to analyzing data. This course is focused on learning analytics for big data using real-world scenarios.

Contents:

This course covers advanced topics in big data analytics including association analysis, nearest neighbor search in high dimensional data, link analysis, page rank, dimensionality reduction, mining stream data and working with very large graphs. The risks of data analytics, the types of data analytics that are out there in the world, the benefits of using data analytics and also the real-world examples that show how to take this knowledge and apply it to everyday life. Big data modeling and big data management systems also make part of the contents.

Text Books:

- Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting by EMC Education Services (Editor), Data 1st Edition, 2015. ISBN-13: 978-1118876138

- Big Data Science & Analytics: A Hands-On Approach Paperback by Arshdeep Bahga, Vijay Madisetti, Published by Arshdeep Bahga, 1st Edition, 2016. ISBN: 978-0996025539

Reference Books:

- Data Analytics: Become a Master in Data Analytics by Richard Dorsey , Publisher: Eric Morrison, 2nd Edition, 2016.

- Big Data Analytics by Venkat Ankam, Published by Venkat Ankam, 1st Edition, 2016.

Course Title: Machine Learning

Course Code: CSC-719

Pre-Requisite: None

Objectives:

This course is an overview of concepts and techniques in machine learning, beginning with topics such as classification and linear regression and ending up with more recent topics such as boosting, support vector machines, hidden Markov models, and Bayesian networks. The course will give the student the basic ideas behind modern machine learning methods.

Contents:

Introduction to Machine Learning, Concept learning, Decision tree learning, Linear models for regression, Linear models for classification, Artificial neural networks, Kernel methods, Sparse kernel machines, Mixture models and the EM algorithm, Evaluation, Combining multiple learners, Support vector machines, Bayesian networks.

Text Books:

- Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies by John D. Kelleher, Brian Mac Namee, Aoife D'Arcy, MIT Press, 1st Edition, 2017. ISBN-13: 978-0262029445

- Machine Learning: Fundamental Algorithms for Supervised and Unsupervised Learning With Real-World Applications by Joshua Chapmann (Author, Publisher), 2nd Edition, 2017.

Reference Books:

- Machine Learning with R - Second Edition by Brett Lantz, Packt Publishing, Second Edition, 2015. ISBN: 978-1-78439-390-8
- Machine Learning with Python: Understanding Machine Learning with Python in the World of Data Science by Robert Wilson (Author, Publisher), 1st Edition, 2016.

Course Ti: Deep Learning

Course Code: DSC-704

Pre-Requisite: Machine Learning

Objectives:

The objective of this course is to acquaint the students with the state-of-the-art deep learning techniques to solve different learning problems. Students will learn to design as well as implement deep neural network architectures (through hands-on tasks) to solve various recognition problems.

Contents:

Introduction to neural networks, activation functions and back-propagation; Convolutional Neural Networks: History, Convolution, Pooling, CNNs for classification, Deep learning Software, CNN Architectures; Sequence Modeling: Recurrent and Recursive Nets: Long-Short Term Memory models and variants, Language modeling and image captioning, Unsupervised learning: Restricted Boltzmann Machines and Auto-encoders; Case Studies.

Books:

Text Book:

- Deep Learning (Adaptive Computation and Machine Learning series), Ian Goodfellow, Yoshua Bengio, Aaron Courville, The MIT Press, 2016.

Reference Books:

- Deep Learning: A Practitioner's Approach, Josh Patterson and Adam Gibson, O'Reilly Media, 2017.
- Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms, Nikhil Buduma and Nicholas Locascio, O'Reilly Media, 2017
- Deep Learning with Python, Francois Chollet, O'Reilly Media, 2017.

Course Title: Data Visualization

Course Code: DSC-702

Pre-Requisite: None

Objectives:

The course is about effective data analysis involves learning how to synthesize data, especially big data, into a story and present that story in a way that resonates with the audience. The course is about how to analyze large amounts of data, communicate complex data in a meaningful way, and quickly slice data into various views. Also, explain how to automate redundant reporting and analyses, create eye-catching visualizations, and use statistical graphics and thematic cartography.

