Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

SYLLABUS

DR VISHWANATH KARAD
MIT - WORLD PEACE UNIVERSITY

SCHOOL OF COMPUTER ENGINEERING & TECHNOLOGY

M.TECH. DATA SCIENCE & ANALYTICS

< BATCH – 2018-19 >

PROGRAMME STRUCTURE

***Preamble**:*

The Computer Science and Engineering is the most sought after branch of Engineering in today's world. With the advancements in hardware and software technologies, there is huge scope for development of a wide range of applications. The Internet and allied technologies had connected the world so immensely that the world is now a "Global Village" The students of MITWPU will be tomorrow's global leaders, researcher, entrepreneurs and change-makers - MITWPU has the objective to make them competent for global scenarios.

The M.Tech (CSE) curriculum offers a varied range of subjects that fall into the core, specialization and basic computer science stream. The course also has provisions for pursuing Seminar, Projects, Internships, interdisciplinary projects as a prudential aspect of the course curriculum. The value based education is ensured by offering Peace related subjects and Yoga practice.

The curriculum is based on the theme of "Continuous Evaluation". Theory and laboratory components are given appropriate importance. The communication skills are enhanced through seminar component. Industry exposure is given through internships / projects, and development of latest tools / technologies is cached through the components of "Add-on skills".

The curriculum will transform the students into winning personalities.

Dr. M. V. Bedekar
Chairman, BoS in Computer Engineering & Technology
Professor & Head, School of Computer Engineering & Technology

Dr. L. K. Kshirsagar
Dean,
Engineering & Technology

Vision and Mission of the Programme

Vision

- To build a value based academic centre of excellence in Computer Engineering.

Mission

- To create an ambience nurturing integrity, discipline, technical knowledge and research in the emerging areas of Computer Engineering.

Programme Educational Objectives

- Adapt to rapidly changing technical scenario.
- Lead teams of multidisciplinary professionals with the sense of integrity, discipline and social responsibility.
- Design and develop systems in various domains.
- Demonstrate Innovative and Entrepreneurial spirit in their professional careers.

Programme Specific Outcomes

- Ability to model, design and develop computer based systems to solve real life problems by applying knowledge of Mathematical Foundations, Algorithmic Principles and Computer Science theory.
- Ability to learn and use advancements in tools and technologies on the basis of software and hardware fundamentals, logic building and analytical skills; so as to cope up with the rapidly changing technical scenarios.
- The ability to perform information processing using high performance computation leading to innovative solutions to solve societal problems.

Foundation / Orientation Programme

The students admitted to the Computer Science and Engineering programme belong to varied backgrounds and possess different levels of technical awareness. Thus, a one day orientation programme is conducted for the students to acquaint them with the overall programme structure and its prospectus. Students are also briefed about the current technical trends and domains, project prospectus, co-curricular, extra-curricular activities along with recruitment opportunities and avenues. The orientation programme helps the students to choose their specialization subject interests towards the fulfilment of their carrier goals and future education.

Programme Structure *:*

(a) Programme duration: 2 Year

(b) System followed: Trimester

(c) Credits System:

Year	Trimester I	Trimester II	Trimester III	Total
FY M Tech	11	11	13	35
SY M Tech	11	10	10	31
				Total : 35+31 = 66

(d). Credits for activities other than academics - *Not Applicable*

(e) Internship: *Not Applicable*

(f) Assessment Criteria:

As per norms of MIT-WPU

(g) Branches or Specialisations: Network Management & Cyber Security

(h) Mandatory Attendance to appear for examination: <90>% (*Para 13.1. of AO : 2017*)

(j) Medium of Instruction and Examination: <*English*>(*Para 9. of AO : 2017*)

(k) Eligibility criteria for admission to the programme <*as per para 4 of AO 2017*)

M. Tech Computer Science & Engineering

2018-19

A. Definition of Credit:-

3 Hr. Lecture / Tutorial per week	*2credit*
2HoursPractical(Lab) per week	*1credit*

B. Credits:-

Total number of credits for Two year postgraduate <***M. Tech***>Programme would be 66.

C. Structure of Credits for Postgraduate<*M. Tech*>program:-

S. No.	Category	Suggested Breakup of Credits(Total 66)
1	*Humanities and Social Sciences and Peace Programmes including Management courses*	*08*
2	*Engineering Science courses*	*04*
3	*Professional core courses*	*22*
4	*Professional Elective courses relevant to chosen specialization/branch*	*10*
5	*Project work, seminar*	*22*
	Total	***66***

D. Course code and definition:-

Course code	Definitions
L	*Lecture*
T	*Tutorial*
ES	*Engineering Science Courses*
WP	*Humanities and Social Sciences and Peace Programs including Management courses*
ME	*Mechanical Engineering Courses*
EC	*Electronics and Communication*
EE	*Electrical Engineering*
CH	*Chemical Engineering*
CS	*Computer Science and Engineering*
PO	*Polymer Engineering*
CE	*Civil Engineering*
PE	*Petroleum Engineering*

E. Grading Scheme:

(*According to Para 12.1 of Academic Ordinances 2017*)

M.Tech. (First Year) (Batch 2018-19)

Trimester – I

Sr. No.	Course Code	Name of Course	Type	Weekly Workload, Hrs			Credits		Assessment, Marks			
				Theory	Tutorial	Lab	Th	Lab	CCA*	LCA*	End Term Test	Total
1	CSD511	Research Methodology	Engg. Science	3	-	--	2	-	50	-	50	100
2	CSD512	Algorithms	Core	3	-	--	2	-	50	-	50	100
3	CSD513	Data Pre-processing & Data Warehousing	Core	3	-	--	2	-	50	-	50	100
4	WPC1	World Famous Philosophers, Sages/Saints and Great Kings	Humanities-WPC	3	-	-	2	-	70	-	30	100
5	CSD514	Lab Practice-I	Core	-	-	6	-	3	-	50	50	100
6	WPC3	Yoga- For Winning Personality	WPC	-	-	-	-	-	-	-	-	-
		Total :		12	-	6	8	3	220	50	230	500

Type: (Refer Para 11 of Academic Ord. 2017)

****Assessment Marks are valid only if Attendance criteria are met**

Weekly Teaching Hours: 18

Total Credits: First Year *M.Tech.* Trimester I: 11

***** CCA : Class Continuous Assessment

***** LCA : Laboratory Continuous Assessment

Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

M.Tech. (First Year) (Batch 2018-19)

Trimester – II

Sr. No.	Course Code	Name of Course	Type	Weekly Workload, Hrs			Credits		Assessment Marks **			
				Theory	Tutorial	Lab	Th	Lab	CCA*	LCA*	End Term Test	Total
1	CSD521	Probability & Statistics	Core	3	-	--	2	-	50	-	50	100
2	CSD522	Foundations of Data Mining	Core	3	-	--	2	-	50	-	50	100
3	CSD523	Elective-I	Departmental Elective	3	-	--	2	-	50	-	50	100
4	WPC4	Philosophy of Science and Religion/Spirituality	WPC	3	-	-	2	-	70	-	30	100
5	CSD524	Lab Practice-II	Core	-	-	6	-	3	-	50	50	100
6	WPC3	Yoga-For Winning Personality	WPC	-	-	-	-	-	-	-	-	-
		Total :		**12**	-	**6**	**8**	**3**	**220**	**50**	**230**	**500**

Type: (Refer Para 11 of Academic Ord. 2017)

****Assessment Marks are valid only if Attendance criteria are met**

Weekly Teaching Hours: 18

***** CCA : Class Continuous Assessment

Total Credits: First Year *M.Tech.* Trimester II: 11

***** LCA : Laboratory Continuous Assessment

Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

M.Tech. (First Year) (Batch 2018-19)
Trimester – III

Sr. No.	Course Code	Name of Course	Type	Weekly Workload, Hrs			Credits		Assessment Marks**			
				Theory	Tutorial	Lab	Th	Lab	CCA*	LCA*	End Term Test/ Oral	Total
1	CSD531	Machine Learning	Core	3	-	--	2	-	50	-	50	100
2	CSD532	Elective – II	Departmental Elective	3	-	--	2	-	50	-	50	100
3	CSD533	Elective – III	Departmental Elective	3	-	--	2	-	50	-	50	100
4	WPC2	Study of Languages, Peace in Communication and Human Dynamics	WPC	3	-	--	2	-	70	-	30	100
5	CSD534	Lab Practice-III	Core	-	-	6	-	3	-	50	50	100
6	CSD535	Seminar-I	Inter-disciplinary	-	-	4	-	2	-	50	50	100
	WPC3	Yoga-For Winning Personality	WPC	-	-	-	-	-	-	-	-	-
		Total :		**12**		**10**	**8**	**5**	**220**	**100**	**280**	**600**

Weekly Teaching Hours: 22

Total Credits: First Year *M.Tech.* Trimester III: 13

Total First Year *M.Tech.*Credits: 11+11+13=35

***** CCA : Class Continuous Assessment

***** LCA : Laboratory Continuous Assessment

Dr. Siddhivinayak Kulkarni
PG Coordinator (DSA)

Dr. M. V. Bedekar
Professor & Head

M.Tech. (Second Year) (Batch 2018-19)

Trimester – IV

Sr. No.	Course Code	Name of Course	Type	Weekly Workload, Hrs			Credits		Assessment Marks**			
				Theory	Tutorial	Lab	Th	Lab	CCA*	LCA*	End Term Test/ Oral	Total
1	CSD611	Business Intelligence	Core	3	-	-	2	-	50	-	50	100
2	CSD612	Web Mining	Core	3	-	-	2	-	50	-	50	100
3	CSD613	Elective – IV Data Security and Data Privacy	Interdisciplinary	3	-	-	2	-	50	-	50	100
4	CSD614	Project Stage-I Seminar	Core	-	-	4	-	2	-	50	50	100
5	CSD615	Lab Practice-IV	Core	-	-	6	-	3	-	50	50	100
6	WPC3	Yoga- for Winning Personality	WPC	-	-	-	-	-	-	-	-	-
		Total :		**9**	-	**10**	**6**	**5**	**150**	**100**	**250**	**500**

Type: (Refer Para 11 of Academic Ord. 2017)

Weekly Teaching Hours: 19

Total Credits: Second Year *M.Tech.* Trimester I: 11

****Assessment Marks are valid only if Attendance criteria are met**

* CCA: Class Continuous Assessment

* LCA: Laboratory Continuous Assessment

M.Tech. (Second Year) (Batch 2018-19)

Trimester – V

Sr. No.	Course Code	Name of Course	Type	Weekly Workload, Hrs			Credits		Assessment Marks**			
				Theory	Tutorial	Lab	Th	Lab	CCA*	LCA*	End Term Test	Total
1	CSD621	Elective – V	Self Study	2		-	1	-	-	-	50	50
2	CSD622	Project Stage-II Seminar	Core	-		18	-	9	-	100	50	150
3	WPC3	Yoga-For Winning Personality	WPC	-		-	-	-	-	-	-	-
		Total :		**2**		**18**	**1**	**9**	**-**	**100**	**100**	**200**

Type: (Refer Para 11 of Academic Ord. 2017)

Weekly Teaching Hours: 20

Total Credits: Second Year *M.Tech.* Trimester II: 10

****Assessment Marks are valid only if Attendance criteria are met**

* CCA : Class Continuous Assessment

* LCA : Laboratory Continuous Assessment

M.Tech. (Second Year) (Batch 2018-19)

