MOHAN BABU UNIVERSITY

Sree Sainath Nagar, Tirupati – 517 102

SCHOOL OF ENGINEERING

M.Tech. Computer Science and Engineering (Artificial Intelligence and Data Science)

CURRICULUM AND SYLLABUS
(For 2022-23 Admitted Students)

FULLY FLEXIBLE CHOICE BASED CREDIT SYSTEM (FFCBCS)

MOHAN BABU UNIVERSITY
Vision

To rise as one of the greatest hubs of innovation and entrepreneurship in the country, wherein students empower themselves with the best of knowledge, unleash their potential to the fullest, and soar high to attain a brighter future for themselves and the nation.

Mission

❖ To provide relevant knowledge founded on the spirit of curiosity, compassion, courage and commitment.

❖ To uphold novelle wings of leadership and excellence under expert mentors who guide students towards wisdom and knowledge.

❖ To create a dynamic learning environment that empowers learners with the right blend of passion and purpose to build a glorious tomorrow.

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

VISION

To become a Centre of Excellence in Artificial Intelligence, Machine Learning and Data Science by imparting high quality education through teaching, training and research

MISSION

❖ To impart quality education in Computer Science and Engineering with specializations in Artificial Intelligence, Machine Learning and Data Science by disseminating knowledge through contemporary curriculum, competent faculty and effective teaching-learning methodologies.

❖ Nurture research, innovation and entrepreneurial skills among students and faculty to contribute to the needs of industry and society.

❖ Inculcate professional attitude, ethical and social responsibilities for prospective and promising Engineering profession.

❖ Encourage students to engage in life-long learning by creating awareness of the contemporary developments in Computer Science and Engineering, Artificial Intelligence, Machine Learning and Data Science.

M.Tech. COMPUTER SCIENCE AND ENGINEERING(ARTIFICIAL INTELLIGENCE AND DATA SCIENCE)

PROGRAM EDUCATIONAL OBJECTIVES

After few years of graduation, the graduates of M.Tech.CSE (AI and DS) will:

PEO1. Able to carry out research in the frontier areas of Computer Science and Engineering Artificial Intelligence, Data Science and develop innovative solutions to meet the opportunities and challenges in the society.

PEO2. Employed in academia, software development, Government organizations or would have established startup companies.

PEO3. Able to demonstrate effective communication and leadership skills, gain knowledge of contemporary and global issues and strive for continuouslearning and practice their profession with high regard to legal and ethical responsibilities.

PROGRAM OUTCOMES

On successful completion of the Program, the graduates of M.Tech. CSE (AI and DS) Program will be able to:

PO1. Demonstrate knowledge to select, learn and apply appropriate techniques, skills and modern engineering tools to solve engineering problems appropriate to the Computer Science and Engineering, Artificial Intelligence and Data Science discipline**. (Knowledge, Skills, Tools)**

PO2. Analyze engineering problems critically,, design, implement, evaluate and manage scientific/technological solutions in the context of Computer Science and Engineering, Artificial Intelligence and Data Science discipline. (**Analyze, Design, Implement, Evaluate, Manage)**

PO3. Practice ethical principles and norms of engineering to assess societal, environmental, health, safety, legal and cultural issues pertaining to Computer Science and Engineering, Artificial Intelligence and Data Science problems. **(Professionalism, Society, Environment)**

PO4. Independently carry out research/investigation and development of solutions to solve practical Computer Science and Engineering, Artificial Intelligence and Data Science problems. **(Research)**

PO5. Function effectively as an individual and in a team to acquire knowledge and recognize opportunities for career progression in Computer Science and Engineering, Artificial Intelligence and Data Science discipline. **(Individual and Team Work)**

PO6. Communicate effectively in professional practice through verbal and written formats. **(Communication)**

PO7. Recognize the need for self-motivated pursuit of knowledge to show commitment and competence in the broadest context of technological change. **(Self-Learning)**

M.Tech. COMPUTER SCIENCE AND ENGINEERING (ARTIFICIAL INTELLIGENCE AND DATA SCIENCE)

Basket Wise - Credit Distribution

S. No.	Basket	Credits (Min.- Max.)
1	SCHOOL CORE	31-34
2	PROGRAM CORE	21-24
3	PROGRAM ELECTIVE	12-18
4	UNIVERSITY ELECTIVE	9-12
	TOTAL CREDITS	**Min. 70**

School Core (31-34 Credits)

Couse Code	Title of the Course	Lecture	Tutorial	Practical	Project based Learning	Credits	Pre-requisite
		L	**T**	**P**	**S**	**C**	
22EE201001	Research Methodology	3	-	-	-	3	-
22EE201002	Innovation and Intellectual Property Rights	2	-	-	-	2	-
22AI211001	Internship	-	-	-	-	2	-
22AI209001	Project Work-Phase I	-	-	-	-	10	-
22AI210001	Project Work-Phase II	-	-	-	-	14	-
Mandatory Non-Credit Courses (Min. 4 Credits) Earned Credits will not be considered for CGPA							
22AI207601	Statistics with R	2	-	-	-	2	-
22LG207601	Technical Report Writing	2	-	-	-	2	-
22MG207601	Project Management	2	-	-	-	2	-
22MG207602	Essentials of Business Etiquettes	2	-	-	-	2	-

Program Core (21-24 Credits)

Couse Code	Title of the Course	Lecture	Tutorial	Practical	Project based Learning	Credits	Pre-requisite
		L	T	P	S	C	
22AI202001	Programming for Data Science	3	-	3	-	4.5	-
22AI201001	Artificial Intelligence	3	-	-	-	3	-
22AI202002	Machine Learning	3	-	3	-	4.5	-
22AI202003	Big Data Frameworks	3	-	3	-	4.5	-
22AI201007	Data Mining	3	-	-	-	3	-
22AI202005	Feature Engineering for Machine Learning	3	-	3	-	4.5	22AI202002
22CS201001	Advanced Data Structures And Algorithms	3	-	-	-	3	-

Program Elective (12-18 Credits)

Couse Code	Knowledge Area	Title of the Course	Lecture	Tutorial	Practical	Project based Learning	Credits	Pre-requisite
			L	T	P	S	C	
22AI202006	Artificial Intelligence	Deep Learning	3	-	3	-	4.5	22AI202002
22AI202007		Computer Vision	3	-	3	-	4.5	22AI202002
22AI201002		Reinforcement Learning	3	-	-	-	3	22AI202002
22AI201003		Machine Learning for Security	3	-	-	-	3	22AI202002
22AI201004		Predictive Analytics	3	-	-	-	3	22AI202002
22AI201005		Text Analytics	3	-		-	3	22AI202002
22AI202008	Data Science	Data Engineering	3	-	3	-	4.5	22AI202003
22AI201006		Data Analytics	3	-	-	-	3	-
22AI202009		Data Visualization	3	-	3	-	4.5	-
22AI201008		Stream Processing and Analytics	3	-	-	-	3	22AI202003
22AI201009		Malware Analysis	3	-	-	-	3	22AI202002
22AI202010		Business Intelligence Tools	3	-	3	-	4.5	22AI202009

University Elective(6 Credits)

Course Code	Title of the Course	Lecture	Tutorial	Practical	Project based Learning	Credits
		L	T	P	S	C
22AI201701	Business Analytics	3	-	-	-	3
22CM201701	Cost Management of Engineering Projects	3	-	-	-	3
22CE201701	Disaster Management	3	-	-	-	3
22SS201701	Value Education	3	-	-	-	3
22SS201702	Pedagogy Studies	3	-	-	-	3
22LG201701	Personality Development through Life Enlightenment Skills	3	-	-	-	3

Note:

1. If any student has chosen a course or equivalent course from the above list in their regular curriculum then, he/she is not eligible to opt the same course/s under University Elective.
2. The student can choose courses from other disciplines offered across the schools of MBU satisfying the pre-requisite other than the above list.

SCHOOL CORE

Course Code	Course Title	L	T	P	S	C
22EE201001	**RESEARCH METHODOLOGY**	3	-	-	-	3

Pre-Requisite --
Anti-Requisite --
Co-Requisite --

COURSE DESCRIPTION:
The course is developed for the students' to understand the underlying concepts of research methodology and a systematic approach for carrying out research in the domain of interest. The course is emphasised on developing skills to recognise and reflect the strength and limitation of different types of research; formulation of the research hypothesis and its systematic testing methods. The course also emphasises on interpreting the findings and research articulating skills along with the ethics of research.

COURSE OUTCOMES: *After successful completion of the course, students will be able to:*
CO1. Demonstrate the underlying concepts of research methodology, types of research and the systematic research process.
CO2. Demonstrate the philosophy of research design, types of research design and develop skills for a good research design.
CO3. Demonstrate the philosophy of formulation of research problem, methods of data collection, review of literature and formulation of working hypothesis.
CO4. Analyse the data and parametric tests for testing the hypothesis.
CO5. Interpret the findings and research articulating skills along with the ethics of research.

CO-PO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	-	-	-	3	-	-	-
CO2	-	-	-	3	-	-	-
CO3	-	-	-	3	-	-	-
CO4	-	-	-	3	-	-	-
CO5	-	-	-	-	3	-	-
Course Correlation Mapping	**-**	**-**	**-**	**3**	**3**	**-**	**-**

Correlation Levels: **3: High;** **2: Medium;** **1: Low**

COURSE CONTENT

Module 1:   Introduction to Research Methodology   (08 Periods)

Meaning of Research, Objectives of Research, Motivation in Research, Types of Research, Research Approaches, Significance of Research, Research Methods versus Methodology, Research and Scientific Method, Research Process, Criteria of Good Research.

Module 2:   Research Design   (08 Periods)

Research design—Basic Principles, Need of research design, Features of good design, Important concepts relating to research design, Different research designs, Basic principles of experimental designs, Developing a research plan.

Module 3:   Research Formulation   (08 Periods)

Defining and formulating the research problem - Selecting the problem - Necessity of defining the problem - Importance of literature review in defining a problem – Data collection – Primary and secondary sources; Critical literature review – Identifying gap areas from literature review; Hypothesis— Types of hypothesis, Development of working hypothesis.

Module 4:   Analysis of Data and Hypothesis Testing   (14 Periods)

Quantitative Tools: Testing and Significance of Measures of Central Tendency, Dispersion; correlation, Principles of least squares—Regression; Errors-Mean Square error, Mean absolute error, Mena absolute percentage errors.
Testing of Hypothesis: Hypothesis Testing Procedure, Types of errors, Parametric testing (t, z and F), Chi-Square Test as a Test of Goodness of Fit; Normal Distribution- Properties of Normal Distribution; Analysis of Variance.

Module 5:   Interpretation and Report Writing   (07 Periods)

Interpretation: Meaning of interpretation; Techniques of interpretation; Precautions in Interpretation.
Report Writing –Significance, Different Steps, Layout, Types of reports, Mechanics of Writing a Research Report, Precautions in Writing Reports; Research ethics—Plagiarism, Citation and acknowledgement.

Total Periods: *45*

EXPERIENTIAL LEARNING

1. Should conduct a survey based on a hypothesis, analyze the data collected and draw the inferences from the data.

2. Should review the literature on the given topic and should identify the scope/gaps in the literature and develop a research hypothesis.

3. Should study a case, formulate the hypothesis and identify an appropriate testing technique for the hypothesis.

4. Study an article and submit a report on the inferences and should interpret the findings of the article.

TEXT BOOKS:
1. C.R. Kothari, Research Methodology: Methods and Techniques, New Age International Publishers, 2nd revised edition, New Delhi, 2004.
2. Garg, B.L., Karadia, R., Agarwal, F. and Agarwal, *An introduction to Research Methodology*, RBSA Publishers, 2002.

REFERENCE BOOKS:
1.   R. Panneerselvam, Research Methodology, PHI learning Pvt. Ltd., 2009.
2.   Singh, Yogesh Kumar. *Fundamental of research methodology and statistics*. New Age International, 2006.

VIDEO LECTURES:
1.   https://nptel.ac.in/courses/121106007
2.   https://onlinecourses.nptel.ac.in/noc22_ge08/preview
3.   https://www.youtube.com/watch?v=VK-rnA3-41c

Web Resources:
1.   https://www.scribbr.com/category/methodology/
2.   https://leverageedu.com/blog/research-design/
3.   https://prothesiswriter.com/blog/how-to-formulate-research-problem
4.   https://www.formpl.us/blog/hypothesis-testing
5.   https://www.datapine.com/blog/data-interpretation-methods-benefits-problems/
6.   https://leverageedu.com/blog/report-writing/

SCHOOL CORE

Course Code	Course Title	L	T	P	S	C
22EE201002	INNOVATION AND INTELLECTUAL PROPERTY RIGHTS	2	-	-	-	2

Pre-Requisite   --

Anti-Requisite   --

Co-Requisite   --

COURSE DESCRIPTION:

The course is designed to provide comprehensive knowledge to the students regarding the general principles of innovation and intellectual property rights, significance of innovation and steps for innovation, Concept and Theories, Criticisms of Intellectual Property Rights, International Regime Relating to IPR. The course provides an awareness on how to protect ones unique creation, claim ownership, knowledge of what falls under the purview of someone's rights and what doesn't, and safeguard their creations and gain a competitive edge over the peers.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Understand the significance of innovation and steps for innovative thinking, and the concepts of intellectual property right and avenues for filling intellectual property rights.

CO2. Understand the legislative practices and protocols for acquisition of trademark and the judicial consequences for violating laws of trademark protection.

CO3. Understand the legislative practices and protocols for acquisition of copyrights and the judicial consequences for violating laws of copyrights protection.

CO4. Understand the fundamentals of patent laws, legislative practices and protocols for acquisition of trade secrets and the judicial consequences for violating laws of trade secrets protection.

CO5. Understand the latest developments and amendments in protection and filling of intellectual rights at international level.

CO-PO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	-	-	-	-	-	3	-
CO2	-	-	-	-	-	3	-
CO3	-	-	-	-	-	3	-
CO4	-	-	-	-	-	3	-
CO5	-	-	-	-	-	3	-
Course Correlation Mapping	-	-	-	-	-	3	-

Correlation Levels: 3: High;   2: Medium;   1: Low

COURSE CONTENT

Module 1:   Introduction to Innovation and IPR   (06 Periods)
Innovation: Difference between Creativity and Innovation – Examples of innovation; Being innovative; Identify Blocks for creativity and innovation – overcoming obstacles; Steps for Innovation
Intellectual property rights: Need for intellectual property rights (IPR); types of intellectual property- Design, Geographical Indication; International organizations, agencies and treaties.

Module 2:   Trademarks   (06 Periods)
Introduction to trademark, Purpose and function of trademarks, acquisition of trade mark rights, protectable matter, selecting and evaluating trade mark, trade mark registration processes.

Module 3:   Law of Copyrights   (06 Periods)
Fundamental of copy right law, originality of material, rights of reproduction, rights to perform the work publicly, copy right ownership issues, copy right registration, notice of copy right, international copy right law.
Law of patents: Foundation of patent law, patent searching process, ownership rights and transfer.

Module 4:   Trade Secrets   (06 Periods)
Trade secrete law, determination of trade secrete status, liability for misappropriations of trade secrets, protection for submission, trade secrete litigation.
Unfair competition: Misappropriation right of publicity, false advertising.

Module 5:   New Development of Intellectual Property   (06 Periods)
New developments in: trade mark law, copy right law, patent law, intellectual property audits. International overview on intellectual property; international - trade mark law, copy right law, international patent law, international development in trade secrets law.

Total Periods: *30*

EXPERIENTIAL LEARNING

1. Should conduct a survey based on the real scenario, where IPR is misused or unethically used and present an article.
2. Prepare an article on the registration processes of IPR practically (copy right/trade mark/ patents).
3. Should study a case of conflict on trademarks/patents and should produce an article mentioning the circumstances and remedial measures.
4. Prepare an article on the latest development in the international intellectual property rights.
5. Refining the project, based on the review report and uploading the text

RESOURCES

TEXT BOOKS:

1. Deborah, E. Bouchoux, *Intellectual property: The law of Trademarks, Copyright, Patents, and Trade Secrets,* Cengage learning, 4th Edition, 2013.

2. Prabuddha Ganguli, *Intellectual property right - Unleashing the knowledge economy*, McGraw Hill Education, 1st Edition, 2017.

3. Tom Kelley & Jonathan Littman, *The Art of Innovation*, Profile Books Ltd, UK, 2008

REFERENCE BOOKS:

1. Neeraj P., & Khusdeep D, *Intellectual Property Rights,* PHI learning Private Limited, 1st Edition, 2019.

2. Nithyananda, K V. *Intellectual Property Rights: Protection and Management*, Cengage Learning India Private Limited, 2019

3. Edward debone, *How to have Creative Ideas*, Vermilon publication, UK, 2007.

VIDEO LECTURES:

1. https://nptel.ac.in/courses/110105139

2. https://www.youtube.com/watch?v=bEusrD8g-dM

3. https://www.youtube.com/watch?v=LS7TTb23nzU

Web Resources:

1. http://www.bdu.ac.in/cells/ipr/docs/ipr-eng-ebook.pdf

2. https://www.wipo.int/edocs/pubdocs/en/intproperty/489/wipo_pub_489.pdf

3. http://cipam.gov.in/

4. https://www.wipo.int/about-ip/en/

5. http://www.ipindia.nic.in/

SCHOOL CORE

Course Code	Course Title	L	T	P	S	C
22AI207601	**STATISTICS WITH R**	2	-	-	-	2

Pre-Requisite -

Anti-Requisite -

Co-Requisite -

COURSE DESCRIPTION: This course introduces the basic concepts of statistics using R language. The course also deals with various types of sampling methods and its impact in the scope of inference through the computation of confidence intervals. The topics covered in the course also includes descriptive statistics, marginal and conditional distribution, statistical transformations, chi-squared test and ANOVA.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Import, manage, manipulate, structure data files and visualize data using R programming.

CO2. Identify trends and patterns in data using Marginal, Conditional distributions andStatistical transformations.

CO3. Analyse data using sampling and probability distribution methods and compute confidence intervals for statistical inference.

CO4. Apply chi-squared goodness-of-fit test, Pearson's χ 2-statistic and ANOVA to investigatethe distribution of data.

CO-PO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	3	2	-	-	-	-	-
CO2	3	2	-	-	-	-	-
CO3	2	2	-	-	-	-	-
CO4	3	2	-	-	-	-	-
Course Correlation Mapping	3	2	-	-	-	-	-

Correlation Levels: 3: High;   2: Medium;   1: Low

COURSE CONTENT

Module 1: INTRODUCTION   *(05 Periods)*
Data, R's command line, Variables, Functions, The workspace, External packages, Data sets, Data vectors, Functions, Numeric summaries, Categorical data.

Module 2:   BIVARIATE AND MULTIVARIATE DATA   *(07 Periods)*
Lists, Data frames, Paired data, Correlation, Trends, Transformations, Bivariate categorical data, Measures of association, Two-way tables, Marginal distributions, Conditionaldistributions, Graphical summaries, Multivariate data - Data frames, Applying a function overa collection, Using external data, Lattice graphics, Grouping, Statistical transformations.

Module 3 POPULATIONS   *(06 Periods)*
Populations, Discrete random variables, Random values generation, Sampling, Families of distributions, Central limit theorem, Statistical Inference - Significance tests, Estimation, Confidence intervals, Bayesian analysis.

Module 4 CONFIDENCE INTERVALS   *(06 Periods)*
Confidence intervals for a population proportion, p - population mean, other confidence intervals, Confidence intervals for differences, Confidence intervals for the median, Significance test - Significance test for a population proportion, Significance test for the mean (t-tests), Significance tests and confidence intervals, Significance tests for the median.

Module 5 GOODNESS OF FIT   *(06 Periods)*
The chi-squared goodness-of-fit test, The multinomial distribution, Pearson's χ 2-statistic, chi-squared test of independence and homogeneity, Goodness-of-fit tests for continuous distributions, ANOVA - One-way ANOVA, Using *lm* for ANOVA.

Total Periods: 30

EXPERIENTIAL LEARNING

1. The data set baby boom (Using R) contains data on the births of 44 children in a one-day period at a Brisbane, Australia, hospital. Compute the skew of the wt variable, which records birth weight. Is this variable reasonably symmetric or skewed? The variable running.time records the time after midnight of each birth. The command diff(running.time) records the differences or inter-arrival times. Is thisvariable skewed?

2. An elevator can safely hold 3, 500 pounds. A sign in the elevator limits the passenger count to 15. If the adult population has a mean weight of 180 pounds with a 25-pound standard deviation, how unusual would it be, if the central limit theorem applied, thatan elevator holding 15 people would be carrying more than 3, 500 pounds?