Contents:

Students learn the skills of data visualization in parallel with the soft skills of communicating with a non-technical audience and core Data Science leadership skills. Emphasis is placed on enabling students to listen to articulated business needs or problem cases and learn how to propose as well as execute Data Science solutions to effectively meet these needs.

Text Books:

- Data Visualisation: A Handbook for Data Driven Design by Andy Kirk (Author), SAGE Publishing, 1st Edition, 2017. ISBN: 978-14739-1213-7

- Effective Data Visualization: The Right Chart for the Right Data by Stephanie D. H. Evergreen (Author), SAGE Publishing, 1st Edition, 2016. ISBN: 978-1506303055

Reference Books:

- Data Visualization with Python and JavaScript: Scrape, Clean, Explore & Transform Your Data by Kyran Dale (Author), 2nd Edition, O'Reilly Media, 2017. ISBN: 978-1491920510

- Machine Learning for Absolute Beginners: A Plain English Introduction by Oliver Theobald (Author, Publisher), 1st Edition, 2017.

Course Title: Distributed Data Engineering

Course Code: DSC-703

Pre-Requisite: None

Objectives:

The course is about well-designed conceptual and logical data models that the design has been built with flexibility and extensibility leading to high application agility and low maintenance costs. A detailed data flow diagrams means a concrete understanding of the business' value chain exists and is documented. The wish to understand how we think means excellent team dynamics while analyzing, designing, and building the application.

Contents:

This course will introduce students to working with distributed systems for efficiently collecting and analyzing large quantities of varied data. This is a survey-style course covering common data platforms and analysis patterns including Postgres (SQL), Hadoop (MapReduce), Spark, Kafka (logs), Lambda Architecture (streaming), and Cassandra (NoSQL).

Text Books:

- Data Engineering Perfect by Brian Shive, Technics Publications, LLC, 2nd Edition, 2017. ISBN: 978-1935504603

- Measurement and Data Analysis for Engineering and Science by Patrick F. Dunn (Author), CRC Press, 3rd Edition, 2016. ISBN: 978-1439825686

Reference Books:

- Data Visualization with Python and JavaScript: Scrape, Clean, Explore & Transform Your Data by Kyran Dale (Author), 2nd Edition, O'Reilly Media, 2017. ISBN: 978-1491920510

- Machine Learning with Python: Understanding Machine Learning with Python in the World of Data Science by Robert Wilson (Author, Publisher), 1st Edition, 2016.

Course Title: Unstructured Data Processing

Course Code: DSC-705

Pre-Requisite: None

Objectives:

This course is aimed at extracting useful information (usually) from huge unstructured datasets by employing techniques from information retrieval, natural language processing and data mining. The objective of this module is to get a good understanding of the basic text mining techniques and study some of its applications as well.

Contents:

Essential Data Science skills involved in working with unstructured data include transforming it into structured data types able to be analysed, processed and used for Machine Learning and Information Retrieval algorithms. The focus is on Natural Language Processing and classification techniques used in Text Mining. Dealing with information overload and information overlook, unstructured vs. (semi-) structured data, evolving information needs and knowledge management issues, the business case for text mining. The text mining pipeline: information retrieval, information extraction and data mining. Fundamentals of natural language processing: linguistic foundations, levels of linguistic analysis.

Text Books:

- Text mining handbook: advanced approaches in analyzing unstructured data, Feldman, Ronen and James Sanger, Cambridge University Press, Edition: 3rd ,2015
- Text mining: classification, clustering and applications, Srivastava, Ashok and Mehran Sahami, Chapman & Hall, Edition: 1st , 2009

Reference Books:

- Mining Text Data by Charu C. Aggarwal, ChengXiang Zhai – 2012
- Machine Learning for Absolute Beginners: A Plain English Introduction by Oliver Theobald (Author, Publisher), 2017

Course Title: Decision Support Systems

Course Code: CSC - 518

Pre-Requisite: None

Objectives:

This course should enable a student to understand managerial decisions, to participate in the decision making process, and to be able to develop models and systems to support the decision making. This course focuses on the use and application of information systems to support the decision-making process. Different types of systems are discussed as a basis for designing and developing highly effective decision support systems. Data models, interactive processes, knowledge-based approaches and integration with database systems are also described. Theoretical concepts would be applied to real-world applications.