Trimester – VI

Sr. No.	Course Code	Name of Course	Type	Weekly Workload, Hrs			Credits		Assessment Marks**			
				Theory	Tutorial	Lab	Th	Lab	CCA*	LCA*	End Term Test	Total
1	CSD631	Elective-VI	Skill / Web-based	2	-	-	1	-	-	50	50	100
2	CSD632	Project Stage-III Seminar	Core	-	-	18	-	9	-	100	100	200
3	WPC3	Yoga-For Winning Personality	WPC	-	-	-	-	-	-	-	-	-
		Total :		**2**		**18**	**1**	**9**	**-**	**150**	**150**	**300**

Type: (Refer Para 11 of Academic Ord. 2017)

Weekly Teaching Hours: 20

Total Credits: Second Year *M.Tech.* Trimester III: 10

Total Second Year *M.Tech.*Credits: 11+10+10=31

Total M.Tech Credits: = 35 +31 = 66 Credits

****Assessment Marks are valid only if Attendance criteria are met**

* CCA : Class Continuous Assessment

* LCA : Laboratory Continuous Assessment

Dr. Siddhivinayak Kulkarni
PG Coordinator (DSA)

Dr. M. V. Bedekar
Professor & Head

Elective Courses:

	Data Science and Analytics		Data Science and Analytics	
	Code	Title	Code	Title
Elect I	CSD523A	Data Storage and Management		
Elect II	CSD532A	Data Analytics with R	CSD532B	Data Analytics with Python
Elect III	CSD533A	Cognitive Computing		
Elect IV	CSD613A	Data Security & Data Privacy	CSD613B	Data Encryption & Data Compression
Elect V	CSD621A	Big Data Analytics	CSD621B	Natural Language Processing & Information Retrieval
Elect VI	CSD631A	Ubiquitous Computing	CSD631B	Deep Learning

Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

COURSE STRUCTURE

Course Code	CSD511			
Course Category	Engineering Science			
Course Title	**Research Methodology**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**3**	**--**	**--**	**3**

Pre-requisites: -

Course Objectives:

1. Knowledge (i) To understand basic concepts of research and research methodology

(ii) To learn the methodology to conduct the literature survey

2. Skills (i) To use tools and techniques for carrying out research

(ii) To learn the effective way of research report writing skills and documentation

3. Attitude (i) To become aware of ethics in research and plagiarism

Course Outcomes:

1. Carry out the research work in a methodological way

2. Apply data analysis methods to generate results & drawing conclusion

3. Apply appropriate tools for research report/ paper writing

4. Write research report and research proposal

Course Contents:

- Introduction to Research Methodology
- Literature Review, Data Collection and Sampling
- Modeling and Statistical Analysis
- Research Reports

Learning Resources:

Reference Books:

1.C.R. Kothari, Research Methodology Methods and Techniques, 3nd edition, New Age, 2014
2. Ranjit Kumar, Research Methodology: A Step by Step Guide for Beginners, 4th Edition. Sage Publication 2014
3. Briony J. Oates, Researching information systems and computing, Sage Publications, 2006
4. Stuart Melville and Wayne Goddard, Research methodology: An introduction for science &engineering students. 2nd edition, 2001

Supplementary Reading:

Web Resources:

Weblinks: *http://www.bbamantra.com/research-methodology/*
https://study.com/academy/topic/introduction-to-research-methods.html
https://www.openlearning.com/courses/introduction-to-research-methodology

MOOCs: *https://www.mooc-list.com/course/understanding-research-methods-coursera*

Pedagogy:

Powerpoint Slides
Board-marker
Two Teacher
Video lectures

Assessment Scheme:

Class Continuous Assessment (CCA) *(with % weights)*

Attendance	Assignments	Mid Term Exam	Total
10	20	20	50

Laboratory Continuous Assessment (LCA) *(with % weights)*

Design and Implementation	Performance of Experiment	Result analysis and Reporting	Attendance/ Discipline/ Initiative/ Behavior	Total
15	15	10	10	50

Term End Examination: *50%*

Syllabus:

Module No.	Contents	Workload in Hrs		
		Theory	*Lab*	*Assess*
1	**Introduction to Research Methodology** Meaning, scope and significance of research, objectives of research, motivation for research, types of research, research methods Vs methodology, research process, definition of good research problem, case study on IEEE & ACM code of Ethics	**8**		
2	**Literature Review, Data Collection and Sampling** Reviewing the literature, types of publications, research article reading, identifying gaps/areas, formulation of a research problem, Hypothesis- meaning, types, development of hypothesis and its testing, developing research plans, data collection types and methods, sampling methods	**8**		
3	**Modeling and Statistical Analysis** Types of analysis, statistics in research, Processing and analysis of data, Data analysis skills, Graphical representation and mapping of data using SPSS/GNU PSPP, Summarizing Data with Histograms and Descriptive Statistics, Distributions, Statistical and multivariate analysis, Correlation and regression, Error analysis	**8**		
4	**Research Reports** Significance of report writing, Different steps of writing report, Layout of the Research report, Types of reports, Mechanisms and tools for writing research report, Detect and avoid plagiarism, Research proposal preparations, Case study of various research grants and funding resources, publishing research work	**6**		

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

COURSE STRUCTURE

Course Code	CSD512			
Course Category	Core			
Course Title	**Algorithms**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**3**	**--**	**--**	**3**

Pre-requisites: -

Course Objectives:

1. **Knowledge (i)** To learn a familiarity with major algorithms and data structures.
2. **Skills (i)** To Analyze the asymptotic performance of algorithms
3. **Attitude (i)** Synthesize efficient algorithms in common engineering design situations

Course Outcomes:

1. Compare between different data structures. Pick an appropriate data structure for a design situation.
2. Analyze worst-case running times of algorithms using asymptotic analysis.
3. Describe the divide-and-conquer, dynamic-programming, parallel paradigms and explain when an algorithmic design situation calls for it.
4. Analyze randomized algorithms. Explain what an approximation algorithm is, and the benefit of using approximation algorithms

Course Contents:

- **Review of Basic Concept**
- **Algorithm Design techniques**
- **Parallel Algorithms**
- **Approximation Algorithms**

Learning Resources:

Reference Books:

1. T. H. Cormen, C. E. Lieserson, R. L. Rivest, and C. Stein, Introduction to Algorithms, Third Edition, MIT Press, 2009
2. Horowitz, Sahni, Galgotia, Fundamentals of Computer Algorithms, 2nd Edition, University Press, 1997
3. S. Dasgupta, C. Papadimitrou, U Vazirani, Algorithms, Mc Graw Hill, 2006
4. Fayez Gebali, Algorithms and Parallel Computing, Willy, 2011
5. J. Kleinberg and E. Tardos, Algorithm Design, Pearson Education Limited, 2006
6. G. Bressard and P. Bratley, Fundamentals of Algorithms, PHI, 1988

Supplementary Reading:

Web Resources:

Weblinks: *http://faculty.simpson.edu/lydia.sinapova/www/cmsc250/LN250_Weiss/L28-Design.htm*
http://cgi.csc.liv.ac.uk/~ped/teachadmin/algor/algor_complete.html

MOOCs: *http://nptel.ac.in/courses/106101060/5*
https://www.coursera.org/specializations/algorithms

Pedagogy:

Two Teacher
Videos
Powerpoint Slides
Marker and Board

Assessment Scheme:

Class Continuous Assessment (CCA) *(with % weights)*

Attendance	Assignments	Mid Term Exam	Total
10	20	20	50

Laboratory Continuous Assessment (LCA) *(with % weights)*

Design and Implementation	Performance of Experiment	Result analysis and Reporting	Attendance/ Discipline/ Initiative/ Behavior	Total
15	15	10	10	50

Term End Examination : *50%*

Syllabus :

Module No.	Contents	Workload in Hrs		
		Theory	*Lab*	*Assess*
1	**Review of Basic Concept** Asymptotic analysis: upper and lower bounds, Standard complexity classes, Analyzing recursive algorithms using recurrence relations (Master's Theorem and Substitution method). Analysis of recursive and non-recursive algorithms.	8		
2	**Algorithm Design techniques** Divide & Conquer, Greedy Algorithm, Dynamic Programming, Backtracking, Branch& Bound. NP hard and NP complete theory to be included with one example on NP Complete Proof.	8		
3	**Parallel Algorithms** Model for parallel computation, Basic Techniques, Complete Binary Tree, Pointer Doubling, Prefix Computation, Selection, Merging, Sorting, Graph Problems.	8		
4	**Approximation Algorithms** Approximation ratio, approx..schemes, Vertex Cover and TSP using approx. algorithm Randomized Algorithms: Closest pair, pattern matching, MST	8		

Prepared By

Prof. Anita Thengade
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

COURSE STRUCTURE

Course Code	CSD513			
Course Category	Core			
Course Title	Data Pre-processing & Data Warehousing			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**3**	**--**	**--**	**3**

Pre-requisites: -

Course Objectives:

1. Knowledge (i) To understand the types and characteristics of Data.

(ii) To understand the necessity of data-pre-processing.

2. Skills (i) To learn the applicability of the techniques in pre-processing

(ii) To understand the usage of data warehouse.

3. Attitude (i) To relate the pre-processing techniques to real-world applications.

Course Outcomes:

1. Recognize the differences between Structured, Unstructured and Semi-structured data, as well as the opportunities surrounding Linked-Data.
2. Characterize and criticality assess good and bad data in the context of data-driven decision making.
3. Apply the pre-processing techniques to real-life examples.

Course Contents:

- **Introduction**
- **Data Pre-preprocessing**
- **Data warehouse**
- **Building a data warehouse**

Learning Resources:

Reference Books:

1. M. Shron, O'Reilly, Thinking with Data: How to Turn Information into Insights, Publisher: O'Reilly Media, 2014
2. T. Fawcett and F. Provost , Data Science for Business: What you Need to Know about Data Mining and Data Analytic Thinking, Publisher: O'Reilly Media, 2013
3. Ralph Kimball, Margy Ross, The Data Warehouse Toolkit, 3rd edition, Publisher: Wiley, 2013
4. J. Han, M. Kamber and J. Pei, Morgan Kaufmann, Data Mining, Concepts and Techniques, Publisher: Elsevier, 2006
5. Ramesh Sharda, Dursun Delen, Efraim Turban, Business Intelligence and Analytics: Systems for Decision Support, Publisher: Pearson/Prentice Hall, 2014

Supplementary Reading:

Web Resources:

Weblinks: *http://www.prolearninghub.com/courses/data-warehouse-concepts-design-data-integration/*
https://en.wikipedia.org/wiki/Data_pre-processing

MOOCs: *https://www.mooc-list.com/tags/data-warehouse*

Pedagogy:

- Chalk and Board
- PPT
- Two Teacher Method

Assessment Scheme:

Class Continuous Assessment (CCA) *(with % weights)*

Attendance	Assignments	Mid Term Exam	Total
10	20	20	50

Laboratory Continuous Assessment (LCA) *(with % weights)*

Design and Implementation	Performance of Experiment	Result analysis and Reporting	Attendance/ Discipline/ Initiative/ Behavior	Total
15	15	10	10	50

Term End Examination : *50%*

Syllabus:

Module No.	Contents	Workload in Hrs		
		Theory	*Lab*	*Assess*
1	**Introduction** Data objects and attribute types, Data Characteristics, Types of Data, Structured, Unstructured, Semi-structured, Discrete, Continuous, Ordinal, Time series data, Geographical data, Big data, etc.., Data Visualization and Manipulation, Data storage, security, processing, governance of Data, data scraping, data cleansing and data deduplication, Measuring Data Similarity and Dissimilarity.	**8**		
2	**Data Pre-preprocessing** Data Pre-preprocessing, Data Cleaning, Data Integration, Data reduction techniques – Principal component analysis, attribute subset selection, introduction to sampling, clustering, aggregation concepts, data transformation, strategies, normalization techniques, discretization – binning, histogram analysis, decision tree and correlation analysis, concept hierarchy for nominal data.	**8**		
3	**Data warehouse** Data warehouse concepts, Data warehouse modeling, Data Cube and OLAP, schemas for multidimensional data models, concept hierarchy, measures, indexing techniques. Data warehouse – design and usage, implementation, architectural components, Role of Metadata, Dimensional Modeling, Data Extraction, Transformation and Loading, Data Quality.	**8**		
4	**Building a data warehouse** Steps in building a data warehouse, building a data warehouse using a relational DBMS, Data Warehouse Implementation and Deployment, Growth and Maintenance. Data Warehousing and the Web, Current trends in Data Warehousing (Distributed Warehousing).	**8**		

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

COURSE STRUCTURE

Course Code	**CSD521**			
Course Category	**Engineering Science**			
Course Title	**Probability and Statistics**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**2**	**-**	**--**	**2**

Pre-requisites:-

Course Objectives:

1. To understand the concepts of correlation and regression used in data analysis.
2. To learn discrete and continuous probability distributions models.
3. To learn about parametric and non-parametric tests used in hypothesis testing.
4. To understand the random process and Markov chains.

Course Outcomes:

On completion of the course, students should be able to

1. use statistical methodology and tools in the engineering problem-solving process.
2. analyze the given probabilistic model of the problem.
3. use parametric and non-parametric tests for hypothesis testing.
4. solve problems related to stochastic processes.

Course Contents:

- Statistical Methods
- Probability
- Tests of Hypotheses
- Stochastic Process

Learning Resources:

Reference Books:

1. Trivedi Kishor S., "Probability & Statistics with Reliability, Queuing and Computer Science Applications", Second Edition, Wiley Student Edition, 2012.
2. Ross Sheldon M., "Introduction to Probability and Statistics for Engineers and Scientists", Fifth Edition, 2014.
3. Gupta S. C. and Kapoor V. K., "Fundamentals of Applied Statistics", Third Edition, S. Chand and Sons, New Delhi, 1987.
4. DeGroot Morris H. and Schervish Mark J., "Probability and Statistics", Fourth Edition, Pearson New International Edition, 2010

Supplementary Reading:

1. Montgometry Douglas C.andRunger George C., "Applied Statistics and Probability for Engineers", Third Edition, Wiley Student Edition, 2008.

Web Resources:

Web links:

https://www.tutorialspoint.com/statistics/probability.htm

MOOCs:

http://nptel.ac.in/courses/111105090/16
http://nptel.ac.in/courses/111105090/46
http://nptel.ac.in/courses/111105090/76

Pedagogy:

- Chalk and Board
- PPT

- Two Teacher Method

Assessment Scheme:

Class Continuous Assessment (CCA)- 50 Marks (100%)

Attendance	Assignments	Mid Term Exam	Total
20%	40%	40%	100%

Term End Examination: 50 marks (100%)

Syllabus:

Theory:

Module No.	Contents	Workload in Hrs		
		Theory	Lab	Assess
1	**Statistical Methods** Measure of central tendency, dispersion, Measure of Skewness and Kurtosis, Correlation, Partial Correlation, regression analysis, Linear and nonlinear regression, standard error estimate.	8	--	--
2	**Probability** Basic probability theory, probability axioms, conditional probability, Bayes' theorem, Discrete and continuous random variables, probability mass function, probability density function, distribution function, mathematical expectation, probability distributions, Binomial, Poisson, Geometric and Normal distributions.	8	-	-
3	**Tests of Hypotheses** Parameterand Statistic, Hypothesis testing, Null and Alternative hypothesis, Level of Significance, Parametric and non-parametric tests, Chi-square Test for goodness of fit and independence of attributes. Student'st- test for small samples.	8	-	-
4	**Stochastic Process** Introduction and classification of stochastic processes, Markov process, Transition probability, Transition probability matrix, First order and higher order Markov process, Markov chain, Steady state condition.	8	-	-

Prepared By	Checked By	Approved By
Dr. N. D. Kankane	Dr. S. A. Kulkarni	Dr. M. V. Bedekar
Course Coordinator	**Course Advisor**	**Chairman BOS**

COURSE STRUCTURE

Course Code	CSD522			
Course Category	**Engineering Core**			
Course Title	**Foundations of Data Mining**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**3**	**--**	**--**	**2**

Pre-requisites:

- **Discrete Structure**
- **Data Communication**
- **Computer Networks**

Course Objectives:

1. Knowledge: (i) To Identify the scope and necessity of Data Mining

(ii) To learn what are data mining algorithms and when and how to apply them.

2. Skills: (i) To explore how algorithm parameters and data properties affect the effectiveness of data mining methods used.

3. Attitude: (i) To solve real world problems using Data Mining algorithms

Course Outcomes:

On completion of the course, students should be able to,

1. Understand the data mining algorithms and the how to apply them.
2. Understand the application of data mining algorithm and its suitability for a given problem.
3. Apply the techniques of clustering, classification, association finding, feature selection and visualization to real world data.

Course Contents:

Introduction to Data Mining

Association Rule Mining And Classification

Clustering And Applications

Regression and Other Methods

Learning Resources:

Reference Books:

1. Jiawei Han and Micheline Kamber, "Data Mining Concepts and Techniques", Third Edition, Elsevier, 2012.
2. Alex Berson and Stephen J. Smith, "Data Warehousing, Data Mining & OLAP", Tata McGraw – Hill Edition, Tenth Reprint 2007.
3. Pang-Ning Tan, Michael Steinbach and Vipin Kumar, "Introduction to Data Mining", Person Education, Third Edition 2007.

Supplementary Reading:

Web Resources:

https://en.wikipedia.org/wiki/Association_rule_learning
https://www.investopedia.com/terms/r/regression.asp
https://en.wikipedia.org/wiki/Cluster_analysis

Weblinks:

MOOCs:

https://www.mooc-list.com/course/cluster-analysis-data-mining-coursera
https://www.coursera.org/specializations/data-mining

Pedagogy:

Powerpoint Slides

Board-marker

Two Teacher

Assessment Scheme:

Class Continuous Assessment (CCA)- 50 Marks (100%)

Attendance	Assignments	Mid Term Exam	Total
20%	40%	40%	100%

Term End Examination: 50 marks (100%)

Syllabus:

Module No.	Contents	Workload in Hrs		
		Theory	Lab	Assess
1	**Introduction to Data Mining** Introduction to data Mining, Data, Stages of Data Mining, Functionalities of data mining, Interestingness of patterns, Data mining task primitives, Integration of data mining with data warehousing, Predictive and Descriptive Data Summarization, Mining Methods, Frequent Pattern Mining, Association and Correlations, Issues in data mining, Multi Dimensional Measure of Data Quality, Data Mining Applications	7		

2	**Association Rule Mining And Classification** Association, Association Algorithm, Classification and Prediction, Basic Concepts, Decision Tree induction, Bayesian Classification, Bayesian Theorem, Towards Naïve Bayesian Classifier, Rule Based Classification, Classification by back propagation, Support vector machines, Associative Classification, Lazy learners, Understanding Prediction	**7**		
3	**Clustering And Applications** Cluster analysis, Categorization of Major Clustering Methods, K means, Hierarchical Methods, Density-Based Methods, Grid Based Methods, Model-Based Clustering Methods, Clustering High Dimensional Data, Constraint – Based Cluster Analysis, Cluster Quality, Outlier Analysis, Distance Based Agglomerative and Divisible Regression, Other Clustering Methods	**8**		
4	**Regression and Other Methods** Linear regression; Nonparametric regression; Evaluation for regression; Model selection and over-fitting .Introduction to Recommendation Systems: Collaborative Filtering, Text mining, Neural Nets, Text DM, Web Data Mining, Graph Mining, Mining Data Streams, Mining Time Series and Multimedia Data Mining, Comparison of Data Mining Techniques	**8**		

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

COURSE STRUCTURE

Course Code	**CSD523A**			
Course Category	**Engineering Core**			
Course Title	**Elective-I Data Storage and Management**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**3**	**--**	**--**	**2**

Pre-requisites:

- **Data Pre-processing and Data Management**

Course Objectives:

1. Knowledge: (i) Students should understand the general scope of data storage and management. (ii) Students should understand logical and physical components of a storage infrastructure.

2. Skills: (i) Students should be able to evaluate storage architectures, including storage subsystems.

3. Attitude: (i) To apply novel techniques for minimizing data storage and apply it for various problems

Course Outcomes:

On completion of course, students should be able to

1. Identify correct storage solutions as per problems needs.
2. Identify appropriate logical and physical components required to develop and deploy the storage solution.
3. Select and define appropriate storage architecture solutions.

Course Contents:

Introduction to Information Storage Management

Storage Systems Architecture

Backup, Recovery, Archival

Business Continuity & Cloud based Storage

Learning Resources:

Reference Books:

1. EMC Corporation, "Information Storage and Management", Wiley India, 2nd Edition, 2012.
2. Robert Spalding, "Storage Networks: The Complete Reference", Tata McGraw Hill, Osborne, 2003.
3. Marc Farley, "Building Storage Networks", Tata McGraw Hill, Osborne, 2nd Edition, 2000.

Supplementary Reading:

Web Resources:

Weblinks:

http://www.storagesearch.com/auspexart.html
https://www.commvault.com/solutions/by-function/data-protection-backup-and-recovery
http://perspectives.avalution.com/2015/what-you-need-to-know-cloud-computing-and-business-continuity/

MOOCs:

https://www.mooc-list.com/course/introduction-data-storage-and-management-technologies-edx
https://www.mooc-list.com/tags/data-storage

Pedagogy:

Powerpoint Slides
Board-marker
Two Teacher
Video lectures

Assessment Scheme:

Class Continuous Assessment (CCA)- 50 Marks (100%)

Attendance	Assignments	Mid Term Exam	Total
20%	40%	40%	100%

Term End Examination: 50 marks (100%)

Syllabus:

Module No.	Contents	Workload in Hrs		
		Theory	Lab	Assess
1	**Introduction to Information Storage Management** Review data creation and the amount of data being created and understand the value of data to a business, challenges in data storage and data management, Solutions available for data storage. Core elements of a data center infrastructure, role of each element in supporting business activities. Intelligent Storage System, Components of an Intelligent Storage System, Types of Intelligent Storage Systems.	8		
2	**Storage Systems Architecture** Physical and logical components of a connectivity environment, Major physical components of a disk drive and their function, logical constructs of a physical disk, access characteristics, and performance Implications. RAID and its components, Various RAID levels. Integrated v/s modular storage systems, High-level architecture and working of an intelligent storage system. SAN and Its Evolution. NAS Devices, Benefits of NAS File Systems, Components of NAS.	8		
3	**Backup, Recovery, Archival** Backup Purpose, Backup Considerations, Backup Granularity, Backup Methods, Backup Architectures, Backup and Restore Operations, Recovery Considerations, Archival considerations.	7		
4	**Business Continuity & Cloud based Storage** Business Continuity, BC Terminology, BC Planning Life Cycle, Failure Analysis, Business Impact Analysis, Information Availability. Cloud Enabling Technologies, Characteristics of Cloud Storage, Benefits of Cloud Storage, Cloud Service Models, Cloud Deployment models, Cloud Storage Infrastructure, Cloud Challenges.	7		

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

COURSE STRUCTURE

Course Code	**CSD524**			
Course Category	**Engineering Core**			
Course Title	**Lab Practice-II**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**--**	**--**	**6**	**3**

Co-requisites:

- Probability & Statistics
- Foundations of Data Mining
- Data Storage and Management

Course Objectives: (i) To learn and apply theoretical concepts of probability and statistics, data mining and data storage-management for assignments

Course Outcomes: (i) To apply skills for solving new problems in related areas.