3. The data set MLB Attend (Using R) contains attendance data for Major League Baseball between the years 1969 and 2000. Use lm to perform a t-test on attendance for the two levels of league. Is the difference in mean attendance significant? Compare your results to those provided by t-test.

RESOURCES

TEXT BOOKS:

1. John Verzani, *Using R for Introductory Statistics*, CRC Press, 2nd Edition, 2014.
2. Sudha G Purohit, Sharad D Gore, Shailaja R Deshmukh, *Statistics Using R*, Narosa Publishing house, 2nd Edition, 2021.

REFERENCE BOOKS:

1. Francisco Juretig, *R Statistics Cookbook*, Packt Publishing, 1st Edition, 2019.
2. Prabhanjan N. Tattar, Suresh Ramaiah, B. G. Manjunath, *A Course in Statistics with R*, Wiley, 2018.

VIDEO LECTURES:

1. https://onlinecourses.nptel.ac.in/noc21_ma76/preview
2. https://onlinecourses.nptel.ac.in/noc19_ma33/preview
3. https://youtu.be/WbKiJe5OkUU?list=PLFW6lRTa1g83jjpIOte7RuEYCwOJa-6Gz
4. https://youtu.be/svDAkvh6utM?list=PLFW6lRTa1g83jjpIOte7RuEYCwOJa-6Gz
5. https://nptel.ac.in/courses/111104120

WEB RESOURCES:

1. https://www.geeksforgeeks.org/r-statistics/
2. https://www.geeksforgeeks.org/r-programming-exercises-practice-questions-and-solutions/
3. https://www.w3schools.com/r/r_stat_intro.asp
4. https://www.w3schools.com/r/r_stat_intro.asp
5. https://statsandr.com/blog/descriptive-statistics-in-r/

SCHOOL CORE

Course Code	Course Title	L	T	P	S	C
22LG207601	**TECHNICAL REPORT WRITING**	2	-	-	-	2

Pre-Requisite -

Anti-Requisite -

Co-Requisite -

COURSE DESCRIPTION: This course deals withpreparing effective technical documents for both written and digital media, with particular emphasis on technical memos, problem-solving and decision-making reports, and organizational, product-support, and technical-information webs.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Demonstrate knowledge of Technical Report Writing and structures with a scientific attitude.
CO2. Analyze the process of writing in preparing effective reports.
CO3. Demonstrate styles of writing for Publication in a Scientific Journal.
CO4. Apply the process of referencing and editing techniques for effective communication in written documents.
CO5. Analyze the strategies in the technical report presentation.

CO-PO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	-	-	-	-	3	-	-
CO2	-	-	-	-	3	-	-
CO3	-	-	-	-	3	-	-
CO4	-	-	-	-	3	-	-
CO5	-	-	-	-	3	-	-
Course Correlation Mapping	**-**	**-**	**-**	**-**	**3**	**-**	**-**

Correlation Levels: **3: High;**   **2: Medium;**   **1: Low**

COURSE CONTENT

Module 1:   INTRODUCTION TO TECHNICAL REPORT WRITING   *(06 Periods)*

Concepts of Technical Report, Types of Reports, Planning Technical Report Writing, Components of a Technical Report, Report Writing in Science and Technology, Selecting and Preparing a Title, Language Use in Report Writing.

Module 2:   PROCESSOF WRITING   *(06 Periods)*

Writing the 'Introduction', Writing the 'Materials and Methods, Writing the Findings/Results, Writing the 'Discussion', Preparing and using "Tables".

Module 3:   STYLE OF WRITING   *(06 Periods)*

Preparing and using Effective 'Graphs', Citing and Arranging References-I, Citing and Arranging References –II, Writing for Publication in a Scientific Journal.

Module 4:   REFERENCING   *(06 Periods)*

Literature citations, Introductory remarks on literature citations, Reasons for literature citations, Bibliographical data according to ISO standards, Citations in the text, Copyright, and copyright laws, the text of the Technical Report, Using a word processing and desktop publishing (DTP) systems, Document or page layout, hints on editing Typographic details, Cross-references.

Module 5:   PRESENTATION   *(06 Periods)*

Presentation with appropriate pointing, Dealing with intermediate questions, Review and analysis of the presentation, Rhetoric tips from A to Z.

Total Periods: 30

EXPERIENTIAL LEARNING

1. Prepare a report on technologies of modern times that enriched the originality of research works and their impacts on society concerning plagiarism.

2. Make PowerPoint presentations on the various style of writing academic reports.

3. Error-free Reports are so important for successful communication and sharing of information. Prepare a detailed chart on proofreading techniques to make a report effective and error-free.

4. Design a logo for a company and write down the copy-right laws for that.

5. Read research articles from any international journal of science and technology and differentiate research writing from other academic and non-academic writings.

6. Write an organizational memo Include a heading, introduction, and summary at the beginning of your memo, and present the details of your discussion in a logical order. Use headings and topic or main-idea sentences to clarify the organization.

7. Prepare an appraisal report on the staff performance of your company.

8. Prepare a PowerPoint presentation on the annual performance report of a company.

9. Critically review and write a report on any one of the recently released products.

10. Read the newspaper and write a detailed report about the content coverage and analyse the factors for the popularity of the newspaper.

RESOURCES

TEXTBOOK

1. RC Sharma and Krishna Mohan, "*Business Correspondence and Report Writing"*, McGraw-Hill Publishing, 3rd Edition, 2005 (reprint).

2. Patrick Forsyth, "*How to Write Reports and Proposals",* The Sunday Times, Kogan Page, New Delhi, Revised 2nd Edition, 2010.

REFERENCE BOOKS:

3. John Seely, *"The Oxford Writing & Speaking",* Oxford University Press, Indian Edition

4. Anne Eisenberg, "*A Beginner's Guide to Technical Communication"*, McGraw-Hill Education (India) Private Limited, New Delhi, 2013.

VIDEO LECTURES:

1. https://vimeo.com/143714818
2. https://digitalmedia.sheffield.ac.uk/media/002.+The+Anatomy+of+a+Technical+Report/1_u8wntcge

Web Resources:

1. http://www.resumania.com/arcindex.html
2. http://www.aresearchguide.com/writing-a-technical-report.htm
3. http://www.sussex.ac.uk/ei/internal/forstudents/engineeringdesign/studyguides/tec report writing

SCHOOL CORE

Course Code	Course Title	L	T	P	S	C
22MG207601	**PROJECT MANAGEMENT**	2	-	-	-	2

Pre-Requisite -
Anti-Requisite -
Co-Requisite

COURSE DESCRIPTION: To understand the importance of decision-making while implementing any project and interpret and discuss the results of qualitative and quantitative analysis

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1 Understand the basic introduction to project management
CO2 Apply the methods of project identification and selection.
CO3 Understand project allocation methods and evaluation.
CO4 Analyse the techniques for project time, review, and cost
CO5 Understand the factors of risk and quality of a project.

CO-PO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	2	1	2	1	-	-	-
CO2	1	1	2	2	-		2
CO3	2	2	1	2	1	-	-
CO4	3	1	2	2	1	-	-
CO5	2	2	1	2	1	1	-
Course Correlation Mapping	**2**	**2**	**2**	**2**	**1**	**1**	**2**

Correlation Levels: **3: High;** **2: Medium;** **1: Low**

COURSE CONTENT

Module 1: Introduction *(05 Periods)*
Concept of project management, project definition and key features of projects, project life cycle phases, typical project management issues, basic project activities

Module 2: Project Identification and Selection *(06 Periods)*
Identification and screening (brainstorming, strength and weakness in the system, environmental opportunities and threats), Project evaluation methods- Payback period, Net present value, Internal rate of return and project evaluation under uncertainty.

Module 3: Project Resource Management *(07 Periods)*
Scheduling resources, resource allocation methods, project crashing and resource leveling, working of systems, design of systems, project work system design, project execution plan, project procedure manual project control system, planning scheduling and monitoring

Module 4:   Time and Cost Management   *(05 Periods)*

Time Management-Network diagram, forward and backward pass, critical path, PERT and CPM, AOA and AON methods, tools for project network, Cost management-earned value method

Module 5:   Risk and Quality Management   *(07 Periods)*

Risk identification, types of risk, risk checklist, risk management tactics, risk mitigation and contingency planning, risk register, communication management, Quality assurance and quality control, quality audit, methods of enhancing quality

Total Periods: 30

EXPERIENTIAL LEARNING

1.   Refer to any video lecture on project evaluation methods and give a brief seminar using PPT

2.   Select any company wherein you will get the details of activities and time and draw the project network diagram and submit a report.

3.

Activity	Predecessor Activity	Normal Time (Weeks)	Crash Time (Weeks)	Normal Cost (Rs.)	Crash Cost (Rs.)
A	-	4	3	8,000	9,000
B	A	5	3	16,000	20,000
C	A	4	3	12,000	13,000
D	B	6	5	34,000	35,000
E	C	6	4	42,000	44,000
F	D	5	4	16,000	16,500
G	E	7	4	66,000	72,000
H	G	4	3	2,000	5,000

Determine a crashing scheme for the above project so that the total project time is reduced by 3 weeks

4.   Collect any case study that discusses the process of probability calculation of success of the project and submit a report

RESOURCES

TEXT BOOKS:

1.   R.Panneerselvam and P.Senthil Kumar (2013), Project Management, PHI Learning Private Limited.
2.   Prasanna Chandra (2014), Projects: Planning, Analysis, Selection, Financing, implementation, and Review.

REFERENCE BOOKS:

1.   A Guide to the Project Management Body of Knowledge: (PMBOK Guide) by Project Management Institute, 2013.
2.   Gopala Krishnan & Rama Murthy, A Text book of Project Management, McMillan India.
3.   S. Choudhary (2004), Project Management, Tata McGraw Hill Publication.

VIDEO LECTURES:

1.   https://onlinecourses.nptel.ac.in/noc19_mg30/preview
2.   https://archive.nptel.ac.in/courses/110/104/110104073/

Web Resources:

1.   https://www.pmi.org/about/learn-about-pmi/what-is-project-management
2.   https://www.manage.gov.in/studymaterial/PM.pdf

SCHOOL CORE

Course Code	Course Title	L	T	P	S	C
22MG207602	**ESSENTIALS OF BUSINESS ETIQUETTES**	2	-	-	-	2

Pre-Requisite

Anti-Requisite

Co-Requisite   -

COURSE DESCRIPTION: This course is designed for learners who desire to improve their Business etiquette and professionlaism.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1.   learn the principles of business etiquettes and professional behavior

CO2.   understand the etiquettes for making business correspondence effective

CO3.   Develop awareness of dining and multicultural etiquettes

CO4.   Demonstrate an understanding of professionalism in terms of workplace behaviors and workplace relationships.

CO5.   Understand attitudes and behaviors consistent with standard workplace expectations.

CO-PO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	1	1	-	-	-	1	-
CO2	1	1	2	1	-	1	-
CO3	2	-	2	-	1	-	-
CO4	1	2	-	1	-	-	-
CO5	1	2	1	-	-	-	-
Course Correlation Mapping	**2**	**2**	**2**	**1**	**1**	**1**	**-**

Correlation Levels: 3: High;   2: Medium;   1: Low

COURSE CONTENT

Module 1:   Business Etiquettes- An Overview   *(06 Periods)*

Significance of Business Etiquettes in 21st Century- Professional Advantage; Need and Importance of Professionalism; Workplace Etiquette: Etiquette for Personal Contact-Personal Appearance, Gestures, Postures, Facial Expressions, Eye-contact, Space distancing

Module 2:   Communication Skills   *(06 Periods)*

Understanding Human Communication, Constitutive Processes of Communication, Language as a tool of communication, Barriers to Effective communication, and Strategies to Overcome the Barriers.

Module 3:   Teamwork and Leadership Skills   *(06 Periods)*

Concept of Teams; Building effective teams; Concept of Leadership and honing Leadership skills. Personality: Meaning & Definition, Determinants of Personality, Personality Traits, Personality and Organisational Behaviour Motivation: Nature & Importance, Herzberg's Two Factor theory, Maslow's Need Hierarchy theory, Alderfer's ERG theory

Module 4:   Interview Skills   *(06 Periods)*

Interview Skills: in-depth perspectives, Interviewer and Interviewee, Before, During and After the Interview, Tips for Success. Meeting Etiquette: Managing a Meeting-Meeting agenda, Minute taking,; Duties of the chairperson and secretary; Effective Meeting Strategies - Preparing for the meeting, Conducting the meeting, Evaluating the meeting

Module 5:   Decision-Making and Problem-Solving Skills   *(06 Periods)*

Decision-Making and Problem-Solving Skills: Meaning, Types and Models, Group and Ethical Decision-Making, Problems and Dilemmas in application of these skills. Conflict Management: Conflict - Definition, Nature, Types and Causes; Methods of Conflict Resolution.

Total Periods:30

EXPERIENTIAL LEARNING

LIST OF EXPERIMENTS:

1. Collect the case studies related to successful leaders and their traits.

2. Conduct a mock interview showcasing interview skills.

3. The case studies will be collected as Assignments and the same will be evaluated.

RESOURCES

TEXT BOOKS:

1. Barbara Pachter, Marjorie Brody. Complete Business Etiquette Handbook. Prentice Hall, 2015.

2. Mahanand, Anand. English for Academic and Professional Skills. Delhi: McGraw, 2013. Print.

REFERENCE BOOKS:

1. Pease, Allan and Barbara Pease. The Definitive Book of Body Language. New Delhi: Manjul Publish

2. Rani, D Sudha, TVS Reddy, D Ravi, and AS Jyotsna. A Workbook on English Grammar and Compos

VIDEO LECTURES:

1. https://www.youtube.com/watch?v=NqlfZOPMqjA
2. http://www.nitttrc.edu.in/nptel/courses/video/109104107/L24.html

Web Resources:

1. http://elibrary.gci.edu.np/bitstream/123456789/685/1/BM-783%20The%20Essential%20Guide%20to%20Business%20Etiquette%20by%20Lillian%20H.%20Chaney%2C%20Jeanette%20S.%20Martin.pdf
2. The Essentials of Business Etiquette: How to Greet, Eat, and Tweet Your Way to Success by Barbara Pachter (Ebook) - Read free for 30 days (everand.com)

PROGRAM CORE

Course Code	Course Title	L	T	P	S	C
22AI202001	**PROGRAMMING FOR DATA SCIENCE**	3	-	3	-	4.5

Pre-Requisite -

Anti-Requisite -

Co-Requisite -

COURSE DESCRIPTION: This course provides a detailed discussion and hands-on experience on Basic Terminologies of data science, Computation using NumPy, Data exploration using Pandas, Data transformation, Data wrangling, Plotting and visualization using Matplotlib, Data aggregation, Statistical thinking and Time series analysis.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Demonstrate efficient storage and data handling methods in Numpy to perform mathematical computations vital for data science.

CO2. Apply Data Preparation and Exploration methods using Pandas to gain insights about raw data and transform quality data to perform analysis.

CO3. Create data visualization using charts, plots and histograms to identify trends, patterns and outliers in data importing Matplotlib and Seaborn.

CO4. Develop methods to analyze and interpret time series data to extract meaningful statistics using statistical distributions.

CO5. Work independently to solve problems with effective communication.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	3	2	2	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	2	2	3	-	-	-	-
CO4	3	2	3	-	-	-	-
CO5	3	3	2	-	-	-	-
Course Correlation Mapping	3	2	3	-	-	-	-

Correlation Levels: 3: High; 2: Medium;1: Low

COURSE CONTENT

Module 1: INTRODUCTION TO DATA SCIENCE *(09 Periods)*

Basic terminologies of data science, Types of data, five steps of data science, Arrays and vectorized computation using NumPy - The NumPy ndarray: A multidimensional array object, Universal functions: Fast element-wise Array functions, Array-oriented Programming with arrays, File input and output with arrays, Linear algebra, pseudorandom number generation.

Module 2: DATA EXPLORATION WITH PANDAS *(09 Periods)*

Process of exploring data, Pandas data structures – Series, Data frame, Index objects; Essential functionality, Summarizing and computing descriptive statistics - Correlation and covariance, Unique values, Value counts and membership; Data loading, Storage, and file formats - Reading and writing data in text format, Binary data formats, Interacting with web APIs, Interacting with databases.

Module 3: DATA CLEANING, PREPARATION AND DATA WRANGLING *(09 Periods)*

Handling missing data, Data transformation, String manipulation - String object methods, Regular expressions, Vectorized string functions in Pandas; Data wrangling: join, Combine and reshape - Hierarchical indexing, Combining and merging datasets, Reshaping and pivoting.

Module 4: DATA VISUALIZATION WITH MATPLOTLIB *(09 Periods)*

Plotting and visualization- A brief matplotlib API primer, Plotting with Pandas and Seaborn, Other python visualization tools; Data aggregation and Group operations- GroupBy mechanics, Data aggregation, Apply: General split-apply-combine, Pivot tables and Cross-tabulation.

Module 5: STATISTICAL THINKING AND TIME SERIES ANALYSIS *(09 Periods)*

Statistical Thinking: Distributions – Representing and plotting histograms, Outliers, Summarizing distributions, Variance, Reporting results; Probability mass function; Cumulative distribution functions - Representing CDFs, Percentile based statistics, Random numbers, Comparing percentile ranks.

Time Series Analysis: Date and Time data types and tools, Time series basics, Date ranges, Frequencies, and shifting, Time zone handling, Periods and period arithmetic.

Total Periods: 45

EXPERIENTIAL LEARNING

LIST OF EXERCISES:

1. Array Computations using NumPy
   a. Perform arithmetic operationsusing array.
   b. Perform slicing and indexing on multi-dimensional arrays.
   c. Perform computations on multi-dimensional array using universal functions (ufunc).
   d. Compute arithmetic mean, standard deviation, variance, percentile, minimum and maximum, cumulative sum and product using statistical functions in NumPy.
   e. Perform set theory operations such as union, intersection, symmetric difference and fetching unique values.

2. Linear Algebra and Random Number generation using linalg and random module in NumPy
   a. Compute dot product, vector product and inner product of two arrays.
   b. Perform matrix operations such as multiplication, determinant, sum of diagonal elements and inverse.
   c. Compute eigenvalues, eigenvectors and singular value decomposition for a square matrix.
   d. Generate random samples from uniform, normal, binomial, chi-square and Gaussian distributions using numpy.random functions.
   e. Implement a single random walk with 1000 steps using random module and extract the statistics like minimum and maximum value along the walk's trajectory.

3. Data Manipulation using pandas
   a. Create DataFrame from List, Dict, List of Dicts, Dicts of Series and perform operations such as column selection, addition, deletion and row selection,addition and deletion.
   b. Create a DataFrame and perform descriptive statistics functions such as sum, mean, median, mode, standard deviation, skewness, kurtosis, cumulative sum, cumulative product and percent changes.
   c. Implement the computation of correlation and covariance by considering the DataFrames of stock prices and volumes obtained from Yahoo Finance! Using pandas-datareader package.

4. Working with different data formats using pandas
   a. Perform reading and writing data in text format using read_csv and read_table considering any online dataset in delimited format (CSV).
   b. Perform reading, writing and parsing data in JSON (Javascript Object Notation) format using read_json.
   c. Perform readingand writing of Microsoft Excel Files (xslx) using read_excel.

5. Interacting with Web APIs and Databases
   a. Predict the last 30 GitHub issues for pandas using request and response object's json method. Move the extracted data to DataFrame and extract fields of interest. (Use url: 'https://api.github.com/repos/pandas-dev/pandas/issues')
   b. Connect to any relational database using corresponding SQL drivers and perform operations such as table creation, populating the table, selecting data from table, moving data from table to DataFrame, updating records and deleting records in a table.

6. Data Cleaning and Preparation
   a. Perform data cleaning by creating a DataFrame and identifying missing data using NA(Not Available) handling methods, filter out missing data using dropna function, fill the missing data using fillna function and remove duplicates using duplicated and drop_duplicates functions.
   b. Perform data transformation by modifying set of values using map and replace method and create transformed version of original dataset without modification using rename method.
   c. Create a DataFrame with normally distributed data using random sampling and detect possible outliers.
   d. Perform text manipulation with regular expression by applying relevant regular expression methods to split a string with a variable number of whitespace characters (tabs, spaces, and newlines) and get a list of all patterns matching.