Contents:

Decision support systems overview, Decision Making, Systems, Modelling, and Support, business intelligence, Data Management, Modelling and Analysis, Decision Support System Development, Fundamentals of Expert Systems and Intelligent Systems, Collaborative Computing Technologies, Knowledge Management.

Text Books:

- Efraim Turban and Jay E. Aronson, Decision Support Systems and Intelligent Systems, Seventh Edition, Prentice Hall Pub. M 2004.
- Decision Support Systems and Business Intelligence Systems. 9e. by E. Turban & J. Aronson, 2010

Reference Books:

- Machine Learning with Python: Understanding Machine Learning with Python in the World of Data Science by Robert Wilson (Author, Publisher), 1st Edition, 2016.

Course Title: Intelligent Agents

Course Code: CSC-715

Pre-Requisite: None

Objectives:

The primary objective of this course is to provide an introduction to the basic principles and applications of intelligent agents. The emphasis of the course is on teaching the fundamentals, and not on providing a mastery of specific commercially available software tools or programming environments. Students will be presented with a wide range of theories of relevance to their research and development to model agent's knowledge representation and learning. Emphasis will be placed on understanding concepts of thinking, planning and learning aspects of intelligent agents and using them to model and build relevant agent-based systems.

Contents:

Agent, Environment, Interaction, Solving Problem by Search Algorithms, Informed Search, Constraint Satisfaction Problem, Logical Agents , Theorem Proving Algorithms (propositional logic, predicate logic), Partial Order Planning, Graph Plan, BDI Agents, Decision trees , Neural Networks, Reinforcement learning, Q- learning, Temporal Difference Learning, Monte Carlo Methods.

Text Books:

- Stuart Russel and Peter Norvig, Artificial Intelligence, A modern Approach, 3rd Edition
- Michael J. Wooldridge, Reasoning about Rational Agents.

Reference Books:

- Jack Minker, Logic Based Artificial Intelligence.
- Steven Michael LaValle, Planning Algorithms.

Course Title: Advanced Natural Language Processing

Course Code: CSC-741

Pre-Requisite: None

Objectives:

This course is intended to introduce the students to the fundamental concepts and ideas in natural language processing (NLP). Students will be acquainted with the algorithms available for the processing of linguistic information as well as the underlying computational properties of natural languages. By the end of this course the student should be able to carry out independent work with modern techniques for processing of texts.

Contents:

Introduction to NLP and its applications, Grammar checkers, dictation, document generation, NL interfaces, The different analysis levels used for NLP, Markup, Finite state automata, Recursive and augmented transition networks, Lexical level: Error-tolerant lexical processing (spelling error correction), Transducers for the design of morphologic analyzers, Part-of-speech tagging, Representations for linguistic resources, Syntactic level: Grammars (e.g. Formal/Chomsky hierarchy, DCGs, systemic, case, unification, stochastic), Parsing (top-down, bottom-up, chart (Earley algorithm), CYK algorithm), Semantic level: Logical forms, Ambiguity resolution, Semantic networks and parsers, Procedural semantics, Montague semantics, Vector Space approaches, Pragmatic level: Knowledge representation, Reasoning, Plan/goal recognition, Speech acts/intentions, Natural language generation.

Text Books:

- Handbook of Natural Language Processing, Nitin Indurkhya and Fred J. Damerau, Chapman & Hall/Crc, Second Edition, 2010.
- Natural Language Processing and Text Mining, Anne Kao and Steve R. Poteet, Springer, 2010.

Reference Books:

- Speech and Language Processing, Daniel Jurafsky and James H. Martin, Pearson Prentice Hal, 2nd Edition, 2008.
- Foundations of Statistical Natural Language Processing, Christopher D. Manning, Hinrich Schuetze, The MIT Press; 1st edition, 1999.