Course Contents: **Part A : Probability & Statistics**

Part B : Foundations of Data Mining

Part C : Data Storage and Management

Assessment Scheme:

Laboratory Continuous Assessment (LCA) 50 marks (100%)

Design and Implementation	Performance of Experiment	Result analysis and Reporting	Attendance/ Discipline/ Initiative/ Behavior	Total
30%	30%	20%	20%	100%

Term End Examination: 50 marks (100%)

Syllabus:

Module No.	Contents	Workload in Hrs		
		Theory	**Lab**	**Assess**
Part A : Probability & Statistics				
1	To determine the coefficient of skewness and kurtosis for the given data.		2	
2	To find correlation between variables in bivariate data and estimation of parameters.		2	

3	To test the goodness of fit using Chi square distribution.		2	
4	Test the given hypothesis using Student's t distribution for small samples.		2	
5	Use of Poisson distribution in real world application with a small chance of success.		2	
6	To find mean and variance in normally distributed data.		2	
Part B : Foundations of Data Mining				
1	Demonstration of preprocessing on given dataset.		2	
2	Demonstration of Association rule process on given dataset using apriori algorithm.		2	
3	Demonstration of classification rule process on given dataset using id3/ j48 algorithm.		2	
4	Demonstration of classification rule process on given dataset using naïve bayes algorithm.		2	
5	Demonstration of clustering rule process on given dataset using simple k-means.		2	
Part C : Data Storage and Management				
1	To deploy and study the usage of IO Meter tool.		2	
2	To deploy and study usage of IO Zone tool.		2	
3	Examine various types RAID levels by analyzing their applicability and configuration.		2	
4	To install and document the use of Vnxe Simulator with respect to a NAS and SAN configuration.		2	
5	Analyze the purpose of backup and develop a configuration for backup system for banking system.		2	

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman, BOS

COURSE STRUCTURE

Course Code	**CSD531**			
Course Category	**Engineering Core**			
Course Title	**Machine Learning**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**3**	**--**	**--**	**2**

Pre-requisites:

-

Course Objectives:

1. Knowledge: (i) To understand machine learning, different types of machine learning and their applications
   (ii) To learn neural networks, Genetic algorithms, Hidden Markov Models
2. **Skills:** (i) To use tools and techniques for carrying out machine learning research
3. **Attitude:** (i) To develop new and efficient algorithms for machine learning

Course Outcomes:

On completion of course, students should be able to

1. Implement different learning techniques and carry out the research work in machine learning
2. Apply appropriate tools for research
3. Use neural networks, Genetic algorithms, Hidden Markov Models as per requirements

Course Contents:

Introduction to Machine Learning

Supervised and Semi-supervised Learning

Unsupervised and Reinforcement Learning

Neural Networks and Genetic Algorithms

Learning Resources:

Reference Books:

1. T. Mitchell, Machine Learning, McGraw-Hill, 1997.
2. E. Alpaydin, Introduction to Machine Learning, PHI, 2004.

Supplementary Reading:

Web Resources:
https://www.coursera.org/learn/machine-learning
https://www.toptal.com/machine-learning/machine-learning-theory-an-introductory-primer

Weblinks:
https://www.kdnuggets.com/2016/07/top-machine-learning-moocs-online-lectures.html

MOOCs:
https://www.mooc-list.com/course/machine-learning-regression-coursera

Pedagogy:

Powerpoint Slides
Board-marker
Two Teacher
Video lectures

Assessment Scheme:

Class Continuous Assessment (CCA)- 50 Marks (100%)

Attendance	Assignments	Mid Term Exam	Total
20%	40%	40%	100%

Term End Examination: 50 marks (100%)

Syllabus:

Module No.	Contents	Workload in Hrs		
		Theory	Lab	Assess
1	**Introduction to Machine Learning** Learning, Need and characteristics of Machine Learning, Types of learning, Unsupervised learning, Supervised learning, Semi-supervised learning, Reinforcement learning: concept with reward and punishment, Applications of all types of learning, Parametric and non-parametric learning	8		

2	**Supervised and Semi-supervised Learning** Supervised Learning: Vapnik-Chervonekis (VC) Dimension, Probably Approximately Correct (PAC) Learning, Model selection and Generalization, Dimensions of a Supervised Machine Learning Algorithm, cross-validation and Re-sampling Methods, Assessing a Classification Algorithm's Performance, Comparing Classification Algorithms, Semi-supervised learning and different methods	**8**		
3	**Unsupervised and Reinforcement Learning** Unsupervised Learning: Mixture densities, Expectation Maximization Algorithm, Mixtures of Latent Variable Models, Reinforcement learning, Temporal Difference Learning, Q Learning	**7**		
4	**Neural Networks and Genetic Algorithms** Bayesian Learning, Neural Networks and various types, Introduction to Deep Learning, Genetic Algorithms, Hidden Markov Models (HMM)	**7**		

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

COURSE STRUCTURE

Course Code	**CSD532A**			
Course Category	**Departmental Elective**			
Course Title	**Data Analytics with R**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**3**	**--**	**--**	**2**

Pre-requisites:

- **Foundations of Data Mining**

Course Objectives:

1. Knowledge: (i) To understand basic concepts of data analytics
(ii) To learn to apply the appropriate technique for data analysis
2. Skills: (i) To use tools and techniques for analyzing the data

3. Attitude: (i) To learn the effective way of assessing the statistical significance of results using R

Course Outcomes:

On completion of course, students should be able to

1. Gain the principle concepts and foundational understanding of data analytics
2. Apply data analysis techniques to generate results & draw effective conclusion
3. Apply appropriate tools for analyzing the data
4. Solve problems and provide better business decisions

Course Contents:

Introduction to Data Analytics

Statistics for Data Analysis

Basics of R

Data Visualization

Learning Resources:

Reference Books:

1. Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani. An Introduction to Statistical Learning: With Applications in R: Springer-Verlag New York, first edition, 2013
2. Maindonald J H and Braun W J, Data Analysis and Graphics Using R – An Example-Based Approach, 2nd edn. Cambridge University Press, 2007
3. Braun-W.J.-Murdoch-D.J.-A-First-Course-in-Statistical-Programming-with-R, Cambridge Press, First edition, 2007.

Supplementary Reading:

Web Resources:
https://www.udemy.com/data-analysis-with-r/
https://www.analyticsvidhya.com/blog/2016/02/complete-tutorial-learn-data-science-scratch/

Weblinks:
https://www.datacamp.com/tracks/data-analyst-with-r

MOOCs:
https://alison.com/course/r-for-data-analysis

Pedagogy:

Powerpoint Slides
Board-marker
Two Teacher
Video lectures

Assessment Scheme:

Class Continuous Assessment (CCA)- 50 Marks (100%)

Attendance	Assignments	Mid Term Exam	Total
20%	40%	40%	100%

Term End Examination: 50 marks (100%)

Syllabus:

Module No.	Contents	Workload in Hrs		
		Theory	Lab	Assess
1	**Introduction to Data Analytics** Introduction, need and significance of data analysis, objectives of data analysis, types of data analysis, application of data analysis.	8		
2	**Statistics for Data Analysis** Qualitative vs. Quantitative analysis, Types of Measurements (Nominal, Ordinal, Interval, Ratio), Central Tendencies - mean, median, standard deviation, variance, correlation, covariance.	8		
3	**Basics of R** Data Structures in R (Vectors, Matrices, Arrays, Lists), Data Manipulation in R, Functions, Data Relationships	7		
4	**Data Visualization** Exploring and visualizing data, Graphics in R, Bar Charts, Pie Charts, Box Plots	6		

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

COURSE STRUCTURE

Course Code	**CSD532B**			
Course Category	**Departmental Elective**			
Course Title	**Data Analytics with Python**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**3**	**--**	**--**	**2**

Pre-requisites:

- Foundations of Data Mining

Course Objectives:

1. **Knowledge**: (i) To understand basic concepts of data analytics
   (ii) To learn to apply the appropriate technique for data analysis
2. **Skills:** (i) To use tools and techniques for analyzing the data
   (ii) To use Python Programming and toolkits for Data Analytics
3. **Attitude:** (i) To learn the effective way of assessing the statistical significance of results using Python

Course Outcomes:

On completion of course, students should be able to

1. Carry out Data Analytics using Python.
2. Model different Data Analytics Cases.
3. Solve data analytics applications using python.

Course Contents:

Data Aggregation and Group Operations

Data Wrangling: Clean, Transform, Merge, Reshape

Introduction

Time Series

Learning Resources:

Reference Books:

1. Mark Lutz, "Programming Python", O'Reilly Media, 4th edition, 2010.
2. Magnus Lie Hetland, "Beginning Python: From Novice to Professional", Apress, Second Edition, 2005.
3. Shai Vaingast, "Beginning Python Visualization Crafting Visual Transformation Scripts", Apress, 2
4. Wes Mc Kinney, "Python for Data Analysis", O'Reilly Media, 2012.

Supplementary Reading:

Web Resources:
https://www.datacamp.com/tracks/data-analyst-with-python
http://bigdata-madesimple.com/step-by-step-approach-to-perform-data-analysis-using-python/

Weblinks:
http://www.data-analysis-in-python.org/

MOOCs:
https://www.coursera.org/learn/python-data-analysis

Pedagogy:

Powerpoint Slides
Board-marker
Two Teacher
Video lectures

Assessment Scheme:

Class Continuous Assessment (CCA)- 50 Marks (100%)

Attendance	Assignments	Mid Term Exam	Total
20%	40%	40%	100%

Term End Examination: 50 marks (100%)

Syllabus:

Module No.	Contents	Workload in Hrs		
		Theory	Lab	Assess
1	**Introduction** Python Programming Basics, Numpy Basics, Python Pandas Library, Introduction to Data Analytics, Data Analytics Examples	8		
2	**Data Wrangling: Clean, Transform, Merge, Reshape** Handling Missing Data, Data Cleaning and Preparation, Combining and Merging Datasets , Reshaping and Pivoting, Data Transformation ,String Manipulation, Regular Expressions. Case Study	8		
3	**Data Aggregation and Group Operations** Group By Mechanics, Data Aggregation, Group-wise Operations and Transformations, Pivot Tables and Cross-Tabulation, Plotting and Visualization, Example and Case Study	7		
4	**Time Series** Date and Time Data Types and Tools, Time Series Basics, Date Ranges, Frequencies, and Shifting, Time Zone Handling, Periods and Period Arithmetic, Resampling and Frequency Conversion, Time Series Plotting, Moving Window Functions, Performance and Memory Usage Notes. Time Series Case Study.	6		

Prepared By

Prof. Sushila Aghav
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

COURSE STRUCTURE

Course Code	**CSD533A**			
Course Category	**Departmental Elective**			
Course Title	**Cognitive Computing**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**3**	**--**	**--**	**2**

Pre-requisites:

Course Objectives:

1. Knowledge: (i) Appealing new model or paradigm for application development using cognitive computing

2. Skills: (i) To identify and evaluate patterns and complex relationships in large and unstructured data sets

(ii) Evaluate data in context and presenting relevant findings along with the evidence that justifies the answers

3. Attitude: (i) To evaluate IBM's Watson question-answering technology

Course Outcomes:

On completion of course, students should be able to

1. Understand and discuss what cognitive computing is, and how it differs from traditional approaches
2. Analyse the business implications of cognitive computing
3. Apply natural language technologies to business problems
4. Develop applications for Watson.

