MANIPAL
ACADEMY of HIGHER EDUCATION
(Deemed to be University under Section 3 of the UGC Act, 1956)

Manipal School of Information Sciences (MSIS)

Manipal Academy of Higher Education, Manipal

Outcome Based Education (OBE) Framework

Two Year full time Postgraduate Program

Master of Engineering - ME (Big Data and Data Analytics)

MANIPAL
ACADEMY of HIGHER EDUCATION
(Deemed to be University under Section 3 of the UGC Act, 1956)

TABLE OF CONTENTS

Contents

NATURE AND EXTENT OF THE PROGRAM

An engineering graduate skillset requirement is changingwith invent of the new technologies. In particular theimpact of Big data andits transformative technologies like Data Analytics provide a high employability in the industry. Big Data and Data Analytics are playing an important role in business, government, healthcare and education. Big Data technologies provide efficient solutions for acquiring and processing large scale data. Data Analytics combines principles and techniques from mathematics, computer science and machine learning for offering predictive and prescriptive solutions

ME (Big Data and Data Analytics) Program is a comprehensive two-year postgraduate program, which aims to provide hands-on experience to prepare industry-readyBig Data and Data Analytics professionals.The program ME (BDA) helps engineering graduates to learn, understand, and practice big data analytics and machine learning approaches, which include the study of modern computing big data technologies and scaling up machine learning techniques focusing on industry applications.

ME (Big Data and Data Analytics) postgraduate degree would welcome graduates from any discipline with 50% mark in qualifying exam. Students after successfully completing the program will get career opportunities as anBig Data Architect, Data Analyst, Database Administrator, Data Scientists, Big Data Engineer and Business Analyst.

PROGRAM EDUCATION OBJECTICE (PEO)

The overall objectives of the Learning Outcomes-based Curriculum Framework (LOCF) for **ME (Big Data and Data Analytics) program are as follows.**

PEO No	Education Objective
PEO 1	Develop in depth understanding of the key technologies in data engineering, data science and business analytics.
PEO 2	Practice problem analysis and decision-making using machine learning techniques.
PEO 3	Gain practical, hands-on experience with statistics, programming languages and big data tools through coursework and applied research experiences.

GRADUATE ATTRIBUTES

S No.	Attribute	Description
1	**Scholarship of Knowledge**	Acquire in-depth knowledge of specific discipline or professional area, including wider and global perspective, with an ability to discriminate, evaluate, analyse and synthesise existing and new knowledge, and integration of the same for enhancement of knowledge.
2	**Critical Thinking**	Analyse complex engineering problems critically, apply independent judgement for synthesising information to make intellectual and/or creative advances for conducting research in a wider theoretical, practical and policy context.
3	**Problem Solving**	Think laterally and originally, conceptualise and solve engineering problems, evaluate a wide range of potential solutions for those problems and arrive at feasible, optimal solutions after considering public health and safety, cultural, societal and environmental factors in the core areas of expertise.
4	**Research Skill**	Extract information pertinent to unfamiliar problems through literature survey and experiments, apply appropriate research methodologies, techniques and tools, design, conduct experiments, analyse and interpret data, demonstrate higher order skill and view things in a broader perspective, contribute individually/in group(s) to the development of scientific/technological knowledge in one or more domains of engineering.
5	**Usage of modern tools**	Create, select, learn and apply appropriate techniques, resources, and modern engineering and IT tools, including prediction and modelling, to complex engineering activities

		with an understanding of the limitations.
6	**Collaborative and Multidisciplinary work**	Possess knowledge and understanding of group dynamics, recognise opportunities and contribute positively to collaborative-multidisciplinary scientific research, demonstrate a capacity for self-management and teamwork, decision-making based on open-mindedness, objectivity and rational analysis in order to achieve common goals and further the learning of themselves as well as others.
7	**Project Management and Finance**	Demonstrate knowledge and understanding of engineering and management principles and apply the same to one's own work, as a member and leader in a team, manage projects efficiently in respective disciplines and multidisciplinary environments after consideration of economical and financial factors.
8	**Communication**	Communicate with the engineering community, and with society at large, regarding complex engineering activities confidently and effectively, such as, being able to comprehend and write effective reports and design documentation by adhering to appropriate standards, make effective presentations, and give and receive clear instructions.
9	**Life-long Learning**	Recognise the need for, and have the preparation and ability to engage in life-long learning independently, with a high level of enthusiasm and commitment to improve knowledge and competence continuously.
10	**Ethical Practices and Social Responsibility**	Acquire professional and intellectual integrity, professional code of conduct, ethics of research and scholarship, consideration of the impact of research outcomes on professional practices and an understanding of responsibility to contribute to the community for sustainable development

		of society.
11	**Independent and Reflective Learning**	Observe and examine critically the outcomes of one's actions and make corrective measures subsequently, and learn from mistakes without depending on external feedback.

QUALIFICATIONS DESCRIPTORS

1. Demonstrate

   (i) A systematic, extensive and coherent knowledge and understanding of an academic field of study as a whole and its applications, and links to related disciplinary areas/subjects of study; including a critical understanding of the established theories, principles and concepts, and of a number of advanced and emerging issues in the field of Big Data;

   (ii)Procedural knowledge that creates different types of professionals related to the Big Data, including research and development, teaching and government and public service;

   (iii) Professional and communication skills in the domain of machine learning, distributed computing, real time streaming, natural language and text processing, including a critical understanding of the latest developments, tools in the domain of big data and data analytics.
2. Demonstrate comprehensive knowledge about materials, including current research, scholarly, and/or professional literature, relating to essential and advanced learning areas pertaining to the bigdata and analytics, and techniques and skills required for identifying problems and issues related.
3. Demonstrate skills in identifying information needs, collection of relevant quantitative and/or qualitative data drawing on a wide range of sources, analysis and interpretation of data.
4. Methodologies as appropriate to the subject(s) for formulating evidence based solutions and arguments
5. Use knowledge, understanding and skills for critical assessment of a wide range of ideas and complex problems and issues relating to the chosen field of study.

6. Communicate the results of studies undertaken in an academic field accurately in a range of different contexts using the main concepts, constructs and techniques of the bigdata and data analytics studies.

7. Address one's own learning needs relating to current and emerging areas of study, making use of research, development and professional materials as appropriate, including those related to new frontiers of knowledge.

8. Apply one's disciplinary knowledge and transferable skills to new/unfamiliar contexts and to identify and analyse problems and issues and seek solutions to real-life problems.

PROGRAM OUTCOMES

After successful completion of ME (Big Data and Data Analytics), Students will be able to:

PO No	Attribute	Competency
PO 1	**Scholarship of Knowledge**	Acquire in-depth knowledge of BDA domain, with an ability to discriminate, evaluate, analyze, synthesize the existing and new knowledge, and integration of the same for enhancement of knowledge.
PO 2	**Critical Thinking**	Analyze complex Big Data and Data Analytics Eco System critically, apply independent judgement for synthesizing information to make intellectual and/or creative advances for conducting research in a wider theoretical, practical and policy context.
PO 3	**Problem Solving**	Think laterally and originally, conceptualize and solve Big Data problems, evaluate a wide range of potential solutions for those problems and arrive at feasible, optimal solutions after considering public health and safety, cultural, societal and environmental factors in the core areas of expertise.
PO 4	**Research Skill**	Extract information pertinent to unfamiliar problems through literature survey and experiments, apply appropriate research methodologies, techniques and tools, design, conduct experiments, analyze and interpret data, demonstrate higher order skill and view things in a broader perspective, contribute individually/in group(s) to the development of scientific/technological knowledge in one or more domains of engineering.
PO 5	**Usage of modern tools**	Create, select, learn and apply appropriate techniques, resources, and modern engineering and IT tools, including

		prediction and modelling, to complex engineering activities with an understanding of the limitations.
PO 6	**Collaborative and Multidisciplinary work**	Possess knowledge and understanding of group dynamics, recognize opportunities and contribute positively to collaborative-multidisciplinary scientific research, demonstrate a capacity for self-management and teamwork, decision-making based on open-mindedness, objectivity and rational analysis in order to achieve common goals and further the learning of themselves as well as others.
PO 7	**Project Management and Finance**	Demonstrate knowledge and understanding of engineering and management principles and apply the same to one's own work, as a member and leader in a team, manage projects efficiently in respective disciplines and multidisciplinary environments after consideration of economical and financial factors
PO 8	**Communication**	Communicate with the engineering community, and with society at large, regarding complex engineering activities confidently and effectively, such as, being able to comprehend and write effective reports and design documentation by adhering to appropriate standards, make effective presentations, and give and receive clear instructions.
PO 9	**Life-long Learning**	Recognize the need for and have the preparation and ability to engage in life-long learning independently, with a high level of enthusiasm and commitment to improve knowledge and competence continuously.
PO 10	**Ethical Practices and Social Responsibility**	Acquire professional and intellectual integrity, professional code of conduct, ethics of research and scholarship, consideration of the impact of research outcomes on professional practices and an understanding of responsibility

		to contribute to the community for sustainable development of society.
PO 11	**Independent and Reflective Learning**	Observe and examine critically the outcomes of one's actions and make corrective measures subsequently and learn from mistakes without depending on external feedback.

COURSE STRUCTURE, COURSEWISE LEARNING OBJECTIVE, AND COURSE OUTCOMES (COS)

FIRST YEAR:

Semester: 1 **Semester: 2**

Subject Code	Subject Title	L	T	P	C	Subject Code	Subject Title	L	T	P	C
BDA 601	Fundamentals of Machine Learning	3	-	-	3	BDA 605	Machine Learning for Big Data	3	-	-	3
BDA 602	Algorithms and Data Structures for Big Data	3	-	-	3	BDA 616	Modern Databases for Big Data	3	-	-	3
BDA 623	Architecture of Big Data Systems	3	-	-	3	MCL 602	Advanced Applications of Probability and Statistics	3	-	-	3
MCL 601	Applied Probability and Statistics	3	-	-	3	BDA 618	Multimedia Analytics	3	-	-	3
	Elective - 1	3	-	-	3		Elective - 2	3	-	-	3
BDA 601L	Fundamentals of Machine Learning Lab	-	-	3	1	BDA 605L	Machine Learning for Big Data Lab	-	-	3	1
BDA 602L	Algorithms and Data Structures for Big Data Lab	-	-	3	1	BDA 616L	Modern Databases for Big Data Lab	-	-	3	1
BDA 623L	Architecture of Big Data Systems Lab	-	-	3	1	MCL 602L	Advanced Applications of Probability and Statistics	-	-	3	1
MCL 601L	Applied Probability and Statistics Lab	-	-	3	1	BDA 618L	Multimedia Analytics Lab	-	-	3	1
	Elective - 1 Lab	-	-	3	1		Elective - 2 Lab	-	-	3	1
BDA 695	Mini Project - 1	-	-	4	-	BDA 696	Mini Project - 2	-	-	-	4
BDA 697	Seminar - 1	-	-	1	-	BDA 698	Seminar - 2	-	-	-	1

Total	**15**	**-**	**15**	**25**	**Total**	**15**	**-**	**15**	**25**

SECOND YEAR (FINAL YEAR):

III and IV Semester		
BDA 799	Project Work	25
Total Number of Credits to Award Degree		75

List of Electives(Theory)

Elective - 1		**Elective - 2**	
Code	**Subject**	**Code**	**Subject**
CSE-625	Mobile Web Application Development	CDC-607	DevOps for Cloud
BDA-622	Principles of Data Visualization	BDA-621	Natural Language and Text Processing
		ENP-601	Entrepreneurship

List of Electives(Lab)

Elective - 1		**Elective - 2**	
Code	**Subject**	**Code**	**Subject**
CSE-625L	Mobile Web Application Development Lab	CDC-607L	DevOps for Cloud Lab
BDA-622L	Principles of Data Visualization Lab	BDA-621L	Natural Language and Text Processing Lab
		ENP-601L	Entrepreneurship Lab

Name of the Institution / Department: Manipal School of Information Sciences

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Fundamentals of Machine Learning
Course Code: BDA-601	**Course Instructor:**
Academic Year: 2020 - 2021	**Semester:** First Year, Semester 1
No of Credits: 3	**Prerequisites:** Basic Programming – preferably Python
Synopsis:	This Course provides insight on 1. This course provides the concept of machine learning, applications, techniques, design issues and approaches to machine learning. 2. This course provides the fundamental knowledge about concept learning, hypothesis and bias. 3. To implement machine learning algorithms such as Decision Tree learning, Probably Approximately Correct (PAC) learning, Bayesian learning, Instance-based learning, Principal Component Analysis (PCA) and Ensemble methods in real time data set for various analysis.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Identify the goals, applications, types and design issues of machine learning techniques.
CO 2:	Relate concept learning and hypothesis space.
CO 3:	Apply PCA learning approach to reduce the dimension.
CO 4:	Analyse different machine learning algorithms.
CO 5:	Design ensemble methods.

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*										
CO 2		*									
CO 3			*								
CO 4				*							
CO 5				*							

Course content and outcomes:	
Content	***Competencies***
Unit 1: Introduction	
Definition of Machine Learning, Goals and applications of machine learning, Basic design issues and approaches to machine learning, Types of machine learning techniques	At the end of the topic student should be able to: 1. Define Machine Learning (C1) 2. Describe about any three applications for which machine learning approaches seem appropriate. (C2) 3. Illustrate different types of machine learning techniques (C3)
Unit 2: Inductive Classification	
The concept learning task, Concept learning as search through a hypothesis space, General-to-specific ordering of hypotheses, Finding maximally specific hypotheses, Version spaces and the candidate elimination algorithm, Inductive bias.	1. Relate concept learning and hypothesis space (C4). 2. Apply different algorithms to obtain most general and most specific hypotheses from the training examples. (C3)
Unit 3: Decision Tree learning	
Representing concepts as decision trees, Recursive induction of decision trees, Picking the best splitting attribute, Entropy and information gain, Searching for simple trees and computational complexity.	1. Apply decision tree algorithm to find the hypothesis space (C3) 2. Construct decision tree machine learning algorithm (C5) 3. Explain the method of choosing training examples and target function in the design of a machine learning system (C2) 4. Explain different validation technique to find the accuracy in training and testing of data set (C5)

Unit 4: Computational learning theory	
Models of learnability: learning in the limit, Probably Approximately Correct (PAC) learning, Sample Complexity: quantifying the number of examples needed to PAC learn, Computational complexity of training. Sample complexity for finite hypothesis spaces, Noise Learning Multiple Classes, Bias-variance trade-off, under-fitting and over-fitting concepts	1. Define various terms related to computational learning approach (C1). 2. Describe different models learning in the limit (C2) 3. Calculate the number of training examples required in different types of learning approaches (C4).
Unit 5: Bayesian learning	
Probability theory and Bayes rule, Naive Bayes learning algorithm - Parameter smoothing, Generative vs. discriminative training, Logistic regression, Bayes nets and Markov nets for representing dependencies	1. Write the applications of Bayes theorem (C3) 2. Describe the use of Logistic Regression in Machine Learning (C2) 3. Predict the target value for the new instance using Naïve Bayes classifier. (C3)
Unit 6: Instance-based learning	
Constructing explicit generalizations versus comparing to past specific examples, K-Nearest Neighbour learning algorithm, Case-based reasoning (CBR) learning	1. Construct explicit generalizations (C5) 2. Discriminate Instances Based and Case-based learning (C4) 3. Explain K-nearest neighbour learning (C5)
Unit 7: Continuous Latent Variables	
Principal Component Analysis (PCA), Applications of PCA	1. Describe use of Principal Component Analysis for the complex data set (C2). 2. Apply PCA to choose principal components for the given data set (C3)
Unit 8: Ensemble methods (bagging and boosting)	