7. Data Wrangling
   a. Perform hierarchical indexing by creating a series with a list of lists (or arrays) as the index, select subsets of data at outer and inner levels using partial indexing.
   b. Rearrange the tabular data with hierarchical indexing using unstack and stack method.
   c. Create two different DataFrames and merge them using index as merge key and combine data with overlap using combine_first method.

8. Perform Data Visualization with Matplotlib and SeaBorn considering online dataset for processing.
    a. Create a Line Plot by setting the title, axis labels, ticks, ticklabels , annotations on subplots and save to a file.
    b. Create Bar Plots using Series and DataFrame index.
        i. Create bar plotswith a DataFrame to group the values in each row together in a group in bars side by side for each value.
        ii. Create stacked bar plots from a DataFrame.
    c. Create Histogram to display the value frequency and Density Plot to generate continuous probability distribution function for observed data.
    d. Create Scatter Plot and examine the relationship between two one-dimensional data series.
    e. Create Box plots to visualize data with many categorical variables.

9. Data Aggregation
    a. Create a tabular dataset as a DataFrame and split data into groups using groupby method including single key and multiple key values. Select group by considering single and multiple columns.
    b. Compute summary statistics such as sum, mean and standard deviation for the grouped data using aggregate method.
    c. Use groupby function to split data into groups based on one column, miultiple columns, compute summary statistics and perform exploratory data analysis. Consider any online dataset for processing.

10. Time Series Analysis
    a. Create time series using datetime object in pandas indexed by timestamps.
    b. Use pandas.date_range to generate a DatetimeIndexwith an indicated length.
    c. Generate data ranges by setting time zone, localize time zone and convert to particular time zone using tz_convert and combine two different time zones.
    d. Perform period arithmetic such as adding and subtracting integers from periods and construct range of periods using period_range function.
    e. Convert Periods and PeriodIndexobjects to another frequency with asfreqmethod.
    f. Convert Series and DataFrame objects indexed by timestamps to periods with the to_periodmethod.

11. Implement Gaussian distribution to detect anomalies of credit card transaction. Consider credit card fraud detection data from Kaggle.

12. Web Scraping using Beautiful Soup
    a. Extract product reviews from Amazon website and save to a file.
    b. Perform Exploratory Data Analysis on extracted product reviews.
        i. Generate WordCloud for all reviews, positive and negative reviews.
        ii. Plot the distribution of stopwords, numerics, wordcount values, charactercount values and average wordlength.
        iii. Display the sentiment value using nltk and vader.
        iv. Create Scatter Intensity Plot of Sentiments.

RESOURCES

TEXT BOOKS:

1. Wes McKinney, Python for Data Analysis, O'Reilly, 2nd Edition, 2017.

2. Alen B. Downey, Think Stats: Exploratory Data Analysis, O'Reilly Publications, 2nd Edition, 2015.

REFERENCE BOOKS:

1. Sinan Ozdemir, Principles of Data Science, Packt Publishers, 2nd Edition, 2018.

2. Rachel Schutt, Cathy O'Neil,*Doing Data Science: Straight Talk from the Frontline*, O'Reilly, 2014.

SOFTWARE/TOOLS:

1. Python 3.8
2. Python Libraries –NumPy,Pandas,Matplotlib, Seaborn,Beautiful Soup, Vader
3. Anaconda Framework

VIDEO LECTURES:

1. https://swayam.gov.in/nd1_noc19_cs60/preview
2. https://onlinecourses.nptel.ac.in/noc21_cs78/preview
3. https://www.youtube.com/watch?time_continue=16&v=8JfDAm9y_7s
4. https://www.youtube.com/watch?v=kPRqJ-RHqZE&t=1s
5. https://www.youtube.com/watch?time_continue=17&v=1BxeY1Q8Akw
6. https://freevideolectures.com/course/4041/nptel-introduction-to-learning-analytics/11
7. https://www.youtube.com/watch?v=Og8hFZoACuE&feature=emb_logo

WEB RESOURCES:

1. https://towardsdatascience.com/
2. https://www.w3schools.com/datascience/
3. https://github.com/jakevdp/PythonDataScienceHandbook
4. https://www.kaggle.com
5. https://www.askpython.com/python/examples/probability-distributions
6. https://www.stefanjaspers.com/python-statistics-probability-sample-distribution/
7. https://www.geeksforgeeks.org/statistics-with-python/
8. https://data-flair.training/blogs/python-probability-distributions/

PROGRAM CORE

Course Code	Course Title	L	T	P	S	C
22AI201001	ARTIFICIAL INTELLIGENCE	3	-	-	-	3

Pre-Requisite -

Anti-Requisite -

Co-Requisite -

COURSE DESCRIPTION: This course provides an over view on the foundation blocks of AI, concepts of agents and its realization. It discusses on the problem-solving approach using searching technique up to search in terms of reasoning and also discusses on distributed AI and its applications along with brief knowledge on expert systems.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Architect intelligent agents using artificial intelligence techniques and principles.

CO2. Analyze and interpret the problem, identify suitable solutions using heuristic functions, optimization algorithms and search algorithms.

CO3. Select and apply appropriate knowledge representation to build Bayesian network models to reason under uncertainty.

CO4. Demonstrate the knowledge on the diverse concepts of parallel and distributed AI and in reasoning of Expert system and ethical implications of intelligent machines for providing privacy, trust, security and safety.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	3	2	2	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	2	2	3	-	-	-	-
CO4	3	2	3	-	-	-	-
Course Correlation Mapping	3	2	3	-	-	-	-

Correlation Levels: 3: High; 2: Medium;1: Low

COURSE CONTENT

Module 1:   INTRODUCTION TO ARTIFICIAL INTELLIGENCE   *(09 Periods)*

Foundations of artificial intelligence, History of artificial intelligence, State of the art, Risks and benefits of AI, Intelligent agents – Agents and environments, The concept of rationality, Structure of agents.

Module 2:   PROBLEM SOLVING BY SEARCHING   *(09 Periods)*

Problem solving agents, Search algorithms, Uninformed search strategies, Informed search strategies–Greedy best-first search, A* search; Heuristic functions

Module 3   SEARCH IN COMPLEX ENVIRONMENTS   *(09 Periods)*

Local search algorithms and optimization problems – Hill-climbing search, Simulated annealing, Local beam search, Evolutionary algorithms; Optimal decisions in games–The minimax search algorithm, Optimal decisions in multiplayer games, Alpha-Beta pruning, Moveordering; Monte Carlo tree search.

Module 4   PROBABILISTIC REASONING   *(10 Periods)*

Representing Knowledge in an uncertain domain, Semantics of Bayesian networks, Probabilistic reasoning over time – Time and uncertainty, Inference in temporal models, Hidden Markov models

Module 5   EXPERT SYSTEM AND REASONING   *(08 Periods)*

Parallel and Distributed AI: Psychological Modelling, Parallelism in Reasoning system. Distributed Reasoning System.
Expert System: Representing and using domain knowledge, Expert System shells, Explanations, Knowledge Acquisition.
**Ethics and Safety in AI:** Limits of AI, Ethics of AI – Surveillance, security and privacy, Fairness and bias, Trust and transparency, AI safety.

Total Periods: 45

EXPERIENTIAL LEARNING

The following is the sample. Faculty shall frame according to the course domain.
1.   Design and implement solution for 8-puzzle problem using search algorithms.

8-puzzle problem: Given a 3×3 board with 8 tiles (every tile has one number from 1 to 8) and one empty space. The objective is to place the numbers on tiles to match the final configuration using the empty space. We can slide four adjacent (left,right, above, and below) tiles into the empty space.
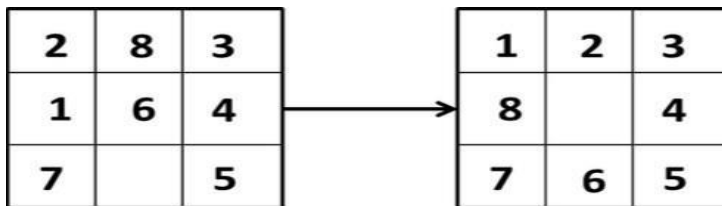
Initial State   **Goal State**

2.   Solve Tic-Tac-Toe and relevant gaming using minimax search algorithm.

Tic-Tac-Toe: The game is to be played between two people (in this program between HUMAN and COMPUTER). One of the player chooses 'O' and the other 'X' to mark their respective cells. The game starts with one of the players and the game ends when one of the players has one whole row/ column/ diagonal filled with his/her respective character ('O' or 'X').If no o

ne wins, then the game is said to be draw.

O	X	O
O	X	X
X	O	X

3.

Compare top 10 open source chatbot platforms based on the components such as interface, chatbot elements, messages, AI and NLP, channels, integration, usage, analytics and optimization.

RESOURCES

TEXT BOOKS:

1. Stuart Russel, Peter Norvig, Artificial Intelligence A Modern Approach, 4th Edition, Pearson Education, 2020.

2. Elaine Rich, Kevin Knight and Shiv Shankar B. Nair, Artificial Intelligence, 3rd edition, Tata McGraw Hill, 2009.

REFERENCE BOOKS:

1. Ela Kumar, "Artificial Intelligence", I.K.International Publishing House Pvt.Ltd, 2008.

2. Wolfgang Ertel," Introduction to Artificial Intelligence", Second Edition, Springer, 2017.

3. Deepak Khemani, A First Course in Artificial Intelligence, McGraw Hill Education, 2017.

4. Saroj Kaushik, Artificial Intelligence, Cengage Learning, 2011.

VIDEO LECTURES:

1. https://nptel.ac.in/courses/106105077

2. https://nptel.ac.in/courses/106105078

3. https://nptel.ac.in/courses/106106182

4. https://www.youtube.com/watch?time_continue=1&v=zx_oMonGdK4

5. https://www.youtube.com/watch?v=dtGRmhZ6Cuo&list=PLEdmaLOtiZFswiGRQAr9iVVucY3L3jvgA&index=9&t=70s

6. https://www.youtube.com/watch?v=yMcZvZayJUA&t=4s

WEB RESOURCES:

1. http://cs.wwc.edu/~cs_dept/KU/PR/Prolog.html

2. http://www.cs.may.ie/~jpower/Courses/PROLOG/

3. http://www.csupomona.edu/~jrfisher/www/prolog_tutorial/contents.html

4. https://www.coursera.org/learn/python-for-applied-data-science-ai

5. https://www.coursera.org/learn/python-for-applied-data-science-ai

PROGRAM CORE

Course Code	Course Title	L	T	P	S	C
22AI202002	**MACHINE LEARNING**	3	-	3	-	4.5

Pre-Requisite -

Anti-Requisite -

Co-Requisite -

COURSE DESCRIPTION: This course introduces the essential methods that are at the core of modern machine learning. It covers theoretical foundations as well as key supervised and unsupervised learning algorithms. Classes on theoretical and algorithmic aspects are augmented by laboratory exercises.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Perform mathematical analysis of machine learning algorithms using computational complexity theory concepts.

CO2. Demonstrate how Bayesian Methods can be used to solve real world problems.

CO3. Build predictive models by constructing hierarchical and compact representation of the available observations.

CO4. Evaluate the performance of the trained classification models and implement necessary measures for improving its performance.

CO5. Select and Apply appropriate clustering algorithms to solve real-time problems and optimize the models learned.

CO6. Work independently to solve problems with effective communication.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	3	2	2	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	2	2	3	-	-	-	-
CO4	3	2	3	-	-	-	-
CO5	3	3	2	-	-	-	-
CO6	-	-	-	-	3	3	-
Course Correlation Mapping	3	2	3	-	3	3	-

Correlation Levels: 3: High; 2: Medium;1: Low

COURSE CONTENT

Module 1: INTRODUCTION   *(09 Periods)*

Machine Learning, Applications, Supervised Learning - Learning a class from examples, Vapnik Chervonenkis (VC) dimension, Probably Approximately Correct (PAC) learning, Noise, Learning multiple classes, Regression, Model selection, and generalization

Module 2:   BAYESIAN DECISION THEORY AND LINEAR   *(09 Periods)*
             DISCRIMINATION

Bayesian Decision Theory: Classification, Losses, and risks, Discriminant functions, Utility theory, Value of information, Bayesian networks, Influence diagrams, Association rules.

Linear Discrimination: Generalizing linear model, Geometry of the linear discriminant, Pairwise separation, Parametric discrimination, Gradient descent, Support vector machines.

Module 3:   DECISION TREES AND DIMENSIONALITY   *(09 Periods)*
             REDUCTION

Decision trees: Univariate trees, Pruning, Rule extraction from trees, Learning rules from data, Multivariate trees.

Dimensionality Reduction: Subset Selection, Principal Component Analysis, Feature Embedding, Factor Analysis, Singular Value Decomposition and Matrix Factorization, Multidimensional Scaling.

Module 4:   PERFORMANCE EVALUATION OF CLASSIFICATION   *(09 Periods)*
             ALGORITHMS

Cross-validation and resampling methods, Measuring error, Interval estimation, Hypothesis testing, Assessing a classification algorithm's performance, and Comparing two classification algorithms. Combining Multiple Learners- Rationale, Voting, Bagging, Boosting, The mixture of experts revisited, Stacked generalization, Cascading

Module 5:   CLUSTERING   *(09 Periods)*

Introduction, Mixture Densities, k-Means Clustering, Expectation, Maximization Algorithm, Mixtures of Latent Variable Models, Spectral Clustering, Hierarchical Clustering, Choosing the Number of Clusters.

Total Periods: 45

EXPERIENTIAL LEARNING

1. Solve the house price prediction problem using linear regression analysis method. Optimize the parameters of the regression function using gradient descent method.

2. Learn the best association rule using the Apriori algorithm with reference to a customer's transaction dataset.

3. Using Optdigits from the UCI repository, implement PCA. For various number of eigenvectors, reconstruct the digit images and calculate the reconstruction error.

4. Build a binary classifier using Support vector Machine and evaluate the performance of the different kernels using a grid search method. (Use breast cancer dataset)

5. Visualize the decision tree built for solving diabetes disease prediction problem and measure the impurity of nodes created via Decision Tree Analysis.

6. Investigates whether the means of two independent data samples differ from one another using a two-sample T-test.

7. Checks whether a sample mean differs from the population mean using One-Sample T-test.

8. Evaluate the gradient boosting ensemble technique for regression problem. Calculate the error residuals and fine-tune the hyperparameters of the model.

9. Using stacked ensembling approach build sub-models and aggregator model to make predictions on the Sonar dataset.

10. Cluster the handwritten digits data using the EM algorithm with a principle components step within each maximization.

11. Use the k-means clustering for color image compression.

12. Generate new handwritten digits from the standard digits corpus using the Gaussian Mixture Model. Identify the optimal number of components Akaike Information Criteria.

RESOURCES

TEXT BOOKS:

1. Ethem Alpaydin, Introduction to Machine Learning, MIT Press (Adaptive Computation and Machine Learning Series), 3rd Edition, 2014.

REFERENCE BOOKS:

1. Stephen Marsland, Machine Learning: An Algorithmic Perspective, Chapman and Hall/CRC Machine Learning and Pattern Recognition Series, 2nd Edition, 2014.

2. Richard O. Duda, Peter E. Hart, David G. Stork, Pattern Classification, Wiley, 2nd Edition, 2012.

SOFTWARE TOOLS:

1. Environment: Google CoLab

2. Programming Language: Python 3.8

3. Machine Learning Library: Scikit-learn, Tensor Flow 2.1 and Keras

VIDEO LECTURES:

1. https://onlinecourses.nptel.ac.in/noc21_cs24/preview

2. https://www.youtube.com/watch?v=_M7Km1XZERU

3. https://www.youtube.com/watch?v=UudeDPTtMos&list=PL1xHD4vteKYVpaIiy295pg6_SY5q znc77&index=18

4. https://www.youtube.com/watch?v=OGLxVh1J-xk

5. https://onlinecourses.nptel.ac.in/noc21_cs85/preview

WEB RESOURCES:

1. https://data-flair.training/blogs/machine-learning-tutorial/
2. https://www.w3schools.com/python/python_ml_getting_started.asp
3. https://www.kaggle.com/dansbecker/learn-machine-learning
4. https://www.youtube.com/watch?v=H9yACitf-KM
5. https://www.youtube.com/watch?v=7oNiqPoKD8

PROGRAM CORE

Course Code	Course Title	L	T	P	S	C
22AI202003	**BIG DATA FRAMEWORKS**	3	-	3	-	4.5

Pre-Requisite -

Anti-Requisite -

Co-Requisite -

COURSE DESCRIPTION: This course provides a detailed discussion and Hands-On-Experience on HDFS Concepts, Hadoop File System, Hadoop Operations, Develop a MapReduce Application, Yarn Architecture, Mesos Setup, Spark Standalone Architecture, Spark Execution Model, Spark Cache, Job Execution and Fault Tolerance.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Demonstrate Knowledge on big data frameworks and Hadoop Distributed File System.

CO2. Develop data processing and analytics applications using Map Reduce Workflows and spark streaming

CO3. Build a spark cluster and manage resources dynamically using YARN or Mesos

CO4. Apply several techniques such as shuffling, partitioning, and caching mechanisms to improve the performance of Spark jobs

CO5. Explore various job scheduling methods and use recommended configurations to build a fault tolerant application.

CO6. Work independently to solve problems with effective communication.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	3	2	2	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	2	2	3	-	-	-	-
CO4	3	2	3	-	-	-	-
CO5	3	3	2	-	-	-	-
CO6	-	-	-	-	3	3	
Course Correlation Mapping	3	2	3	-	3	3	-

Correlation Levels: 3: High; 2: Medium;1: Low

COURSE CONTENT

Module 1:   INTRODUCTION TO HADOOP   *(09 Periods)*

Data, Data Storage and Analysis, Querying all your data, Beyond Batch, Comparison with Other Systems, The Design of HDFS, HDFS Concepts, The Command-Line Interface, Hadoop File Systems, The Java Interface, Data Flow, Parallel Copying with distcp.

Module 2:   HADOOP OPERATIONS AND DEVELOPING A MAP REDUCE   *(11 Periods)*
APPLICATION

Hadoop Operations: Cluster Specification, Cluster Setup and Installation, Hadoop Configuration, Security, Benchmarking a Hadoop Cluster.
MapReduce Application: The Configuration API, Setting up the Development Environment, Writing a Unit Test With MRUnit, Running Locally on Test Data, Running on a Cluster, Tuninga Job, MapReduce Workflows.

Module 3   CLUSTER MANAGEMENT   *(08 Periods)*

Background, Spark Components-Driver, Workers and Executors, Configuration; Spark Standalone-Architecture, Single-NodeSetup scenario, Multi Node Setup; YARN-Architecture, Dynamic Resource Allocation, Scenario; Mesos-Setup, Architecture, Dynamic Resource Allocation, Basic Setup Scenario.

Module 4   PERFORMANCE TUNING   *(08 Periods)*

Spark Execution Model, Partitioning, Shuffling Data – Shuffling and Data Partitioning, Operators and Shuffling; Serialization, Spark Cache, Memory Management, Shared Variables, Broadcast Variable, Accumulators; Data Locality.

Module 5   JOB EXECUTION AND FAULT TOLERANCE   *(09 Periods)*

Job Execution: Life Cycle of a Spark Job-Spark Master, Spark Driver, Spark Worker, Job Life Cycle; Job Scheduling- Scheduling Within an Application, Scheduling with External Utilities.
Fault Tolerance: Fault Tolerance-Internal and External Fault Tolerance, Service Level Agreements, Resilient Distributed Datasets, Batch versus Streaming, Testing Strategies, Recommend Configurations.

Total Periods: 45

EXPERIENTIAL LEARNING

LIST OF EXERCISES:

1. Build and set up Hadoop framework to run in single-node and multi-node setup.

2. Write and implement a text file processing program using MapReduce model.

3. Implement a MapReduce program suing pySpark for identifying potential customers.

4. Consider the dataset (attached Ex.4.txt) and write a mapper and reducer program for finding the cost of the item that is most expensive, for each location.

    Link to download dataset:

    https://drive.google.com/file/d/109Lp4_BgQPilDNDduO7638YCDCd7MGY2
    /view?usp=sharing

5. Install Apache Spark and configure it to run on a single machine. Create workers on different machines and configure a multi-node setup.