Course Contents:

Foundations of Cognitive Computing

Design Principles of Cognitive Systems

Natural Language Processing-Support of Cognitive System

Watson as a Cognitive System

Learning Resources:

Reference Books:

1. Judith Hurwitz, Marcia Kaufman, Adrian Bowles, Cognitive Computing and Big Data Analytics, 1st Edition, Wiley Publisher, 2015, ISBN: 978-1-118-89662-4
2. Hurwitz, Kaufman, and Bowles, Cognitive Computing and Big Data Analytics, Wiley, Indianapolis, IN, 2005, ISBN: 978-1-118-89662-4
3. Peter Finger, Cognitive Computing: A Brief Guide for Game Changers, Meghan Kiffler Press, 1st Edition, 2015, ISBN: 973-0-92965251-1
4. Kai Hwang, Cloud Computing for Machine Learning and Cognitive Applications, MIT Press Publishers, June 2017 | ISBN: 9780262341110

Supplementary Reading:

Web Resources: http://www.redbooks.ibm.com/redpapers/pdfs/redp4955.pdf
http://bigdata-madesimple.com/what-exactly-is-cognitive-computing/

Weblinks: https://cognitivecomputingconsortium.com/definition-of-cognitive-computing/
https://www.ibm.com/developerworks/learn/cognitive/index.html

MOOCs: https://cognitiveclass.ai/
https://cognitivecomputingchallenge.uchicago.edu/page/learning-resources

Pedagogy:

Powerpoint Slides
Board-marker
Two Teacher
Video lectures

Assessment Scheme:

Class Continuous Assessment (CCA)- 50 Marks (100%)

Attendance	Assignments	Mid Term Exam	Total
20%	40%	40%	100%

Term End Examination: 50 marks (100%)

Syllabus:

Module No.	Contents	Workload in Hrs		
		Theory	Lab	Assess
1	**Foundations of Cognitive Computing** Cognitive computing as new generation, Uses of cognitive systems, what makes system cognitive, Gaining insights from data, Artificial intelligence-the foundation, Understanding cognition, Understanding complex relationships, The elements of cognitive systems	8		
2	**Design Principles of Cognitive Systems** Components of cognitive systems, Building the Corpus, Bringing data into the cognitive system, Machine learning, Hypothesis generation and scoring, Presentation and visualization services	8		
3	**Natural Language Processing-Support of Cognitive System** The role of NLP in a cognitive system, Understanding linguistics, Phonology, morphology, lexical analysis, syntax and syntactic analysis, importance of Hidden Markov models, Semantic Web, Applying natural language technologies to business problems, enhancing shopping experience, fraud detection	7		
4	**Watson as a Cognitive System** Watson defined, Advancing research with a "Grand Challenge", Preparing Watson for jeopardy, commercial applications, components of deepQA architecture, Question analysis, hypothesis generation, scoring and confidence generation	7		

Prepared By

Dr. V.Y.Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

COURSE STRUCTURE

Course Code	**CSD534**			
Course Category	**Core**			
Course Title	**Lab Practice-III**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**-**	**-**	**6**	**03**

Course Contents:

Laboratory Practice-III (LP III) is companion course of theory courses (core and elective) in Trimester III. It consists of the assignments/mini-projects/study project tutorial of the subjects taught during the Trimester III. The course faculty should frame the suitable assignments/problem statements based on the concern theory subject.

There will be continuous evaluation of these assignments during the Trimester. Student has to submit a Journal/report consisting of suitable write up in the prescribed format. Softcopy of journal/report and code is to be maintained at department/institute in digital repository. Faculty advisor/ laboratory instructor suggested language/platform/framework is to be used for completing assignments/mini-project.

Guidelines for Term Work Assessment

Continuous assessment of laboratory work is done based on performance of student. Each assignment/ mini project assessment to be done based on parameters with appropriate weightage. Faculty should do the overall assessment as well as mini project assessment be based on the suggested parameters.

The laboratory work will be based on completion of following assignments/experiments-

Assessment Scheme:

Laboratory Continuous Assessment (LCA) 50 marks (100%)

Design and Implementation	Performance of Experiment	Result analysis and Reporting	Attendance/ Discipline/ Initiative/ Behavior	Total
30%	30%	20%	20%	100%

Term End Examination: 50 marks (100%)

Syllabus:

Laboratory:

Module No.	Contents	Workload in Hrs		
		Theory	Lab	Assess
Part A : Machine Learning				
1	Take a benchmark dataset for classification and run any three different supervised learning algorithms and compare their performance.		4	
2	Implement expectation maximization algorithm.		4	
3	Implement any two semi-supervised learning algorithms and compare their performance.		4	
4	Implement Deep Neural Network for any application. You can use tensorflow or any other framework.		4	
Part B : Elective-II				
Data Analytics with R				
1	Study of R-data analytics tools and libraries:		4	
2	Perform Data Wrangling: Clean, Transform, Merge, Reshape with R		4	
3	Read Stock market data from yahoo finance. Analyze it for recommendation of purchase and sale. [Apply all python data analytics method to perform analysis]		2	
4	A Shop has number of items for sale. Build a required Database to develop an application for considering one aspect of growth to the business: such as organizing [arranging] products based on demand and patterns. Use python programming. [Apriori Algorithm for frequent Pattern and Rules]		4	
Data Analytics with Python				
1	Study of python-data analytics tools and libraries: jupyter, notebooks, Ipython, pandas, Numpy, matplotlib, scikit learn, SciPy.		2	
2	Installation and Setup of ANACONDA with Python.		2	
3	Reading and writing files with different format: CSV, Text, xls, JSON, XML, HTML.		2	
4	Use Pandas to handle missing Values.		2	
5	Perform Data Wrangling: Clean, Transform, Merge, Reshape with Python.		2	

6	Read Stock market data from yahoo finance. Analyze it for recommendation of purchase and sale. [Apply all python data analytics method to perform analysis]		4	
7	A Shop has number of items for sale. Build a required Database to develop an application for considering one aspect of growth to the business: such as organizing [arranging] products based on demand and patterns. Use python programming. [Apriori Algorithm for frequent Pattern and Rules]		4	
Part B : Elective-III				
Web Mining				
1	Recommender System based on usage analysis.		2	
2	Identifying the Social Network connectivity of a user based on the link structure, to identify potential friends.		2	
3	Information Extraction, Sentiment Analysis of web pages.		2	
4	Implementing a system to analyze the effectiveness of a Web site by comparing the site structure to the navigational behavior of users, analyzing site and user e-metrics, and predict user behavior for individual or segments of users.		2	
5	Performing data mining on Web usage (or e-commerce) data from a particular Web site in order to analyze the behavior of users, including various site metrics, user metrics, user segments, associations, and opportunities for personalization. Processing with WEKA.		2	
6	Recommender System based on usage analysis.		2	
Cognitive Computing				
1	Discuss the purpose of hypothesis testing		2	
2	Analyze Watson Experience Manager (WEM) and perform following tasks: a. Assign user roles, b. Manage Corpus, c) Train Watson, d) Configure Watson, e) Test Watson		6	
3	Assignment based on natural language processing for content summary generation		4	

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

COURSE STRUCTURE

Course Code	**CSD535**			
Course Category	**Interdisciplinary**			
Course Title	**Seminar-I**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs		**-**	**4**	**2**

Course Objectives:

- Improve oral and written communication skills.
- To provide opportunity to use and understand the basic principles of communication and active, empathetic listening and speaking in front of audience.
- To Identify, understand and discuss real-world issues, new technologies, products, algorithms and services.

Course Outcomes:

After completion of this course students will be able to-

Present the multiple thinking strategies to examine real-world issues and explore creative avenues of expression

Acquire, articulate, create and convey intended meaning using verbal and non-verbal method of communication

Improve communication skills, presentation skills and other soft skills

Develop and enhance leadership skills

Improve technical writing skills

Assessment Scheme:

The students will have to deliver the seminar I in trimester III on any technical state-of-the-art topic approved by the guide. The presentation should cover introduction, motivation, literature survey, mathematical modeling, data-table discussion (if applicable) and conclusion and future work.

It is appreciated and strongly recommended that the student should select the domain of his/her seminar and identify the literature confined to the domain. Thorough literature study based on the broad identified topic has to be carried out. Selection of seminar topic in multidisciplinary domain will be strongly recommended and supported.

To bring the quality and appropriateness of the seminar work it is mandatory for the seminar guides to maintain a progressive record of the meetings. During meeting with the seminar guides, it is expected that it should include the discussion agenda, weekly outcomes achieved, corrective actions and comments on the progress report as per the plan submitted by the students. During trimester, it is mandatory to keep log book of these discussions with the guides and need to be shown to the examiner at the end term examination.

The reports should be prepared using LaTeX application tool and submitted in the department.

Laboratory Continuous Assessment (LCA) 50 marks (100%)

Literature Survey	Presentation	Attendance/ Discipline/ Initiative/ Behavior	Total
40%	40%	20%	100%

Term End Assessment: 50 marks (100%)

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

COURSE STRUCTURE

Course Code	CSD611			
Course Category	Core			
Course Title	**Business Intelligence**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**3**	**--**	**--**	**2**

Pre-requisites: -

Course Objectives:

1.Knowledge: (i) Identify the major frameworks of computerized decision support: decision support systems (DSS), data analytics and business intelligence (BI).

2.Skills: (i) To identify and evaluate mathematical models for business intelligence
(ii) Evaluate performance of various methods used in business intelligence

3. Attitude: (i) To study applications of business intelligence for real world problems in society

Course Outcomes:

1. Apply theoretical concepts of the course materials (e.g., textbook, journal articles, etc) to the decision-making and BI processes and technologies

2. Undertake systematic investigation/research related to the decision support and BI systems and technologies for today's dynamic business environment.

3. Develop professional attitudes in students in relation to the team work, interpersonal communication, and business ethics.