Using committees of multiple hypotheses, Bagging, Boosting, DECORATE, Active learning with ensembles	1. Choose a suitable method of ensemble learning approach (C3). 2. Explain various ensemble techniques (C5)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	30	60
Quiz	02	04
Small Group Discussion (SGD)	02	02
Self-directed learning (SDL)	-	04
Problem Based Learning (PBL)	02	04
Case Based Learning (CBL)	-	-
Revision	02	-
Assessment	06	-
TOTAL	**44**	**74**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Sessional Examination 1	*	*			
Sessional Examination 2			*	*	
Assignment/Presentation	*	*	*	*	
End Semester Examination	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. T. Mitchell, "Machine Learning", McGraw-Hill, 1997. 2. E. Alpaydin, "Machine Learning", MIT Press, 2010. 3. C. Bishop," Pattern Recognition and Machine Learning", Springer, 2006. 4. E. Hart, R. Duda and D. Stork, "Pattern Classification", Wiley-Interscience, 2000. 5. T. Hastie, R. Tibshirani and J. Friedman, "The Elements of Statistical Learning: Data Mining, 6. Inference and Prediction", Springer, 2nd Edition, 2009. 7. Jason Bell, "Machine Learning for Big Data", Wiley Big Data Series, 2016. 8. 7. Rama Murthy G," Multidimensional Neural Networks Unified Theory", New Age International, 2008.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Algorithms and Data Structures for Big Data
Course Code: BDA 602	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 1
No of Credits: 3	**Prerequisites:** Programming in Python, C
Synopsis:	This Course provides insight on 1. This course introduces students to elementary data structures and design of algorithms. 2. Students learn how to design optimal algorithms with respect to time and space 3. Students learn how to implement link list, stack, queues, searching and sorting techniques, sets, trees and graphs. 4. Students learn to implement string and text processing techniques. 5. Students learn to implement Data stream algorithms.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Analyse recursive programs, solve a general class of recurrence relations (C4)
CO 2:	Design programs for implementation of linked lists, stack, queues, binary search tree, sorting and searching (C4)
CO 3:	Design programs for dictionary, hash tables, graphs and shortest path techniques. (C4)
CO 4:	Design string and text processing programs. (C4)

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*	*	*						*		
CO 2	*	*	*						*		
CO 3	*	*	*						*		
CO 4	*	*	*	*							

Course content and outcomes:	
Content	***Competencies***
Unit 1: Algorithm specification and analysis techniques	
Analysis of recursive programs. Solving recurrence equations. General solution for a large class of recurrences.	At the end of the topic student should be able to: 1. Define recursive programs (C2) 2. Design simple recursive programs (C4) 3. Solve recurrence relations (C4)
Unit 2: Elementary data structures	
Implementation of lists, stacks, queues.	1. Design singly linked list (C3) 2. Design doubly linked list(C3) 3. Explain the concepts of array-based stacks (C2) 4. Explain the concepts of pointer-based stacks (C2) 5. Design and implement Queues. (C4)
Unit 3: Sorting and Searching Techniques	
Quick sort, heap sort, merge sort. Linear search and binary search.	1. Design applications with suitable sorting and searching techniques. (C4)
Unit 4: Hashing and Dictionaries	
Hashing and Dictionaries	1. Design various hash functions and implement suitable hash tables (C4)
Unit 5: Binary search trees	
Construction. Inorder, preorder and postorder traversals.	1. Understand and implement BST and its various traversal techniques (C2)
Unit 6: Graphs	
Representation of graphs. Depth First Searching. Breadth First Searching. Minimum cost spanning tree. Single source shortest paths and all-	1. Define graphs (C2) 2. Design data structure for graphs (C6) 3. Formulate an algorithm to solve minimum cost spanning tree(C6)

pairs shortest path.	4. Formulate an algorithm to solve Single source shortest path (C6) 5. Formulate an algorithm to solve All- pair shortest path(C6)
Unit 7: String and text processing techniques	
Pattern-Matching Algorithms. Text Compression. Tries.	1. Design applications with suitable pattern matching algorithms (C4).
Unit 8: Data stream algorithms	
Sampling, Random Projections, Basic Algorithmic Techniques, Group Testing, Tree Method and Graph sketching.	1. Implement suitable data streaming algorithms (C3).

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	30	60
Quiz	02	04
Small Group Discussion (SGD)	02	02
Self-directed learning (SDL)	-	04
Problem Based Learning (PBL)	02	04
Case Based Learning (CBL)	-	-
Revision	02	-
Assessment	06	-
TOTAL	**44**	**74**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos				
Nature of assessment	CO 1	CO 2	CO 3	CO 4
Sessional Examination 1	*	*		
Sessional Examination 2		*	*	
Assignment/Presentation				*
End Semester Examination	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Introduction to Algorithms - Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest. MIT Press. 2. Data Structures and Algorithms - Aho, Hopcroft and Ulmann. Pearson Publishers. 3. Data Structures and Algorithms in Python - Michael T. Goodrich, Roberto Tamassia, and Michael H. Goldwasser. John Wiley & Sons. 4. Data Streams: Algorithms and Applications - S. Muthukrishnan. Foundations and Trends in Theoretical Computer Science archive, Volume 1 Issue 2, August 2005, Pages 117 – 236.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Architecture of Big Data Systems
Course Code: BDA 623	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 1
No of Credits: 3	**Prerequisites:** Programming in Python, Java
Synopsis:	This Course provides insight on 1. Students learn the concept of big data characteristics, batch and lambda architecture. 2. This course introduces students to basics file systems in Big Data 3. This course helps the student to understand the concepts of Hadoop framework, Spark framework and their internals. 4. This course helps the students to learn Map-reduce programming, Spark programming. 5. Students learn the different layers with use cases demonstrations.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Examine the type of data in big data (C3)
CO 2:	To design applications based with Hadoop framework (C5)
CO 3:	To design applications based with spark architecture (C5)
CO 4:	To build applications based on the Big Data Architecture platforms and analyse the results based on the outcome of the applications used (C6)

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*	*	*			*					
CO 2	*	*	*		*		*			*	
CO 3	*	*	*		*		*			*	
CO 4	*	*	*		*	*	*			*	

Course content and outcomes:	
Content	***Competencies***
Unit 1: Classifying Big Data Characteristics	
Analysis type - real time or batched for later analysis. Processing methodology - predictive, analytical, ad-hoc query, and reporting. Data frequency and size on demand, as with social media data Continuous feed, real-time - weather data, transactional data Time series - time-based data Data type - transactional, historical, master data and metadata. Content formats - structured, unstructured, semi-structured Data sources - Web and social media, humans, machines, transaction data and biometric data.	At the end of the topic student should be able to: 1. Identify different types of Data 2. Identify processing methodology
Unit 2: Big Data processing - the Lambda architecture	
Append-only, immutable data Batch layer Serving layer Speed layer Case study: Druid - A Real-time Analytical Data Store	1. Understand Lambda architecture to handle Big Data (C2). 2. Understand different layers in Lambda Architecture (C2).
Unit 3 Batch layer, Serving layer and Speed layer	
Choosing a storage solution for the batch layer: Distributed file systems, Vertical partitioning.	1. Develop applications to store data in HDFS (C4). 2. Develop applications for batch processing

MapReduce: a paradigm for Big Data computing. Performance metrics for the serving layer Requirements for a serving layer database Computing real time views Storing real time views Challenges of incremental computation Asynchronous versus synchronous updates	using Map Reduce technique (C4). 3. Understand the need of serving layer (C2). 4. Design application to store data for processing in serving layer (C4). 5. Understand the need of Speed layer for data processing (C2).
Unit 4: Spark: Alternatives to MapReduce	
Spark Architecture Spark Session DataFrame Transformations and Actions Spark SQL Resilient Distributed Datasets (RDDs)	1. Understand Spark Architecture for data processing (C2). 2. Design applications using DataFrames and RDDs (C4).
Unit 5: Stream Processing using Spark	
Advantages and challenges of stream processing Stream Processing Design Points Streaming APIs Structured Stream Processing	1. Understand different stream processing techniques (C2). 2. Design applications for handling real time data using Structured Streaming (C4).
Unit 6: Machine Learning using Spark	
High level M-Lib concepts M-Lib in Action	1. Understand different libraries and packages for machine learning in Spark (C2). 2. Design machine learning model using

	Spark (C4).

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	30	60
Quiz	02	04
Small Group Discussion (SGD)	02	02
Self-directed learning (SDL)	-	04
Problem Based Learning (PBL)	02	04
Case Based Learning (CBL)	-	-
Revision	02	-
Assessment	06	-
TOTAL	**44**	**74**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos				
Nature of assessment	CO 1	CO 2	CO 3	CO 4
Sessional Examination 1	*	*		
Sessional Examination 2		*	*	
Assignment/Presentation				*
End Semester Examination	*	*	*	*
Laboratory examination	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Big Data: Principles and best practices of scalable real-time data systems - Nathan Marz and James Warren. Manning Publisher. 2. Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale – Tom White, O'Reilly Publication 4th Edition. 3. Spark: The Definitive Guide: Big Data Processing Made Simple – Bill Chambers, MateiZaharia, O'Reilly Publication 1st Edition. 4. http://static.druid.io/docs/druid.pdf, http://druid.io/docs/0.8.0/design/design.html 5. Big data architecture and patterns - IBM developerWorks. 6. http://www.ibm.com/developerworks/library/bd-archpatterns1/ 7. Big Data and Analytics -IBM developerWorks. http://www.ibm.com/developerworks/analytics/ 8. http://lambda-architecture.net/ 9. Apache HBase - http://hbase.apache.org/ 10. Apache Spark Streaming - https://spark.apache.org/streaming/ 11. SummingbirdMapReduce library - https://github.com/twitter/summingbird

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Applied Probability and Statistics
Course Code: MCL 601	**Course Instructor:**

Academic Year: 2020-2021	**Semester:** First Year, Semester 1
No of Credits: 3	**Prerequisites:** Basic algebra and calculus
Synopsis:	This course provides an introduction to fundamental concepts in probability and statistics that are essential for data science applications.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Understand and apply the basic principles of sampling.
CO 2:	Model random phenomena using random variables.
CO 3:	Calculate & interpret probability as a measure of quantifying uncertainty.
CO 4:	Construct Bayesian models for analysing practical problems.
CO 5:	Use sample information and perform hypothesis-test analysis using an appropriate statistical technique to explain attributes of a population.

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*										
CO 2	*	*	*								
CO 3	*	*	*	*				*			
CO 4		*	*	*		*		*			
CO 5		*	*	*		*				*	

Course content and outcomes:	
Content	***Competencies***
Unit 1: Counting, probability concepts, and conditional probability	
Multiplication rule; permutation; combination - Sampling: with/without replacement and order matters/does not matter - Binomial & multinomial coefficients - Distribution problems	At the end of the topic student should be able to: 1. Understand and apply the basic principles of sampling (C1, C3). 2. Understand and apply the basic principles of probability (C1, C3).

Set theory; sample space; outcomes; events - Frequency based definition of probability - Equally likely vs. not equally likely outcomes - Axioms of probability Conditional probability; probability tree model; chain rule - Decomposition and the law of total probability - Bayes' rule - intuition, dependence/independence of events.	3. Differentiate and relate frequency-based interpretation of probability to classical approach (C4). 4. Apply Bayesian principle for modelling practical problems (C5).
Unit 2: Random variables	
Modelling using discrete random variables: Bernoulli, geometric, binomial, negative binomial, hypergeometric, and Poisson distributions - Probability mass function and cumulative distribution function - Expectation and variance: discrete case - Modelling using continuous random variables: uniform, normal, log-normal, exponential, and beta distributions; probability density function - Expectation and variance: continuous case - Functions of random variables.	1. Understand and differentiate discrete and continuous random variables of practical interest (C2, C4). 2. Gain solid foundation in the mathematical aspects of random variables (C2). 3. Understand how to use random variables to model random phenomena (C4). 4. Compare and contrast practical applicability of random variables (C6).
Unit 3: Sampling and parameter estimation	
Population and sample - Statistic & sampling distribution - Sample mean	1. Differentiate population and sample (C4). 2. Describe population parameters using

and variance - Central limit theorem – intuition and applications Point estimation - Standard error - Interval estimation: interpretation of confidence interval - Hypothesis testing: p-values, significance level and their interpretations, application to analysis of one- /two-sample mean and paired data.	inferences drawn from a sample (C6). 3. Design and apply appropriate hypothesis tests for practical problems (C3). 4. Communicate and explain the results of hypothesis testing (C6).

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	30	60
Quiz	02	04
Small Group Discussion (SGD)	02	02
Self-directed learning (SDL)	-	04
Problem Based Learning (PBL)	02	04
Case Based Learning (CBL)	-	-
Revision	02	-
Assessment	06	-
TOTAL	**44**	**74**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination

Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Sessional Examination 1	*	*			
Sessional Examination 2		*	*	*	
Assignment/Presentation	*	*	*	*	*
End Semester Examination	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Introduction to Probability, Charles M. Grinstead, American Mathematical Society; 2nd Revised Edition 1997. Available online at https://open.umn.edu/opentextbooks/textbooks/introduction-to-probability 2. A First Course in Probability, Sheldon Ross, 9th Edition, Pearson Education India; 9th Edition, 2013. 3. Biostatistics Open Learning textbook – Online resource from University of Florida available at https://bolt.mph.ufl.edu/6050-6052/ 4. 4. All of Statistics: A Concise Course in Statistical Inference, Larry Wasserman – Springer.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Fundamentals of Machine Learning Lab
Course Code: BDA-601L	**Course Instructor:**

Academic Year: 2020 - 2021	**Semester:** First Year, Semester 1
No of Credits: 1	**Prerequisites:** Basics of Programming
Synopsis:	This Course provides insight on 1. This course provides the concept of machine learning, applications, techniques, design issues and approaches to machine learning. 2. This course provides the fundamental knowledge about concept learning, hypothesis and bias. 3. To implement machine learning algorithms such as Decision Tree learning, Probably Approximately Correct (PAC) learning, Bayesian learning, Instance-based learning, Principal Component Analysis (PCA) and Ensemble methods in real time data set for various analysis.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Identify the software and tools for designing machine learning applications.
CO 2:	Apply concept learning and hypothesis space.
CO 3:	Apply machine learning approach to reduce the dimension.
CO 4:	Analyse different machine learning algorithms.
CO 5:	Design ensemble methods.