6. Install Mesos and configure Spark to run with Mesos and perform dynamic resource allocation.

7. Using pySpark in an interactive mode,perform the following tasks:
   i. load data from a CSV file (CRAN package download logs http://cran-logs.rstudio.com/)

   ii. display the first n rows of the data

   iii. transform each row of data into an array

   iv. use countbykey method to find the number of downloads for a package

8. Implement the following Spark Dataframe and SQL operations;
   i. Create a spark dataframe from python list and RDD

   ii. Change the data frame properties

   iii. filter and aggregate the data
   iv. transform a dataframe column v. build a view with the Spark DataFrame

9. Write a pySpark program to count the number of occurrences of words in a text and use explicit caching.

10. Analyze the impact of number of worker cores on a parallelized operation and use Caching to reduce computation time.

11. Perform image classification using Spark Deep Learning.

12. Build a linear regression model using Apache spark's spark.ml for stock price prediction.

RESOURCES

TEXT BOOKS:

1. Tom White, Hadoop:The Definitive Guide,O'Reilly,4th Edition,2015.

2. Llya Ganelin,Ema Or.hian, Kai Sasaki, Brennon York, Spark: Big Data Cluster Computing in Production,1st Edition,2016.

REFERENCE BOOKS:

1. Nick Pentreath,Machine Learning with Spark,PacktPublishing,2015.

2. Mohammed Guller, Big Data Analytics with Spark, Apress,2015.

3. Donald Miner, Adam Shook, Map Reduce Design Pattern, O'Reilly, 2012

4. Frank J. Ohlhorst, Big Data Analytics: Turning Big Data into Big Money, Wiley Publication,December, 2012.

5. Kevin Roebuck, Big Data: High-Impact Strategies - What You Need to Know: Definitions, Adoptions, Impact, Benefits, Maturity, Vendors, Tebbo Publisher, 2011.

6. Alex Holmes, Hadoop in Practice, Manning Publications Publisher, 2012.

SOFTWARE/TOOLS:

1. APACHE HADOOP, APACHE SPARK, JDK

VIDEO LECTURES:

1. https://onlinecourses.nptel.ac.in/noc20_cs92/preview
2. https://archive.nptel.ac.in/courses/106/104/106104189/
3. https://www.youtube.com/watch?v=g7Qpnmi0Q-s
4. https://www.youtube.com/watch?v=cYL42BBL3Fo
5. https://www.youtube.com/watch?v=x-PCNX4prLA
6. https://www.youtube.com/watch?v=QaoJNXW6SQo
7. https://www.youtube.com/watch?v=mafw2-CVYnA

WEB RESOURCES:

1. https://github.com/rshah204/MapReduce -Implementation-in-Python/blob/master/MapReduce.ipynb
2. https://github.com/rshah204/MapReduce-Implementation-in-PySpark
3. https://github.com/XD-DENG/Spark-practice#load-data
4. https://github.com/XD-DENG/Spark-practice/blob/master/Spark%20DataFrames%20%26%20SQL%20 -%20Basics.ipynb
5. https://github.com/tirthajyoti/Spark-with-Python/blob/master/Word_Count.ipynb
6. https://github.com/tirthajyoti/Spark-with-Python/blob/master/SparkContext_Workers_Lazy_Evaluations.ipynb
7. https://medium.com/linagora -engineering/making-image-classification-simple-with-spark-deep-learning-f654a8b876b8
8. https://drive.google.com/file/d/109Lp4_BgQPilDNDduO7638YCDCd7MGY2/view?usp=sharing

PROGRAM CORE

Course Code	Course Title	L	T	P	S	C
22AI201007	**DATA MINING**	3	-	-	-	3

Pre-Requisite -
Anti-Requisite -
Co-Requisite -

COURSE DESCRIPTION: This course provides a detailed discussion on the concepts, principles, methods and applications of data mining with the focus on major data mining techniques such as Association pattern mining, Clustering, Classification and Outlier analysis. The course will also cover algorithms and methods related to data preprocessing, Frequent pattern mining, Clustering and Classification with relevant case studies.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Demonstrate knowledge on Data Mining and Data Preparation process to gather, combine and organize data to make informed decision.

CO2. Identify frequent patterns, associations using frequent Itemset Mining algorithms for further analysis of data mining tasks and to support business decision-making process.

CO3. Select and Apply appropriate Clustering and Classification methods to solve real-life problems.

CO4. Analyze and Apply suitable Outlier detection methods in diverse domains such as Quality Control, Fault detection, Financial Fraud and Intrusion detection.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	3	2	-	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	3	3	3	-	-	-	-
CO4	3	3	3	-	-	-	-
Course Correlation Mapping	3	3	3	-	-	-	-

Correlation Levels: 3: High; 2: Medium; 1: Low

COURSE CONTENT

Module 1: INTRODUCTION TO DATA MINING AND DATA *(09 Periods)*
PREPARATION

Introduction: The Data Mining Process, The Basic Data Types, The Major building blocks, Scalability Issues and the Streaming Scenario.

Data Preparation: Feature Extraction, Data Type Portability, Data Cleaning, Data Reduction and Transformation – Sampling, Feature Subset Selection, Dimensionality Reduction with Axis Rotation.

Module 2: ASSOCIATION PATTERN MINING *(09 Periods)*

The Frequent Pattern Mining Model, Association Rule Generation Framework, Frequent Item set Mining Algorithms, Alternative Models: Interesting Patterns, Meta-algorithms, Pattern Summarization, Pattern Querying.

Applications: Classification, Clustering, Outlier detection, Market Basket Analysis, Recommendations and Collaborative Filtering, Web log analysis.

Module 3: CLUSTER ANALYSIS *(09 Periods)*

Feature selection for Clustering, Representative based algorithm-k-Means algorithm; Hierarchical Clustering algorithm- Agglomerative Methods; Probabilistic model based algorithm-Relationship of EM to k-Means and other representative methods, Grid and Density based methods-DBSCAN, DENCLUE, CLIQUE; Cluster Validation, Clustering Categorical Data, Cluster Ensembles.

Applications: Applications to other Data Mining problems, Text Applications, Multimedia Applications, Social Network Analysis.

Module 4: DATA CLASSIFICATION *(09 Periods)*

Feature selection for Classification, Decision Trees, Rule based Classifiers, Probabilistic Classifier, Support Vector Machines, Instance based Learning**,** Classifier Evaluation, Ensemble Methods.

Module 5: OUTLIER ANALYSIS *(09 Periods)*

Extreme Value Analysis, Probabilistic Models, Clustering for Outlier detection, Distance based Outlier Detection, Density-Based Methods, Information-Theoretic Models, Outlier Validity, Outlier Detection with Categorical Data, High-Dimensional Outlier Detection, Outlier Ensembles.

Applications: Quality Control and Fault detection, Financial Fraud and Anomalous Events, Intrusion Detection Application.

Total Periods: 45

EXPERIENTIAL LEARNING

1. **Prediction of Adult Income based on Census Data available in kaggle:**

   Apply classification algorithm to predict the income level of an individual that exceeds 50K based on the census data available at the online repository. The dataset consists

of variables such as age, type of work, working hours, sex and many more. Perform data analysis to understand the standard of living of the city, benefit of setting up the business or bank loan eligibility. Also, to understand the real estate preferences by average income of the people residing in the area, figure out the type of tourist places that people from other countries would like to travel.

2. **Product and Price Comparing tool**:

With the increase in popularity of e-commerce portals, shopping websites are magnifying to a great extent to enable online shoppers to purchase anything with just one click and get it delivered at your doorstep. To purchase an item, people tend to spend quite a lot of time in searching a product and comparing it with other websitesby themselves. Automate the task of comparing product and price of a product to help customer buy cheap and best deal available. Also, track the consumer demand and inform when the commodity price is lowest and notify consumers proactively.

3. **Comparative Study of Clustering Algorithms:**

Implement clustering algorithms k-means , hierarchical clustering and Expectation Maximization algorithm using python and compare based on the metrics :

- number of clusters
- size of dataset
- performance of the algorithm
- accuracy of the algorithm

RESOURCES

TEXT BOOKS:

1. Charu C.Aggarwal,*Data Mining*,The Text book, Springer,2015.
2. Jiawei Han and Micheline Kamber,*Data Mining Concepts and Techniques*, 3rd edition, Elsevier, 2012.

REFERENCE BOOKS:

1. Alex Berson and Stephen J.Smith, ―*Data Warehousing, Data Mining & OLAP‖*, Tata McGraw – Hill Edition, 35th Reprint 2016.
2. K.P. Soman, Shyam Diwakar and V. Ajay, ― *Insight into Data Mining Theory and Practice,* Eastern Economy Edition, Prentice Hall of India, 2006.

VIDEO LECTURES:

f1.   https://nptel.ac.in/courses/106105174
2.   https://nptel.ac.in/courses/110107092
3.   https://freevideolectures.com/course/4566/nptel-data-mining/9
4.   https://www.youtube.com/channel/UCdaj0DjWHoy7hFLtZtjtZ7g
5.   https://www.youtube.com/watch?v=pBAbMbgaBAk
6.   https://www.youtube.com/watch?v=esmzYhuFnds&t=6s
7.   https://www.youtube.com/watch?v=qg_M37WGKG8

WEB RESOURCES:

1.   https://www.tutorialandexample.com/outlier-analysis-in-data-mining
2.   https://www.mygreatlearning.com/academy/learn-for-free/courses/data-mining1

3. https://www.researchgate.net/publication/49616224_Data_mining_techniques_and_applications

4. https://www.geeksforgeeks.org/data-mining/

5. https://www.tutorialspoint.com/data_mining/dm_pdf_version.htm

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

PROGRAM CORE

Course Code	Course Title	L	T	P	S	C
22AI202005	**FEATURE ENGINEERING FOR MACHINE LEARNING**	3	-	3	-	4.5

Pre-Requisite   22AI202002-Machine Learning

COURSE DESCRIPTION: This course provides discussion and hands-on experience on machine learning pipeline, Feature Scaling, Selection, Text parsing, Vectorization, Classification, Hashing, Dimensionality Reduction, non-linear Featurization, Feature Extraction, End-to-End learning with Deep Neural architectures.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Perform pre-processing of datasets to build efficient machine learning models.

CO2. Transform text and image data in to a more suitable form for further processing and building machine learning model.

CO3. Analyze the given problem and design suitable feature engineering strategy.

CO4. Reduce the dimensionality of the data and select optimal subset of the feature set.

CO5. Implement automated feature extraction using deep learning architectures.

CO6. Work independently to solve problems with effective communication.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	3	2	2	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	2	2	3	-	-	-	-
CO4	3	2	3	-	-	-	-
CO5	3	3	2	-	-	-	-
CO6	-	-	-	-	3	3	-
Course Correlation Mapping	3	2	3	-	3	3-	-

Correlation Levels: 3: High; 2: Medium; 1: Low

COURSE CONTENT

Module 1: INTRODUCTION *(09 Periods)*

Machine Learning Pipeline, Scalars, Vectors, and Spaces, Log transformation, Feature Scaling or Normalization, Feature Selection, Text Data - Bag-of-X: Turning Natural Text into Flat Vectors, Filtering for Cleaner Features, Parsing and Tokenization, Collocation Extraction for Phrase Detection.

Module 2: FEATURE SCALING *(09 Periods)*

Term Frequency & Inverse Document frequency, Scaling Bag-of-Words with Tf-Idf Transformation, Classification with Logistic Regression, Tuning Logistic Regression with Regularization, Encoding Categorical Variables, Feature Hashing, Bin Counting.

Module 3 DIMENSIONALITY REDUCTION *(09 Periods)*

Intuition, Derivation, Linear Projection, Variance and Empirical Variance, Principal Components: First Formulation, Principal Components: Matrix-Vector Formulation, General Solution of the Principal Components, Transforming Features, Whitening and ZCA, Considerations and Limitations of PCA.

Module 4 NONLINEAR FEATURIZATION *(09 Periods)*

K-Means Clustering, Clustering as Surface Tiling, k-Means, Featurization for Classification, Alternative Dense Featurization.

Module 5 IMAGE FEATURE EXTRACTION AND DEEP LEARNING *(09 Periods)*

Feature Extraction- SIFT and HOG, Image Gradients, Gradient Orientation Histograms, SIFT Architecture; Learning Image Features with Deep Neural Networks -Fully Connected Layers, Convolutional Layers;

Total Periods: 45

EXPERIENTIAL LEARNING

LIST OF EXERCISES:

1. Visualize the distribution of review counts of yelp business reviews before and after log transform, and Box-Cox transformed counts.

   Dataset: https://www.kaggle.com/datasets/omkarsabnis/yelp-reviews-dataset

2. Visualize the distribution of data with different feature scaling methods on online news popularity dataset for article word count.

   Dataset: https://www.kaggle.com/datasets/deepakshende/onlinenewspopularity

3. Plot the histograms for original and scaled data from Experiential Learning 2 and compare the scaling results.

4. Build a linear regression model to predict the chance of admission for each student without normalization and with normalization and compare the accuracy.

Dataset: https://www.kaggle.com/datasets/faressayah/college-data

5. Predict heart diseases using logistic regression on UCI dataset.

   Dataset:https://www.kaggle.com/arviinndn/heart-disease-prediction-uci-dataset/data

6. Perform parameter tuning using regularization for a Logistic regression on any online dataset.

7. Perform Principal component analysis (PCA) of the scikit-learn digits dataset (a subset of the MNIST dataset) and visualize the results.

8. Anomaly detection in time series data using Principal component analysis (PCA).

9. Cluster the following eight points (with (x, y) representing locations) into three clusters:

   A1(2, 10), A2(2, 5), A3(8, 4), A4(5, 8), A5(7, 5), A6(6, 4), A7(1, 2), A8(4, 9)

10. Use k-means algorithm to cluster headlines into different categories using python.

11. Perform image matching on real-time images using scale-invariant feature transform (SIFT) algorithm.

12. Detect multiple faces with dlib using (Histogram of Oriented Gradients) HOG and CNN on any online dataset.

RESOURCES

TEXT BOOKS:

1. Alice Zheng, Amanda Casari,*Feature Engineering for Machine Learning*, O'Reilly Media, Inc. (ISBN: 9781491953242), 2018.

REFERENCE BOOKS:

1. Kjell Johnson and Max Kuhn , *Feature Engineering and Selection: A Practical Approach for Predictive Models,* CRC Press, 2020.

SOFTWARE/TOOLS:

1. Environment: Google CoLab
2. Programming Language: Python 3.8,OpenCV
3. Machine Learning Library: Pandas, numpy, scikit-learn

VIDEO LECTURES:

1. https://www.datacamp.com/courses/feature-engineering-for-machine-learning-in-python
2. https://www.coursera.org/learn/feature-engineering
3. https://www.udemy.com/course/feature-engineering-for-machine-learning/
4. https://www.datacamp.com/courses/feature-engineering-with-pyspark?irclickid=xHmUT903Ez6CRdrTwkRSeVn8UkDzXNV5dSZS0w0&irgwc=1&utm_medium=affiliate&utm_source=impact&utm_campaign=2355712
5. https://www.coursera.org/learn/feature-engineering-matlab?irclickid=WCW2O%3AQXWxyIUq2WaWTSN2NBUkDzXNVRdSZS0w0&irgwc=1&utm_medium=partners&utm_source=impact&utm_campaign=3310965&utm_content=b2c

WEB RESOURCES:

1. https://neptune.ai/blog/feature-engineering-tools
2. https://www.analyticsvidhya.com/blog/2021/03/step-by-step-process-of-feature-engineering-for-machine-learning-algorithms-in-data-science/
3. https://www.udemy.com/course/feature-engineering-for-machine-learning/
4. https://developers.google.com/machine-learning/data-prep
5. https://www.analyticsvidhya.com/blog/2021/05/feature -scaling-techniques-in-python-a-complete-guide/
6. https://medium.com/@dennisndungu68/text-classification-using-k-means-33bea24e4a94

PROGRAM CORE

Course Code	Course Title	L	T	P	S	C
22CS201001	**ADVANCED DATA STRUCTURES AND ALGORITHMS**	3	-	-	-	3

Pre-Requisite -

Anti-Requisite -

Co-Requisite -

COURSE DESCRIPTION: This course provides a detailed discussion on Data Structure Concepts, Stacks, Queues, Trees, Graphs, Skip Lists, Heaps, Hashing, and Real-Time Applications of Data Structures. This course also examines Design and Analysis of Algorithmic Concepts on Divide and Conquer, Greedy Method, Dynamic Programming, Backtracking, NP-Completeness and advanced algorithms Approximation algorithms and Number-Theoretic algorithms.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Analyze linear data structures such as linked lists, stacks, and queues for efficient data organization and manipulation.

CO2. Analyze data structures such as trees, graphs, heaps and hash tables for efficient storage, search and retrieval of data.

CO3. Design and evaluate the performance of efficient techniques such as Divide-and-conquer, Greedy Method, Dynamic Programming, Backtracking, to solve engineering problems.

CO4. Apply NP-Hard and NP-complete, Approximation and Number theoretic algorithms to solve complex problems.

CO5. Apply knowledge to select appropriate data structures and algorithms for solving Societal applications.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	3	2	-	-	-	-	-
CO2	3	2	-	-	-	-	-
CO3	3	2	-	-	-	-	-
CO4	3	2	-	-	-	-	-
CO5	2	3	1	-	-	-	-
Course Correlation Mapping	**3**	**2**	**1**	**-**	**-**	**-**	**-**

Correlation Levels: 3: High; 2: Medium; 1: Low

COURSE CONTENT

Module 1: ELEMENTARY DATA STRUCTURES AND TREES   *(08 Periods)*

Elementary Data Structures: Introduction, Time and space complexity, Asymptotic notations, Stacks, Queues - Circular Queue, Dequeue; Linked lists.
Trees: Binary trees, Binary search trees, AVL Trees, Red Black Trees.

Module 2: GRAPHS AND SKIP LISTS   *(09 Periods)*

Graphs: Graph traversal techniques, Minimum spanning trees, Maximum Bipartite Matching, Minimum cost flow.
Skip Lists: Need for Randomizing data structures and algorithms, Search and update operations on skip lists, Probabilistic analysis of skip lists, Deterministic skip lists.

Module 3: HEAPS AND HASHING   *(09 Periods)*

Heaps: Heap Implementation, priority queues, Applications, Binomial heaps, operations on binomial heaps, Fibonacci heaps, Mergeable heap operations.
Hashing: Hash functions and problems, Collision resolution techniques, Universal hashing, Applications.

**Module 4: DIVIDE AND CONQUER, GREEDY METHOD, DYNAMIC   *(10 Periods)*
                    PROGRAMMING AND BACK TRACKING**

***Divide and Conquer*:** General method, Binary search.
***Greedy Method*:** General method, Job sequencing with deadlines.
***Dynamic Programming*:** 0/1 Knapsack problem, Traveling salesperson problem.
***Backtracking*:** N–Queen's problem.

**Module 5: NP-COMPLETENESS, APPROXIMATION ALGORITHMS   *(09 Periods)*
                    AND NUMBER-THEORETIC ALGORITHMS**

***NP-Completeness*:** Polynomial-time, Polynomial-time verification, NP-completeness and reducibility, NP-complete problems.
***Approximation Algorithms*:** The vertex-cover problem, Traveling salesman problem, Subset-sum problem.
Number Theoretic Algorithms: Elementary number-theoretic notions, Greatest common divisor.

                                                            Total Periods: 45

EXPERIENTIAL LEARNING

1. Perform a study on the challenges, existing solutions, applications related to solving the layout problem and submit a detailed report with your findings on several optimal and heuristic algorithms that provide useful insight into the layout problem to provide near-optimal solutions. The following research papers can be studied. Other relevant papers can also be studied.

    a) https://www.taylorfrancis.com/chapters/mono/10.1201/9781003285090-12/advanced-algorithms-layout-problem-sunderesh-heragu

    b) https://doi.org/10.3182/20060517-3-FR-2903.00208

RESOURCES

TEXT BOOKS:

1. Ellis Horowitz, Sartaj Sahni, and S Rajasekaran, Fundamentals of Computer Algorithms, Universities Press, 2nd Edition, 2008.
2. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein, Introduction to Algorithms, PHI Learning, 3rd Edition, 2009.
3. Goodrich, Tamassia, Goldwasser, Data structures & Algorithms in Java, 6th Edition, Wiley, 2014.