Course Contents:

Introduction to Business Intelligence and Decision Support System

Decision Making Concepts

Knowledge Management and BI

Business intelligence applications

Laboratory Exercises / Practical:

1. To build a decision support system for health care industry.
2. Classification of image data based on the low level features of the image.
3. Lab assignment based on clustering applications
4. Business case studies

Learning Resources:

Reference Books:

1.Turban, Sharda,Delen, Decision Support and Business Intelligence Systems, Ninth Edition, Pearson
2.Carlo Vercellis, Business Intelligence: Data Mining and Optimization for Decision Making, John Wiley & Sons, Ltd, First Edition, 2009, ISBN: 9780470511381.
3. Rick Sherman, Business Intelligence Guidebook : From Data Integration to Analytics, Elsevier Science and Technology Publisher, First Edition, 2014, ISBN10 012411461X
4. Cindi Howson, Successful Business Intelligence, Second Edition: Unlock the Value of BI & Big Data, Second Edition, 2014, ISBN-10: 007180918X

Supplementary Reading:

Web Resources: http://searchbusinessanalytics.techtarget.com/definition/business-intelligence-BI

https://data-warehouses.net/guide/introduction.html

Weblinks: https://www.learningtree.com/training-directory/business-intelligence-training/

http://mahedee.net/basic-concepts-of-business-intelligence-bi/

MOOCs: https://www.coursera.org/specializations/data-warehousing

https://www.coursera.org/learn/business-intelligence-tools

Pedagogy:

Two Teacher

Video Lectures

Case Studies

PPT Slides

Assessment Scheme:

Class Continuous Assessment (CCA) *(with % weights)*

Attendance	Assignments	Mid Term Exam	Total
10	20	20	50

Laboratory Continuous Assessment (LCA) *(with % weights)*

Design and Implementation	Performance of Experiment	Result analysis and Reporting	Attendance/ Discipline/ Initiative/ Behavior	Total
15	15	10	10	50

Term End Examination : *50%*

Syllabus :

Module No.	Contents	Workload in Hrs		
		Theory	*Lab*	*Assess*
1	**Introduction to Business Intelligence and Decision Support System** Effective and timely decisions, The role of mathematical models, Business intelligence architectures, Ethics and business intelligence, Definition of system, Representation of the decision-making process, Evolution of information systems, Development of a decision support system	8		
2	**Decision Making Concepts** Concepts of Decision Making, Techniques of Decision Support System (DSS), Development of Decision Support System (DSS), Applications of DSS, Role of Business Intelligence in DSS. Data warehouse Modelling, data warehouse design, data-ware-house technology, Distributed data warehouse, and materialized view, Use of DW in DSS	8		
3	**Knowledge Management and BI** Knowledge Definition, Characteristics of Knowledge management, Explicit and Tactic Knowledge, Knowledge management activities, Approaches to Knowledge Management, Knowledge Engineering, Process of Knowledge Acquisition, IT in knowledge management , Technologies for KM	8		
4	**Business intelligence applications** Marketing models, Relational marketing, Salesforce	8		

	management, Business case studies, Logistic and production models, Supply chain optimization, Optimization models for logistics planning, Revenue management systems, Business case studies			

Prepared By

Prof. Sushila Aghav
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

COURSE STRUCTURE

Course Code	**CSD612**			
Course Category	**Core**			
Course Title	**Web Mining**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**3**	**--**	**--**	**2**

Pre-requisites:

- **Foundations of Data Mining**

Course Objectives:

1. Knowledge: (i) To learn various techniques to mine the Web and other information networks, social networks, and social media
(ii) To learn to apply the appropriate technique for data analysis

2. Skills: (i) To learn how to extract knowledge from web scale datasets by various techniques.

3. Attitude: (i) Understand emerging areas in the ever evolving Web.

Course Outcomes:

On completion of course, students should be able to

1. Be familiar with classic and recent developments in Web search and web data mining,
2. Acquire statistical techniques to analyze complex information and social networks,
3. Interpret emergent features such as the structure and evolution of the Web graph, its traffic patterns, and the spread of information.

Course Contents:

Introduction to Web Data Mining

Web Usage Mining

Web Structure Mining

Web Content Mining

Learning Resources:

Reference Books:

1. Bing Liu, Web Data Mining: Exploring Hyperlinks, Content, and Usage Data, 2nd Edition, Springer, 2011
2. Soumen Chakrabarti, Mining the Web, Morgan-Kaufmann, first edition, 2002

Supplementary Reading:

Web Resources:
https://www.kdnuggets.com/2014/09/most-viewed-web-mining-lectures-videolectures.html

Weblinks:
https://www.cs.uic.edu/~liub/WebContentMining.html

MOOCs:
https://www.coursera.org/specializations/data-mining

Pedagogy:

Powerpoint Slides
Board-marker
Two Teacher
Video lectures

Assessment Scheme:

Class Continuous Assessment (CCA)- 50 Marks (100%)

Attendance	Assignments	Mid Term Exam	Total
20%	40%	40%	100%

Term End Examination: 50 marks (100%)

Syllabus:

Module No.	Contents	Workload in Hrs		
		Theory	Lab	Assess
1	**Introduction to Web Data Mining** Need, Importance, Applications of Web Data mining. Capturing users web activities, Client side v/s middleware v/s server side data and usage logging. Web Mining and its types, Web Usage Mining, Web Structure Mining, Web Content Mining	**8**		
2	**Web Usage Mining** Learning from Browser, Server Logs, Identifying frequent item sets, pattern identification, representing patterns in form of relations or Graphs, Understanding web application or website usage, Heat maps. Using statistical tools for usage analysis and machine learning for prospective improvements.	**8**		
3	**Web Structure Mining** Understanding link structure of the web, Static v/s dynamic linking, representing the link structure as graphs, identifying most / least used links, paths, Categorizing links based on required attributes, Clustering links based on required attributes. Web as a graph, identifying nodes, edges, in-degree, out-degree, HITS Algorithm PageRank algorithm.	**7**		
4	**Web Content Mining** Storing web content as text, database, various document types, generating meta-information of web documents, labeling, tagging, identifying feature sets. Representing web documents, Vector Space Model. TF-IDF, web-page summarization, tokenization, n-gram analysis, Categorizing web pages based on required attributes, Clustering web pages based on required attributes.	7		

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

COURSE STRUCTURE

Course Code	CSD613A			
Course Category	Interdisciplinary			
Course Title	**Data Security and Data Privacy**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**3**	**--**	**--**	**2**

Pre-requisites:

Course Objectives:

1. **Knowledge** (i) To understand the concept data privacy and data security

2. **Skills** (i) To analyse the privacy preserving techniques and its importance for data security

3. **Attitude** (i) To apply data security techniques for real world examples

(ii) To solve privacy issues using biometric authentication techniques

Course Outcomes:

- Demonstrate privacy preserving techniques for data mining applications
- Analyze the current problems in data security
- Implementation of biometrics techniques for real world examples

Course Contents:

- Introduction to Data Privacy
- Privacy Preserving Data Mining
- Introduction to Data Security
- Identity Security

Learning Resources:

Reference Books:

1. Natraj Venkatraman and Ashwin Shriram, Data Privacy: Principles and Practice, First edition, CRC Press, 2017
2. Richard M. Thompson II, Emily C. Barbour and Alison M. Smith, Understanding Privacy and Data Protection: What You Need to Know, Nova Publishers New York, 2014
3. Jonathan LeBlanc, Tim Messerschmidt, Identity and Data Security for Web Development Best Practices, Publisher: O'Reilly Media, 2016
4. Jay Jacobs and Bob Rudis, Data-Driven Security, First edition, Wiley and Sons Publishers, 2014

Supplementary Reading:

Web Resources: https://www.bsigroup.com/en-GB/our-services/training-courses/Data-Protection/
https://www.udemy.com/data-security/

Weblinks: https://iapp.org/certify/cipp/
https://www.liaison.com/blog/2017/10/23/big-data-machine-learning-data-security/

MOOCs:
https://risk.thomsonreuters.com/en/compliance-training-courses/data-privacy-and-security-training.html
https://teamtreehouse.com/library/introduction-to-data-security

Pedagogy:

Powerpoint Slides
Board-marker
Two Teacher
Video lectures

Assessment Scheme:

Class Continuous Assessment (CCA)- 50 Marks (100%)

Attendance	Assignments	Mid Term Exam	Total
20%	40%	40%	100%

Term End Examination: 50 marks (100%)

Syllabus:

Module No.	Contents	Workload in Hrs		
		Theory	*Lab*	*Assess*
1	**Introduction to Data Privacy** What is data privacy and its importance, methods for protecting data, balancing between data privacy and utility, anonymization of design principles, nature of data in enterprise, static data anonymization, classification of data in multidimensional data set, Group based anonmization, privacy preserving graph data, privacy preserving time series data	**8**		
2	**Privacy Preserving Data Mining** Data Mining: key functional areas of data mining, association rule mining, clustering, test data fundamentals, utility of test data, privacy preserving of test data, quality of test data, insufficiencies of anonymized test data	**8**		
3	**Introduction to Data Security** Current problems in security, understanding entropy in password security, standards for identity, what data should be protected? Password encryption, hashing, salting, password attack vectors, password hashing function, key stretching, recomputing hashes	**8**		
4	**Identity Security Fundamentals** Identity types, Enhancing User Experience by Utilizing Identity, Introducing Trust Zones, Browser Fingerprinting, Location-Based Tracking, Device Fingerprinting (Phone/Tablet), Device Fingerprinting (Bluetooth Paired Devices), Implementing Identity, Device and Browser Fingerprinting, Two-Factor Authentication and n-Factor Authentication, Biometrics as Username Instead of Password, How to Rate Biometric Effectiveness Revision	**8**		

Prepared by **Approved by**

Dr. S. A. Kulkarni Dr. M. V. Bedekar

Course Coordinator **Chairman BOS**

Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

COURSE STRUCTURE

Course Code	CSD613B			
Course Category	Interdisciplinary			
Course Title	**Data Encryption and Compression**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**3**	**--**	**--**	**3**

Pre-requisites: -

Course Objectives:

1. Knowledge (i) To understand general scope of data encryption and compression

2. Skills **(i)** To identify the overall process of data encryption and compression

(ii) To analyse the primary characteristics of data encryption and compression

3. Attitude **(i)** To apply new techniques of data compression and encryption techniques for business organizations

Course Outcomes:

1. Understand basic concepts of data encryption and compression
2. Identify appropriate techniques for data encryption and compression
3. Select and define appropriate methods for data encryption and compression

Course Contents:

- Data Encryption Techniques
- Key based Cryptography
- Introduction to data compression
- Data Compression Techniques

Learning Resources:

Reference Books:

1. William Stallings, Cryptography and Network Security, Seventh edition, Pearson, 2016
2. Adam Elbirt, nderstanding and Applying Cryptography and Data Security, Auerbach Publications; first edition
3. Mark Nelson and Jean-loup Gailly, The Data Compression Book, Second Edition, 1995

Supplementary Reading:

Web Resources: https://onlinecourses.nptel.ac.in/noc18_cs07/preview

Weblinks: https://www.coursera.org/learn/machine-learning/lecture/0EJ6A/motivation-i-data-compression

MOOCs: https://www.coursera.org/learn/algorithms-part2/lecture/OtmHU/introduction-to-data-compression
https://www.coursera.org/learn/crypto

Pedagogy:

Powerpoint Slides
Board-marker
Two Teacher
Video lectures

Assessment Scheme:

Class Continuous Assessment (CCA) *(with % weights)*

Attendance	Assignments	Mid Term Exam	Total
10	20	20	50

Laboratory Continuous Assessment (LCA) *(with % weights)*

Design and Implementation	Performance of Experiment	Result analysis and Reporting	Attendance/ Discipline/ Initiative/ Behavior	Total
15	15	10	10	50

Term End Examination : *50%*

Syllabus :

Module No.	Contents	Workload in Hrs		
		Theory	*Lab*	*Assess*
1	**Data Encryption Techniques** Plaintext, Cipher Text, Substitution, Transportation Techniques, Encryption and Decryption, Types of Attacks	**8**		
2	**Key based Cryptography** Key range and size, Algorithm types and modes, DES, IDEA and different and linear Cryptanalysis RSA, Symmetric and asymmetric key together, digital signature, Knapsack algorithm	**8**		
3	**Introduction to data Compression** Need for data compression, fundamental concepts of data compression, coding Communication model, compression ratio, requirements of data compression	**8**		
4	**Data Compression Techniques** Lossy compression, Lossless compression Entropy encoding, statistical encoding, Huffman coding, arithmetic coding, source encoding, vector quantization, differential encoding, predictive coding, pulse modulation, transform based coding discrete cosine coding, fractal compression	**6**		

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

COURSE STRUCTURE

Course Code	**CSD 614**			
Course Category	**Interdisciplinary**			
Course Title	**Project Stage-I Seminar**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs		**-**	**4**	**2**

Course Objectives:

- To develop problem solving abilities using mathematics;
- To apply algorithmic strategies while solving problems;
- To develop software engineering documents and testing plans;
- To use algorithmic solutions using distributed, embedded, concurrent and parallel environments.
- To encourage and expose students for participation in National/ International paper presentation activities.
- Exposure to Learning and knowledge access techniques using Conferences, Journal papers and participation in research activities.