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*										
CO 2		*									
CO 3			*								
CO 4				*							
CO 5				*							

Course content and outcomes:	
Content	***Competencies***
Unit 1: Introduction	
Definition of Machine Learning Goals and applications of machine learning Basic design issues and approaches to machine learning Types of machine learning techniques	At the end of the topic student should be able to: 1. Identify programming environments available for the machine learning (C1) 2. Classify the pros and cons of various environments for ML coding (C2)
Unit 2: Inductive Classification	
The concept learning task. Concept learning as search through a hypothesis space. General-to-specific ordering of hypotheses. Finding maximally specific hypotheses. Version spaces and the candidate elimination algorithm. Inductive bias.	1. Design a machine learning model to get a Maximally Specific Hypothesis for the given training examples (C5). 2. Construct a machine learning model to obtain most general and most specific hypotheses for the given training examples (C5)
Unit 3: Decision Tree learning	
Representing concepts as decision trees. Recursive induction of decision trees. Picking the best splitting attribute Entropy and information gain. Searching for simple trees and computational complexity.	1. Develop a machine learning classifier using decision tree and random forest (C5) 2. Examine different applications of decision tree and random forest (C4)
Unit 4: Computational learning theory	
Models of learnability: learning in the limit. Probably Approximately Correct (PAC)	1. Design a learning method to determine the sample complexity of training examples (C5)

learning. Sample complexity: quantifying the number of examples needed to PAC learn. Computational complexity of training. Sample complexity for finite hypothesis spaces. Noise. Learning Multiple Classes. Bias-variance trade-off, under-fitting and over-fitting concepts.	2. Analyse bias-variance trade-off, under-fitting and over-fitting concepts (C4)
Unit 5: Bayesian learning	
Probability theory and Bayes rule. Naive Bayes learning algorithm - Parameter smoothing. Generative vs. discriminative training Logistic regression. Bayes nets and Markov nets for representing dependencies	1. Design a machine learning model using Bayes learning (C5). 2. Develop a machine learning classifier models using different approach (C5) 3. Design Bayes nets and Markov nets for representing dependencies (C5)
Unit 6: Instance-based learning	
Constructing explicit generalizations versus comparing to past specific examples. K-Nearest Neighbour learning algorithm. Case-based reasoning (CBR) learning.	1. Design machine learning models to classify the instances using K-NN and CBR approaches (C5).
Unit 7: Continuous Latent Variables	
Principal Component Analysis (PCA), Applications of PCA	1. Apply PCA for different complex applications (C3)
Unit 8: Ensemble methods (bagging and boosting)	
Using committees of multiple	1. Design a Bayesian Networks (C5)

hypotheses. Bagging Boosting DECORATE Active learning with ensembles.	2. Develop machine learning models using Ensemble models. (C5)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	12	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	-	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	03	-
Clinic	-	-
Practical	24	-
Revision	03	-
Assessment	06	-
TOTAL	**48**	**-**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos

Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Sessional Examination 1	*	*			
Sessional Examination 2			*	*	
Assignment/Presentation	*	*	*	*	*
Laboratory Examination	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Machine Learning, T. Mitchell, McGraw-Hill, 1997 2. Machine Learning, E. Alpaydin, MIT Press, 2010 3. Pattern Recognition and Machine Learning, C. Bishop, Springer, 2006 4. Pattern Classification, R. Duda, E. Hart, and D. Stork, Wiley-Interscience, 2000 5. T. Hastie, R. Tibshirani and J. Friedman, The Elements of Statistical Learning: Data Mining, 6. Inference and Prediction. Springer, 2nd Edition, 2009 7. Machine Learning for Big Data, Jason Bell, Wiley Big Data Series 8. Multidimensional Neural Networks Unified Theory, Rama Murthy G 9. Current literature

Name of the Program:	ME (Big Data and Data Analytics)

Course Title:	Algorithms and Data Structures for Big Data Lab
Course Code: BDA 602L	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First year, First semester
No of Credits: 1	**Prerequisites:** Programming in C or Python
Synopsis:	1. Students learn how to design optimal algorithms with respect to time and space 2. Students learn how to implement link list, stack, queues, searching and sorting techniques, sets, trees and graphs. 3. Students learn to implement string and text processing techniques. 4. Students learn to implement Data stream algorithms.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Evaluate the performance of Algorithms
CO 2:	Develop applications using suitable data structures
CO 3:	Design applications using Data streaming and pattern matching algorithms

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*	*	*						*		
CO 2	*	*	*			*					
CO 3	*	*	*		*	*					

Course content and outcomes:

Content	***Competencies***
Unit 1: Elementary Data Structures	
Linked List, Stacks, Queues, Sorting and Searching Techniques	At the end of the topic student should be able to: 1. Implement Linked list, Stacks, Queues (C4). 2. Design applications using various

	searching and sorting techniques.
Unit 2: Tree, Sets and Hash Table	
Binary Tree, Binary search tree Sets and Hash Tables	1. Implement Binary Tree and BST (C4) Design applications using Hash Tables
Unit 3: Graph	
Representation of Graph BFS and DFS Shortest path algorithms	1. Implement Graph and its traversals (BFS, DFS) (C4). 2. Design applications with shortest path algorithms (C4)
Unit 4: Pattern Matching and Data streaming	
	1. Implement pattern matching algorithms.

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	12	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	-	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	03	-
Clinic	-	-
Practical	24	-
Revision	03	-
Assessment	06	-
TOTAL	**48**	**-**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos			
Nature of assessment	CO 1	CO 2	CO 3
Sessional Examination 1	*	*	
Sessional Examination 2		*	*
Assignment/Presentation	*	*	*
End Semester Examination	*	*	*
Laboratory Examination	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Data Structures and Algorithms in Python - Michael T. Goodrich, Roberto Tamassia, and Michael H. Goldwasser. John Wiley & Sons. 2. Data Streams: Algorithms and Applications - S. Muthukrishnan. Foundations and Trends in Theoretical Computer Science archive, Volume 1 Issue 2, August 2005, Pages 117 – 236.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Architecture of Big Data Systems Lab
Course Code: BDA 623L	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First year, First semester 1
No of Credits: 1	**Prerequisites:** Programming in Python, Java
Synopsis:	1. This course helps the student to understand the concepts of Hadoop framework, Spark framework and their internals. 2. This course helps the students to learn Map-reduce programming, Spark programming. 3. This course helps the students to build machine learning model using Spark framework.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Install and develop applications using Hadoop and its ecosystems
CO 2:	Build applications using Spark frame work
CO 3:	Build Machine Learning models using Spark

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*	*	*		*	*			*	*	
CO 2	*	*	*		*	*			*	*	
CO 3	*	*	*		*	*			*	*	

Course content and outcomes:

Content	***Competencies***
Unit 1: Hadoop ecosystem	
Installation and configuring Hadoop ecosystem	At the end of the topic student should be able to: 1. Practice applications in HDFS and YARN. (C3)

	2. Practice applications using Sqoop, Hive, PIG. (C3) 3. Compute programs using MapReduce. (C3)
Unit 2: Spark Framework	
Spark tool chain – RDD, DataFrame, SQL and Streaming	1. Develop applications using Spark DataFrame and SQL (C4). 2. Design real time applications using Spark Streaming (C4).
Unit 3: Machine Learning using Spark	
MLIB	1. Compute machine learning models using Spark. (C3)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	12	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	-	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	03	-
Clinic	-	-
Practical	24	-
Revision	03	-
Assessment	06	-
TOTAL	**48**	**-**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos			
Nature of assessment	CO 1	CO 2	CO 3
Sessional Examination 1	*		
Sessional Examination 2		*	*
Assignment/Presentation	*	*	*
End Semester Examination	*	*	*
Laboratory Examination	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Big Data: Principles and best practices of scalable real-time data systems - Nathan Marz and James Warren. Manning Publisher. 2. Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale – Tom White, O'Reilly Publication 4th Edition. 3. Spark: The Definitive Guide: Big Data Processing Made Simple – Bill Chambers, MateiZaharia, O'Reilly Publication 1st Edition.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Applied Probability and Statistics Lab
Course Code: MCL 601L	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 1
No of Credits: 1	**Prerequisites:** MCL 601
Synopsis:	This course provides a hands-on introduction to fundamental concepts in probability and statistics that are essential for data science applications using the R programming language.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Apply the basic principles of sampling to practical problems.
CO 2:	Visualize probability concepts through frequency-based interpretations.
CO 3:	Simulate discrete and continuous random variables for modelling random phenomena.
CO 4:	Design and apply hypothesis tests followed by interpretation of results.
CO 5:	Interpret statistical results and communicate them unambiguously and effectively.

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*	*	*		*						
CO 2		*	*		*						
CO 3	*	*	*	*	*						
CO 4		*	*	*	*	*					
CO 5				*	*	*		*		*	

Course content and outcomes:	
Content	***Competencies***
Unit 1: Counting, probability concepts, and conditional probability	

Multiplication rule; permutation; combination - Sampling: with/without replacement and order matters/does not matter - Binomial & multinomial coefficients - Distribution problems Set theory; sample space; outcomes; events - Frequency based definition of probability - Equally likely vs. not equally likely outcomes - Axioms of probability Conditional probability; probability tree model; chain rule - Decomposition and the law of total probability - Bayes' rule - intuition, dependence/independence of events.	At the end of the topic student should be able to: 1. Understand the basic principles of the R programming language (C1). 2. Develop short code snippets to understand the basic principles of sampling and probability (C1, C3). 3. Visualise and interpret probability concepts through a frequency-based approach (C6). 4. Program and analyse Bayesian models for practical problems (C4).
Unit 2: Random variables	
Modelling using discrete random variables: Bernoulli, geometric, binomial, negative binomial, hypergeometric, and Poisson distributions - Probability mass function and cumulative distribution function - Expectation and variance: discrete case - Modelling using continuous random variables: uniform, normal, log-normal, exponential, and beta distributions; probability density function - Expectation and variance: continuous case - Functions of random variables.	1. Understand and apply R functions to simulate discrete and continuous random variables (C3). 2. Using sampling, compute and interpret different attributes of random variables (C4). 3. Visualise and interpret histograms and probability mass/density functions of random variables using state of the art visualisation libraries in R (C4). 4. Develop codes to model random phenomena using appropriate random variables (C5).

Unit 3: Sampling and parameter estimation	
Population and sample - Statistic & sampling distribution - Sample mean and variance - Central limit theorem – intuition and applications Point estimation - Standard error - Interval estimation: interpretation of confidence interval - Hypothesis testing: p-values, significance level and their interpretations, application to analysis of one- /two-sample mean and paired data	1. Visualise sample data through histograms (C3). 2. Compute estimates of population parameters using samples and communicate the uncertainty in the estimates (C4). 3. Use R in-built functions for performing hypothesis tests (C4). 4. Interpret and communicate the results of hypothesis tests (C6).

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	12	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	-	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	03	-
Clinic	-	-
Practical	24	-
Revision	03	-
Assessment	06	-
TOTAL	**48**	**-**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Sessional Examination 1	*	*			
Sessional Examination 2			*	*	
Assignment/Presentation	*	*	*	*	*
Laboratory examination	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Introduction to Probability, Charles M. Grinstead, American Mathematical Society; 2nd Revised Edition 1997. Available online at https://open.umn.edu/opentextbooks/textbooks/introduction-to-probability 2. A First Course in Probability, Sheldon Ross, 9th Edition, Pearson Education India; 9th Edition, 2013. 3. Biostatistics Open Learning textbook – Online resource from University of Florida available at https://bolt.mph.ufl.edu/6050-6052/ 4. All of Statistics: A Concise Course in Statistical Inference, Larry Wasserman – Springer.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Mini Project - 1
Course Code: BDA 695	**Course Instructor:**
Academic Year: 2020 - 2021	**Semester:** First Year, Semester 1
No of Credits: 4	**Prerequisites:** Any programming language and circuit basics

Synopsis:	Students are expected to select a problem in the area of their interest and the area of their specialization that would require an implementation in hardware / software or both in a semester
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Apply the objectives of the project work and provide an adequate background with a detailed literature survey
CO 2:	Breakdown the project into sub blocks with sufficient details to allow the work to be reproduced by an independent researcher
CO 3:	Compose hardware/software design, algorithms, flowchart, methodology, and block diagram
CO 4:	Evaluate the results
CO 5:	Summarize the work carried out

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1				*							
CO 2					*			*			
CO 3							*			*	
CO 4						*					*
CO5:							*				

Course content and outcomes:	
Content	***Competencies***
Phase 1	
Problem identification, synopsis submission, status submission, mid evaluation.	At the end of the topic student should be able to: 1. Identify the problem/specification (C1) 2. Discuss the project (C2) 3. Prepare the outline (C3) 4. Describe the status of the project (C2) 5. Prepare a mid-term project presentation report (C3) 6. Prepare and present mid-term project presentation slides (C3, C5) 7. Develop project implementation in hardware/software or both in chosen platform (C5)
Phase 2	
Status submission, final evaluation.	1. Prepare the progress report (C3) 2. Prepare the final project presentation report (C3) 3. Prepare and present final project presentation slides (C3, C5) 4. Modify and Develop implementation in hardware/software or both in chosen platform (C3, C5) 5. Justify the methods used and obtained results (C6)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	-	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	48	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	-	-
Clinic	-	-
Practical	-	-
Revision	-	-
Assessment	03	-
TOTAL	**51**	**09**

Assessment Methods:	
Formative:	**Summative:**
Project Problem Selection	Mid-Term Presentation
Synopsys review	Second status review
First status review	Demo & Final Presentation

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Mid Presentation	*	*			
Presentation	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	Particular to the chosen project

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Seminar - 1
Course Code: BDA 697	**Course Instructor:**
Academic Year: 2020 - 2021	**Semester:** First Year, Semester 1
No of Credits: 1	**Prerequisites:** Communication Skill
Synopsis:	1. To select, search and learn technical literature. 2. To Identify a current and relevant research topic. 3. To prepare a topic and deliver a presentation. 4. To develop the skill to write a technical report. 5. Develop ability to work in groups to review and modify technical content.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Show competence in identifying relevant information, defining and explaining topics under discussion.
CO 2:	Show competence in working with a methodology, structuring their oral work, and synthesizing information.
CO 3:	Use appropriate registers and vocabulary, and will demonstrate command of voice modulation, voice projection, and pacing.
CO 4:	Demonstrate that they have paid close attention to what others say and can respond constructively.
CO 5:	Develop persuasive speech, present information in a compelling, well-structured, and logical sequence, respond respectfully to opposing ideas, show depth of knowledge of complex subjects, and develop their ability to synthesize, evaluate and reflect on information.

Mapping of COs to POs											
COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*							*	*		*
CO 2	*							*	*		*
CO 3	*							*	*		*
CO 4	*							*	*		*
CO5:	*							*	*		*

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	-	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	14	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	-	-
Clinic	-	-
Practical	-	-
Revision	-	-
Assessment	-	-
TOTAL	**14**	**-**

Assessment Methods:	
Formative:	**Summative:**
Seminar Topic Selection	
Synopsys review	
PPT Review	

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Presentation	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	Particular to the chosen Seminar

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Mobile Web Application Development
Course Code: CSE-625	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 1
No of Credits: 3	**Prerequisites:** Basic knowledge of OOP's concepts, Java programming language
Synopsis:	1. This course would provide fundamental knowledge about android platform. 2. The course will also provide skill sets to design and develop secured mobile web applications. 3. This course will provide basic knowledge about programming for technologies available on smart phones.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Discuss the challenges of mobile web application development.
CO 2:	Apply HTML5, CSS, javascript and DOM API's in web application development.
CO 3:	Use of programming for technologies available on smart phones.
CO 4:	Design and develop secure mobile web applications.