REFERENCE BOOKS:

1. John R. Hubbard, Programming with Java, 2nd Edition, McGraw Hill, 2009.
2. Debasis Samanta, Classic Data Structures, 2nd Edition, Prentice Hall, 2009.
3. Robert Lafore, Data Structures & Algorithms in Java, 2nd Edition, Pearson, 2007.
4. Mark Allen Weiss, Data Structures and Algorithm Analysis in C++, Pearson, 4 th Edition, 2014.
5. Michael T. Goodrich, Roberto Tomassia, Algorithm Design: Foundations, Analysis and Internet Examples, Wiley, 2002

VIDEO LECTURES:

1. https://nptel.ac.in/courses/106102064
2. http://nptel.ac.in/courses/106106127
3. https://www.youtube.com/watch?v=RBSGKlAvoiM
4. https://www.youtube.com/watch?v=xZKqH7ZcS_Y
5. https://www.youtube.com/watch?v=8hly31xKli0
6. https://www.youtube.com/watch?v=P5IH4lqCJSk

WEB RESOURCES:

1. https://www.cs.auckland.ac.nz/software/AlgAnim/red_black.html
2. https://www.geeksforgeeks.org/maximum-bipartite-matching/
3. https://www.gatevidyalay.com/job-sequencing-with-deadlines/
4. https://www.tutorialandexample.com/collision-resolution-techniques-in-data-structure
5. https://developers.google.com/optimization/cp/queens

PROGRAM ELECTIVE

Course Code	Course Title	L	T	P	S	C
22AI202006	**DEEP LEARNING**	3	-	3	-	4.5

Pre-Requisite	22AI202002-Machine Learning
Anti-Requisite	-
Co-Requisite	-

COURSE DESCRIPTION: This course provides an overview of deep learning, a field of machine learning that deals with the creation and implementation of modern neural networks such as Recurrent Neural Networks, Convolutional Neural Networks. This course also deals with advanced deep learning models like Convolutional autoencoders, Attention mechanisms for machine translation; Generative adversarial networks (GANs).

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Acquire working knowledge with deeper theoretical understanding of neural networks and explore the different parameters of the network.

CO2. Construct a generative model for learning probability distribution using Restricted Boltzman Machines.

CO3. Analyze temporal sequential input data using gated memory based neural units.

CO4. Utilize Convolutional Neural Network for analyzing visual imagery and utilize transfer learning approaches for reducing the training efforts.

CO5. Apply encoder-decoder architecture for image denoising, and learning representation of a set of data.

CO6. Work independently to solve problems with effective communication.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	3	2	2	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	2	2	3	-	-	-	-
CO4	3	2	3	-	-	-	-
CO5	3	3	2	-	-	-	-
CO6	-	-	-	-	3	3	-
Course Correlation Mapping	3	2	3	-	3	3	-

Correlation Levels: 3: High; 2: Medium; 1: Low

COURSE CONTENT

Module 1: INTRODUCTION TO NEURAL NETWORKS   *(09 Periods)*

Basic architecture of neural networks, Training a neural network with backpropagation, Practical issues in neural network training- Overfitting, Vanishing and exploding gradient problems, Difficulties in convergence, Local and spurious optima, Computational challenges.

Module 2: RESTRICTED BOLTZMANN MACHINES   *(09 Periods)*

Hopfield networks, Boltzmann machine, Restricted boltzmann machines- Training the RBM, contrastive divergence algorithm, Practical issues and improvisations; Stacking restricted boltzmann machines- Deep boltzmann machines and deep belief networks;.

Module 3   RECURRENT NEURAL NETWORKS   *(09 Periods)*

Architecture of recurrent neural networks, Challenges of training recurrent networks- Layer normalization; Echo-state networks, Long short-term memory (LSTM), Gated recurrent units (GRUs).

Module 4   CONVOLUTIONAL NEURAL NETWORKS   *(09 Periods)*

Structure of a convolutional network, Training a convolutional network, Case studies of convolutional architectures- AlexNet, ZFNet, VGG, GoogLeNet, ResNet, Effects of depth, Pretrained models.

Module 5   ADVANCED TOPICS IN DEEP LEARNING   *(09 Periods)*

Convolutional autoencoders, Attention mechanisms- Recurrent models of visual attention, Attention mechanisms for machine translation; Generative adversarial networks (GANs).

Total Periods: 45

EXPERIENTIAL LEARNING

LIST OF EXERCISES:

1. Perform splitting of data for training, testing, and validation using k-fold cross validation.
2. Construct and implement multi-layer feed forward neural network for hand written digit classification problem.
3. Solve the overfitting problem in a neural architecture using DropOut technique.
4. Implement image reconstruction using Restricted boltzmann machines (RBM).
5. Solve a time series forecasting (stock prediction) using LSTM RNN.
6. Solve a seq2seq problem (machine translation) using LSTM Recurrent Neural Architecture.
7. Implement a binary and multi class image classification using Convolution Neural Network.
8. Perform hyper parameter tuning using Bayesian optimization technique for a Convolution Neural Network.
9. Study the efficiency of the transfer learning approach for a classification problem on the following architectures; VGG-16, Alexnet, and Inception-V3.

10. Implement the image dimensionality reduction problem using a AutoEncoder architecture.

RESOURCES

TEXT BOOKS:

1. Charu C. Aggarwal, Neural Networks and Deep Learning, Springer, 2018.

REFERENCE BOOKS:

1. Ian Goodfellow, Yoshua Bengio, Aaron Courville,Deep Learning, 4th Edition, MIT Press, 2016.
2. Deng & Yu, Deep Learning: Methods and Applications, Now Publishers, 2013.

SOFTWARE/TOOLS:

1. Environment: Google CoLab
2. Programming Language: Python 3.8
3. Machine Learning Library: Tensor Flow 2.1 and Keras

VIDEO LECTURES:

1. https://nptel.ac.in/courses/106106184
2. https://www.coursera.org/specializations/deep-learning?utm_source=deeplearningai&utm_medium=institutions&utm_campaign=WebsiteCoursesDLSTopButton
3. https://www.simplilearn.com/tutorials/deep-learning-tutorial/what-is-deep-learning
4. https://www.udemy.com/course/data-science-logistic-regression-in-python/
5. https://www.udemy.com/course/data-science-deep-learning-in-python/

WEB RESOURCES:

1. https://www.youtube.com/watch?v=aPfkYu_qiF4
2. https://www.guru99.com/deep-learning-tutorial.html
3. https://www.coursera.org/professional-certificates/tensorflow-in-practice
4. https://www.mathworks.com/discovery/deep-learning.html
5. https://www.techtarget.com/searchenterpriseai/definition/deep-learning-deep-neural-network#:~:text=Deep%20learning%20is%20a%20type,includes%20statistics%20and%20predictive%20modeling.

PROGRAM ELECTIVE

Course Code	Course Title	L	T	P	S	C
22AI202007	**COMPUTER VISION**	3	-	3	-	4.5

Pre-Requisite	22AI202002-Machine Learning
Anti-Requisite	-
Co-Requisite	-

COURSE DESCRIPTION: This course provides an introduction to computer vision including fundamentals of image formation, camera models and geometry, stereo, object tracking, image clustering, classification and scene understanding. Methods for depth recovery from stereo images, camera calibration, automated alignment, tracking, boundary detection, and recognition are explored in detail.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Demonstrate knowledge on various image representation and transformation techniques suited for digital image processing tasks.

CO2. Analyze stereo images, and apply camera calibration methods for estimation of disparity maps.

CO3. Design image retrieval systems or search engines using image indexing and visual words as image representations.

CO4. Apply various image clustering, classifications and segmentation methods for solving real world problems such as OCR, face recognition, and image understanding.

CO5. Utilize advanced techniques and algorithms for processing video frames and provide solutions for object tracking problems.

CO6. Work independently to solve problems with effective communication.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	3	2	-	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	3	3	3	-	-	-	-
CO4	3	3	3	-	-	-	-
CO5	3	3	3	-	-	-	-
CO6	-	-	-	-	3	3	-
Course Correlation Mapping	3	3	3	-	3	3	-

Correlation Levels: **3: High;**   **2: Medium;1: Low**

COURSE CONTENT

Module 1:   IMAGE DESCRIPTORS AND IMAGE MAPPING   *(09 Periods)*

Image Descriptors – Harris Corner Detector, SIFT (Scale-Invariant Feature Transform), Matching Geotagged Images; Image Mapping - Homographies, Warping Images, Creating Panoramas

Module 2:   CAMERA MODELS AND MULTIPLE VIEW GEOMETRY   *(09 Periods)*

Camera Models - Pin-hole Camera Model, Camera Calibration, Pose Estimation, Augmented Reality; Multiple View Geometry - Epipolar Geometry, Computing with cameras and 3D structure, Multiple view reconstruction, Stereo Images.

Module 3   CLUSTERING AND IMAGE SEARCHING   *(09 Periods)*

Clustering Images - K-means clustering, Hierarchical clustering, Spectral clustering; Searching Images - Image Retrieval, Visual Words, Indexing Images, Ranking results using Geometry.

Module 4   CLASSIFICATION OF IMAGES   *(09 Periods)*

Classifying Image Content - K-Nearest Neighbors, Bayes Classifier, Support Vector Machines, Optical Character Recognition.

Module 5   IMAGE SEGMENTATION AND OBJECT TRACKING   *(09 Periods)*

Graph Cuts, Segmentation using Clustering, Variational Methods - Processing Video, Object Tracking- Optical Flow, Lucas-Kanade algorithm.

Total Periods: 45

EXPERIENTIAL LEARNING:

LIST OF EXERCISES:

1. Identify the internal corners of an image using the Harris corner detector algorithm.
2. Execute feature matching using Scale Invariant Feature Transform Technique.
3. Estimate the homographies between the images (using RANSAC) and then warp all images to a common image plane.
4. Compute disparity maps from a stereo image pair with normalized cross-correlation.
5. Reconstruct a 3D scene using two images taken with a camera with known calibration.
6. Place a computer graphics model on an object present in a scene using camera parameters computed from feature matches.
7. Solve the problem of recognizing numbers in images of printed sudokus.
8. Implement image segmentation using the normalized cuts algorithm.
9. Compute image gradients on the first normalized cuts eigenvectors. Combine these gradient images to detect image contours of objects.
10. Track an object present in the video frames using Lucas-Kanade Optical flow algorithm.
11. Perform image classification using Naive Bayes Classifier.
12. Cluster images using features extracted from PCA, and HoG using an unsupervised learning approach.

RESOURCES

TEXT BOOKS:

1. Jan Eric Solem, "Programming Computer Vision with Python, O'Reilly Media", 2012.

REFERENCE BOOKS:

1. Mark Nixon and Alberto S. Aquado, "Feature Extraction & Image Processing for Computer Vision‖", Academic Press, 3rd Edition,2012.
2. Simon J. D. Prince, "Computer Vision: Models, Learning, and Inference" Cambridge University Press, 2012.

SOFTWARE/TOOLS:

1. Python
2. OpenCV

VIDEO LECTURES:

1. https://www.youtube.com/watch?v=715uLCHt4jE

2. https://nptel.ac.in/courses/108103174

3. https://www.youtube.com/watch?v=Z_YNkw65gp8

4. https://www.youtube.com/playlist?list=PL7v9EfkjLswLfjcI-qia-Z-e3ntl9l6vp

WEB RESOURCES:

1. https://nptel.ac.in/courses/108103174

2. https://github.com/jbhuang0604/awesome-computer-vision#books

3. https://medium.com/readers-writers-digest/beginners-guide-to-computer-vision-23606224b720

4. https://doc.lagout.org/science/0_Computer%20Science/2_Algorithms/An%20Introduction%20to%203D%20Computer%20Vision%20Techniques%20and%20Algorithms%20%5BCyganek%20%26%20Siebert%202009-02-09%5D.pdf

5. https://www.udacity.com/course/introduction-to-computer-vision--ud810

PROGRAM ELECTIVE

Course Code	Course Title	L	T	P	S	C
22AI201002	**REINFORCEMENT LEARNING**	3	-	-	-	3

Pre-Requisite 22AI202002-Machine Learning

Anti-Requisite -

Co-Requisite -

COURSE DESCRIPTION: This course provides an overview on the concepts of Reinforcement learning, a very active research sub-field of Machine Learning. This course also deals with Tabular solution methods, Planning and Learning with Tabular Methods, Approximate solution methods, Applications and case studies related to reinforcement learning.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Demonstrate the basic concepts of Reinforcement learning and Multi-arm Bandits.

CO2. Formalize problems as Markov Decision Process and solve using dynamic programming.

CO3. Analyze Monte Carlo method, Temporal difference learning for prediction and control.

CO4. Utilize planning and control to perform Trajectory sampling and Heuristic search.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	3	2	2	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	2	2	3	-	-	-	-
CO4	3	2	3	-	-	-	-
Course Correlation Mapping	3	2	3	-	-	-	-

Correlation Levels: 3: High; 2: Medium;1: Low

COURSE CONTENT

Module 1:   INTRODUCTION TO REINFORCEMENT LEARNING AND   *(09 Periods)*
**   MULTI-ARM BANDITS**

Reinforcement Learning: Introduction, Elements of Reinforcement Learning, Limitations and Scope, Tic-Tac-Toe, History of Reinforcement Learning

Multi-Arm Bandits : An n-Armed Bandit Problem, Action-Value Methods, Incremental Implementation, Tracking a Nonstationary Problem, Optimistic Initial Values, Upper Confidence Bound Action Selection, Gradient Bandits, Associative Search (Contextual Bandits).

Module 2:   FINITE MARKOV DECISION PROCESS   *(09 Periods)*

The Agent– Environment Interface, Goals and Rewards, Returns and Episodes, UnifiedNotation for Episodic and Continuing Tasks, The Markov Property, Markov Decision Processes,Value Functions, Optimal Value Functions, Optimality, and Approximation.

Module 3   DYNAMIC PROGRAMMING   *(09 Periods)*

Policy Evaluation, Policy Improvement, Policy Iteration, Value Iteration, Asynchronous Dynamic Programming Generalized Policy Iteration, Efficiency of Dynamic Programming.

Module 4   MONTE CARLO METHODS AND TEMPORAL-DIFFERENCE   *(09 Periods)*
**   LEARNING**

Monte-Carlo Methods: Monte Carlo Prediction - Monte Carlo Estimation of Action Values, Monte Carlo Control, Monte Carlo Control without Exploring Starts, Off-policy Prediction via Importance Sampling, Incremental Implementation, Off-Policy Monte Carlo Control, Importance Sampling on Truncated Returns

Temporal-Difference Learning: TD Prediction - Advantages of TD Prediction Methods, Optimality on TD(0), Sarsa: On-Policy TD Control, Q-Learning: Off-PolicyTD Control.

Module 5   PLANNING AND LEARNING   *(09 Periods)*

Models and Planning, Integrating Planning, Acting, and Learning, Model Incorrectness, Prioritized Sweeping, Full vs. Sample Backups, Trajectory Sampling, Heuristic Search, Monte Carlo Tree Search.

                                                                        Total Periods: 45

EXPERIENTIAL LEARNING

The following is the sample. Faculty shall frame according to the course domain.

1.   Identify suitable deep reinforcement learning techniques that can be used to analyze the stock trading market, and provide proper investment reports.

2.   **Traffic Signal Control Using Reinforcement Learning:** The main challenges are scalability, signal coordination, data feasibility, etc. To address these challenges, design the RL agents utilizing the 'pressure' concept to achieve signal coordination at the region-level; and show that implicit coordination could be achieved by individual

control agents with a well-crafted reward design thus reducing the dimensionality.

3. **Applying Reinforcement Learning for Self Driving Cars :** Parking can be achieved by learning automatic parking policies. Lane changing can be achieved using Q-Learning while overtaking can be implemented by learning an overtaking policy while avoiding collision and maintaining a steady speed thereafter.

RESOURCES

TEXT BOOKS:

1. Richard S. Sutton and Andrew G. Barto, Reinforcement Learning an Introduction, The MIT Press Cambridge, Massachusetts London, England, 2nd edition 2018.

REFERENCE BOOKS:

1. Phil Winder,Reinforcement Learning Industrial Applications of Intelligent Agents, O'Reilly Media, Inc., First Edition , 2020.
2. Aske Plat , Learning to Play: Reinforcement Learning and Games, Springer,2020.
3. A, Taweh Beysolow , Applied Reinforcement Learning with Python: With OpenAI Gym, Tensorflow, and Keras , 2019.

VIDEO LECTURES:"

1. https://nptel.ac.in/courses/106/106/106106143/
2. https://onlinecourses.nptel.ac.in/noc19_cs55/preview
3. https://www.coursera.org/specializations/reinforcementlearning?utm_source=bg&utm_medium=sem&utm
4. https://www.youtube.com/watch?v=LzaWrmKL1Z4

WEB RESOURCES:

1. https://neptune.ai/blog/how-to-structure-organize-track-and-manage-reinforcement-learning-rl-projects
2. https://livetalent.org/elearning-course/reinforcement-learning/
3. https://www.guru99.com/reinforcement-learning-tutorial.html
4. https://enjeeneer.io/sutton_and_barto/rl_notes.pdf

PROGRAM ELECTIVE

Course Code	Course Title	L	T	P	S	C
22AI201003	**MACHINE LEARNING FOR SECURITY**	3	-	-	-	3

Pre-Requisite   - 22AI202002-Machine Learning

Anti-Requisite   -

Co-Requisite   -

COURSE DESCRIPTION: The course provides theoretical and conceptual knowledge about the application of machine learning principles in security using Python. Topics covered in the course will be user access control, user log, malware packets in communication, building / simulating firewalls and not limited to the above.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1.   Analyze methods of Security, Attack and Mitigation.

CO2.   Develop algorithms for adversary and vulnerability analysis.

CO3.   Apply algorithms for adversary and vulnerability analysis..

CO4.   Select and Apply methods of machine learning to estimate the vulnerability.

CO5.   Recognize / identify the potential of machine learning in security breach discovery.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	3	2	2	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	2	2	3	-	-	-	-
CO4	3	2	3	-	-	-	-
CO5	3	3	2	-	-	-	-
Course Correlation Mapping	3	2	3	-	-	-	-

Correlation Levels: 3: High; 2: Medium;1: Low

COURSE CONTENT

Module 1:   MACHINE LEARNING AND SECURITY   *(09 Periods)*

Introduction: Cyber Threat Landscape, Adversaries in Machine Learning, Machine Learning Problems and Approaches.

Algorithms and Considerations: Training Algorithms, Classification Algorithms, Practical Considerations in Classification, Clustering.

Module 2:   ANOMALIES   *(09 Periods)*

Detection: Anomaly Detection – Methods – Intrusion Detection – Host Intrusion Detection, Network Intrusion Detection, Web Application Intrusion Detection.

Algorithms and Considerations: Anomaly Detection with Data and Algorithms - Forecasting, Statistical Measures, Unsupervised Machine Learning Algorithms; Challenges of using Machine Learning in Anomaly Detection, Response and Mitigation, Practical System Design Concerns.

Module 3:   MALWARE   *(09 Periods)*

Malware Analysis: Definition, Background and Classification of Malwares – Feature Generation.

Feature Engineering: Feature Classification, Sampling and Labelling Malware*.*

Module 4   NETWORK TRAFFIC ANALYSIS   *(09 Periods)*

Network Defense: Machine Learning and Network Security, Intrusion Detection, Access Control and Authentication – Detecting "In-Network Attackers" – Data Centric Security, Honeypots, Botnets.

Algorithms and Considerations: Building a Predictive Model to Classify Network Attacks – Preparing, Exploring and Classification of Network Data.

Module 5   PROTECTION AND ADVERSARIES   *(09 Periods)*

Protection: Protecting Consumer Web – Types of Abuse - Supervised Learning for Abuse Problems - Collecting and Labelling Data.

Adversarial Machine Learning: Importance, Vulnerabilities, Model Poisoning, Evasion Attack.

Total Periods: 45

EXPERIENTIAL LEARNING
1. Building a Intelligent Firewall that employs methods of Machine Learning – Filters Packets and finds the source IP where the malware – packets are originated and blocks the group of IPs.
2. Build a tool in Python to identify and learn data types in malware packets. Which is an alarm-repository for the firewall that blocks the communication containing specific types of data.

RESOURCES

TEXT BOOKS:

1. Clarence Chio, David Freeman, Machine Learning and Security - Protecting Systems with Data and Algorithms, O'Reilly Media, 2018.

2. Emmanuel Tsukerman, Machine Learning for Cybersecurity Cookbook, Packt Publishing, 2019.

REFERENCE BOOKS:

1. Matthew Kirk, Thoughtful Machine Learning with Python - A TEST-DRIVEN APPROACH, O'Reilly Media, Inc., 2017.

2. TJO Connor, Violent Python, A Cookbook for Hackers, Forensic Analysts, Penetration Testers and Security Engineers, Syngress imprint of Elsevier, 2013.