Course Title: Advanced Digital Image Processing

Course Code: CEN-745

Pre-Requisite: None

Objectives:

This course will provide mathematical foundations and practical techniques for digital manipulation of images, image acquisition, pre-processing, and segmentation. The course will expose the students to the basic theory and algorithms widely used in digital image processing. After the completion of this course the students will be able to understand the basic concepts behind the processing of digital images as well as various techniques of filtering/processing images in spatial as well as in frequency domain. The course will serve as the basis for more advance topics in Computer Vision.

Contents:

Introduction to Digital Image Processing Computer Vision and Pattern Recognition, Fundamentals Element of visual Perception, Image Sensing and Acquisition Image Sampling and Quantization. Pixel operations, linear & non-linear operations, Image Enhancement in spatial Domain: Background, Grey level Transformations, Filtering in spatial domain. Image Enhancing in Frequency Domain: Frequency domain, Fourier Transform, Filtering in frequency domain, Color Image Processing, Fundamentals of Image Compression, Lossless and lossy compression, Image Compression standards, Image Segmentation: Detection of Discontinuities, Edge and Boundary detection, Thresholding, Region Based segmentation, Morphological image processing, Representation schemes: Boundary and region descriptors.

Text Books:

- Digital Image Processing, R. C. Gonzalez and R. E. Woods, Addison Wesley, 3rd Edn., 2007.
- Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab, Chris Solomon and Toby Breckon, 2011.

Reference Books:

- Machine Learning with Python: Understanding Machine Learning with Python in the World of Data Science by Robert Wilson (Author, Publisher), 1st Edition, 2016.

Course Title: Advanced Neural Networks and Fuzzy Logic

Course Code: CSC-749

Pre-Requisite: Artificial Intelligence

Objectives:

This course presents an overview of the theory and applications of artificial neural network and fuzzy systems to computer science and software engineering applications. The objective of this course is on the understanding of various neural network and fuzzy systems models and the applications of these models to solve computing/software engineering problems.

Contents:

Artificial Intelligence Artificial Neural Network overview, Supervised Learning: Single-Layer Networks , Perceptrons , Adalines Supervised Learning: Multi-Layer Networks, Multi-Layer Perceptrons (MLPs) , Backpropagation , Conjugate Gradient method , Levenberg-Marquardt (LM) method , Madalines , Radial-Basis Networks , Cascade-Correlation Networks , Polynomial Networks , Recurrent Networks (Time series , Backpropagation through time , Finite Impulse Response (FIR) MLP), Temporal Differences method (TD). Unsupervised Learning, Simple Competitive Networks: Winner-take-all | Hamming network , Learning Vector Quantization (LVQ), Counterpropagation Networks (CPN) , Adaptive Resonance Theory (ART) , Kohonen Self-Organizing Maps (SOMs) , Principal Component Analysis networks (PCA), Associative Models, Linear Associative Memory (LAM) , Hopfield Networks , Brain-State-in-a-Box , BSB) , Boltzmann Machines and Simulated Annealing , Bi-Directional Associative Memory (BAM), Optimization Problems, Neural Network Approaches, Evolutionary Programming , Fuzzy logic and its connection to NNs

Text Books:

- Neural networks: methodology and applications, by G. Dreyfus-computers-, 2005
- Evolving Fuzzy Systems - Methodologies, Advanced Concepts and Applications, By Edwin Lughofer, 2011.

Reference Books:

- Neural Networks: A Comprehensive Foundation, Simon Haykin, Prentice Hall, Upper Saddle River, NJ, SECOND EDITION, 1999
- Artificial neural networks: an introduction, by Kevin L. Priddy, Paul E. Keller-Technology & Engineering-2005

Course Title: Pattern Recognition

Course Code: CSC-751

Pre-Requisite: None

Objectives:

The goal of this course is to provide an introduction to the fundamental concepts of machine learning and pattern recognition with examples from several application areas. The students will be acquainted with real world regression and classification problems and the models and classifiers to solve these problems. Students will also be introduced to dimensionality reduction and feature selection concepts. Additionally, students will be exposed to various clustering techniques. A key objective to this course is for the students to also acquire hands-on experience related to classification and clustering tasks.