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*	*PO 12*
CO 1	*	*										
CO 2		*	*		*							
CO 3		*	*		*							
CO 4		*	*	*	*							

Course content and outcomes:	
Content	*Competencies*
Unit 1: Challenges of mobile Web application development	
The limitations of mobile networks. Reducing the page weight – the amount of markup and external elements. Avoiding useless network usage. Understanding the "mobile-first" design principles. Limitations imposed by battery life.	At the end of the topic student should be able to: 1. Discuss limitation of mobile networks (C2) 2. Explain way of reducing the page weight (C2) 3. Discuss limitations imposed by battery life (C2)
Unit 2: Setting up a personal Web site	
Setting free VMs - micro-instances - on AWS. Installing and configuring NGINX on AWS micro instances. Working with routing and reverse proxies HTTP and REST APIs	1. Set up VMs on AWS (C4) 2. Discuss HTTP and REST API's (C2) 3. Solve issues on installation and configuration NGIX on AWS (C3)
Unit 3: HTML5 and CSS for mobile devices.	
Media queries for handling mobile form-factors. Principles and practice of responsive design. Mobile UX, Viewport, Fluid design and responsive images	1. Discuss principles of responsive design (C2) 2. Use of media queries for handling mobile form-factors (C3)
Unit 4: Programming with JavaScript and DOM APIs	
Accessing document fragments. Using jQuery and other light-weight libraries. AJAX and asynchronous programming	1. Explain jQuery and other light weight libraries (C2) 2. Discuss AJAX and asynchronous programming (C2)

Unit 5: Architecture of Android applications	
Android application framework, core libraries, android runtime, Linux kernel	1. Discuss android application framework (C2) 2. Explain android runtime system and core libraries (C2)
Unit 6: Programming for technologies available on smart phones	
Introduction to PhoneGap, Handling Touch events. Making use of the accelerometer and the Location APIs. Accessing camera and media devices	1. Use of PhoneGap in mobile web applications (C3) 2. Practice accelerometer and location APIs (C3)
Unit 7: Developing offline facilities in mobile web applications	
Localstorage and IndexDB APIs	1. Discuss Localstorage and IndexDB APIs (C2)
Unit 8: Designing and developing secure mobile web applications	
Understanding the single-origin policy, Dangers of Cross-site scripting Principles of the secure socket layer and HTTPS Practical encryption for client-server communication in Web applications. Best practices in developing secure client-side code	1. Discuss different encryptions techniques used in client-server communication (C2) 2. Practice best practices in developing secure client-side code (C3)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	30	60
Quiz	02	04
Small Group Discussion (SGD)	02	02
Self-directed learning (SDL)	-	04
Problem Based Learning (PBL)	02	04
Case Based Learning (CBL)	-	-
Revision	02	-
Assessment	06	-
TOTAL	**44**	**74**

Assessment Methods:	
Formative:	**Summative:**
Theory Assignment	Sessional Examination
Lab Assignment	University End Semester Examination
Lab Test	Viva
Viva	

Mapping of assessment with Cos				
Nature of assessment	CO 1	CO 2	CO 3	CO 4
Sessional Examination 1	*	*		
Sessional Examination 2			*	*
Assignment/Presentation		*	*	*
End Semester Examination	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Learning Web App Development (Build Quickly with Proven JavaScript Techniques) - Semmy Purewal. O'Reilly Media. 2014. MSOIS, MAHE, Manipal 13 2. The Browser Security Handbook. Michal Zalewski. https://code.google.com/p/browsersec/wiki/Main 3. High Performance Responsive Design - Tom Barker. O'Reilly publisher. 2014. 4. Apple UI Design Basics. https://developer.apple.com/library/ios/documentation/UserExperience/Conceptual/MobileHIG/i ndex.html 5. Android Design Principles. https://developer.android.com/design/index.html 6. Android Application Development Reference. https://developer.android.com/develop/index.html

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	**Principles of Data Visualization**
Course Code: BDA 622	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 1
No of Credits: 3	**Prerequisites:** Programming in Python
Synopsis:	This Course provides insight on 1. This course introduces data visualization, the art and science of turning data into readable graphics. 2. Teach how to design and create data visualizations based on data available and tasks to be achieved. 3. Students learn how do data extraction, data modelling and data processing. 4. Students learn to map data attributes to graphical attributes, and strategic visual encoding based on known properties of visual perception
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Extracting, transforming and storing data from various data sources.
CO 2:	An understanding of the key techniques and theory used in visualization, including data models, graphical perception and techniques for visual encoding and interaction.
CO 3:	Exposure to number of common data domains and corresponding analysis tasks.
CO 4:	Practical experience building and evaluating visualization systems.
CO 5:	The ability to read and discuss research papers from the visualization literature.

Mapping of COs to POs											
COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*		*		*	*					
CO 2	*	*			*						
CO 3	*	*	*								
CO 4	*		*		*			*			
CO 5	*	*	*	*							

Course content and outcomes:	
Content	***Competencies***
Unit 1: Introduction to Web scraping	
Web scraping models and techniques, Case study: BeautifulSoup, Scrapy, Selenium	At the end of the topic student should be able to: 1. Understanding various formats of data. (C1) 2. Design programs to dynamically extract data from web. (C4) 3. Design programs to read data from various data sources. (C4)
Unit 2: Data Analysis	
Data structures for analysis: numpy, pandas Data Wrangling: Clean, Transform, Merge, Reshape Data Aggregation and Group Operations Case study: Exploratory analysis of	1. Understand and integrate various data structures for data analysis process (C2). 2. Create various techniques to clean and handle missing data (C4). 3. Design data filtering and transformation techniques (C4).

public / scrapped datasets	
Unit 3: Data Visualization	
Data Visualization – classification, infographics versus data visualization, visualization for supporting exploratory data analysis, visual art, choosing appropriate visual encodings, rules for visualization - Visualization techniques: time series, statistical distributions, maps - Data visualization for web	1. Describe what is the purpose of Visualization. (C2) 2. Describe various ways of classifying visualization. (C2) 3. Explain what is explorative and explanative visualization. (C2) 4. Differentiate data visualization and visual art. (C2) 5. Create visualization for time series data. (C4) 6. Create visualization for statistical distributions. (C4) 7. Create visualization for maps, Hierarchical data and network data. (C4)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	30	60
Quiz	02	04
Small Group Discussion (SGD)	02	02
Self-directed learning (SDL)	-	04
Problem Based Learning (PBL)	02	04
Case Based Learning (CBL)	-	-
Revision	02	-
Assessment	06	-
TOTAL	**44**	**74**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Sessional Examination 1	*	*			
Sessional Examination 2			*	*	
Assignment/Presentation	*	*	*	*	*
End Semester Examination	*	*	*	*	*
Laboratory examination	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. **Website Scraping with Python: Using BeautifulSoup and Scrapy,** Gábor & Hajba, APRESS Publications, 1st Edition, 2018. 2. **Web Scraping with Python: Collecting More Data from the Modern Web**, Ryan Mitchell Shroff, O'Reilly, 2nd Edition, 2018. 3. **Designing Data Visualizations**, Julie Steele and Noah Iliinsky; O'Reilly Media; 1st Edition, 2011. 4. **Python for Data Analysis**, Wes McKinney; Shroff; O'Reilly; 2nd Edition, 2018.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Mobile Web Application Development Lab
Course Code: CSE-625L	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 1
No of Credits: 1	**Prerequisites:** Basic knowledge of OOP's concepts, Java programming language
Synopsis:	1. This course would provide fundamental knowledge about android platform. 2. The course will also provide skill sets to design and develop secured mobile web applications. 3. This course will provide basic knowledge about programming for technologies available on smart phones.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Solve issues related to mobile web application development
CO 2:	Apply HTML5, CSS, javascript and DOM API's in web application development
CO 3:	Use of programming for technologies available on smart phones
CO 4:	Construct secure mobile web applications

Mapping of COs to POs

Cos	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*	*PO 12*
CO 1		*	*		*							
CO 2		*	*		*							
CO 3		*	*		*							
CO 4		*	*	*	*							

Course content and outcomes:	
Content	*Competencies*
Unit 1: Installation of Android Studio	
Installation of Android Studio, environment setting, Project creation, building a project, running a sample project	At the end of the topic student should be able to: 1. Identify different features in android studio (C1) 2. Explain Android manifest file (C2) 3. Discuss DVM, DDMS, android emulator (C2) 4. Discuss the issues related running and debugging applications (C2)
Unit 2: Introduction to HTML5 and CSS for mobile devices	
Implementation of mobile web applications using HTML5 and CSS	1. Practice to create more responsive mobile web applications (C3) 2. Develop android applications using Mobile UX, Viewport, Fluid design and responsive images (C4)
Unit 3: Programming with JavaScript and DOM APIs	
Using jQuery and other light-weight libraries. AJAX and asynchronous programming	1. Develop mobile web applications using JQuery and other light-weight libraries (C4) 2. Apply AJAX and asynchronous programming techniques in mobile web applications(C3)
Unit 4: Programming for technologies available on smart phones	
Introduction to PhoneGap, Handling Touch events. Making use of the	1. Write mobile web applications using PhoneGap techniques (C3)

accelerometer and the Location APIs. Accessing camera and media devices.	2. Apply touch events in mobile web applications (C3) 3. Demonstrate the use of accelerometer and location APIs in mobile web applications (C3)
Unit 5: Designing and developing secure mobile web applications	
Practical encryption for client-server communication in Web applications. Best practices in developing secure client-side code	1. Use of various encryption techniques for communications in mobile web applications (C3)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	12	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	-	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	03	-
Clinic	-	-
Practical	24	-
Revision	03	-
Assessment	06	-
TOTAL	**48**	**-**

Assessment Methods:	
Formative:	**Summative:**

Theory Assignment	Sessional Examination
Lab Assignment	University End Semester Examination
Lab Test	Viva
Viva	

Mapping of assessment with Cos				
Nature of assessment	CO 1	CO 2	CO 3	CO4
Sessional Examination 1	*	*		
Sessional Examination 2		*	*	*
Assignment/Presentation	*	*		*
End Semester Examination				
Laboratory examination	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Learning Web App Development (Build Quickly with Proven JavaScript Techniques) - Semmy Purewal. O'Reilly Media. 2014. MSOIS, MAHE, Manipal 13 2. The Browser Security Handbook. Michal Zalewski. https://code.google.com/p/browsersec/wiki/Main 3. High Performance Responsive Design - Tom Barker. O'Reilly publisher. 2014. 4. Apple UI Design Basics. https://developer.apple.com/library/ios/documentation/UserExperience/Conceptual/MobileHIG/i ndex.html 5. Android Design Principles. https://developer.android.com/design/index.html 6. Android Application Development Reference. https://developer.android.com/develop/index.html

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Principles of Data Visualization Lab
Course Code: BDA-622L	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First year, semester 1
No of Credits: 1	**Prerequisites:** Programming in Python
Synopsis:	This Course provides insight on 1. This course introduces data visualization, the art and science of turning data into readable graphics. 2. Teach how to design and create data visualizations based on data available and tasks to be achieved. 3. Students learn how do data extraction, data modelling and data processing. 4. Students learn to map data attributes to graphical attributes, and strategic visual encoding based on known properties of visual perception.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Data scrapping from different data sources.
CO 2:	Data Cleaning, transformations and Analysis.
CO 3:	Data Visualization using different techniques, tools and charts.

Mapping of COs to POs											
COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*	*	*		*				*	*	
CO 2	*	*	*		*	*		*	*	*	
CO 3	*	*	*	*	*	*		*		*	

Course content and outcomes:	
Content	***Competencies***
Unit 1: Data Scrapping	
Web scrapping models Installing and configuring tools to handle different data types.	At the end of the topic student should be able to: 1. Identify different types of data sources (C2). 2. Design applications to scrap static data (C4). 3. Design applications to extract data from dynamic web pages (C4).
Unit 2: Data Analysis	
Working with packages like numpy, pandas, sklearn Perform exploratory data analysis.	1. Design scripts to clean, handle missing data (C4). 2. Design scripts to apply required transformations to cleaned data (C4).
Unit 3: Data Visualization	
Creating different types of Visualization. Creating different types of charts.	1. Develop applications for exploratory data visualization (C4). 2. Develop scripts to create static visualization using various visual encodings (C4). 3. Create dynamic visualization for web (C4).

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	12	-
Seminar	-	-
Quiz	-	-

Small Group Discussion (SGD)	-	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	03	-
Clinic	-	-
Practical	24	-
Revision	03	-
Assessment	06	-
TOTAL	**48**	**-**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos			
Nature of assessment	CO 1	CO 2	CO 3
Sessional Examination 1	*		
Sessional Examination 2		*	*
Assignment/Presentation	*	*	*
End Semester Examination	*	*	*
Laboratory Examination	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. **Website Scraping with Python: Using BeautifulSoup and Scrapy,** Gábor & Hajba, APRESS Publications, 1st Edition, 2018. 2. **Web Scraping with Python: Collecting More Data from the Modern Web**, Ryan Mitchell Shroff, O'Reilly, 2nd Edition,

	2018. 3. **Designing Data Visualizations**, Julie Steele and Noah Iliinsky; O'Reilly Media; 1st Edition, 2011. 4. **Python for Data Analysis**, Wes McKinney; Shroff; O'Reilly; 2nd Edition, 2018.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Machine Learning for Big Data
Course Code: BDA 605	**Course Instructor:**
Academic Year: 2020 - 2021	**Semester:** First Year, Semester 2
No of Credits: 3	**Prerequisites:** Programming with Python and Data Visualization
Synopsis:	This Course provides insight on 1. This course provides the concept of neurons and biological motivation, activation functions and threshold units, supervised and unsupervised learning, perceptron network models in Artificial Neural Networks. 2. This course provides the knowledge about learning from unclassified data using clustering techniques. 3. This course provides the concept of Support Vector Machines for linear and non-linear classification. 4. This course provides the concept of Deep Learning and design of convolutional neural network for Deep Learning. 5. This course provide the knowledge about the applications and design of Reinforcement Learning algorithms.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Describe activation functions, weights and threshold units used in artificial neural networks, supervised and unsupervised learning, gradient descent approach, types of perceptron models, overfitting

CO 2:	Explain the concept of hierarchical clustering and non-hierarchical clustering, support vector machine, deep neural networks and reinforcement learning
CO 3:	Demonstrate artificial neural network models, clustering models, support vector classifier models, Deep learning models and reinforcement learning models
CO 4:	Compare and contrast single layer, multilayer and deep neural networks in terms of accuracy in classification
CO 5:	Design back propagation neural network, K-means and agglomerative clustering, deep neural network, reinforcement learning models and selection of a machine learning algorithm for the given data analysis.

Mapping of COs to POs											
COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*										
CO 2		*									
CO 3			*								
CO 4				*							
CO 5				*							

Course content and outcomes:	
Content	***Competencies***
Unit 1: Artificial Neural Networks	
Neurons and biological motivation, Activation functions and threshold units, Supervised and unsupervised learning, Perceptron Model: representational limitation and gradient descent training, Multilayer	At the end of the topic student should be able to: 1. Relate biological neurons with artificial neurons and the motivation for ANN development. (C1) 2. Distinguish Supervised and unsupervised learning (C2).

networks and back propagation, Overfitting	3. Describe about error reduction techniques in used Artificial Neural Networks based learning (C2) 4. Write the usability of different activation functions for ANN learning system. (C3) 5. Describe the architecture of various perceptron networks. (C2)
Unit 2: Clustering	
Learning from unclassified data, Clustering. Hierarchical Agglomerative Clustering, Non- Hierarchical Clustering - k-means partitional clustering, Expectation maximization (EM) for soft clustering, Semi-supervised learning with EM using labelled and unlabelled data.	1. Write the different methods of learning from unclassified data (C3). 2. Explain the operations of various clustering models in machine learning (C5) 3. Describe the methods used for measuring dissimilarity between two clusters. (C2) 4. Apply clustering techniques for data analysis. (C3)
Unit 3: Kernel Methods	
Dual Representations, Design of Kernels .	1. Describe Dual Representations. (C2) 2. Explain the Kernel trick for learning non-linear functions (C5)
Unit 4: Support Vector Machines (SMV)	
Maximum margin linear separators, Quadratic programming solution to finding maximum margin separators, Kernels for learning non-linear functions, Varying length pattern classification using SVM	1. Describe about Maximum Margin and Support Vector Machine. (C2) 2. Examine the advantages of maximum margin linear separators technique in SVM (C4) 3. Explain the Kernel trick for learning non-linear functions (C5)

	4. Show the relation between two forms of representation of a hyperplane (C3)
Unit 5: Deep Learning	
Introduction to Deep Learning, Introduction to convolutional Neural Network (CNN), CNN Architecture and layers, Building simple CNN model for classification, Training and Testing the CNN model	1. Define Deep Learning. (C1) 2. Describe the applications of deep learning. (C2) 3. Explain the architecture of Deep Neural Network and CNN (C5) 4. Design a classifier for the image classification system. (C5)
Unit 6: Reinforcement Learning	
Characteristics, N-arm Bandit Problem, Calculating the Value Function, Associative Learning – Adding States, The Markov Property & Markov Decision Process	1. Explain the concept of Multi-Armed Bandit Problem (MABP). (C2) 2. Write the functions of Upper Confidence Bound (UCB) algorithm. (C3) 3. Outline the learning process and characteristics of reinforcement learning. (C4) 4. Explain about Markov decision process. (C5)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	30	60
Quiz	02	04
Small Group Discussion (SGD)	02	02
Self-directed learning (SDL)	-	04
Problem Based Learning (PBL)	02	04

Case Based Learning (CBL)	-	-
Revision	02	-
Assessment	06	-
TOTAL	**44**	**74**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Sessional Examination 1	*	*	*		*
Sessional Examination 2	*	*	*	*	*
Assignment/Presentation	*	*	*	*	*
End Semester Examination	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. T. Mitchell, “Machine Learning”, McGraw-Hill, 1997. 2. E. Alpaydin, “Machine Learning”, MIT Press, 2010. 3. C. Bishop, “Pattern Recognition and Machine Learning”, Springer, 2006. 4. R. Duda, E. Hart, and D. Stork, “Pattern Classification”, Wiley Interscience, 2000. 5. Satish Kumar, “Neural Networks - A Class Room Approach”, Second Edition, Tata McGraw-Hill, 2013.