3. Benjamin Graham, Security Analysis: Sixth Edition, McGraw Hill Education, 2017.

VIDEO LECTURES:

1. https://www.youtube.com/channel/UC6qyO5J_rObFJewJWs89XRA/videos

2. https://www.youtube.com/watch?v=uPSgfNhd2qY

3. https://www.youtube.com/watch?v=3hig_oEz8Rg

4. https://www.youtube.com/watch?v=RnFGwxJwx-0

5. https://www.youtube.com/watch?v=gdizrw47kvs

WEB RESOURCES:

1. https://github.com/13o-bbr-bbq/machine_learning_security

2. https://github.com/PacktPublishing/Hands-on-Machine-Learning-for-Cyber-Security

3. http://www.secrepo.com/

4. https://www.cisco.com/c/en/us/products/security/machine-learning-security.html

5. https://datascience.foundation/datatalk/machine-learning-and-data-security

PROGRAM ELECTIVE

Course Code	Course Title	L	T	P	S	C
22AI201004	**PREDICTIVE ANALYTICS**	3	-	-	-	3

Pre-Requisite 22AI202002-Machine Learning

Anti-Requisite -

Co-Requisite -

COURSE DESCRIPTION: This course provides a detailed discussion on the concepts of predictive analytics and modelling with few case studies such as Recovering Lapsed Donors Fraud Detection and help desk operations. This course also focuses on the importance of Data Understanding, Data Preparation, Association Rules, Descriptive Modelling, Model assessment and deployment in decision making.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Demonstrate knowledge on the basics of Predictive Analytics and Modelling to predict future outcomes.

CO2. Analyse the insights of data using data visualization and preparation methods for improved decision making.

CO3. Identify patterns in categorical data and build descriptive models using Association rules, Principal Component Analysis and Clustering algorithms.

CO4. Assess and Deploy Predictive Models using batch approach , Regression and Ensemble Models to improve corporate operations.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	3	2	2	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	2	2	3	-	-	-	-
CO4	3	2	3	-	-	-	-
Course Correlation Mapping	3	2	3	-	-	-	-

Correlation Levels: 3: High;   2: Medium;   1: Low

COURSE CONTENT

Module 1: INTRODUCTION *(09 Periods)*

Overview of Predictive Analytics: Analytics, Predictive Analytics, Business Intelligence, Predictive Analytics vs. Statistics, Predictive Analytics vs. Data Mining, Challenges in Using Predictive Analytics, Educational Background. Setting up the Problem: Predictive Analytics Processing Steps, Defining Data for Predictive Modelling, Defining the Target Variable, Defining Measures of Success for Predictive Models, Doing Predictive Modelling Out of Order, and Case Study: Recovering Lapsed Donors and Fraud Detection.

Module 2: DATA UNDERSTANDING *(08 Periods)*

What the Data Looks Like, Single Variable Summaries, Data Visualization in One Dimension, Histograms, Multiple Variable Summaries, Data Visualization, Two or Higher Dimensions, The Value of Statistical Significance, Pulling It All Together into a Data Audit.

Module 3 DATA PREPARATION *(08 Periods)*

Variable Cleaning:Incorrect Values, Consistency in Data Formats, Outliers, Multidimensional Outliers, Missing Values, Fixing Missing Data. Feature Creation: Simple Variable Transformations, Fixing Skew, Binning Continuous Variables, Numeric Variable Scaling, Nominal Variable Transformation, Ordinal Variable Transformations, Date and Time Variable Features, ZIP Code Features, Multidimensional Features, Variable Selection Prior to Modelling, Sampling.

Module 4 ASSOCIATION RULES AND DESCRIPTIVE MODELING *(10 periods)*

Item sets and Association Rules: Terminology, Condition, Left-Hand-Side, Antecedent, Right-Hand-Side, Consequent, Output, Conclusion, Rule (Item Set), Support, Antecedent Support, Confidence, Accuracy, Lift. Parameter Settings, How the Data Is Organized, Measures of Interesting Rules, Deploying Association Rules, Problems with Association Rules, Building Classification Rules from Association Rules.

Descriptive Modelling: Data Preparation Issues with Descriptive Modelling, Principal Component Analysis, Clustering Algorithms.

Module 5 PREDICTIVE MODELING *(10 Periods)*

Predictive Modelling and Assessing Predictive Models: Batch Approach to Model Assessment, Assessing Regression Models. Model Ensembles: Motivation for Ensembles, Bagging, Boosting, Improvements to Bagging and Boosting, Model Ensembles and Occam's razor, Interpreting Model Ensembles. Model Deployment, Help Desk Case Study.

Total Periods: 45

EXPERIENTIAL LEARNING

The following is the sample. Faculty shall frame according to the course domain.
1. **Diabetic Prediction:**

   The National Institute of Diabetes and Digestive and Kidney Diseases has a created a dataset. The objective of the dataset is to diagnostically predict whether or not a patient has diabetes, based on certain diagnostic measurements included in the dataset. Several constraints were placed on the selection of these instances from a

larger database. In particular, all patients here are females at least 21 years old of Pima Indian heritage. The datasets consists of several medical predictor variables and one target variable, Outcome. Predictor variables includes the number of pregnancies the patient has had, their BMI, insulin level, age, and so on. Build a machine learning model to accurately predict whether or not the patients in the dataset have diabetes or not?

2. **Heart Disease Prediction:**

The EHDPS (Effective Heart Disease Prediction System) predicts the likelihood of patients getting heart disease. It enables significant knowledge, eg, relationships between medical factors related to heart disease and patterns, to be established. Predict the chances of having a heart stroke from heart disease dataset available online in kaggle.

3. **Industry Use Cases:**

Many industries use predictive analytics as a core part of their strategy. Modern tools make it increasingly easy to assemble the data and make improvements in the processes. Considering these aspects analyse the importance of using predictive analytics in an organization for improved decision making through the industry use cases such as Fundraising, Real Estate, Healthcare and Wellness, Software Testing, Supply Chain Management, Food Delivery, Insurance and Higher Education.

RESOURCES

TEXT BOOKS:

1. Dean Abbott - Applied Predictive Analytics_ Principles and Techniques for the Professional
   Data Analyst-Wiley,2014.

REFERENCE BOOKS:

1. Siegel, Eric, Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die, Wiley, 2016.
2. Theobald, Oliver, Data Analytics for Absolute Beginners Cengage Learning, 2nd Edition, 2019.
3. Bari, A., Chaouchi, M., Jung, T. Analytics for Dummies, 2nd Edition, 2016.

VIDEO LECTURES:

1. https://onlinecourses.nptel.ac.in/noc21_mg86
2. http://www.nitttrc.edu.in/nptel/courses/video/110104086/L08.html
3. https://nptel.ac.in/courses/110104086
4. https://onlinecourses.swayam2.ac.in/imb20_mg19
5. https://nptel.ac.in/courses/110105089

WEB RESOURCES:

1. https://www.futurelearn.com/courses/predictive-analytics
2. https://www.mygreatlearning.com/academy/learn-for-free/courses/predictive-modeling-and-analytics-regression
3. https://ecm.elearningcurve.com/Fundamentals_of_Predictive_Analytics_Class_p
4. https://www.edx.org/learn/predictive-analytics
5. https://www.kaggle.com/code/aks2411/auto-mpg-prediction/log
6. https://www.ibm.com/in-en/analytics/predictive-analytics

PROGRAM ELECTIVE

Course Code	Course Title	L	T	P	S	C
22AI201005	**TEXT ANALYTICS**	3	-	-	-	3

Pre-Requisite	22AI202002-Machine Learning
Anti-Requisite	-
Co-Requisite	-

COURSE DESCRIPTION: Text Analytics are the methods and techniques used to extract useful knowledge from text to support decision making. This field includes a collection of research from the natural language processing, databases, data mining, and machine learning communities. The aim of this course is to be a primer for text analytics theory. Topics discussed in the course include: Text tokenization, Text normalization, Text syntax and structure, Text classification, Feature extraction, Classification algorithms, Text summarization, Keyphrase extraction, Topic modeling, Text similarity and clustering, Document clustering, Automated document summarization, Semantic analysis, Sentiment analysis.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Demonstrate knowledge on preprocessing techniques for text data.

CO2. Apply various computational, language processing, machine learning techniques to classify and cluster text.

CO3. Build manual and automated text summarizers for getting more insights from the given text data.

CO4. Apply various distance or similarity measures to estimate the degree of similarity between two text documents.

CO5. Design machine learning models using supervised learning approaches and perform sentiment analysis.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	3	2	2	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	2	2	3	-	-	-	-
CO4	3	2	3	-	-	-	-
CO5	3	3	2	-	-	-	-
Course Correlation Mapping	3	2	3	-	-	-	-

Correlation Levels: 3: High; 2: Medium;1: Low

COURSE CONTENT

Module 1:   TEXT PROCESSING   *(10 Periods)*

Text tokenization – Sentence tokenization, Word tokenization; Text normalization –Cleaning, Tokenizing, Removing special characters, Expanding contractions, Case conversions, Removing stopwords, Correcting words, Stemming, Lemmatization; Text syntax and structure– Parts of Speech tagging, Text parsing.

Module 2:   TEXT CLASSIFICATION   *(09 Periods)*

Text classification, Automated text classification, Blueprint of text classification, Feature extraction – Bag of Words model, TF-IDF model, Averaged word vectors, TF-IDF weighted averaged word vectors; Classification algorithms – Naïve Bayes, Support vector machines, Evaluating classification models, Building a multi-class classification system.

Module 3:   TEXT SUMMARIZATION   *(08 Periods)*

Text summarization and information extraction, Singular value decomposition, Keyphrase extraction – Weighted tag–based phrase extraction; Topic modeling – Latent semantic indexing, Latent Dirichlet allocation, Non-negative matrix factorization.

Module 4:   TEXT SIMILARITY AND CLUSTERING   *(09 Periods)*

Text Similarity, Analyzing Term Similarity, Cosine Distance and Similarity, Analyzing Document Similarity, Hellinger-Bhattacharya Distance, Okapi BM25 Ranking, Document Clustering-K Means, Affinity Propagation, Ward's Agglomerative Hierarchical Clustering.

Module 5:   AUTOMATED DOCUMENT SUMMARIZATION, SEMANTI *(09 Periods)*
ANALYSIS

Extraction-based techniques, Abstraction-based techniques, Latent semantic analysis, Text rank, Semantic analysis - Word sense disambiguation, Named entity recognition; Sentiment analysis of movie reviews – Getting and formatting data, Text normalization, Feature extraction, Model performance evaluation, Preparing datasets, Supervised machine learning technique.

Total Periods: 45

EXPERIENTIAL LEARNING

The following is the sample. Faculty shall frame according to the course domain.

1. Perform clustering of popular movies based on their IMDb synopses as raw data using K-Means clustering- Students will find out similarities within groups of people in order to build a movie recommending system for users and they are going to analyze a dataset from Netflix database to explore the characteristics that people share in movies' taste, based on how they rate them.

2. Classify the given texts using Naïve Bayes algorithm and evaluate the classifier performance- Students can Understand and learn how to implement the Naive Bayes Classifier in R and Python.

3. Keep Track of Marketing Trends: Trend analysis is an important aspect of marketing our product or service. We use Google Trends and other tools to look for trending terms. If we don't have the correct instrument to identify patterns, we can miss them. We may discover some hidden gems, such as specialized niche keywords, by doing a keyword analysis on feedback from your target audience across various social media platforms and review sites, and begin taking the necessary steps to be at the forefront of your industry. Use text summarize key phrases extraction approach to analyse trending marketing.

RESOURCES

TEXT BOOKS:

1. Dipanjan Sarkar, Text Analytics with Python, APRESS, 2016.

REFERENCE BOOKS:

1. Steven Bird, Ewan Klein and Edward Loper, Natural Language Processing with Python, O'Reilly, 2018.
2. Charu C. Aggarwal, Machine Learning for Text, Springer, 2018.
3. Benjamin Bengfort , Rebecca Bilbro , Tony Ojeda, Applied Text Analysis with Python, O'Reilly, 2018..

VIDEO LECTURES:

1. http://nptel.ac.in/courses/110107129
2. https://freevideolectures.com/course/4135/nptel-business-analytics-text-mining-modeling-using-python/38
3. https://freevideolectures.com/course/4135/nptel-business-analytics-text-mining-modeling-using-python/39
4. https://freevideolectures.com/course/4135/nptel-business-analytics-text-mining-modeling-using-python/40
5. https://www.youtube.com/watch?v=FtGBzBi51mQ

WEB RESOURCES:

1. https://www.lexalytics.com/technology/text-analytics
2. https://download.e-bookshelf.de/download/0008/3870/64/L-G-0008387064-0017200370.pdf-
3. https://www.linguamatics.com/what-text-mining-text-analytics-and-natural-language-processing
4. https://www.coursera.org/learn/text-mining
5. https://onlinecourses.nptel.ac.in/noc19_mg47/preview

PROGRAM ELECTIVE

Course Code	Course Title	L	T	P	S	C
22AI202008	**DATA ENGINEERING**	3	-	3	-	4.5

Pre-Requisite 22AI202003-Big Data Frameworks

Anti-Requisite -

Co-Requisite -

COURSE DESCRIPTION: This course provides a detailed discussion and hands-on experience on Modern data Engineering tools, data pipelines, data collection stage, data curation stage, data aggregation stage, data engineering challenges and effective deployment strategies. The course focuses on building databases and processing systems.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Analyze Modern data engineering and tools, storage and compute data lakes to store and retrieve data efficiently on Cloud.

CO2. Create and run data pipelines for organizing large volumes of data on Cloud.

CO3. Perform data collection, data curation, data aggregation and demonstrate their use in storage systems on Cloud.

CO4. Deploy and monitor pipelines in production on Cloud and solve data engineering challenges .

CO5. Work independently to solve problems with effective communication.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	3	2	-	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	3	3	3	-	-	-	-
CO4	3	3	3	-	-	-	-
CO5	-	-	-	-	3	3	-
Course Correlation Mapping	3	3	3	-	3	3	-

Correlation Levels: **3: High;**   **2: Medium;1: Low**

COURSE CONTENT

Module 1:   MODERN DATA ENGINEERING AND TOOLS   *(10 Periods)*

Data Engineering and Analytics: Core capabilities of storage and compute resources, Availability of varying datasets, Computing Adoption of cloud computing, Data storytelling, The monetary power of data.

Discovering Storage and Compute Data Lakes: Exploring the benefits of data lakes, Adhering to compliance frameworks, Segregating storage and compute in a data lake, Discovering data lake Architectures, The CAP theorem.

Module 2:   DATA PIPELINES AND DATA COLLECTION STAGE   *(08 Periods)*

Understanding Data Pipelines: Components of a data pipeline, Process of creating a data pipeline, Running a data pipeline

Data Collection Stage: Architecting the Electroniz data lake, Understanding the bronze layer, Configuring data sources, Configuring data destinations, Building the ingestion pipelines.

Module 3   DATA CURATION STAGE   *(09 Periods)*

Understanding Delta Lake: Understanding Delta Lake, Creating a Delta Lake table, Changing data in an existing Delta Lake table, Performing time travel, Performing upserts of data, Understanding isolation levels, Understanding concurrency control.

Data Curation Stage: The need for curating raw data, The process of curating raw data, Developing a data curation pipeline, Running the pipeline for the silver layer, Verifying curated data in the silver layer.

Module 4   DATA AGGREGATION STAGE   *(09 Periods)*

The need to aggregate data, The process of aggregating data, Developing a data aggregation pipeline, Running the aggregation pipeline, Understanding data consumption, Verifying aggregated data in the gold layer.

Module 5   DATA ENGINEERING CHALLENGES AND EFFECTIVE   *(09 Periods)*
                DEPLOYMENT STRATEGIES

Deploying and Monitoring Pipelines in Production: The deployment strategy, Developing the master pipeline, Testing the master pipeline, Scheduling the master pipeline, Monitoring pipelines.

Solving Data Engineering Challenges: Schema evolution, Sharing data, Datagovernance.

Total Periods: 45

EXPERIENTIAL LEARNING

LIST OF EXERCISES:

1. Use an Azure Databricks workspace and perform Structured Streaming with batch jobs by using Delta Lake. You need to complete the exercise within a Databricks Notebook. To begin, you need to have access to an Azure Databricks workspace.

2. Demonstrates the experience of working with Apache Spark in Azure Synapse Analytics. Learn how to use libraries like Hyperspace and MSSparkUtil to optimize the experience of working with Data Lake storage accounts from Spark notebooks.

3. a) Query Parquet data with serverless SQL pools
   b) Create external tables for Parquet and CSV files
   c) Create views with serverless SQL pools
   d) Secure access to data in a data lake when using serverless SQL pools
   e) Configure data lake security using Role-Based Access Control (RBAC) and Access Control Lists (ACLs)

4. a) Use some Databricks notebooks to learn fundamentals concepts and techniques for working with DataFrames.
   b) Build the Azure Databricks DataFrames concepts learned in the previous lab above by exploring some advanced methods data engineers can use to read, write, and transform data using DataFrames.

5. a) Perform Data Exploration in Synapse Studio
   b) Ingest data with Spark notebooks in Azure Synapse Analytics
   c) Transform data with DataFrames in Spark pools in Azure Synapse Analytics
   d) Integrate SQL and Spark pools in Azure Synapse Analytics

6. a) Perform petabyte-scale ingestion with Azure Synapse Pipelines
   b) Import data with PolyBase and COPY using T-SQL
   c) Use data loading best practices in Azure Synapse Analytics

7. a) Execute code-free transformations at scale with Azure Synapse Pipelines
   b) Create data pipeline to import poorly formatted CSV files
   c) Create Mapping Data Flows

8. Create linked services, and orchestrate data movement and transformation in Azure Synapse Pipelines.

9. a) Secure Azure Synapse Analytics supporting infrastructure
   b) Secure the Azure Synapse Analytics workspace and managed services
   c) Secure Azure Synapse Analytics workspace data

10. a) Configure Azure Synapse Link with Azure Cosmos DB
    b) Query Azure Cosmos DB with Apache Spark for Synapse Analytics
    c) Query Azure Cosmos DB with serverless SQL pool for Azure Synapse Analytics

11. a) Use Stream Analytics to process real-time data from Event Hubs
    b) Use Stream Analytics windowing functions to build aggregates and output to Synapse Analytics
    c) Scale the Azure Stream Analytics job to increase throughput through partitioning
    d) Repartition the stream input to optimize parallelization

12. a) Learn the key features and uses of Structured Streaming
    b) Stream data from a file and write it out to a distributed file system
    c) Use sliding windows to aggregate over chunks of data rather than all data
    d) Apply watermarking to remove stale data
    e) Connect to Event Hubs read and write streams

RESOURCES

TEXT BOOKS:

1. Manoj Kukreja, Danil Zburivsky, Data Engineering with Apache Spark, Delta Lake, and Lakehouse: Create Scalable Pipelines that Ingest, Curate, and Aggregate Complex Data in a Timely and Secure Way, Packt Publishing, 2021.

2. Crickard, Paul, Data Engineering with Python: Work with Massive Datasets to Design Data Models and Automate Data Pipelines Using Python, Packt Publishing, 2020.

REFERENCE BOOKS:

1. Eagar, G, Data Engineering with AWS: Learn how to Design and Build Cloud-based Data Transformation Pipelines Using AWS, Packt Publishing, 2021.
2. Housley, Matt, and Reis, Joe, Fundamentals of Data Engineering: Plan and Build Robust Data Systems, O'Reilly Media, Incorporated, 2022.

SOFTWARE/TOOLS:

1. Microsoft Azure, Windows PowerShell
2. Apache Spark, Delta Lake, JSON

VIDEO LECTURES:

1. http://nptel.ac.in/courses/ 106104189
2. http://nptel.ac.in/courses/ 106106179
3. https://www.youtube.com/watch?v=TScSdIJ-_Oo
4. https://www.youtube.com/watch?v=IPkQpBdde5Y&t=1s
5. https://www.youtube.com/watch?time_continue=22&v=cbobqI3ZGuA

WEB RESOURCES:

1. https://github.com/MicrosoftLearning/DP-203-Data-Engineer/tree/master/Instructions/Labs
2. https://docs.microsoft.com/en-us/azure/synapse-analytics/spark/apache-spark-overview
3. https://github.com/PacktPublishing/Data -Engineering-with-Apache-Spark-Delta-Lake-and-Lakehouse
4. https://github.com/PacktPublishing/Data-Engineering-with-Python
5. https://blog.claydesk.com/azure-data-lake-tutorial/

PROGRAM ELECTIVE

Course Code	Course Title	L	T	P	S	C
22AI201006	**DATA ANALYTICS**	3	-	-	-	3

Pre-Requisite -
Anti-Requisite -
Co-Requisite -

COURSE DESCRIPTION: The course provides Introduction to Data Analytics and its Life Cycle, Review of Basic Data Analytic Methods Using R, Advanced Analytical Theory and Methods, Advanced Analytics-Technology and Tools: In-Database Analytics and Communicating and Operationalizing an Analytics Project

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Demonstrate knowledge on the foundations of Analytics and Data Analytics Life Cycle.