Contents:

Introduction to Pattern recognition and Machine learning, Matrices and vectors: Toeplitz and Vendermonde matrices, classification and regression, Bayesian Decision theory, Normal Density and decision functions for normal distribution, Maximum likelihood estimation, Dimensionality reduction – Component analysis, feature selection, Hidden Markov Models and Artificial neural networks, Non-parametric methods, Unsupervised learning and clustering: Clustering techniques.

Text Books:

- The Elements of Statistical Learning, Trevor Hastie, Robert Tibshirani and Jerome Friedman, Springer, 2009.
- Pattern Recognition and Classification: An Introduction, by Geoff Dougherty, S. Theodoridis & K. Koutroumbas, Academic Press, 2012.

Reference Books:

- Pattern recognition and Machine Learning, Christopher M. Bishop, Springer, 2007.
- Introduction to Machine Learning, Ethem Alpaydin, MIT Press, 2004.

Course Title: Computer Vision

Course Code: CSC-764

Pre-Requisite: None

Objectives:

By the end of this course, the students would have developed an understanding of the problems in simulating human perception into machines. Students will have a thorough understanding of the state of the art computer vision methods, algorithms and results. The students will also be able to apply the tools and techniques learned to solve practical vision related problems.

Contents:

Introduction to Computer Vision and related areas along with applications, Image formation and representation: imaging geometry, digitization, cameras and projections, rigid and affine transformations, Filtering: convolution, smoothing. Segmentation: region splitting and merging; quadtree structures for segmentation; Feature detection: edge detection, corner detection, line and curve detection, SIFT and HOG descriptors, shape context descriptors. Model fitting: Hough transform, line fitting, ellipse and conic sections fitting, algebraic and Euclidean distance measures. Camera calibration: camera models; intrinsic and extrinsic parameters; affine, and perspective camera models. Epipolar geometry: introduction to projective geometry; epipolar constraints; the essential and fundamental matrices; Motion analysis: the motion field of rigid objects; motion parallax; optical flow, the image brightness constancy equation, affine flow; differential techniques; feature-based techniques; Motion tracking: the Kalman filter; Object recognition and shape representation.

Text Books:

- Computer Vision: Algorithms and Applications, R. Szeliski, Springer, 2011.
- Computer Vision: A Modern Approach, D. Forsyth and J. Ponce, Prentice Hall, 2nd ed., 2011.

Reference Books:

- Computer Vision: A Modern Approach, By David Forsyth, Jean Ponce, Prentice Hall, 2003.
- Handbook of Mathematical Models in Computer Vision, By Nikos Paragios, Yunmei Chen, Olivier Faugeras, Birkhäuser, 2006

Course Title: Cloud Computing

Course Code: CSC-781

Pre-Requisite: None

Objectives:

Understanding the systems, protocols and mechanisms to support cloud computing, Application architectures for cloud computing, understanding the hardware necessary for cloud computing and design and implementation of cloud computing application

Contents:

This course introduces students to the cloud and the computing on the cloud. Initially, the focus is on the technology context, i.e. multi-core architectures, virtualization, parallel computing models and big data storage. Next, famous cloud computing models including Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS) are studied with the help of Amazon AWS (IaaS), Microsoft Azure (PaaS) and Google App Engine (SaaS). In addition to computing models, Data and computation models, e.g. MapReduce, are an important part of this module. The theoretical concepts are explained with hand-on experience of cloud platforms supported by case studies. The course concludes with an insight into the cloud risk areas including risks with service provider, technical risks, security issues, connectivity issues, etc. and research work in these areas is also discussed.

Text Books:

- Handbook of Cloud Computing, Borko Furht. Springer, 2010.
- Cloud Computing: SaaS, PaaS, IaaS, Virtualization, Business Models, Mobile, Security, and More, Kris Jamsa Jones & Bartlett Publishers, 2012

Reference Books:

- Cloud Computing and SOA: Convergence in your enterprise, David Linthicum. Addison Wesley, 2009
- Distributed File Systems: Hadoop, Lustre, Google File System, Andrew File System, Off system, Distributed File System, Ceph. General books LLC, 2010

Course Title: Advanced Information Theory

Course Code: CSC-554

Pre-Requisite: None

Objectives:

This course presents the advance concepts of Information Theory, that stays at the basis of modern digital communications, data compression, lossy source coding and multiuser networks. Details of to what computer scientists mean by "information", including topics in data compression (such as zip files and mp3), error correcting codes, information entropy, cryptography, and randomness.