	6. T. Hastie, R. Tibshirani and J. Friedman," The Elements of Statistical Learning: Data Mining", Inference and Prediction, Springer, 2nd Edition, 2009. 7. Jason Bell, "Machine Learning for Big Data", Wiley Big Data Series, 2016. 8. J. Shawe-Taylor and N. Cristianini, "Kernel Methods for Pattern Analysis", Cambridge University Press, 2004. 9. S. Haykin, "Neural Networks and Learning Machines", Prentice Hall of India, 2010. 10. Rama Murthy G, "Multidimensional Neural Networks Unified Theory", New Age International, 2008. 11. F. Camastra and A. Vinciarelli, "Machine Learning for Audio, Image and Video Analysis – Theory and Applications", Springer, 2008.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Modern Database for Big Data
Course Code: BDA 616	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 2
No of Credits: 3	**Prerequisites:** Databases with SQL Queries
Synopsis:	This Course provides insight on Basic MapReduce Partitioning and combining, Key-Value Databases, Document Databases, Column-Family Stores, Graph Databases
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Examine different types of data.
CO 2:	Design queries to handle different data types.
CO 3:	Explain different data models.
CO 4:	Explain the concepts of map reduce in handling of data.

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*		*		*						
CO 2	*	*	*		*						
CO 3	*		*								
CO 4	*		*		*						

Course content and outcomes:

Content	***Competencies***

Unit 1: Introduction	
Introduction to growth of traditional and modern database systems	At the end of the topic student should be able to: 1. Analyze traditional and modern database systems (C4)
Unit 2: SQL	
Syntax and Semantics	1. Design various SQL queries (C5)
Unit 3: NoSQL Database	
Why NoSQL? - Data Models	1. Limitations of traditional databases.(C2) Various Data Models to handle huge amount of data. (C2)
Unit 4: Distribution models for scalability	
Horizontal partitioning. - Data sharding. -Master-slave replication. Peer-to-peer replication. - Version stamps – business and system transactions.	1. Understand different distribution models for data Scalability (C3) 2. Describe achieve Data shardinng? (C2)
Unit 5: Consistency Models	
Update consistency, Read Consistency, CAP Theorem	1. Understating Data consistency techniques (C2). 2. Implementing CAP theorem (C3).
Unit 6: MapReduce	
Basic MapReduce Partitioning and combining. - Composing MapReduce calculations. - Two-stage map-reduce example. Incremental MapReduce.	1. Understanding MapReduce technique (C2). 2. Design applications using suitable MapReduce techniques (C4).
Unit 7: Case study	
Key-Value Databases - Document Databases - Column-Family Stores - Graph Databases	1. Design applications using different types of databases (C4).
Unit 8: Beyond NoSQL	

File systems, Event sourcing, Memory Image, Version control, XML Database, Object Database	1. Alternate techniques to NoSQL databases (C4)
Unit 9:	
Choosing your database	1. Design steps to choose proper databases based on user requirements (C4).

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	30	60
Quiz	02	04
Small Group Discussion (SGD)	02	02
Self-directed learning (SDL)	-	04
Problem Based Learning (PBL)	02	04
Case Based Learning (CBL)	-	-
Revision	02	-
Assessment	06	-
TOTAL	**44**	**74**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos				
Nature of assessment	CO 1	CO 2	CO 3	CO 4
Sessional Examination 1	*	*		

Sessional Examination 2		*	*	*
Assignment/Presentation	*	*	*	*
End Semester Examination	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. **Database System Concepts**, Avi Silberschatz, Henry F. Korth, and S. Sudarshan. McGraw Hill, 6th Edition, 2010. 2. **NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence**, Pramod J. Sadalage, Martin Fowler, Addison-Wesley, 2012. 3. **Seven Databases in Seven Weeks: A Guide to Modern Databases and the NoSQL Movement**, Eric Redmond, Jim R. Wilson, Pragmatic Bookshelf. 2012.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Advanced Applications of Probability and Statistics
Course Code: MCL 602	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 2
No of Credits: 3	**Prerequisites:** MCL 601
Synopsis:	This course provides an introduction to advanced applications of probability and statistics for multivariate and time series data.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Compute and interpret descriptive statistics for multivariate data
CO 2:	Apply linear and logistic regression models for practical problems and assess model performance
CO 3:	Interpret the output of principal component analysis (PCA) applied to multivariate data for dimension reduction
CO 4:	Identify multivariate data with mixed data type features and cluster using an appropriate technique
CO 5:	Understand the basics of time series modelling and apply to real-life problems

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*		*								
CO 2	*	*	*	*							
CO 3	*	*	*	*				*			
CO 4		*	*	*	*	*					
CO 5	*	*	*								

Course content and outcomes:	
Content	***Competencies***
Unit 1: Multivariate Distributions	
Mean vector, covariance and correlation – population vs. sample - The multivariate Gaussian – joint-, marginal-, and conditional distributions, Mahalanobis distance and outliers - Properties of the multivariate Gaussian - Parameter estimation: maximum likelihood estimation (MLE) and maximum aposteriori estimation (MAP).	At the end of the topic student should be able to: 1. Understand the organisation of multivariate data (C2). 2. Relate multivariate population and sample parameters (C4). 3. Understand and apply multivariate Gaussian modelling to practical problems (C2, C3). 4. Compare parameter estimation using different probabilistic approaches (C4).
Unit 2: Linear and Logistic Regression	
Simple linear regression – regression model, estimating and interpreting coefficients, accuracy of coefficient estimates and model, ANOVA, R2 statistic - Multiple linear regression – estimating coefficients, qualitative predictors, interaction effects, potential problems - Logistic regression – binary and multinomial logistic regression models, estimating and interpreting coefficients, assessing model calibration and discrimination, area under the ROC curve.	1. Model a linear relationship between input and output variables, and assess model performance (C5). 2. Use different performance metrics to conclude what is a good linear fit to the data (C6). 3. Interpret model coefficients and investigate the effect of input variables on output through sensitivity analysis (C6). 4. Apply logistic regression modelling for binary and multiclass classification and assess model performance (C6).
Unit 3: Principal Component Analysis, Cluster Analysis	
Geometric intuition of principal	1. Understand the mathematical foundation

components - Maximum variance perspective – algebraic setup, eigenvectors and eigenvalues of sample correlation matrix - Interpretation and application of principal components for dimension reduction. Dissimilarity measures for mixed data types - Partition around medoids (PAM) vs. K-means algorithms - Selecting the number of clusters.	of principal component analysis (PCA) (C2). 2. Perform and interpret the output of PCA applied to multivariate data for dimension reduction (C6). 3. Assess when PCA is applicable for clustering multivariate data (C6). 4. Compare and contrast methods for clustering multivariate data with mixed data types (C6).
Unit 4: Bootstrapping, Time Series Analysis	
Time series concepts: stationarity, trend, seasonality, autocorrelation - Autoregressive moving average (ARMA) models - Resampling, smoothing, windowing, and rolling average - First and second order differencing - Validating time series predictions.	1. Understand the basic principles of bootstrapping as an experimental method to estimate the sampling distributions of a statistic (C2). 2. Understand the basic mathematical principles of time series modelling (C2). 3. Apply time series modelling to practical problems (C3). 4. Interpret the results of times series model predictions (C3).

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	30	60
Quiz	02	04
Small Group Discussion (SGD)	02	02
Self-directed learning (SDL)	-	04

Problem Based Learning (PBL)	02	04
Case Based Learning (CBL)	-	-
Revision	02	-
Assessment	06	-
TOTAL	**44**	**74**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Sessional Examination 1	*	*			
Sessional Examination 2		*	*	*	
Assignment/Presentation	*	*	*	*	*
End Semester Examination	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. An Introduction to Statistical Learning with Applications in R, Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, Springer; 1st Edition, 2013, Corr. 7th printing 2017 Edition. 2. An Introduction to Applied Multivariate Analysis with R, Brian Everitt and Torsten Hothorn– Springer Publications,1st Edition, 2011. 3. Machine Learning - A Probabilistic Perspective, Kevin P. Murphy, The MIT Press; 1st Edition, 2012.

	4. Mathematics for Machine Learning, Marc Peter Deisenroth, A Aldo Faisal, and Cheng Soon Ong, Cambridge University Press, 2020. – Online resource from Cambridge University Press available at https://mml-book.github.io/book/mml-book.pdf

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Multimedia Analytics
Course Code: BDA 618	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 2
No of Credits: 3	**Prerequisites:** Basic Programming
Synopsis:	This Course provides insight on 1. Time domain audio processing techniques. 2. Identify different image representation methods. 3. Implementing different image feature extraction methods. 4. Implementing different Video classification models.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Examine different audio encoding techniques.
CO 2:	Illustrate Time domain audio processing techniques.
CO 3:	Identify different image representation methods.
CO 4:	Analyse different image feature extraction methods
CO 5:	Analyse different Video classification models

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*	*	*								
CO 2	*	*	*			*					
CO 3	*	*	*		*	*					
CO 4	*	*	*	*	*					*	

CO 5	*	*	*		*						

Course content and outcomes:	
Content	***Competencies***
Unit 1: Audio Acquisition Representation and Storage	
Sound Physics, Production Perception, Audio Encoding and Storage Format, Time Domain Audio processing Image and Video Acquisition, Representation and Storage	At the end of the topic student should be able to: 1. Understand sound physics (C2). 2. Implement different Audio encoding and storage techniques (C4). 3. Design applications for Image and Video acquisition and Storage (C4).
Unit 2: Image Handling and Processing	
Reading images from files, Simple Image transformations, Matrices, Colors and Filters, Contours and Segmentation, Object detection and recognition	1. Understand different image processing techniques (C2). 2. Design application to handle different filters (C4). 3. Implement object detection and recognition algorithms (C4).
Unit 3: Video	
Video Principles, Standards, Video classification models, Motion Detection, Object Tracking in Video	1. Understand different Video standards (C2). 2. Develop applications using video classification models (C4). 3. Design applications of motion detection and Object tracking (C4).
Unit 4: Case study	
Speech and hand writing recognition -	1. Design applications for – Speech and

Automatic Face recognition, Sign board detection, Lane change detection - Video segmentation and key frame extraction	handwriting recognition (C4). 2. Design application to detect and classify images (C4). 3. Design application to extract information from given Video (C4).

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	30	60
Quiz	02	04
Small Group Discussion (SGD)	02	02
Self-directed learning (SDL)	-	04
Problem Based Learning (PBL)	02	04
Case Based Learning (CBL)	-	-
Revision	02	-
Assessment	06	-
TOTAL	**44**	**74**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Sessional Examination 1	*	*			
Sessional Examination 2			*	*	

Assignment/Presentation	*	*	*	*	*
End Semester Examination	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Machine Learning for Audio, Image and Video Analysis, Francesco Camastra and Alessandro Vinciarelli Springer's Publication, 2nd edition. 2015. 2. Practical Python and OpenCV, An Introductory, Example Driven Guide to Image Processing and Computer Vision, Dr. Adrian Rosebrock, 4th edition, 2019 3. Computer Vision with Python Cookbook: Leverage the power of OpenCV 3 and Python to build computer vision applications, Aleksei Spizhevoi, Aleksandr Rybnikov, Packt Publishing, 1st Edition, 2018.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	DevOps for Cloud
Course Code: CDC-607	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 2
No of Credits: 3	**Prerequisites:** Ubuntu OS , Networking and Software Life Cycle

Synopsis:	This Course provides insight on
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Explain the concept of automation of Product Life Cycle stages.
CO 2:	Demonstrate Continuous Integration / Continuous Testing / Continuous Deployment of Product.
CO 3:	Compare and contrast existing Software Methodologies with Devops Life Cycle stages.
CO 4:	Design and Devops methodologies for Product development and Release.
CO 5:	Explain the concepts of Tools used in each stages of Devops.

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*	*					*				
CO 2			*		*						*
CO 3		*	*			*					

CO 4	*							*	*		
CO 5	*									*	

Course content and outcomes:	
Content	***Competencies***
Unit 1: DevOps Introduction	
Understanding Development - Developement SDLC : WaterFall & Agile - Understanding Operations - Dev vs Ops - DevOps to the rescue - What is DevOps - DevOps SDLC - Continous Delivery model - DevOps tools for DevOps SDLC - DevOps Roles & Responsiblities.	At the end of the topic student should be able to: 1. Explain about the Product Life Cycle Software methodologies (C2) 2. Describe Devops life cycle for Product Development and Release (C2) 3. Explain the stages of Devops (C2) 4. Describe about Continuous Integration / Continuous Deployment pipeline. (C2) 5. Write the significance of automation in Product life cycle management. (C3) 6. Describe different between standard software methodologies and Devops software methodologies. (C2)
Unit 2: Linux	
Linux Introduction, Principles & Linux distro – Booting - Command line utililities & Basic commands - Linux Filesystem - Filters & I/O Redirections - Users & Group administration - File permissions &	1. Explain the evolution of Linux OS (C2) 2. Explain Linux File System (C2) 3. Demonstrate Linux Users and Groups (C3) 4. Describe OS Level Virtualization techniques like Containers (C3) 5. Demonstrate basic Linux Commands (C4)

Ownerships - Sudo - Software Managemen - Useful tools: ssh, telnet, scp, rsync, disk utils, backups etc - Service & Process management - Shell Scripting - Systems and HW stats – Linux Containers (lxc) - Dockers – Kubernetes and Microservices .	
Unit 3: Networking fundamentals	
Components of computer networks - Classification: LAN, WAN, Peer to Peer network, Server based - Switches - Routers - Network Architecture - Protocols - Port numbers - DNS - DHCP - IP Addresses - Ip Addresses & Subnet Masks - IP Address Ranges - Subnetting - Private Vs Public networks - High Availaiblity - Firewalls & NACL - Web Application Architecture - Infrastructure - Network layout - Services & Components - Architecture from a DevOps perspective.	1. Explain Computer network and devices (C2) 2. Demonstrate subnetting and its need (C3) 3. Explain IPV4 Addressing scheme (C2) 4. Demonstrate type of Network Devices like Switches , Hub , Router using Simulator Tools (C4) 5. Describe networking Services like DNS , DHCP , NACL , FTP etc (C4)
Unit 4: Automation, Orchestration & Config Management	
Version control system with Git : What is VCS & why it is needed - DevOps use cases - Setup your own repo with git - Manage your code base/source code with GIT & GITHUB	1. Explain need and types of version control software (C1) 2. Describe architecture of Distributed version control systems (C2) 3. Explain Git and Github as case study (C3)
Unit 5: Continuous Integration with Jenkins	
Introduction to continuous integration.	1. Describe about Continuous Integration /