CO2. Import, Export Data and perform descriptive statistics using R programming to make inferences.

CO3. Identify trends and patterns in data using Data Exploration and Statistical Methods.

CO4. Apply Advanced Analytical Methods to perform Text Analysis and Time–Series Analysis.

CO5. Develop and Operationalize Analytics project creating suitable product deliverables and portray data using visual representation.

CO6. Analyze and Design Data Analytics Applications to solve societal problems.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	3	2	-	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	3	3	3	-	-	-	-
CO4	3	3	3	-	-	-	-
CO5	3	3	3	-	-	-	-
CO6	3	3	2	-	-	-	-
Course Correlation Mapping	3	3	3	-	-	-	-

Correlation Levels: 3: High;   2: Medium;   1: Low

COURSE CONTENT

Module 1:   INTRODUCTION TO DATA ANALYTICS and R   *(09 Periods)*

Practice in Analytics: BI versus Data Science, Current Analytical Architecture, Emerging Big Data Ecosystem and a New Approach to Analytics.

Data Analytics Life Cycle: Key Roles for a Successful Analytics Project Background and Overview of Data Analytics Lifecycle Phases - Discovery Phase, Data Preparation Phase, Model Planning, Model Building, Communicate Results, Operationalize.

Introduction to R: R Graphical User Interfaces, Data Import and Export, Attribute and Data Types, Descriptive Statistics.

Module 2:   BASIC DATA ANALYTICAL METHODS   *(09 Periods)*

Exploratory Data Analysis: Visualization Before Analysis, Dirty Data, Visualizing a Single Variable, Examining Multiple Variables, Data Exploration Versus Presentation.

Statistical Methods for Evaluation: Hypothesis Testing, Difference of Means, Wilcoxon Rank-Sum Test, Type I and Type II Errors, Power and Sample Size, ANOVA, Decision Trees in R, Naïve Bayes in R.

Module 3:   ADVANCED ANALYTICAL TECHNOLOGY AND METHODS   *(09 Periods)*

Time Series Analysis: Overview of Time Series Analysis, Box-Jenkins Methodology, ARIMA Model, Autocorrelation Function (ACF),Autoregressive Models, Moving Average Models , ARMA and ARIMA Models ,Building and Evaluating an ARIMA Model, Reasons to Choose and Cautions.

Text Analysis: Text Analysis Steps, A Text Analysis Example, Collecting Raw Text, Representing Text, Term Frequency—Inverse Document Frequency (TFIDF), Categorizing Documents by Topics, Determining Sentiments, Gaining Insights.

Module 4:   FINAL DELIVERABLES AND VISUALIZATION   *(09 Periods)*

Communicating and Operationalizing an Analytics Project, Creating the Final Deliverables: Developing Core Material for Multiple Audiences, Project Goals, Main Findings, Approach, Model Description, Key Points Supported with Data, Model Details Recommendations, Additional Tips on Final Presentation, Providing Technical Specifications and Code, Data Visualization.

Module 5:   DATA ANALYTICS APPLICATIONS   *(09 Periods)*

Text and Web: Data Acquisition, Feature Extraction, Tokenization, Stemming, Conversion to Structured Data, Sentiment Analysis, Web Mining.

Recommender Systems: Feedback, Recommendation Tasks, Recommendation Techniques, Final Remarks.

Social Network Analysis: Representing Social Networks, Basic Properties of Nodes, Basic and Structural Properties of Networks.

Total Periods: 45

EXPERIENTIAL LEARNING

The following is the sample. Faculty shall frame according to the course domain.

1. **Sentiment Analyzer of Social Media:**
   This is one of the interesting and innovative machine learning projects. As, social media like Facebook, Twitter, and YouTube is the ocean of big data. Therefore, mining these data can be beneficial in a number of ways to understand user sentiments and opinions. Perform sentiment analysis considering   social media such as Twitter, facebook or

Amazon product reviews or movie reviews to understand the opinion or reaction for a product or service of a customer.

2. **Music Recommendation System:**

Are you a lover of music? Always love to listen to your favorite one? Then, you will be glad to know about this interesting machine learning project idea. This can also be an innovative project. The goal of this project is to recommend music based on user listening history.

3. **Human Activity Recognition System:**

The human activity recognition system is a classifier model that can identify human fitness activities. To develop this project, you have to use a smart phone dataset, which contains the fitness activity of 30 people which is captured through smart phones. This project will help you to understand the solving procedure of the multi-classification problem.

4. **Uber Data Analysis:**

Perform data visualization on the uber data. The dataset contains 4.5 million of uber pickups in the new york city. This much data needs to be represented beautifully in order to analyze the rides so that further improvements in the business can be made.

RESOURCES

TEXT BOOKS:

1. EMC Education Services, Data Science and Big Data Analytics – Discovering, Analyzing, Visualizing and Presenting Data, John Wiley and Sons, 2015.

2. Joao Moreira, Andre Carvalho, André Carlos Ponce de Leon Ferreira Carvalho, Tomás Horvath - A General Introduction to Data Analytics - John Wiley and Sons, 1st Edition,2019·

REFERENCE BOOKS:

1. Anil Maheshwari, Data Analytics Made Accessible, Lake Union Publishing, 1st Edition, 2017

2. Richard Dorsey, Data Analytics: Become a Master in Data Analytics, Create Space Independent Publishing Platform, 2017.

VIDEO LECTURES:

1. https://nptel.ac.in/courses/106/107/106107220/

2. https://nptel.ac.in/courses/106/104/106104021/

3. https://nptel.ac.in/courses/106/106/106106093/

4. https://nptel.ac.in/courses/106/106/106106095/

5. https://nptel.ac.in/courses/106/104/106104135/

WEB RESOURCES:

1. https://extension.harvard.edu/academics/programs/data-science-graduate-certificate/

2. https://www.simplilearn.com/pgp-data-science-certification-bootcamp-program

3. https://ischoolonline.berkeley.edu/data-science/

4. https://developers.google.com/machine-learning/crash-course/

5. https://www.mastersindatascience.org/learning/what-is-data-analytics/

PROGRAM ELECTIVE

Course Code	Course Title	L	T	P	S	C
22AI202009	**DATA VISUALIZATION**	3	-	3	-	4.5

Pre-Requisite -
Anti-Requisite -
Co-Requisite -

COURSE DESCRIPTION: This course provides a detailed discussion and hands-on experienceon Introduction to data visualization, Data exploration and plots, Numpy, Pandas, Introductionto Matplotlib, Simplifying visualizations using Seaborn and Geoplotlib, Making things interactive with Plotly, Folium and Bokeh.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Demonstrate knowledge on data visualization methods to identify suitable factors for the adoption of data visualization.

CO2. Apply data exploration using Numpy, Pandas and Plots to perform data presentation.

CO3. Design layouts using Matplotlib forT image presentation of data.

CO4. Construct effective data visualizations using Seaborn, Geoplotlib to perform statistical graphics.

CO5. Develop interactive applications using Plotly, Folium, Bokeh to explore diverse data like COVID-19 dataset analysis, Newyork City, Stock Price and Airbnb.

CO6. Work independently to solve problems with effective communication.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	3	2	-	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	3	3	3	-	-	-	-
CO4	3	3	3	-	-	-	-
CO5	3	2	3	-	-	-	-
CO6	-	-	-	-	3	3	-
Course Correlation Mapping	3	3	3	-	3	3	-

Correlation Levels: 3: High;   2: Medium; 1: Low

COURSE CONTENT

Module 1: DATA VISUALIZATION, DATA EXPLORATION AND PLOTS   *(09 Periods)*

Data Visualization: What is data visualization, Key elements, Importance, The power of visual storytelling, Examples, Benefits, Types of charts and graphs used in data visualization.
Data Exploration: NumPy, NumPy operations, Pandas, Advantages of Pandas over NumPy, Disadvantages of Pandas, Operations of Pandas, Series.
Plots: Comparison plots, Relation plots, Composition plots, Distribution plots, Geo plots, Aspects of good visualization.

Module 2: VISUALIZATION USING MATPLOTLIB   *(09 Periods)*

Overview of plots in Matplotlib, Pyplot basics, Basic text and legend functions, Basic plots – Bar chart, Pie chart, Stacked bar chart, Stacked area chart, Histogram, Box plot, Scatter plot, Bubble plot; Layouts, Images – Basic image operations, Plotting multiple images in a grid.

Module 3: VISUALIZATION USING SEABORN AND GEOPLOTLIB   *(09 Periods)*

Seaborn: Advantages of Seaborn, Controlling figure aesthetics, Color palettes, Bar plots, Kernel density estimation, Plotting bivariate distributions, Visualizing pairwise relationships, Violin plots, Multi-plots in Seaborn, Regression plots, Squarify.
Plotting Geospatial Data: Design principles of Geoplotlib, Geospatial visualizations, Tile providers, Custom layers.

Module 4: VISUALIZATION USING PLOTLY, FOLIUM   *(09 Periods)*

Types of visualizations using Plotly, Plotly chart types, Bar charts, 3D scatter chart, 3D line chart, Plotly for COVID-19 dataset analysis, Plotly animation, Scatter matrix, Treemap, Geographic visualizations using Folium.

Module 5: VISUALIZATION USING BOKEH   *(09 Periods)*

Bokeh: Concepts of Bokeh, Interfaces in Bokeh, Bokeh server, Presentation, Integrating, Adding widgets.
Case Studies: Implementing Matplotlib and Seaborn on New York city database, Visualizing Bokeh stock prices, Analyzing Airbnb data with geoplotlib.

Total Periods: 45

EXPERIENTIAL LEARNING

LIST OF EXERCISES:

1. **Data Exploration using NumPy and Pandas:**
   a) Compute arithmetic mean, median, standard deviation and variance using statistical functions in NumPy.
   b) Perform operations indexing, slicing, splitting, iterating, computing mean of elements by loading any online dataset in delimited format (CSV) using genfromtxt method in NumPy.
   c) Compute mean of elements for each column and for complete dataset by loading any online dataset in delimited format (CSV) using read_csv method in Pandas.
   d) Perform operations filtering, sorting, combining and reshaping with any online dataset using NumPy and Pandas.

2. **Visualization basics with Matplotlib:**
   a) Plot the data pairs (x, y) as circles connected with line segments (1,1), (2,3), (4,4). Visualize and save the plot using plt.savefig() method.
   b) Insert axis labels, title, text, annotations and legends for the created plot.
   c) Create a bar plot setting the title, axis labels and legends loading any online dataset.
   d) Create a pie chart to visualize water usage with title and labels. Consider online water usage dataset for processing.

3. **Data Visualization using Matplotlib:**
   a) Create stacked bar Plot and stacked area chart setting the title, axis labels and legends considering any online dataset for processing.
   b) Use histogram and box plot with title and labels to visualize the intelligence quotient by considering online IQ Scores dataset for processing.
   c) Create scatter plot and examine the relationship between variables in the data set. Consider any online dataset for processing.

4. **Image Operations using Matplotlib:**
   a) Create a scatter plot with marginal histograms including labels and title. consider any online dataset for processing.
   b) Import necessary modules and load sample images. Visualize the images in a 2x2 grid and represent the image in array. Remove the axes and give each image a label.

5. **Simplified Visualization using Seaborn:**
   a) Create box plot and violin plot using Seaborn to compare intelligence quotient (IQ) Scores for different test groups by considering online IQ Scores dataset for processing.
   b) Use bar plot to compare the movie scores of five different movies. Consider online movie scores dataset for processing.
   c) Implement linear regression by visualizing linear relationship between maximum longevity and body mass in the regression plot using the regplot() function in Seaborn. Consider online animal dataset for processing.

6. **Plotting Geospatial Data using Seaborn:**
   a) Plot the geospatial data on a map and find the densely populated areas for cities in Europe that have a population of more than 100k, Display using dot-density visualization. Consider online world cities population dataset for processing.
   b) Create a custom layer that will allow to display geospatial data and animate the data points over time.
   c) Create a choropleth plot with GeoJSON Data.

7. **Data Visualization using Plotly:**
   a) Perform COVID-19 dataset analysis using Plotly.

8. **Data Visualization using Folium:**
   a) Create a New Delhi map using featuregroup, layer control, and marker configuration in Folium to visualize key attractions in New Delhi, India and around.
   b) Create a Folium heatmap with time animation.

9. **Interactive Visualization with Bokeh:**
   a) Import necessary modules and load word population dataset using pandas. Extract the required data and create a plot using Bokeh figure method with labels and title.
   b) Use grid plot to display the country plots next to each other and arrange the visualizations vertically.
   c) Create line plots to compare the global mean population density with that of Japan.
   d) Create models interface-based plot to display the lines and axes.
   e) Create a basic visualization with stock price dataset. Include interactive visualization using widgets to dynamically change the stock that is displayed in graph.

10. **Data Visualization on New York City Database:**
    a) Import necessary modules and load datasets "New York Population Records" and "New York Housing Unit Records". The first dataset contains information about the New York population and the second dataset contains information about housing units. Understand the structure of the dataset by looking at the features.
    b) Perform data wrangling for median housing income, wage by gender for different occupation categories, wage distribution.
    c) Create a plot containing multiple subplots that visualize information with regard to New York City wages.
        i. Visualize the median household income for the US, New York, New York City, and its districts.
        ii. Visualize the average wage by gender for the given occupation categories for the population of New York City.
        iii. Visualize the wage distribution for New York and New York City. Use the yearly wage intervals: 10k steps between 0 and 100k, 50k steps between 100k and 200k, and >200k.
    d) Use a tree map to visualize the percentage for the given occupation subcategories and for the population of New York City.

11. **Interactive Visualization of Stock Prices using Bokeh:**
    a) Import necessary modules and load stock prices data. Understand the structure of the dataset by looking at the features.
    b) Create a new column that holds the formatted short version of the date value and update the dataset.
    c) Create an interactive visualization that displays a Candlestick Plot to handle stock price data.
    d) Compare two stocks to each other by selecting them from dropdowns. Use RangeSlider to restrict the displayed date range in the requested year.
    e) Display the volume of the selected stock based on the selection of visualization either Candlestick or Line Plot.

12. **Data Analysis on Airbnb data using Geoplotlib:**
    a) Import necessary modules and load Airbnb data. Understand the structure of the dataset by looking at the features.
    b) Create a simple dot map from the points to get a glance of the data.
    c) Create a custom layer and plot each point on the map with a color that is defined by the currently selected attribute, either price or rating
    d) Display the hotspots for the most expensive and best-rated accommodation across New York

RESOURCES

TEXT BOOKS:

1. Mario Dobler, Tim Grobmann, *Data Visualization with Python*, Packt Publishing, 2019.

2. Kalilur Rahman, *Python Data Visualization Essentials Guide*, bpb Publications, 2021.

REFERENCE BOOKS:

1. Abha Belorkar, Sharath Chandra Guntuku, Shubhangi Hora, Anshu Kumar, *Interactive Data Visualization with Python*, 2nd Edition, Packt Publishing, 2020.

2. Claus O.Wilke, *Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures*, O'Reilly, 2019.

3. Danyel Fisher, Miriah Meyer, *Making Data Visual: A Practical Guide to Using Visualization for Insight,* O'Reilly, 2017.

SOFTWARE/TOOLS:

1. Python 3.9, Python Libraries – NumPy, Pandas, Matplotlib, Seaborn, Plotly, Folium, Bokeh.

2. IDLE/Jupyter Notebook/JupyterLab/Pycharm/Google Colab.

VIDEO LECTURES:

1. https://youtu.be/eFByJkA3ti4

2. https://youtu.be/JhK2qVi5dC4

3. https://www.youtube.com/watch?v=qdnM8Fpvdqc

4. https://www.youtube.com/watch?v=MUP2m46uw8I

5. https://www.youtube.com/watch?v=7kPqESo1vRw

6. https://www.youtube.com/watch?v=Nt84_TzRkbo

WEB RESOURCES:

1. https://plotly.com/python/

2. https://datavizcatalogue.com/

3. https://www.tableau.com/learn/articles/data-visualization

4. https://www.analyticsvidhya.com/blog/2021/10/interactive -plots-in-python-with-plotly-a-complete-guide/

5. https://www.digitalvidya.com/blog/python-visualization/

6. https://d3js.org/

PROGRAM ELECTIVE

Course Code	Course Title	L	T	P	S	C
22AI201008	**STREAM PROCESSING AND ANALYTICS**	3	-	-	-	3

Pre-Requisite   22AI202003-Big Data Frameworks

Anti-Requisite   -

Co-Requisite   -

COURSE DESCRIPTION: This course covers the fundamental concepts of stream processing the general blueprints of architectures that implements streaming with spark. This course also focuses on structured streaming , its programming model and in similar organization spark streaming , creation of streaming applications, monitoring and performance tuning.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Demonstrate knowledge on the fundamental concepts of stream processing model and architectures to enable efficient data streaming.

CO2. Create custom sinks to write data to systems not supported by the default implementations.

CO3. Process data in a scalable and resource constrained way using time-based stream processing.

CO4. Apply DStreams, execution model in combination with a functional programming model to develop and execute stream applications.

CO5. Investigate several performance tuning strategies and apply an optimal method to tune the overall performance of a job.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	3	2	2	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	2	2	3	-	-	-	-
CO4	3	2	3	-	-	-	-
CO5	3	3	2	-	-	-	-
Course Correlation Mapping	3	2	3	-	-	-	-

Correlation Levels: 3: High; 2: Medium;1: Low

COURSE CONTENT

Module 1:   STREAM PROCESSING MODEL AND ARCHITECTURES   *(09 Periods)*

Stream Processing Model: Introducing Stream Processing-What is Stream processing? Examples of Stream processing, Scaling Up data processing, Distributed Stream processing, Apache Spark; Sources and Sinks, Transformations and Aggregations, Window Aggregations, Stateless and Stateful Processing, Stateful streams, Example:Local stateful computation in Scala, The effect of time.

Streaming Architectures: Components of a data platform, Architectural models, Use of batch processing component in Streaming application, Lambda Architecture, Kappa Architecture, Streaming Algorithms, Batch Algorithms.

Module 2:   STRUCTURED STREAMING   *(09 Periods)*

Steps with Structured Streaming, Batch Analytics, Streaming, Analytics, Initializing Spark, Sources: Acquiring Streaming Data, Transforming Streaming Data, Sinks: Output the Resulting Data, Structured Streaming in Action, Structured Streaming sources-Understanding sources, File Source, Kafka Source, Socket Source.

Module 3   SINKS AND EVENT TIME-BASED STREAM PROCESSING   *(09 Periods)*

Structured Streaming Sinks: Understanding Sinks, SinkAPI, File Sink, Kafka Sink, Memory Sink, Console Sink, Foreach Sink.

Event Time-Based Stream Processing: Understanding Event Time in Structured Streaming, Time-Based Window Aggregations; Monitoring Structured Streaming Applications,Continuous Processing.

Module 4   SPARK STREAMING   *(09 Periods)*

Need for Spark Streaming, The DStream Abstraction, The Structure of a Spark Streaming Application, Simple Streaming Application using DStream,The Spark Streaming Programming Model, The Spark Streaming Execution Model- The Bulk-Synchronous Architecture, The Receiver Model; Spark Streaming Sources, Spark Streaming Sinks, Time-Based Stream Processing, Working with Spark SQL- Accessing Spark SQL Functions from Spark Streaming, Dealing with Data at Rest, Join Optimizations, Updating Reference Datasets in a Streaming Application.

Module 5   MONITORING AND PERFORMANCE TUNING   *(09 Periods)*

Monitoring Spark Streaming: The Streaming UI, Understanding Job Performance Using the Streaming UI, The Monitoring REST API, The Metrics Subsystem, The Internal Event Bus, Interacting with the Event Bus.

Performance Tuning: The Performance Balance of Spark Streaming, External Factors that Influence the Job's Performance, How to Improve Performance? Tweaking the Batch Interval, Backpressure, Dynamic Throttling, Caching, Speculative Execution, Streaming Classification with Naive Bayes.

Total Periods: 45

EXPERIENTIAL LEARNING

The following is the sample. Faculty shall frame according to the course domain.