Course Outcomes:

After completion of this course students will be able to-

- To write problem solutions for projects using mathematical modeling, FOSS programming tools and devices or commercial tools;
- To write Software Requirement Specifications (SRS) and other software engineering documents in the project report using mathematical models
- To write a conference/journal paper;
- To practice presentation, communication and report writing skills.

Assessment Scheme:

Laboratory Continuous Assessment (LCA) 50 marks (100%)

Project Synopsys/Review	Presentation/Report Submission	Attendance/ Discipline/ Initiative/ Behavior	Total
20%	60%	20%	100%

Term End Examination/Oral: 50 marks (100%)

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

COURSE STRUCTURE

Course Code	**CSD615**			
Course Category	**Core**			
Course Title	**Lab Practice-IV**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**-**	**-**	**6**	**03**

Course Contents:

Laboratory Practice-IV (LP IV) is companion course of theory courses (core and elective) in Trimester IV. It consists of the assignments/mini-projects/study project tutorial of the subjects taught during the Trimester IV. The course faculty should frame the suitable assignments/problem statements based on the concern theory subject.

There will be continuous evaluation of these assignments during the Trimester. Student has to submit a Journal/report consisting of suitable write up in the prescribed format. Softcopy of journal/report and code is to be maintained at department/institute in digital repository. Faculty advisor/ laboratory instructor suggested language/platform/framework is to be used for completing assignments/mini-project.

Guidelines for Term Work Assessment

Continuous assessment of laboratory work is done based on performance of student. Each assignment/ mini project assessment to be done based on parameters with appropriate weightages. Faculty should do the overall assessment as well as mini project assessment based on the suggested parameters.

The laboratory work will be based on completion of following assignments/experiments-

Assessment Scheme:

Laboratory Continuous Assessment (LCA) 50 marks (100%)

Design and Implementation	Performance of Experiment	Result analysis and Reporting	Attendance/ Discipline/ Initiative/ Behavior	Total
30%	30%	20%	20%	100%

Term End Examination: 50 marks (100%)

Syllabus:

Laboratory:

Module No.	Contents	Workload in Hrs		
		Theory	Lab	Assess
Part A : Business Intelligence				
1	Study of R/Python tool		2	
2	For managing the day to day facilities in Maternity Hospital, Observe and analyze the birth data set births2006.smpl.using R/Python .		2	
3	Conclude with suggestions which can be adopted for decision making on Resource management in Hospital.		4	
4	Use R/Python to predict the student's placement status based on past data.		4	
5	Observe the student placement data. Conclude with some decision to improve on student's placement.		4	
6	Demonstrate customer behavior on Retail Market Data. Identify their likelihood item purchase. Conclude with decisions to improve the sales.		4	
Part B : Elective-II				
Data Security and Data Privacy				
1	Analyse Data Privacy acts in any three developed countries and discuss the tools and techniques that can be used for implementation.		4	
2	Implement steganography, [https://en.wikipedia.org/wiki/Steganography] Input – an image file which contains the hidden message a. implement using a picture / image file. b. Implement using an audio file		6	
3	Develop entropy based multimodal password strength approach by combining alphanumeric characters and image		4	
4	Implement Two – Factor authentication 2FA [https://en.wikipedia.org/wiki/Multi-factor_authentication] Using Wi-Fi and Bluetooth.		6	
Web Mining				
1	Web Usage Mining – Assignments List – Install and Configure a web server.		6	

	Create web-site Enable log generation in common log format a. Analyze website usage based on the log generated. b. Identify most-used, least-used pages, c. Analyze webpage accession based on keyboard and mouse movements. Use standard open source applications for - web-server, log analysis, graph generator, machine learning, database, Implement using Python / Java / C++			
2	Web Structure Mining Install and Configure a web server. Create web-site Enable log generation in common log format a. Perform link analysis and find the in-degree, out-degree of web-pages b. Check for connectivity, traversal patterns, loops in web-structure c. Identify most-traversed, least-traversed links.		6	
3	Web Content Mining Install and Configure a web server. Create web-site Enable log generation in common log format a. Analyze webpage reading pattern based on the log generated. b. Generate meta-data for web-pages by labeling, tokenization, etc. c. Identify usefulness / relevancy of web-pages and recommend related / similar web-pages to the user based o content.		8	

Prepared By **Approved By**

Dr. S. A. Kulkarni Dr. M. V. Bedekar

Course Coordinator **Chairman BOS**

COURSE STRUCTURE

Course Code	CSD621A			
Course Category	Self- Study			
Course Title	**Big Data Analytics**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**2**	**--**	**--**	**1**

Pre-requisites: -

Foundations of Data Mining

Course Objectives:

1. **Knowledge** (i) To understand the concepts of big data and its challenges

   (ii) To understand storage techniques and processing techniques for big data

2. **Skills** (i) To apply skills and tools to manage and analyse big data

3. **Attitude** (i) To apply big data analysis techniques to solve business oriented issues

Course Outcomes:

1. Understand the concept and challenge of big data and why existing technology is inadequate to analyze the big data;
2. Collect, manage, store, query, and analyze various form of big data; and
3. Gain hands-on experience on large-scale analytics tools to solve some open big data problems; and
4. Understand the impact of big data for business decisions and strategy.

Course Contents:

- Fundamentals of Big Data
- Enterprise Technologies and Big Data Business Intelligence
- Big Data Storage and Processing Concepts
- Big Data Analysis Techniques

Learning Resources:

Reference Books:

1. Thomas Erl, Wajid Khattak, Paul Buhler, Big Data Fundamentals: Concepts, Drivers and Techniques, First edition, Prentice Hall, 2016
2. EMC Education Services (Author, Editor), Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data 1st Edition, 2015
3. D. Umesh Kumar, Business Analytics: The science of data driven decision making, Wile India, first edition, 2017

Supplementary Reading:

Web Resources: https://www.jigsawacademy.com/full-stack-big-data-analytics/

https://www.edx.org/course/big-data-analytics-adelaidex-analyticsx

Weblinks: https://www.simplilearn.com/big-data-and-analytics/

MOOCs: https://www.coursera.org/specializations/big-data

Pedagogy:

Powerpoint Slides

Board-marker

Two Teacher

Video lectures

Assessment Scheme:

Term End Examination : *50%*

Syllabus:

Module No.	Contents	Workload in Hrs		
		Theory	*Lab*	*Assess*
1	**Fundamentals of Big Data** Concepts and Technology, Datasets, Analysis Types, Key performers indicators, Big Data characteristics, Structured and Unstructured Data, Business architecture, Business process management, Digitization, Data procurement, privacy, security, big data lifecycle	**8**		
2	**Enterprise Technologies and Big Data Business Intelligence** Online Transaction processing, Online analytical processing, Extract transform load, Data warehouse, Data mart, traditional BI, Dashboards, Big data BI, Data visualization for big data,	**8**		

	case study			
3	**Big Data Storage and Processing Concepts** Clusters, file system and distributed file system, NoSQL, Sharding, replication, master-slave, peer-to-peer, sharding and replication, combining sharding and master-salve replication, combining sharding and peer-to-peer replication, CAP theorem, ACID, BASE	**8**		
4	**Big Data Analysis Techniques** Quantitative analysis, Qualitative analysis, Data mining, statistical analysis, A/B testing, correlaton, regression, machine learning, classification, outlier detection, filtering, semantic analysis, visual analysis, heat maps, time series plots, network graphs	**8**		

Prepared by

Dr. S. A. Kulkarni
Course Coordinator

Approved by

Dr. M. V. Bedekar
Chairman BOS

Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

COURSE STRUCTURE

Course Code	CSD621B			
Course Category	Self- Study			
Course Title	**Natural Language Processing and Information Retrieval**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**2**	**--**	**--**	**1**

Pre-requisites: -

Course Objectives:

2. **Knowledge** (i) To understand linguistic phenomena and to explore the linguistic features relevant to each NLP task
   (ii) To understand semantic representation, inference and knowledge representations

2. **Skills** (i) To revise the basic concepts of Information Retrieval and learn various storage techniques

3. **Attitude** (i) To study and compare various Classification Algorithms

Course Outcomes:

1. Understand basic concepts of data encryption and compression
2. Identify appropriate techniques for data encryption and compression
3. Select and define appropriate methods for data encryption and compression

Course Contents:

- Natural Language Processing
- Semantics
- Information Retrieval
- Text Mining

Learning Resources:

Reference Books:

1. Christopher D. Manning, Hinrich Schiitze Foundations of Statistical Natural Language Processing, The MIT Press Cambridge, Massachusetts
2. Yates & Neto, "Modern Information Retrieval", Pearson Education, ISBN 81-297-0274-6
3. C.J. Rijsbergen, "Information Retrieval", (www.dcs.gla.ac.uk)

Supplementary Reading:

Web Resources: http://web.stanford.edu/class/cs276/
https://swayam.gov.in/course/4185-information-storage-and-retrieval

Weblinks:
https://ep.jhu.edu/programs-and-courses/605.744-information-retrieval

MOOCs: https://www.class-central.com/tag/natural%20language%20processing
https://www.coursera.org/learn/nlp

Pedagogy:

Powerpoint Slides
Board-marker
Two Teacher
Video lectures

Assessment Scheme:

Term End Examination : *50%*

Syllabus:

Module No.	Contents	Workload in Hrs		
		Theory	*Lab*	*Assess*
1	**Natural Language Processing** Text Categorization for NLP - Information Extraction, Question/Answering, Text Summarization. The role of NLP in IR, NLP for Text Retrieval, Markov Models, Lexical Acquisition, Part-of-Speech Tagging- Rule-based Part-of-speech Tagging, Stochastic Part-of-speech Tagging, Sequence Labeling: POS-tagging	**8**		