- Build & Release and relation with DevOps - Understanding development and developers - Why Continuous integration Jenkins introduction and setup - Jenkins projects/jobs - Jenkins plugins | Jenkins administration: Users - Nodes/slaves - Managing plugins - Managing software versions - Introduction - Phases - Java builds - Build and Release job/project setup | Nexus: Intro & Setup - Software versioning & Hosted repository - Integration with Jenkins - Continuous integration job/project setup | Complete Jenkins project: Packinging Artifacts - Static code Analysis - Tomcat setup Staging & productions - Artifacts deployments to webservers from Jenkins - Build Pipeline - Jenkins not just CI tool anymore - More DevOps use cases of Jenkins	Continuous Deployment pipeline. (C2) 2. Write the significance of automation in Product life cycle management. (C3) 3. Describe different between standars software methodologies and Devops software methodologies. (C2) 4. Give examples for Automation of stages of Product development using Devops . (C2) 5. Write the limitation of a Current Software methodologies for Product Development. (C3) 6. Describe the architecture of Continuous Integration server. (C2) 7. Apply Devops methodologies for Product Development and Release(C3)
Unit 5: Ansible	
Configuration Management & Automation - What is Ansible & its features - Ansible setup on local & cloud - Understanding Ansible architecture & Execution - Inventory | Ad hoc commands: Automating change Management with Ad Hoc	1. Write the steps in Automation of Testing in Web development. (C3) 2. Explain the operations Continuous Testing. (C5) 3. Write the taxonomy of Continuous Integration / Continuous Delivery / Continuous Deployment (C3)

commands - Playbook Introduction - Ansible configuration with ansible.cfg - Ansible documentation - Modules, modules & lots of modules - Writing playbook for webserver & DB server deployments - Tasks - Variables - Templates - Loops - Handlers - Conditions - Register - Debugging - Ansile Roles - Identify server roles - Roles structure - Creating, Managing and executing roles - Ansible Galaxy - Exploring Roles from Galaxy - Download Galaxy roles and integrate with your code - Ansible Advanced Execution - Improving execution time - Limiting and selecting tasks - Troubleshooting and Testing.	4. Design a Workflow for Automation of Product life cycle using Devops (C5, P3). 5. Construct a Continuous Integration / Continuous Deployment pipeline (C5) 6. Compare Standard Software methodologies vs Devops methodologies for Product Development. (C6, P2) 7. Describe about Containers and Container Orchestration Services. (C2) 8. Examine the advantages of using Containers in Web development(C4) 9. Describe Container orchestration services architecture(C2) 10. Show the function of Container orchestration services(C3) 11. Define Configuration Management tools and its need. (C1) 12. Describe the features of Configuration Management. (C2) 13. Explain the architecture of Configuration Management (C5) 14. Design a Configuration Management Codes to administrate infrastructure of organization (C5) 15. Explain the need of Continuous Monitoring tools (C5) 16. Design an Architecture Continuously Monitor infrastructure. (C4)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	30	60
Quiz	02	04
Small Group Discussion (SGD)	02	02
Self-directed learning (SDL)	-	04
Problem Based Learning (PBL)	02	04
Case Based Learning (CBL)	-	-
Revision	02	-
Assessment	06	-
TOTAL	**44**	**74**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Sessional Examination 1	*	*			
Sessional Examination 2			*	*	*
Assignment/Presentation				*	*
End Semester Examination	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Eric Foster-Johnson, John C. Welch, Micah Anderson, Beginning Shell Scripting (Programmer to Programmer), Wrox Publications 2. Randal K. Michael "Mastering Unix Shell Scripting: Bash, Bourne, and Korn Shell Scripting for Programmers, System Administrators, and UNIX Gurus", 2nd Edition, Wiley Publications 3. Bintu Harwani, "UNIX & Shell Programming", Oxford Publications, 2013 4. John Ferguson Smart, "Jenkins: The Definitive Guide",O'reilly Publications 5. Mitesh Soni, "Jenkins Essentials", Packt Publications 6. Rafal Leszko, "Continuous Delivery with Docker and Jenkins", Packt Publications 7. Veselin Kantsev, "Implementing DevOps on AWS", Packt Publications 8. Randall Smith, "Docker Orchestration", Packt Publications 9. Alan Berg, "Jenkins Continuous Integration Cookbook", Packt Publications 10. Kumaran S., Senthil, " Practical LXC and LXD Linux Containers for Virtualization and Orchestration", Apress Publications 11. Konstantin Ivanov, " Containerization with LXC" , Packt Publications 12. Karl Matthias, Sean Kane, "Docker: Up & Running:Shipping

	Reliable Containers in Production",O'Reilly Media

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Natural Language and Text Processing
Course Code: BDA-621	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 2
No of Credits: 3	**Prerequisites:** Programming in Python
Synopsis:	This course introduces fundamental concepts in natural language and text processing.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Understand syntax and semantics of text.
CO 2:	Perform text processing by implementing lexical analysis, word stemming, word stop, and term selection.
CO 3:	Perform categorizing and tagging of words.
CO 4:	Classification and information extraction from text.
CO 5:	Design models for sentiment and semantic analysis from text.

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*										
CO 2	*	*									
CO 3	*	*									
CO 4		*	*	*							
CO 5		*	*	*	*						

Course content and outcomes:	
Content	***Competencies***
Unit 1: Natural Language Basics; Accessing, Processing and Understanding Text; Categorizing and Tagging Words	
Natural Language, Linguistics, Language Semantics, Text Corpora. Accessing Text Corpora, from the Web and from Disk, Conditional Frequency Distributions, Regular Expressions for Detecting Word Patterns, Tokenization. Using a Tagger, Tagged Corpora, Automatic Tagging, N-Gram Tagging, Transformation-Based Tagging.	At the end of the topic student should be able to: 1. Understand the basic principles of organizing textual data (C2). 2. Understand how to access text corpora from different media (C2). 3. Understand how words can be used as building blocks for textual analysis (C2). 4. Understand different types of tagging (C2).
Unit 2: Classification and Information Extraction; Text Similarity and Clustering	
Automated Text Classification, TF-IDF Model, Advanced Word Vectorization Models, Classification Algorithms - Multinomial Naïve Bayes, Support Vector Machines. Text Summarization and Information Extraction - Text Normalization, Feature Extraction, Keyphrase Extraction, Topic Modelling, Automated Document Summarization. Term Similarity, Analysing Document	1. Understand the mathematical principles of word vectorization (C2). 2. Compare and contrast different classification algorithms for text analysis (C6). 3. Understand how to perform feature extraction for text analysis (C3). 4. Understand how to compare and cluster text documents (C3).

Similarity, Document Clustering.	
Unit 3: Semantic and Sentiment Analysis	
Exploring WordNet, Word Sense Disambiguation, Named Entity Recognition, Analysing Semantic Representations, Sentiment Analysis.	1. Understand how to access the WordNet lexical database (C3). 2. Understand how to perform semantic analysis of natural language expressions (C3). 3. Understand how to perform sentiment analysis of text documents (C3).

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	30	60
Quiz	02	04
Small Group Discussion (SGD)	02	02
Self-directed learning (SDL)	-	04
Problem Based Learning (PBL)	02	04
Case Based Learning (CBL)	-	-
Revision	02	-
Assessment	06	-
TOTAL	**44**	**74**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Sessional Examination 1	*	*			
Sessional Examination 2		*	*	*	
Assignment/Presentation	*	*	*	*	*
End Semester Examination	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Text Analytics with Python: A Practitioner's Guide to Natural Language Processing, Dipanjan Sarkar; Publisher: Apress, 2nd Edition, 2019. 2. Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit, by Steven Bird, Ewan Klein, Edward Loper, O'Reilly Media, Inc, 1st edition 2009. 3. Hands-On Natural Language Processing with Python: A practical guide to applying deep learning architectures to your NLP applications, Rajesh Arumugam, Rajalingappaa Shanmugamani, Packt Publishing Limited, 2018.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Entrepreneurship
Course Code: ENP-601	**Course Instructor:**
Academic Year: 2020 - 2021	**Semester:** First Year, Semester 2
No of Credits: 3	**Prerequisites:** -
Synopsis:	This course introduces students to the theory of entrepreneurship and its practical implementation. It focuses on different stages related to the entrepreneurial process, including business model innovation, monetization, small business management as well as strategies that improve performance of new business ventures. Centered on a mixture of theoretical exploration as well as case studies of real-world examples and guest lectures, students will develop an understanding of successes, opportunities and risks of entrepreneurship. This course has an interdisciplinary approach and is therefore open to students from other Majors.
Course Outcomes (COs):	On successful completion of this course, students will be able to:
CO 1:	To impart knowledge on the basics of entrepreneurial skills and competencies to provide the participants with necessary inputs for creation of new ventures.
CO 2:	To familiarize the participants with the concept and overview of entrepreneurship with a view to enhance entrepreneurial talent
CO 3:	To appraise the entrepreneurial process starting with pre-venture stage
CO 4:	To Create and exploit innovative business ideas and market opportunities
CO 5:	To Build a mind-set focusing on developing novel and unique approaches to market opportunities
CO 6:	To explore new vistas of entrepreneurship in 21st century environment to

	generate innovative business ideas through case studies.

Mapping of COs to POs											
COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*										
CO 2				*							
CO 3			*								
CO 4						*					
CO 5								*			
CO 6										*	

Course content and outcomes:	
Content	***Competencies***
Unit 1: Introduction to Entrepreneurship	
Meaning and Definition of Entrepreneurship - Employment vs Entrepreneurship, Theories of Entrepreneurship, approach to entrepreneurship, Entrepreneurs VS Manager	At the end of the topic student should be able to: 1. Explain the meaning of Entrepreneurship (C1) 2. Discuss the theories of Entrepreneurship (C1) 3. Discuss the approaches to Entrepreneurship (C1)
Unit 2: Entrepreneurial Traits	
Personality of an entrepreneur, Types of Entrepreneurs	1. Discuss the Personality traits of entrepreneurs. (C2)
Unit 3: Process of Entrepreneurship	
Factors affecting Entrepreneurship	1. Identify the fundamentals and

process	responsibilities of entrepreneurship (C2) 2. Exemplify one's capabilities in relation to the rigors of successful ventures (C3) 3. Identify and differentiates the different characteristics and competencies of an entrepreneurs (C2)
Unit 4: Business Start-up Process	
Idea Generation, Scanning the Environment, Macro and Micro analysis	1. Explain the Process of Business start up (C1) 2. Develop creativity and critical thinking in identifying opportunities (C5) 3. Apply innovative approaches in envisioning ones entrepreneurial career (C3)
Unit 5: Business Plan writing	
Points to be considered, Model Business plan	1. Identify different business models (C3) 2. Describe different parts of a business plan(C2)
Unit 6: Case studies	
Indian and International Entrepreneurship	1. Perform self-assessment and analyse entrepreneurial personal traits and competencies (C4) 2. Evaluate oneself and plan courses of action to help develop one's entrepreneurial characteristics and competencies. (C5)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	30	60
Quiz	02	04

Small Group Discussion (SGD)	02	02
Self-directed learning (SDL)	-	04
Problem Based Learning (PBL)	02	04
Case Based Learning (CBL)	-	-
Revision	02	-
Assessment	06	-
TOTAL	**44**	**74**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos						
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5	CO 6
Sessional Examination 1	*	*				
Sessional Examination 2			*	*		
Assignment/Presentation					*	*
End Semester Examination	*	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. NVR Naidu and T. Krishna Rao, "Management and Entrepreneurship", IK International Publishing House Pvt. Ltd 2008. 2. Mohanthy Sangram Keshari, "Fundamentals of Entrepreneurship", PHI Publications, 2005 3. Butler, D. (2006). Enterprise planning and development. USA: Elsevier Ltd. Gerber, M.E. (2008) Awakening the

	entrepreneur within. NY: Harper Collins.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Machine Learning for Big Data Lab
Course Code: BDA 605L	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 2
No of Credits: 1	**Prerequisites:** Programming with Python and Data Visualization
Synopsis:	This Course provides insight on
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Demonstrate activation functions, weights and threshold units in artificial neural networks
CO 2:	Demonstrate Artificial Neural Network, Clustering, Support Vector Machine, Deep Neural Network and Reinforcement Learning models
CO 3:	Analyse Artificial Neural Network, Clustering, Support Vector Machine, Deep Neural Network and Reinforcement Learning models
CO 4:	Compare and contrast single layer, multilayer and deep neural networks in terms of accuracy in classification
CO 5:	Design different types of artificial neural network models, clustering models, deep neural network models, reinforcement learning models

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*										
CO 2		*									
CO 3			*								
CO 4				*							
CO 5				*							

Course content and outcomes:	
Content	***Competencies***
Unit 1: Artificial Neural Networks	
Neurons and biological motivation. Activation functions and threshold units. Supervised and unsupervised learning Perceptron Model: representational limitation and gradient descent training. Multilayer networks and back propagation. Overfitting.	At the end of the topic student should be able to: 1. Demonstrate activation functions, weights and threshold units in artificial neural networks (C3) 2. Demonstrate ANN models (C3) 3. Design of ANN models for classification (C5) 4. Analyse the performance issues (C4)
Unit 2: Clustering	
Learning from unclassified data. Clustering. Hierarchical Aglomerative Clustering. Non-Hierarchical Clustering - k-means partitional clustering. Expectation maximization (EM) for soft clustering. Semi-supervised learning with EM using labeled and unlabled data.	1. Demonstrate various clustering models in machine learning (C3) 2. Design different types of clusters (C5) 3. Analyse the performance of clustering techniques on different data (C4) 4. Apply clustering techniques for data analysis. (C3)
Unit 3 Kernel Methods	
Dual Representations Design of Kernels	1. Design of different kernel techniques (C5)
Unit 4: Support Vector Machines (SMV)	

Maximum margin linear separators. Quadractic programming solution to finding maximum margin separators. Kernels for learning non-linear functions. Varying length pattern classification using SVM	1. Demonstrate Maximum margin linear separators. (C3) 2. Design SVM classifiers (C5) 3. Analyse the performance of SVM (C4)
Unit 5: Deep Learning	
Introduction to Deep Learning Introduction to convolutional Neural Network (CNN) CNN Architecture and layers Building simple CNN model for classification Training and Testing the CNN model	1. Develop Deep Neural Network/ CNN (C5) 2. Design a classifier for the image classification system. (C5) 3. Compare performance of CNN and ANN for image classification (C4)
Unit 6: Reinforcement Learning	
Characteristics N-arm Bandit Problem Calculating the Value Function Associative Learning – Adding States The Markov Property & Markov Decision Process	1. Apply reinforcement learning model using different principles (C3) 2. Analyse various reinforcement learning techniques (C4) 3. Design of reinforcement learning models (C5)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	12	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	-	-

Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	03	-
Clinic	-	-
Practical	24	-
Revision	03	-
Assessment	06	-
TOTAL	**48**	**-**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test - yes	Sessional examination
Theory Assignments	End semester examination - yes
Lab Assignment & Viva - yes	Viva

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Sessional Examination 1	*	*			
Sessional Examination 2			*	*	
Assignment/Presentation			*	*	*
Laboratory examination			*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Machine Learning, T. Mitchell, McGraw-Hill, 1997 2. Machine Learning, E. Alpaydin, MIT Press, 2010 3. Pattern Recognition and Machine Learning, C. Bishop, Springer, 2006

4. Pattern Classification, R. Duda, E. Hart, and D. Stork, Wiley-Interscience, 2000
5. Neural Networks - A Class Room Approach, Satish Kumar, Second Edition, Tata McGraw-Hill, 2013
6. The Elements of Statistical Learning: Data Mining, Inference and Prediction, T. Hastie, R. Tibshirani and J. Friedman, Springer, 2nd Edition, 2009
7. Machine Learning for Big Data, Jason Bell, Wiley Big Data Series
8. Kernel Methods for Pattern Analysis, J. Shawe-Taylor and N. Cristianini, Cambridge University Press, 2004
9. Neural Networks and Learning Machines, S. Haykin, Prentice Hall of India, 2010
10. Multidimensional Neural Networks Unified Theory, Rama Murthy G
11. F.Camastra and A.Vinciarelli, Machine Learning for Audio, Image and Video Analysis – Theory and Applications, Springer, 2008

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Modern Databases for Big Data Lab
Course Code: BDA 616L	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First year, Second semester
No of Credits: 1	**Prerequisites:** Programming in Java, SQL, Python
Synopsis:	This Course provides insight on 1. This course helps students to write SQL queries to work on data. 2. Deals with different data models 3. Discuss the distributed architecture to handle data which is scalable. 4. Students work with different types of NoSQL databases.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Design queries to extract required data.
CO 2:	Experiment with different types NoSQL databases to handle Big Data
CO 3:	Analyse proper databases which are fault tolerant and scalable.