Contents:

Asymptotic Equipartition Theorem, types, and typical sequences, Information measures and their properties: entropy, Kullback-Leibler divergence, mutual information, source coding theorem, channel coding theorem, rate distortion theory, quantization, maximum entropy principle Typical sequences and typical sets, error exponents in: hypothesis testing, source coding, and channel coding, information theory and estimation, rudiments of network information theory.

Text Books:

- T.M. Cover and J.A. Thomas, Elements of Information Theory, 2nd ed., Wiley, 2006;
- Csisz`ar and J. K¨orner, Information theory: coding theorems for discrete memoryless systems, 2nd ed., Cambridge University Press, 2011.

Reference Books:

- Codes: an introduction to information communication and cryptography by Norman Biggs, 2008

Course Title: Text Mining

Course Code: CSC-747

Pre-Requisite: Data Mining

Objectives:

Text Mining is aimed at extracting useful information (usually) from huge unstructured datasets by employing techniques from information retrieval, natural language processing and data mining. The objective of this module is to get a good understanding of the basic text mining techniques and study some of its applications as well.

Contents:

Dealing with information overload and information overlook, unstructured vs. (semi-) structured data, evolving information needs and knowledge management issues, the business case for text mining. The text mining pipeline: information retrieval, information extraction and data mining. Fundamentals of natural language processing: linguistic foundations, levels of linguistic analysis. Approaches to text mining: rule-based vs. machine learning based vs. hybrid; generic vs. domain specific; domain adaptation. Dealing with real text: text types, document formats and conversion, character encodings, markup, low-level processes (sentence splitting, tokenization, part of speech tagging, chunking). Information extraction: term extraction, named entity recognition, relation extraction, fact and event extraction; partial analysis vs. full analysis. Data mining and visualisation of results from text mining. Evaluation of text mining systems: evaluation measures, role of evaluation challenges, usability evaluation, the U-Compare initiative. Text mining applications and services; case studies.

Text Books:

- Text mining handbook: advanced approaches in analyzing unstructured data, Feldman, Ronen and James Sanger, Cambridge University Press, Edition: 2nd ,2008
- Text mining: classification, clustering and applications, Srivastava, Ashok and Mehran Sahami, Chapman & Hall, Edition: 1st , 2009

Reference Books:

- Mining Text Data by Charu C. Aggarwal, ChengXiang Zhai - 2012

Course Title: Advanced DBMS

Course Code: CSC-752

Pre-Requisite: Database Management System

Objectives:

At the end of this course, the expectation is that the students will gain competence in following areas: Databases beyond relational, Query optimization, Data marts, Data warehousing, XML, OLAP.

Contents:

Object-Oriented Databases, Object-Relational Databases, Mobile Databases, Temporal, Spatial and Geographic Databases, Distributed Database Design, Distributed Multimedia Database Systems, Data Warehouse and OLAP Systems, Business Intelligence, XML Data Models, XML Documents and DTD, XML Query Languages, Current Research and Development Trends of Database Analysis, Design, Modeling and Applications.

Text Books:

- An Advanced Course in Database Systems: Beyond Relational Databases, S. W. Dietrich and S. D. Urban, Prentice Hall, 2005.
- Advanced Database Management Systems by Rini Chakrabarti, Shilbhadra Dasgupta - 2011

Reference Books:

- Database Management Systems, Ramakrishnan R & Gehrke J, 3rd edn, McGraw Hill, New York, 2003.

Course Title: Advanced Data Warehousing

Course Code: CSC-760

Pre-Requisite: None

Objectives:

By the end of this course students will be familiar with concepts of Data Warehousing including: Strategic need of data warehousing, Building blocks of a data warehouse, Data warehouse project management, Business requirements of a data warehouse, Architectural components of a data warehouse, Data warehouse metadata management, Dimensionality Modeling, ETL & Data quality, Online Analytical Processing, as well as the following areas of data mining: Motivation for data mining, Data Preprocessing, Data mining primitives and query languages, Architectures of data mining systems, Major Data Mining Tasks, Cluster Analysis , Statistical measures in large databases, Classifications and Predictions, Anomaly Detection.