2	**Semantics** Probabilistic Context Free Grammars, Lexicalized Parsing, Probabilistic Parsing, Semantics: Representing Meaning, Semantic Analysis, Lexical Semantics Word Sense Disambiguation: Selection Restriction-Based Disambiguation, Limitations of Selection Restrictions, Robust Word Sense Disambiguation, Machine Learning Approaches , Dictionary-Based Approaches	**8**		
3	**Information Retrieval** Automatic Text Analysis: Luhn's ideas, Conflation Algorithm, Indexing and Index Term Weighing, Automatic Classification. Measures of Association, Different Matching Coefficient, The Vector Space Model, Boolean Model, Storage: Inverted file, Suffix trees & suffix arrays, Signature Files, Scatter storage or hash addressing, Searching strategies: Boolean Search, Serial search, cluster based retrieval, Relevance Feedback, retrieval performance evaluation	**8**		
4	**Text Mining** Text Mining-Text classification and Clustering. Document Clustering: Introduction to the problem. Single Pass Algorithm, k-means clustering, Agglomerative clustering, Expectation Maximization (EM), Single Link Algorithm. Feature Selection. Evaluation of classification. Information Extraction and Integration: Semantic Web	**6**		

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

Dr. Vishwanath Karad
MIT WORLD PEACE UNIVERSITY | PUNE
TECHNOLOGY, RESEARCH, SOCIAL INNOVATION & PARTNERSHIPS

COURSE STRUCTURE

Course Code	**CSD622**			
Course Category	**Core**			
Course Title	**Project Stage-II Seminar**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs		**-**	**18**	**9**

Course Objectives:

- To develop problem solving abilities using mathematics;
- To apply algorithmic strategies while solving problems;
- To develop software engineering documents and testing plans;
- To use algorithmic solutions using distributed, Embedded, concurrent and parallel environments.
- To encourage and expose students for participation in National/ International paper presentation activities.
- Exposure to Learning and knowledge access techniques using Conferences, Journal papers and participation in research activities.

Course Outcomes:

After completion of this course students will be able to-

- To write problem solutions in projects using mathematical modeling, using FOSS programming tools and devices or commercial tools;
- To write SRS and other software engineering documents in the project report using mathematical models
- To write a conference/journal paper;
- To practice presentation, communication and report writing skills.

Assessment Scheme:

Laboratory Continuous Assessment (LCA) 50 marks (100%)

Project Synopsys/Review	Presentation/Report Submission	Attendance/ Discipline/ Initiative/ Behavior	Total
20%	60%	20%	100%

Term End Examination: 50 marks (100%)

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

COURSE STRUCTURE

Course Code	**CSD631A**			
Course Category	**Elective**			
Course Title	**Ubiquitous Computing**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**2**	**--**	**--**	**1**

Pre-requisites: -

Course Objectives:

1. **Knowledge** (i) To understand the basic concept of Ubiquitous Computing

2. **Skills** (i) Design and enhance algorithms for localization based on existing positioning techniques and platforms

3. **Attitude** (i) Recognize the important issues and concerns on security and privacy

Course Outcomes:

- Recognize the research issues in ubiquitous computing
- Appreciate positioning techniques and in depth understanding on location-based services and applications
- gain hands-on experiences in building applications that realize the usefulness of ubiquitous computing

Course Contents:

- **Ubiquitous Computing**
- **Ubiquitous standards**
- **Wearable Computing**
- **Service management in Ubiquitous Computing**

Learning Resources:

Reference Books:

1. Asoke K. Talukder, Roopa R. Yavagal, Mobile Computing Technology, Applications, and Service Creation, McGraw-Hill, 2007.
2. Jingyu Zhou, Feilong Tang, Yao Shen, and Minyi Guo. Pervasive Computing: Concepts, Technologies and Applications. CRC Press, 2017.
3. André Perez. Mobile Networks Architecture. Wiley-ISTE, 2013

Supplementary Reading:

Web Resources:
https://www.class-central.com/course/coursera-mobile-and-ubiquitous-computing-2771

Weblinks: http://freevideolectures.com/Course/2341/Embedded-Systems/37

MOOCs: https://www.coursetalk.com/providers/coursera/courses/introduction-to-user-interface-design-part-2

Pedagogy:

Powerpoint Slides
Board-marker
Two Teacher
Video lectures

Assessment Scheme:

Term End Examination : *50%*

Syllabus:-

Module No.	Contents	Workload in Hrs		
		Theory	*Lab*	*Assess*
1	**Ubiquitous Learning Tools and Technologies** Introduction **Ubiquitous Learning:** A Promising Innovative Paradigm, Historical Development of Computing and IT in Education, Past Experience and Issues, Practice and Challenge at Waseda E-School **Emerging Technologies and Systems for Ubiquitous Learning:** Emerging Computing Paradigms for Education, Ubiquitous Learning Support Systems and Technologies, **Integration of Real-World Practice and Experience with Ubiquitous Learning:** Ubiquitous Learning, UPS (Ubiquitous	**8**		

	Personal Study), **Nature of Ubiquitous Learning and Provision of Well-Being in Education:** Ubiquitous and Pervasive, The Possible Trend of Pervasive Technology in Education			
2	**Service Management In Ubiquitous Computing Environments** Introduction **Service Management in Ubiquitous Computing Environments:** Introduction, **Ubiquitous** Computing Environments, Service Management Framework, General Components of a Service Management System, System Support Components, Service Management Challenges, **Techniques for Service Management in PvCE:** Introduction, Classification of Service Discovery Protocols, Service Discovery in Infrastructure-Based Networks, Service Discovery in Infrastructure-Less Networks, Multiprotocol Service Discovery, Service Discovery Approaches **Service Composition:** Service Composition Functions, Survey of Methods in Service Composition Process, Service Composition Approaches	**8**		
3	**Wearable Computing And Sensor Systems For Healthcare** Introduction, The Health Body Area Network, Medical and Technological Requirements of Health Sensors, Wearable Sensors for Vital Signals Monitoring, Wearable Sensors for Activity Recognition, Sensors and Signals for Emotion Recognition **Intra-BAN Communications in Ubiquitous Healthcare Systems: Standards and Protocols:** IEEE 802.15.4 and ZigBee, Bluetooth, Bluetooth Low Energy, Integrated and Additional Solutions for Health BAN Communications	**8**		
4	**Standards And Implementation Of Ubiquitous Computing Applications** Introduction: Pervasiveness and Mobility in Computing and Communications, Context Awareness, Heterogeneity, **Wireless Technologies and Standards:** A Simple Classification of Wireless Networks, Concluding Remarks, **Middleware:** Future Trends: Beyond the Middleware, **Case Studies:** Ubiquitous Computing in Extreme Areas; The Hiker's Personal Digital Assistant, Ubiquitous Computing in Personal Health Systems; The MyHealthService Approach	**8**		

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

COURSE STRUCTURE

Course Code	CSD631B			
Course Category	Elective			
Course Title	**Deep Learning**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs	**3**	**--**	**--**	**2**

Pre-requisites: -

Probability and Statistics, Machine Learning

Course Objectives:

1. **Knowledge** (i) To understand the concept of algorithms in deep learning

2. **Skills** (i) Design and enhance algorithms for localization based on existing positioning techniques and platforms

3. **Attitude** (i) To analyse the real world problems using deep learning algorithms

Course Outcomes:

- Identify the deep learning algorithms which are more appropriate for various types of learning tasks in various domains.
- Implement deep learning algorithms and solve real-world problems.

Course Contents:

- Introduction to Deep Learning
- Convolutional Neural Networks
- CNNs on Graphs
- Deep Unsupervised Learning

Learning Resources:

Reference Books:

1. Ian Goddfellow, Yoshua Bengio, Auron Courville, Deeap Learning (Adaptive Computation and Machine Learning Series), first edition, MIT Press, 2016,
2. Nikhil Budua, Fundamentals of Deep Learning, Designing Next Generation Artificial Intelligence Algorithms, First edition, O'Reilly Publications, 2016
3. Adam Gibson, Josh Patterson, Deep Learning, A Practitioners Approach, First edition, O'Reilly Publications, 2017

Supplementary Reading:

Web Resources:
https://www.edx.org/course/deep-learning-explained-microsoft-dat236x-0
https://www.analyticsvidhya.com/blog/2016/08/deep-learning-path/

Weblinks: https://www.class-central.com/report/deep-learning-online-courses/

MOOCs:
https://www.coursera.org/courses?languages=en&query=deep%20learning

Pedagogy:

Powerpoint Slides
Board-marker
Two Teacher
Video lectures

Assessment Scheme:

Term End Examination: *50%*

Syllabus

Module No.	Contents	Workload in Hrs		
		Theory	*Lab*	*Assess*
1	**Introduction to Deep Learning** Machine Learning Review, Feedforward Deep Networks, Backpropagation, Regularization Methods, Optimization Methods --SGD, Momentum method, Adaptive first order methods, Batch Normalization, Initialization Strategies, Polyak Averaging	**8**		
2	**Convolutional Neural Networks** Visualizing Convolutional Networks, Variants (Locally Connected Networks, Tiled CNNs, Dilated Convolutions),	**8**		

	Motivations: Neuroscientific Motivations: Efficiency, Equivariance, Invariance, parameter tying, Ultra Deep Architectures, Residual Networks, ResNet in ResNet, Highway Networks, Fractal Networks, Some Analysis of Residual Learning, DenseNets, Similarity Learning and Siamese Networks			
3	**CNNs on Graphs** CNNs on Non-Euclidean Domains, Locally Connected Networks Spectral Networks, Graph Embedding: Basic Message Passing Neural Networks, Sequence Learning with NNs, Recurrent Neural Networks, Design Patterns for RNNs, Intro to long term dependencies and gated architectures Challenges with Long Term Dependencies Reservoir Computing (basic idea), Long Short Term Memory and GRU Attention Models	**8**		
4	**Deep Unsupervised Learning** Linear Projections and Linear Autoencoders, Sparse Coding Autoencoders, Distributed versus Localist Representations Autoencoders wrap-up, Revision	**7**		

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS

COURSE STRUCTURE

Course Code	**CSD632**			
Course Category	**Core**			
Course Title	**Project Stage-III Seminar**			
Teaching Scheme and Credits	**L**	**T**	**Laboratory**	**Credits**
Weekly load hrs		**-**	**18**	**9**

Course Objectives:

- To develop problem solving abilities using mathematics;
- To apply algorithmic strategies while solving problems;
- To develop software engineering documents and testing plans;
- To use algorithmic solutions using distributed, Embedded, concurrent and parallel environments.
- To encourage and expose students for participation in National/ International paper presentation activities.
- Exposure to Learning and knowledge access techniques using Conferences, Journal papers and participation in research activities.

Course Outcomes:

After completion of this course students will be able to-

- To write problem solutions in projects using mathematical modeling, using FOSS programming tools and devices or commercial tools;
- To write SRS and other software engineering documents in the project report using mathematical models
- To write a conference/journal paper;
- To practice presentation, communication and report writing skills.

Assessment Scheme:

Laboratory Continuous Assessment (LCA) 50 marks (100%)

Project Synopsys/Review	Presentation/Report Submission	Attendance/ Discipline/ Initiative/ Behavior	Total
20%	70%	10%	100%

Term End Assessment: 100 marks (100%)

Prepared By

Dr. S. A. Kulkarni
Course Coordinator

Approved By

Dr. M. V. Bedekar
Chairman BOS