Mapping of COs to POs											
COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*	*	*		*					*	
CO 2	*	*	*		*	*				*	
CO 3	*	*	*		*	*	*			*	

Course content and outcomes:	
Content	***Competencies***

Unit 1: NoSQL DB	
Installing NoSQL database. Querying NoSQL DB	At the end of the topic student should be able to: 1. Design queries to get data stored in NoSQL database (C4).
Unit 2: Data Distribution and Scalable	
Horizontal partitioning and data shrading. MapReduce in databases.	1. Configure databases for handling data distribution and scalability (C4). 2. Develop applications with MapReduce technique to handle the data (C4).
Unit 3: Case study	
Choose proper database based on need	1. Design applications to handle data using appropriate database (C4).

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	12	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	-	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	03	-
Clinic	-	-
Practical	24	-
Revision	03	-
Assessment	06	-
TOTAL	**48**	**-**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos			
Nature of assessment	CO 1	CO 2	CO 3
Sessional Examination 1	*		
Sessional Examination 2		*	*
Assignment/Presentation	*	*	*
Laboratory Examination	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. **Database System Concepts**, Avi Silberschatz, Henry F. Korth, and S. Sudarshan. McGraw Hill, 6th Edition, 2010. 2. **NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence**, Pramod J. Sadalage, Martin Fowler, Addison-Wesley, 2012. 3. **Seven Databases in Seven Weeks: A Guide to Modern Databases and the NoSQL Movement**, Eric Redmond, Jim R. Wilson, Pragmatic Bookshelf. 2012.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Advanced Applications of Probability and Statistics Lab
Course Code: MCL 602L	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 2
No of Credits: 1	**Prerequisites:** MCL 602
Synopsis:	This course provides an introduction to advanced applications of probability and statistics for analysing multivariate and time series data using the R programming language.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Compute and interpret descriptive statistics for multivariate data
CO 2:	Build and assess linear and logistic regression models for practical problems
CO 3:	Perform principal component analysis (PCA) for dimension reduction in multivariate data
CO 4:	Cluster multivariate data with mixed data types
CO 5:	Apply time series modelling to real-life problems

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*	*	*		*						
CO 2		*	*	*	*			*			
CO 3		*	*	*	*			*			
CO 4		*	*	*	*	*		*			
CO 5	*	*	*								

Course content and outcomes:	
Content	***Competencies***
Unit 1: Multivariate Distributions	
Mean vector, covariance and correlation – population vs. sample - The multivariate Gaussian – joint-, marginal-, and conditional distributions, Mahalanobis distance and outliers - Properties of the multivariate Gaussian - Parameter estimation: maximum likelihood estimation (MLE) and maximum aposteriori estimation (MAP).	At the end of the topic student should be able to: 1. Compute descriptive statistics of multivariate data (C2). 2. Perform exploratory data analysis of multivariate data (C4). 3. Identify outliers in multivariate data (C3). 4. Visualise and understand the properties of multivariate Gaussian data (C3).
Unit 2: Linear and Logistic Regression	
Simple linear regression – regression model, estimating and interpreting coefficients, accuracy of coefficient estimates and model, ANOVA, R2 statistic - Multiple linear regression – estimating coefficients, qualitative predictors, interaction effects, potential problems - Logistic regression – binary and multinomial logistic regression models, estimating and interpreting coefficients, assessing model calibration and discrimination, area under the ROC curve.	1. Use in-built functions in R to build linear models for practical problem (C3). 2. Compute different performance metrics to assess model performance (C6). 3. Interpret model coefficients and investigate the effect of input variables on output through sensitivity analysis (C6). 4. Use in-built functions in R to build logistic regression models for practical binary classification problems and assess model performance (C6).
Unit 3: Principal Component Analysis, Cluster Analysis	
Geometric intuition of principal	1. Visualise the geometric interpretation of

components - Maximum variance perspective – algebraic setup, eigenvectors and eigenvalues of sample correlation matrix - Interpretation and application of principal components for dimension reduction. Dissimilarity measures for mixed data types - Partition around medoids (PAM) vs. K-means algorithms - Selecting the number of clusters.	principal component analysis (PCA) (C3). 2. Use in-built functions in R to perform PCA on multivariate data (C3). 3. Compare and contrast PCA for variance maximization vs. clustering of multivariate data (C6). 4. Cluster multivariate data with mixed data types using in-built functions in R (C3).
Unit 4: Bootstrapping, Time Series Analysis	
Time series concepts: stationarity, trend, seasonality, autocorrelation - Autoregressive moving average (ARMA) models - Resampling, smoothing, windowing, and rolling average - First and second order differencing - Validating time series predictions.	1. Apply bootstrapping on a practical data set and assess performance (C3). 2. Understand and apply in-built functions in R for time series modelling (C3). 3. Apply time series modelling to practical problems (C3). 4. Interpret the results of times series model predictions (C3).

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	12	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	-	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-

Case Based Learning (CBL)	03	-
Clinic	-	-
Practical	24	-
Revision	03	-
Assessment	06	-
TOTAL	**48**	**-**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Sessional Examination 1	*	*			
Sessional Examination 2			*	*	*
Assignment/Presentation	*	*	*	*	*
Laboratory examination	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. An Introduction to Statistical Learning with Applications in R, Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, Springer; 1st Edition, 2013, Corr. 7th printing 2017 Edition. 2. An Introduction to Applied Multivariate Analysis with R, Brian Everitt and Torsten Hothorn– Springer Publications,1st Edition, 2011. 3. Machine Learning - A Probabilistic Perspective, Kevin P. Murphy, The MIT Press; 1st Edition, 2012.

	4. Mathematics for Machine Learning, Marc Peter Deisenroth, A Aldo Faisal, and Cheng Soon Ong, Cambridge University Press, 2020. – Online resource from Cambridge University Press available at https://mml-book.github.io/book/mml-book.pdf

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Multi Media Analytics Lab
Course Code: BDA 618L	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First year, Second semester
No of Credits: 1	**Prerequisites: Programming in Python**
Synopsis:	This Course provides insight on 1. Students understand the physics behind Audio, its encoding techniques. 2. Students learn how to read and represent images. 3. Students learn various image processing and information extraction techniques. 4. Students learn to handle video data and extract information from it.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Experiment applications to extract information from Audio.
CO 2:	Design applications for Image Analysis.
CO 3:	Develop Models for Video analysis.

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*	*	*		*	*					
CO 2	*	*	*	*	*	*			*	*	
CO 3	*	*	*	*	*	*			*	*	

Course content and outcomes:	
Content	***Competencies***
Unit 1: Audio Analysis	
Audio encoding and processing	At the end of the topic student should be able to: 1. Implement different audio encoding techniques (C4). 2. Develop applications to perform audio analysis (C4).
Unit 2: Image Analysis	
Image encoding, filters and transformations.	1. Implement various image storing and reading techniques (C4). 2. Develop image processing techniques (C4).
Unit 3: Video Analysis	
Video encoding and processing techniques.	1. Develop applications to extract object of interest from input video (C4).

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	12	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	-	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	03	-
Clinic	-	-
Practical	24	-

Revision	03	-
Assessment	06	-
TOTAL	**48**	**-**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos			
Nature of assessment	CO 1	CO 2	CO 3
Sessional Examination 1	*	*	
Sessional Examination 2		*	*
Assignment/Presentation	*	*	*
Laboratory Examination	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Machine Learning for Audio, Image and Video Analysis, Francesco Camastra and Alessandro Vinciarelli Springer's Publication, 2nd edition. 2015. 2. Practical Python and OpenCV, An Introductory, Example Driven Guide to Image Processing and Computer Vision, Dr. Adrian Rosebrock, 4th edition, 2019 3. Computer Vision with Python Cookbook: Leverage the power of OpenCV 3 and Python to build computer vision applications, Aleksei Spizhevoi, Aleksandr Rybnikov, Packt

	Publishing, 1st Edition, 2018.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Mini Project - 2
Course Code: BDA 696	**Course Instructor:**
Academic Year: 2020 - 2021	**Semester:** First Year, Semester 2
No of Credits: 4	**Prerequisites:** Any programming language and circuit basics

Synopsis:	Students are expected to select a problem in the area of their interest and the area of their specialization that would require an implementation in hardware / software or both in a semester
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Apply the objectives of the project work and provide an adequate background with a detailed literature survey
CO 2:	Breakdown the project into sub blocks with sufficient details to allow the work to be reproduced by an independent researcher
CO 3:	Compose hardware/software design, algorithms, flowchart, methodology, and block diagram
CO 4:	Evaluate the results
CO 5:	Summarize the work carried out

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1				*							
CO 2					*			*			
CO 3							*			*	

CO 4						*					*
CO5:							*				

Course content and outcomes:	
Content	***Competencies***
Phase 1	
Problem identification, synopsis submission, status submission, mid evaluation.	At the end of the topic student should be able to: 1. Identify the problem/specification (C1) 2. Discuss the project (C2) 3. Prepare the outline (C3) 4. Describe the status of the project (C2) 5. Prepare a mid-term project presentation report (C3) 6. Prepare and present mid-term project presentation slides (C3, C5) 7. Develop project implementation in hardware/software or both in chosen platform (C5)
Phase 2	
Status submission, final evaluation.	1. Prepare the progress report (C3) 2. Prepare the final project presentation report (C3) 3. Prepare and present final project presentation slides (C3, C5) 4. Modify and Develop implementation in hardware/software or both in chosen platform (C3, C5) 5. Justify the methods used and obtained

	results (C6)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	-	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	48	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	-	-
Clinic	-	-
Practical	-	-
Revision	-	-
Assessment	03	-
TOTAL	**51**	**09**

Assessment Methods:	
Formative:	**Summative:**
Project Problem Selection	Mid-Term Presentation
Synopsys review	Second status review
First status review	Demo & Final Presentation

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Mid Presentation	*	*			

Presentation	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	Particular to the chosen project

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Seminar - 2
Course Code: BDA 698	**Course Instructor:**
Academic Year: 2020 - 2021	**Semester:** First Year, Semester 2
No of Credits: 1	**Prerequisites:** Communication Skill
Synopsis:	1. To select, search and learn technical literature. 2. To Identify a current and relevant research topic. 3. To prepare a topic and deliver a presentation. 4. To develop the skill to write a technical report. 5. Develop ability to work in groups to review and modify technical content.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Show competence in identifying relevant information, defining and explaining topics under discussion.
CO 2:	Show competence in working with a methodology, structuring their oral work, and synthesizing information.
CO 3:	Use appropriate registers and vocabulary, and will demonstrate command of voice modulation, voice projection, and pacing.
CO 4:	Demonstrate that they have paid close attention to what others say and can respond constructively.
CO 5:	Develop persuasive speech, present information in a compelling, well-structured, and logical sequence, respond respectfully to opposing ideas, show depth of knowledge of complex subjects, and develop their ability to synthesize, evaluate and reflect on information.

Mapping of COs to POs											
COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*							*	*		*
CO 2	*							*	*		*
CO 3	*							*	*		*
CO 4	*							*	*		*
CO5:	*							*	*		*

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	-	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	14	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	-	-
Clinic	-	-
Practical	-	-
Revision	-	-
Assessment	-	-
TOTAL	**14**	**-**

Assessment Methods:	
Formative:	**Summative:**
Seminar Topic Selection	
Synopsys review	
PPT Review	

Mapping of assessment with Cos

Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Presentation	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	Particular to the chosen Seminar

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Devops for CloudLab
Course Code: CDC-607L	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 2
No of Credits: 1	**Prerequisites:** Ubuntu OS , Networking and Software Life Cycle
Synopsis:	This Course provides insight on
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Explain the concept of automation of Product Life Cycle stages
CO 2:	Design an Devops methodologies for Product development and Release
CO 3:	Demonstrate Continuous Integration / Continuous Testing / Continuous Deployment of Product
CO4:	Explain the concepts of Tools used in each stages of Devops .
CO5:	Demonstrate Continuous Monitoring of Production Environment

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*	*			*						
CO 2	*		*								
CO 3		*	*		*						
CO 4		*	*		*						
CO 5							*	*			

Course content and outcomes:	
Content	***Competencies***
Unit 1: DevOps Introduction	
Understanding Development - Developement SDLC : WaterFall & Agile - Understanding Operations - Dev vs Ops - DevOps to the rescue - What is DevOps - DevOps SDLC - Continous Delivery model - DevOps tools for DevOps SDLC - DevOps Roles & Responsiblities.	At the end of the topic student should be able to: 1. Demonstrate differences between Waterfall and agile software development methodologies (C2)
Unit 2: Linux	
Linux Introduction, Principles & Linux distro – Booting - Command line utililities & Basic commands - Linux Filesystem - Filters & I/O Redirections - Users & Group administration - File permissions & Ownerships - Sudo - Software Managemen - Useful tools: ssh, telnet, scp, rsync, disk utils, backups etc - Service & Process management	1. Design Ubuntu based VM using hypervisor to understand booting process , linux file system , linux networking , Users , Groups and Permissions, tools (ssh , scp etc) (C3) 2. Design a docker environment to containerize web application (C3) 3. Design a Kubernetes cluster to deploy containerized application using Kubernetes deployment and service models (C4)