Contents:

Data Warehouse: Planning and Requirements, Data Warehouse Architecture, Data Warehouse Infrastructure, Dimensional Modeling, Metadata, Extraction, Transformation and Loading, Online Analytical Processing, Data Preparation Techniques: outlier and missing data analysis, Data Reduction Techniques, Introduction to Data Mining, Modeling and Principal Feature Extraction, Clustering, Hierarchical Clustering, Partitional Clustering, Classification , Decision Tree Classification, Bayesian Classification, Nearest Neighbor Classification.

Text Books:

- Data Warehousing Fundamentals for IT Professionals, Paulraj Pooniah, Wiley, 2nd Edition, 2010.

Reference Books:

- Data Mining Concepts & Techniques, Jaiwei Han, Micheline Kamber, 2nd Edition, 2005.

Course Title: Ontology Engineering

Course Code: SEN-764

Pre-Requisite: None

Objectives:

This Course provides students with a theoretical and practical understanding of leading edge solutions for the Semantic Web. It introduces students to the W3C standard Web Ontology Language, OWL, its underlying Description Logics, establishing patterns to avoid the pitfalls in using OWL. The course provides an opportunity to become familiar with a widely used environment for developing and an API for applying OWL ontologies, and making use of reasoning services accessible via both. Ontology provide rich, repressive vocabularies of terms describing a domain (e.g. medicine, astronomy, music, etc.). They are key to information exchange, data integration and search.

Contents:

Introduction to Description Logics and Reasoning, concepts of semantic interoperability, integration and automation; concept of metadata and ontology; RDF and RDFS , Ontology Web Language (OWL) and Ontology Engineering Methodologies.

Text Books:

- Ontology Engineering in a Networked World, by Mari Carmen Suárez-Figueroa, Asunción Gómez-Pérez, Enrico Motta – 2012
- D. Allemang and J. Hendler: Semantic Web for the Working Ontologist. Morgan Kaufmann (2008).

Reference Books:

- Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web. First Edition - Asunción Gómez-Pérez, Mariano Fernandez-Lopez, Oscar Corcho, 2010.

Annex MSDS – Electives

Elective Courses

The list of elective courses has been revised to be more pertinent to Data Science. The following non-relevant courses have been removed from the list of electives for MS Data Science Program.

CSC-701	Computer Supported Cooperative Work
SEN-720	Advanced Human Computer Interaction
SEN-756	Advanced Usability Engineering
CSC-750	Intelligent Tutoring Systems
EET-519	Distributed Networking
EET-556	Mobile Communications and Networking
EET-702	Advanced Network Security
EET-761	Network Protocols and Standards
CSC-720	Advanced Operating Systems
CEN-720	Advanced Computer Architecture
CSC-753	Distributed Databases
CSC-754	Object Oriented Databases
CSC-755	Web based DBMS
CSC-756	Multimedia Databases

Approved Elective Courses – MS Data Science

Elective Courses – MS Data Science			
Sr. No	**Course Code**	**Course Title**	**Credit Hours**
5	DSC-702	Data Visualization	3
6	DSC-703	Distributed Data Engineering	3
2	DSC-705	Unstructured Data Processing	3
3	CSC-518	Decision Support Systems	3

4	CSC-715	Intelligent Agents	3
5	CSC-741	Advanced Natural Language Processing	3
6	CEN-745	Advanced Digital Image Processing	3
7	CSC-749	Advanced Neural Networks and Fuzzy Logic	3
8	CSC-751	Pattern Recognition	3
9	CSC-764	Computer Vision	3
10	CSC-781	Cloud Computing	3
11	CSC-554	Advanced Information Theory	3
12	CSC-747	Text Mining	3
13	CSC-752	Advanced DBMS	3
14	CSC-760	Advanced Data Warehousing	3
15	SEN-764	Ontology Engineering	3
16	ESC-500	Thesis	6