- Shell Scripting - Systems and HW stats – Linux Containers (lxc) - Dockers – Kubernetes and Microservices	
Unit 3: Networking fundamentals	
Components of computer networks - Classification: LAN, WAN, Peer to Peer network, Server based - Switches - Routers - Network Architecture - Protocols - Port numbers - DNS - DHCP - IP Addresses - Ip Addresses & Subnet Masks - IP Address Ranges - Subnetting - Private Vs Public networks - High Availaiblity - Firewalls & NACL - Web Application Architecture - Infrastructure - Network layout - Services & Components - Architecture from a DevOps perspective.	1. Design a College/ University network using packet tracer to understand computer networking devices like Hub , Switches , Routers and Firewalls (C3) 2. Design a Network project using Packet tracer to understand Networking services like DNS , DHCP , FTP etc (C3)
Unit 4: Automation, Orchestration & Config Management	
Version control system with Git : What is VCS & why it is needed - DevOps use cases - Setup your own repo with git - Manage your code base/source code with GIT & GITHUB	1. Create Github account and set up repository and use git commands to Clone , Fork and commit files to Github repositories (C4)
Unit 5: Continuous Integration with Jenkins	
Introduction to continuous integration. - Build & Release and relation with DevOps - Understanding development	1. Design a Continuous Integration server using Jenkins in Master Slave architecture (C3)

and developers - Why Continuous integration Jenkins introduction and setup - Jenkins projects/jobs - Jenkins plugins | Jenkins administration: Users - Nodes/slaves - Managing plugins - Managing software versions - Introduction - Phases - Java builds - Build and Release job/project setup | Nexus: Intro & Setup - Software versioning & Hosted repository - Integration with Jenkins - Continuous integration job/project setup | Complete Jenkins project: Packinging Artifacts - Static code Analysis - Tomcat setup Staging & productions - Artifacts deployments to webservers from Jenkins - Build Pipeline - Jenkins not just CI tool anymore - More DevOps use cases of Jenkins	2. Demonstrate CI/CD for JAVA/PHP/nodejs web application (C4) 3. Design an Eclipse Selenium testing project to automate Web application Testing Process (C4)
Unit 6: Ansible	
Configuration Management & Automation - What is Ansible & its features - Ansible setup on local & cloud - Understanding Ansible architecture & Execution - Inventory | Ad hoc commands: Automating change Management with Ad Hoc	1. Design a Configuration management service using Ansible to administer group of nodes in lab (C2) 2. Demonstrate installation of Software packages like git , Eclipse , Mysql on group of nodes using Ansible (C4) 3. Design a Continuous monitoring server

commands - Playbook Introduction - Ansible configuration with ansible.cfg - Ansible documentation - Modules, modules & lots of modules - Writing playbook for webserver & DB server deployments - Tasks - Variables - Templates - Loops - Handlers - Conditions - Register - Debugging - Ansile Roles - Identify server roles - Roles structure - Creating, Managing and executing roles - Ansible Galaxy - Exploring Roles from Galaxy - Download Galaxy roles and integrate with your code - Ansible Advanced Execution - Improving execution time - Limiting and selecting tasks - Troubleshooting and Testing	using Nagios to monitor group of servers for different dervices like CPU Utilization , RAM Usage , Network Bandwidth , Apache server logs , Database server logs etc (C5)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	12	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	-	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	03	-
Clinic	-	-

Practical	24	-
Revision	03	-
Assessment	06	-
TOTAL	**48**	**-**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos			
Nature of assessment	CO 1	CO 2	CO 3
Sessional Examination 1	*	*	*
Assignment/Presentation		*	*
Laboratory Examination	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Eric Foster-Johnson , John C. Welch , Micah Anderson, Beginning Shell Scripting (Programmer to Programmer), Wrox Publications 2. Randal K. Michael "Mastering Unix Shell Scripting: Bash, Bourne, and Korn Shell Scripting for Programmers, System Administrators, and UNIX Gurus", 2nd Edition, Wiley Publications 3. Bintu Harwani, "UNIX & Shell Programming", Oxford Publications, 2013 4. John Ferguson Smart, "Jenkins: The Definitive Guide",O'reilly Publications

	5. Mitesh Soni, "Jenkins Essentials", Packt Publications 6. Rafal Leszko, "Continuous Delivery with Docker and Jenkins", Packt Publications 7. Veselin Kantsev, "Implementing DevOps on AWS", Packt Publications 8. Randall Smith, "Docker Orchestration", Packt Publications 9. Alan Berg, "Jenkins Continuous Integration Cookbook", Packt Publications 10. Kumaran S., Senthil, " Practical LXC and LXD Linux Containers for Virtualization and Orchestration", Apress Publications 11. Konstantin Ivanov, " Containerization with LXC" , Packt Publications 12. Karl Matthias, Sean Kane, "Docker: Up & Running:Shipping Reliable Containers in Production",O'Reilly Media

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Natural Language and Text Processing Lab
Course Code: BDA 621L	**Course Instructor:**
Academic Year: 2020-2021	**Semester:** First Year, Semester 2
No of Credits: 1	**Prerequisites:** BDA 621
Synopsis:	This course provides an introduction to programming principles for natural language and text processing.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Access text corpora from different media
CO 2:	Use regular expressions for analysing and extracting patterns in text data
CO 3:	Perform text processing using state of the art software libraries
CO 4:	Classify text documents
CO 5:	Implement models for sentiment and semantic analysis from text

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*				*						
CO 2		*	*		*						
CO 3			*		*						
CO 4			*	*	*						
CO 5		*	*	*	*						

Course content and outcomes:

Content	***Competencies***
Unit 1: Natural Language Basics; Accessing, Processing and Understanding Text; Categorizing and Tagging Words	
Natural Language, Linguistics, Language Semantics, Text Corpora.	At the end of the topic student should be able to: 1. Understand how to access text corpora

Accessing Text Corpora, from the Web and from Disk, Conditional Frequency Distributions, Regular Expressions for Detecting Word Patterns, Tokenization. Using a Tagger, Tagged Corpora, Automatic Tagging, N-Gram Tagging, Transformation-Based Tagging.	from different media (C2). 2. Use software libraries for text tokenization (C3). 3. Implement regular expressions for detecting word patterns (C3). 4. Implement different types of word tagging (C3).
Unit 2: Classification and Information Extraction; Text Similarity and Clustering	
Automated Text Classification, TF-IDF Model, Advanced Word Vectorization Models, Classification Algorithms - Multinomial Naïve Bayes, Support Vector Machines. Text Summarization and Information Extraction - Text Normalization, Feature Extraction, Keyphrase Extraction, Topic Modelling, Automated Document Summarization. Term Similarity, Analysing Document Similarity, Document Clustering.	1. Implement different techniques to represent words as vectors, compare and contrast them (C6). 2. Apply classification algorithms for text data (C3). 3. Implement building blocks of text summarization (C4). 4. Perform document clustering using software libraries (C3).
Unit 3: Semantic and Sentiment Analysis	
Exploring WordNet, Word Sense Disambiguation, Named Entity Recognition, Analysing Semantic Representations, Sentiment Analysis.	1. Access the WordNet lexical database (C3). 2. Perform semantic analysis of natural language expressions (C3). 3. Perform sentiment analysis of text documents (C3).

Learning strategies, contact hours and student learning time

Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	-	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	-	-
Self-directed learning (SDL)	4	10
Problem Based Learning (PBL)	4	10
Case Based Learning (CBL)	4	10
Clinic	-	-
Practical	28	56
Revision	2	10
Assessment	6	-
TOTAL	**48**	**96**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test – **yes**	Sessional examination
Theory Assignments	End semester examination – **yes**
Lab Assignment & Viva – **yes**	Viva

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Sessional Examination 1					
Sessional Examination 2					
Assignment/Presentation	*	*	*	*	*
End Semester Examination	*	*	*	*	*
Laboratory examination	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. Text Analytics with Python: A Practitioner's Guide to Natural Language Processing, Dipanjan Sarkar; Publisher: Apress, 2nd Edition, 2019. 2. Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit, by Steven Bird, Ewan Klein, Edward Loper, O'Reilly Media, Inc, 1st edition 2009. 3. Hands-On Natural Language Processing with Python: A practical guide to applying deep learning architectures to your NLP applications, Rajesh Arumugam, Rajalingappaa Shanmugamani, Packt Publishing Limited, 2018.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Entrepreneurship Lab

Course Code: ENP-601L	**Course Instructor:**
Academic Year: 2020 - 2021	**Semester:** First Year, Semester 2
No of Credits: 1	**Prerequisites: -**

Synopsis:	This Course provides insight on This course introduces students to the theory of entrepreneurship and its practical implementation. It focuses on different stages related to the entrepreneurial process, including business model innovation, monetization, small business management as well as strategies that improve performance of new business ventures. Cantered on a mixture of theoretical exploration as well as case studies of real-world examples and guest lectures, students will develop an understanding of successes, opportunities and risks of entrepreneurship. This course has an interdisciplinary approach and is therefore open to students from other Majors.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	Understand the concept of entrepreneurship
CO 2:	To appraise the entrepreneurial process starting with pre-venture stage through group discussion
CO 3:	To Build a mind-set focusing on developing novel and unique approaches to market opportunities by considering case studies and understand the complete flow of entrepreneurship

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1	*					*		*			
CO 2						*					
CO 3								*		*	

Course content and outcomes:

Content	*Competencies*
Unit 1: Introduction to Entrepreneurship	
Meaning and Definition of Entrepreneurship-Employment vs Entrepreneurship, Theories of Entrepreneurship, approach to entrepreneurship, Entrepreneurs VS Manager	At the end of the topic student should be able to: 1. Discuss the theories of Entrepreneurship (C1) 2. Discuss the approaches to Entrepreneurship (C1)
Unit 2: Process of Entrepreneurship	
Factors affecting Entrepreneurship process	1. Exemplify one's capabilities in relation to the rigors of successful ventures (C3) 2. Identify and differentiates the different characteristics and competencies of an entrepreneurs (C2)
Unit 3: Business Plan writing	
Points to be considered, Model Business plan	1. Identify different business models (C3) 2. Describe different parts of a business plan(C2)
Unit 4: Case studies	
Indian and International Entrepreneurship	1. Perform self-assessment and analyse entrepreneurial personal traits and competencies (C4) 2. Evaluate oneself and plan courses of action to help develop one's entrepreneurial characteristics and competencies. (C5)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	12	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	-	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	03	-
Clinic	-	-
Practical	24	-
Revision	03	-
Assessment	06	-
TOTAL	**48**	**-**

Assessment Methods:	
Formative:	**Summative:**
Internal practical Test	Sessional examination
Theory Assignments	End semester examination
Lab Assignment & Viva	Viva

Mapping of assessment with Cos			
Nature of assessment	CO 1	CO 2	CO 3
Sessional Examination 1	*	*	
Sessional Examination 2			*
Assignment/Presentation		*	*
Laboratory Examination	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	1. NVR Naidu and T. Krishna Rao, "Management and Entrepreneurship", IK International Publishing House Pvt. Ltd 2008. *2.* Mohanthy Sangram Keshari, "Fundamentals of Entrepreneurship", PHI Publications, 2005 *3.* Butler, D. (2006). Enterprise planning and development. USA: Elsevier Ltd. Gerber, M.E. (2008) Awakening the entrepreneur within. NY: Harper Collins.

Name of the Program:	ME (Big Data and Data Analytics)
Course Title:	Project Work
Course Code: BDA 799	**Course Instructor:**
Academic Year: 2020 - 2021	**Semester:** Second Year, Semester 3, 4
No of Credits: 25	**Prerequisites:** SDLC, Communication Skills, technical skills.
Synopsis:	The project work aims to challenge analytical, creative ability and to allow students to synthesize, apply the expertise and insight learned in the core discipline. Students build self-confidence, demonstrate independence, and develop professionalism on successfully completion of the project.
Course Outcomes (COs):	On successful completion of this course, students will be able to
CO 1:	To be acquainted with working environment and processes that in place at the relevant Industries.
CO 2:	To familiarize the challenges as relevant professionals.
CO 3:	Review the literature and develop solutions for real time onboard projects.
CO 4:	Write technical report and deliver presentation.
CO 5:	Apply engineering and management principles to achieve project goal.

Mapping of COs to POs

COs	*PO 1*	*PO 2*	*PO 3*	*PO 4*	*PO 5*	*PO 6*	*PO 7*	*PO 8*	*PO 9*	*PO 10*	*PO 11*
CO 1						*	*	*	*	*	*
CO 2					*						
CO 3	*	*	*	*	*						
CO 4	*	*	*	*							
CO5:						*	*	*	*	*	*

Course content and outcomes:	
Content	***Competencies***
Phase 1:	
Problem identification, synopsis submission, status submission, mid evaluation.	At the end of the topic student should be able to: 1. Identify the problem/specification (C1) 2. Discuss the project (C2) 3. Prepare the outline (C3) 4. Prepare a mid-term project presentation report (C3) 5. Prepare and present mid-term project presentation slides (C3, C5) 6. Develop project implementation in hardware/software or both in chosen platform (C5)
Phase 2	
Status submission, final evaluation.	1. Prepare the progress report (C3) 2. Prepare the final project presentation report (C3) 3. Prepare and present final project presentation slides (C3, C5) 4. Modify and Develop implementation in hardware/software or both in chosen platform (C3, C5) 5. Justify the methods used and obtained results (C6)

Learning strategies, contact hours and student learning time		
Learning strategy	*Contact hours*	*Student learning time (Hrs)*
Lecture	-	-
Seminar	-	-
Quiz	-	-
Small Group Discussion (SGD)	14	-
Self-directed learning (SDL)	-	-
Problem Based Learning (PBL)	-	-
Case Based Learning (CBL)	-	-
Clinic	-	-
Practical	-	-
Revision	-	-
Assessment	-	-
TOTAL	**14**	**-**

Assessment Methods:	
Formative:	**Summative:**
Project Problem Selection	Mid-Term Presentation
Synopsys review	Second status review
First status review	Demo & Final Presentation

Mapping of assessment with Cos					
Nature of assessment	CO 1	CO 2	CO 3	CO 4	CO 5
Mid Presentation	*	*			
Presentation	*	*	*	*	*

Feedback Process	• End-Semester Feedback
Reference Material	Particular to the chosen project

PROGRAM OUTCOMES (POS) AND COURSE OUTCMES (COS) MAPPING

Sl.No.	Course Code	Course Name	Credits	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11
1	BDA 601	Fundamentals of Machine Learning	3	*	*	*	*							
2	BDA 602	Algorithms and Data Structures for Big Data	3	*	*	*	*					*		
3	BDA 623	Architecture of Big Data Systems	3	*	*	*		*	*	*			*	
4	MCL 601	Applied Probability and Statistics	3	*	*	*	*		*		*		*	
5	CSE-625	Mobile Web Application Development	3	*	*	*	*	*						
6	BDA-622	Principles of Data Visualization	3	*	*	*	*	*	*		*			
7	BDA 601L	Fundamentals of Machine Learning Lab	1	*	*	*								
8	BDA 602L	Algorithms and Data Structures for Big Data Lab	1	*	*	*		*	*			*		
9	BDA 623L	Architecture of Big Data Systems Lab	1	*	*	*		*	*			*	*	
10	MCL 601L	Applied Probability and Statistics Lab	1	*	*	*	*	*	*		*		*	
11	CSE-625L	Mobile Web Application DevelopmentLab	1	*	*	*	*							
	BDA-622L	Principles of Data Visualization Lab	1	*	*	*	*	*	*		*	*	*	
12	BDA 695	Mini Project - 1	4				*	*	*	*	*		*	*
13	BDA 697	Seminar - 1	1	*							*	*		*
14	BDA 605	Machine Learning for Big Data	3	*	*	*	*							
15	BDA 616	Modern Databases for Big Data	3	*	*	*		*						
16	MCL 602	Advanced Applications of Probability and Statistics	3	*	*	*	*	*	*		*			
17	BDA 618	Multimedia Analytics	3	*	*	*	*	*	*				*	

18	CDC-607	DevOps for Cloud	3	*	*	*		*	*	*	*	*	*	*
19	BDA-621	Natural Language and Text Processing	3	*	*	*	*	*						
20	ENP-601	Entrepreneurship	3	*		*	*		*		*		*	
21	BDA 605L	Machine Learning for Big Data Lab	1	*	*	*	*							
22	BDA 616L	Modern Databases for Big DataLab	1	*	*	*		*	*	*			*	
23	MCL 602L	Advanced Applications of Probability and StatisticsLab	1	*	*	*	*	*	*		*			
24	BDA 618L	Multimedia Analyticslab	1	*	*	*	*	*	*			*	*	
25	CDC-607L	DevOps for CloudLab	1	*	*	*		*		*	*			
26	BDA-621L	Natural Language and Text Processing Lab	1	*	*	*	*	*						
27	ENP-601L	Entrepreneurship Lab	1	*					*		*		*	
28	BDA 696	Mini Project - 2	4				*	*	*	*	*		*	*
29	BDA 698	Seminar - 2	1	*							*	*		*
30	BDA 799	Project Work	25	*	*	*	*	*	*	*	*	*	*	*