1. ➢ Use Stream Analytics to process real-time data from Event Hubs
   ➢ Use Stream Analytics windowing functions to build aggregates and output to Synapse Analytics

> ➤ Scale the Azure Stream Analytics job to increase throughput through partitioning
> ➤ Repartition the stream input to optimize parallelization

2. ➤ Learn the key features and uses of Structured Streaming
   ➤ Stream data from a file and write it out to a distributed file system
   ➤ Use sliding windows to aggregate over chunks of data rather than all data
   ➤ Apply watermarking to remove stale data
   ➤ Connect to Event Hubs read and write streams

RESOURCES

TEXT BOOKS:

1. Gerard Maas, François Garillot, Stream Processing with Apache Spark,O'Reilly,2019.

2. Tomasz Drabas,Denny Lee, Learning PySpark,Packt Publishing,2017.

REFERENCE BOOKS:

1. Anindita Basak, Krishna Venkataraman, Ryan Murphy, Manpreet Singh, Stream Analytics with Microsoft Azure: Real -time data processing for quick insights using Azure Stream Analytics,Packt Publishing,2017.

2. Henrique,Andrade, Buğra Gedik, Deepak S. Turaga, Fundamentals of Stream Processing: Application Design, Systems, and Analytics, Cambridge University Press,2014.

VIDEO LECTURES:

1. https://www.youtube.com/watch?v=Rw9OCYhxzwI

2. https://www.youtube.com/watch?v=vFshGQ2ndeg&t=15s

3. https://www.youtube.com/watch?v=A3Mvy8WMk04&t=6s

4. https://www.youtube.com/watch?v=9B7-z0HW058&list=PL9sbKmQTkW076JzfKvnUpGoKVocjkl1S5&index=6

5. https://www.youtube.com/watch?v=Kb08RTmjnkw&t=12s

WEB RESOURCES:

1. https://github.com/stream-processing-with-spark

2. https://github.com/spark-notebook/spark-notebook

3. https://spark-notebook.io/

4. https://spark.apache.org/docs/latest/streaming-programming-guide.html

5. https://spark.apache.org/docs/3.2.0/api/java/org/apache/spark/mllib/classification/NaiveBayes.html

PROGRAM ELECTIVE

Course Code	Course Title	L	T	P	S	C
22AI201009	**MALWARE ANALYSIS**	3	-	-	-	3

Pre-Requisite -
Anti-Requisite -
Co-Requisite -

COURSE DESCRIPTION: The course provides theoretical and conceptual knowledge about the various aspects of the computer system, particularly MS Windows. Introduction to malware, behaviours of malware, mitigating with malware and understanding anti-techniques to prevent influence of malware.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1.   Demonstrate knowledge about the prevalence of malware in the systems.

CO2.   Develop ideas for reverse-engineering and elucidate the internals of registry.

CO3.   Apply debugging in dynamic analysis using tools.

CO4.   Explore methods for counter measures to prevent influence of malware.

CO5.   Analyze anti-techniques for extended analysis and prevention of malware.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	3	2	2	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	2	2	3	-	-	-	-
CO4	3	2	3	-	-	-	-
CO5	3	3	2	-	-	-	-
Course Correlation Mapping	3	2	3	-	-	-	-

Correlation Levels: 3: High; 2: Medium;   1: Low

COURSE CONTENT

Module 1: BASIC ANALYSIS *(09 Periods)*

Basic Static Techniques - Finding Strings, Packed and Obfuscated Malware, Portable Executable File Format, Linked Libraries and Functions; Malware Analysis in Virtual Machines - Structure of a Virtual Machine, Creating and Using Malware Analysis Machine - Risks of Using VMware for Malware Analysis; Basic Dynamic Analysis - Sandboxes - Running and Monitoring Malware - Comparing Registries.

Module 2: ADVANCED STATIC ANALYSIS *(09 Periods)*

The x86 Architecture - Reverse Engineering - General Registers, Flags, Stack; Recognizing C Code in Assembly - Global vs. Local Variables - Disassembling Arithmetic Operations - Recognizing if, switch and loop statements; Recognizing function call conventions, structures, Windows API, Registry, Networking APIs, Kernel vs. User Mode, Native APIs.

Module 3: ADVANCED DYNAMIC ANALYSIS *(09 Periods)*

Debugging - Source level vs. Assembly level debugging, Kernel vs. User-mode debugging; Exceptions - Modifying Execution with a Debugger, Program Execution; OllyDbg - The OllyDbg Interface, Loading Malware, Memory Map, Viewing Threads and Stacks, Executing Code, Loading DLLs, Exception Handling, Scriptable Debugging.

Module 4: MALWARE ANALYSIS *(09 Periods)*

Malware Functionality - Downloaders, Launchers Backdoors, Privilege Escalation, Process Injection, Hook Injection; Data Encoding - Ciphers, Simple Cryptographic Algorithms, Encodings; Malware Focussed Network Signatures - Network Counter Measures, Content-Based Network Countermeasure, Combining Dynamic and Static Analysis Techniques.

Module 5: ANTI-TECHNIQUES *(09 Periods)*

Anti-Reverse-Engineering - Anti-Disassembly, Defeating Disassembly Algorithms, Anti-Disassembly Techniques, Obscuring Flow Control, Brief topics on Anti-debugging, Anti-Virtual Machines.

Total Periods: 45

EXPERIENTIAL LEARNING

The following is the sample. Faculty shall frame according to the course domain.

1. Develop a tool in Python for custom debug operation on Windows Registry to explore the Key-Value pairs of third-party software.
2. Build a tool in Python to create a registry similar to the Windows operating systems, Integrate the operations of OllyDbg into the custom Python.

RESOURCES

TEXT BOOKS:

1. Michael Sikorski and Andrew Honig, Practical Malware Analysis, no starch press, 2012.

2. Monnappa K. A., *Learning Malware Analysis*, Packt Publishing, 2018.

REFERENCE BOOKS:

1. Alexey Kleymenov and Amr Thabet, Mastering Malware Analysis: The complete malware analyst's guide to combating malicious software, APT, cybercrime, and IoT attacks, Packt Publisher, 2019.

2. Michael Ligh, Steven Stair,*Malware Analyst's Cookbook and DVD: Tools and Techniques for Fighting Malicious Code*, Wiley, 2011.

VIDEO LECTURES:

1. https://www.youtube.com/watch?v=D4pc63SeHxI
2. https://www.youtube.com/watch?v=yorvgUhbEbI
3. https://www.youtube.com/watch?v=JLh1xBv7GVQ
4. https://www.youtube.com/watch?v=VYROU-ZwZX8
5. https://www.youtube.com/watch?v=j3lgxdylktM

WEB RESOURCES:

1. https://github.com/surajr/Machine-Learning-approach-for-Malware-Detection
2. https://github.com/Pyran1/MalwareCollection
3. https://github.com/imamitsingh/Classification-of-malwares-using-machine-learning-algorithms
4. https://towardsdatascience.com/malware-classification-using-machine-learning-7c648fb1da79
5. https://ieeexplore.ieee.org/document/9368268

PROGRAM ELECTIVE

Course Code	Course Title	L	T	P	S	C
22AI202010	**BUSINESS INTELLIGENCE TOOLS**	3	-	3	-	4.5

Pre-Requisite	22AI202009-Data Visualization
Anti-Requisite	-
Co-Requisite	-

COURSE DESCRIPTION: This course provides a detailed discussion and hands-on experience on business intelligence tools and concepts like foundations of BI, Decision Making, DataModelling for BI Tools, ETL Basics and Optimizations, Business Reporting, and Business Performance Management, Moving BI Process to Product Environment.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Demonstrate knowledge on the foundations of Business Intelligence and decision making to improve operational efficiency.

CO2. Create Data Model exploring different possibilities using PowerBI to provide Business intelligence solutions.

CO3. Connect to different data sources and extract data using ETL operations to handle queries and extract meaningful insights for efficient business operations.

CO4. Create Business reports and visualize using PowerBI to provide useful insights about business and improve decision making.

CO5. Deploy Business Intelligence process in a product environment on Cloud.

CO6. Work independently to solve problems with effective communication.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	3	2	-	-	-	-	-
CO2	3	2	3	-	-	-	-
CO3	3	3	3	-	-	-	-
CO4	3	3	3	-	-	-	-
CO5	3	3	3	-	-	-	-
CO6	-	-	-	-	3	3	
Course Correlation Mapping	3	3	3	-	3	3	-

Correlation Levels: 3: High; 2: Medium;1: Low

COURSE CONTENT

Module 1: FOUNDATIONS OF BUSINESS INTELLIGENCE AND DECISION MAKING *(10 Periods)*

Foundations of Business Intelligence: What Is business intelligence?, BI Characteristics, Benefits of BI, BI Platform Components, BI Platform Location, BI Concepts,BI Approaches, BI Capabilities.

Decision Making: Introduction and Definitions, Phases of the Decision-Making Process, The Intelligence Phase, The Design Phase, The Choice Phase, The Implementation Phase, Support for various phases, Decision Support Systems: Capabilities, DSS Classifications, Components of Decision Support Systems.

Module 2: DATA MODELING FOR BI SOLUTIONS *(08 Periods)*

Modeling steps- Business Model, Logical Model, Dimensional Model, Physical Model;Defining our Model-Sales Dimension, Status Table, Currency Dimension, Customer Dimension, Employee Dimension¸ Product Dimension, and Time Dimension; Exploring Data Modeling Possibilities, Data Modeling Tools: Preparing the ETL- Source Systems, Source Tables, Source Fields.

Module 3 ETL BASICS AND OPTIMIZATIONS *(09 Periods)*

Details of the Solution, Open Source ETL Suites, Downloading and Installing Pentaho Data Integration. Understanding ETL Concepts- Repositories and Connections, Transformations, How to Organize a Set of Transformations in a Workflow, Create and Share a Connection, The Global Picture, The Product Parent Category Tables, The Customer and Customer Country Tables, The Employee and Employee Category and Employee Department Tables, The Fact Table, Creating the Time Dimension, Connecting All of It Up, Designing the Job Open Source Alternatives to PDI;ETL Optimizations.

Module 4 BUSINESS REPORTING, AND PERFORMANCE MANAGEMENT *(10 Periods)*

Business Reporting: Business Reporting Definitions and Concepts, Data and Information Visualization, Different Types of Charts and Graphs, The Emergence of Data Visualization and Visual Analytics, BI Tools: Microstrategy Desktop-Tableau, Microsoft Power BI, Qlik Sense.

Performance Management: Performance Dashboards, Closed-Loop BPM Cycle, Performance Measurement, Key Performance Indicator, Balanced Scorecards, Six Sigma as a Performance Measurement System.

Module 5 MOVING BI PROCESS TO PRODUCT ENVIRONMENT *(08 Periods)*

Multienvironment Scenario-Deploying a Productive Environment, Adding Environments, Isolation of Environments, Multienvironment Recommendations; Maintaining Your Environment, Security, Auditing, Moving BI Processes to the Cloud- Deciding our Cloud Provider, Choosing the Right Cloud Provider: Amazon Web Services (AWS), Microsoft Azure, Google Cloud, Vendor-Based Cloud Solutions.

Total Periods: 45

EXPERIENTIAL LEARNING

LIST OF EXERCISES:

1. a) Install the PowerBI desktop using the link specified below:
    b) https://powerbi.microsoft.com/en-us/downloads/
    c) Analyze the functionalities available in PowerBI desktop to connect to the data, edit
    d) query, create relationships, and build reports.

2. **Load and Transform:**
    a) Load the database tables and explore the different options available in Query editor.
    b) Perform data transformation tasks such as group by, transpose, changing data types, splitting columns and review the data model.
    c) Perform Loading of data from an Excel file.
    d) Perform Loading multiple CSV files from a folder.

3. a) Create a static table, enter data and relate with other tables and queries in the model. Review the data model.
    b) Load data from web and display in PowerBI reports.

4. **Data Model:**
    a) Combine the queries using the merge queries option in Query Editor. Use Merge Query option in PowerBI to implement different types of SQL types such as Left outer Join, Right outer Join, Full outer join, Inner join.
    b) Combine queries using Append Queries to implement Union operation in SQL and review the data model.

5. **Data Analysis Expressions (DAX):**
    a) Create new columns and measures using the various DAX functions.
    b) Create the Date Dimension using DAX.
    c) Use the variables to simplify the calculations.

6. a) Apply Information functions and login functions in DAX.
    b) Create dynamic calculations using filter functions in DAX to work on Table and Relationship.

7. **Visualizations in PowerBI:**
    Analyse the components of Report Development Interface and list of visualization in PowerBI.

8. a) Create visualizations using different types of Bar chart such as horizontal, vertical and stacked to compare the values of unique categories.
    b) Apply Filters and slicers, both are used to limit the data in the visualizations

9. a) Create Line chart to display the trends of data over a period of time.
    b) Standardize the visual design of the report pages using report themes in PowerBI

10. a) Create multi-page reports using visualization
    b) Display the total value of a measure using card.

11. a) Display which data category has the highest rank, that is, the Largest value using ribbon chart.
    b) Visualize the relationship between 2 measures using scatter chart.
    c) Display the percentage of total using Donut and Piechatr.

12. **Q&A using PowerBI:**
    Explore the data based on the available data set. Create questions in natural language. Use PowerBI to provide suggestion on the question raised and to select best visualization to answer the question. Select barchart to display the information.

RESOURCES

TEXT BOOKS:

1. Albert Nogus, Juan Valladaraes, Business Intelligence Tools for Small Companies, Apress, Tenth Edition 2017.
2. Chandraish Sinha, Mastering PowerBI, BpB Publications,2022.

REFERENCE BOOKS:

1. Ramesh Sharda, Dursun Delen, Efraim Turban, Business Intelligence and Analytics, Pearson, 10th Edition,2014.
2. Hancoc,;Toren, Practical Business Intelligence with SQL Server 2005, Pearson,2022.

SOFTWARE/TOOLS:

1. Power BI Tool.

VIDEO LECTURES:

1. https://www.coursera.org/courses?query=power%20bi
2. https://nptel.ac.in/courses/110107092
3. https://www.udemy.com/topic/business-intelligence/
4. https://www.edx.org/learn/power-bi
5. https://www.ed2go.com/careers/information-technology/business-intelligence-analyst
6. https://www.udacity.com/blog/search/label/business-intelligence
7. https://mindmajix.com/search?searchString=business%20intelligence&type=course

WEB RESOURCES:

1. www.pentaho.com/
2. https://www.ibm.com/developerworks/library/os-weka2/
3. http://www.saedsayad.com/
4. http://www.cs.ccsu.edu/~markov/ccsu_courses/datamining-3.html
5. https://cognitiveclass.ai/
6. https://data-flair.training/blogs/business-intelligence/
7. https://www.tutorialspoint.com/power_bi/index.htm
8. https://www.educba.com/business-intelligence-tool/
9. https://chartio.com/learn/business-intelligence/10-business-intelligence-tools-to-visualize-and-analyze-your-data

UNIVERSITY ELECTIVE

Course Code	Course Title	L	T	P	S	C
22AI201701	**BUSINESS ANALYTICS**	3	-	-	-	3

Pre-Requisite -

Anti-Requisite -

Co-Requisite -

COURSE DESCRIPTION: This course emphasizes on the basic concepts of Business Analytics. It covers the basic excel skills, Excel look up functions for database queries in business analytics. By the end of this course students will acquire basic knowledge to implement statistical methods for performing descriptive, predictive and prescriptive analytics.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Understand the basic concepts and models of Business Analytics

CO2. Select Suitable basic excel function to perform analytics on spread sheets.

CO3. Apply different statistical techniques and distributions for modeling the data

CO4. Develop user-friendly Excel applications by using statistical models for effectiveness decision making.

CO5. Analyze the performance of different optimization models used in prescriptive analytics on Binary and Categorical data.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	3	-	-	-	-	-	-
CO2	-	3	-	-	-	-	-
CO3	-	3	-	-	-	-	-
CO4	-	-	3	-	-	-	-
CO5	-	3	-	-	-	-	-
Course Correlation Mapping	3	3	-	-	-	-	-

Correlation Levels: 3: High;   2: Medium;   1: Low

COURSE CONTENT

Module 1:   FOUNDATIONS OF BUSINESS ANALYTICS   *(9 Periods)*

Introduction, What is Business Analytics, Evolution of Business Analytics, Scope of Business Analytics, Data for Business Analytics, Applications of Business Analytics, Models in Business Analytics, Problem Solving with Analytics.

Module 2:   ANALYTICS ON SPREADSHEETS   *(9 Periods)*

Basic Excel Skills, Excel Functions, Using Excel Lookup Functions for Database Queries, Spreadsheet Add-Ins for Business Analytics.
Visualizing and Exploring Data: Data Visualization, Creating Charts In Microsoft Excel, Other Excel Data Visualization, Statistical Methods For Summarizing Data, Exploring Data Using Pivot tables.

Module 3: DATA MODELING   *(9Periods)*

Basic concepts of Probability, Random Variables and Probability Distributions, Continuous Probability Distributions.
Statistical Sampling, Estimation population parameters, Sampling Error, Sampling Distributions, Hypothesis Testing, ANOVA, Chi Square Test.

Predictive analytics   *(9 Periods)*
Module 4

Trend lines And Regression Analysis, Modeling Relationships And Trends In Data, Simple Linear Regression, Multiple Linear Regression, Building Good Regression Models,
Strategies for predictive decision modeling, implementing models on spreadsheets, spreadsheet applications in business analytics, developing user-friendly excel applications, analysing uncertainty and model assumptions, model analysis using analytic solver platform

Module 5   Prescriptive analytics   *(9Periods)*

Linear Models: Building Linear Models, Implementing Linear Optimization Models On Spreadsheets, Graphical Interpretation Of Linear Optimization, Linear Optimization Models for prediction and Insight.
Integer Models: Solving models with Integer Variables, Integer Optimization Models with Binary Numbers
Decision Analysis: Formulating Decision Problems, Decision Strategies Without Outcome Probabilities, Decision Trees With Outcome Probabilities, Decision Trees.

Total Periods: 45

EXPERIENTIAL LEARNING

1. **Diabetic Prediction:**

   The National Institute of Diabetes and Digestive and Kidney Diseases has a created a dataset. The objective of the dataset is to diagnostically predict whether or not a patient has diabetes, based on certain diagnostic measurements included in the dataset. Several constraints were placed on the selection of these instances from a larger database. In particular, all patients here are females at least 21 years old of Pima Indian heritage. The datasets consists of several medical predictor variables and one target variable, Outcome. Predictor variables includes the number of pregnancies the patient has had, their BMI, insulin level, age, and so on. Build a machine learning model to accurately predict whether or not the patients in the dataset have diabetes or not?

2. Solve the house price prediction problem using **Linear regression analysis** method. Optimize the parameters of the regression function using gradient descent method.

3. Visualize the decision tree built for solving Heart disease prediction problem and measure the impurity of nodes created via **Decision Tree Analysis**.

   Dataset:https://www.kaggle.com/arviinndn/heart-disease-prediction-uci dataset/data

4. The data set baby boom (Using R) contains data on the births of 44 children in a one- day period at a Brisbane, Australia, hospital. Compute the skew of the wt variable, which records birth weight. Is this variable reasonably symmetric or skewed?

5. Visualize the **Distribution of data** with different feature scaling methods on online news popularity dataset for article word count.

   Dataset:https://www.kaggle.com/datasets/deepakshende/onlinenewspopularity

6. **Human Activity Recognition System:**

   The human activity recognition system is a classifier model that can identify human fitness activities. To develop this system, you have to use a smart phone dataset, which contains the fitness activity of 30 people which is captured through smart phones. This system will help you to understand the solving procedure of the **Multi-classification problem**.

RESOURCES

TEXT BOOKS:

1. James Evans, *Business Analytics*, Pearson Education, 2nd Edition, 2017.

REFERENCE BOOKS:

1. Marc J.Schniederjans, *Business Analytics*, Pearson Education,2015
2. Camm,Cochran, *Essentials of Business Analytics*, Cenage learning, 2015

VIDEO LECTURES:

1. https://nptel.ac.in/courses/110105089
2. https://archive.nptel.ac.in/courses/110/107/110107092/
3. https://nptel.ac.in/courses/110106050

Web Resources:

1. https://www.proschoolonline.com/certification-business-analytics-course/what-is-ba
2. https://michael.hahsler.net/SMU/EMIS3309/slides/Evans_Analytics2e_ppt_01.pdf
3. https://www.guru99.com/business-analyst-tutorial-course.html

UNIVERSITY ELECTIVE

Course Code	Course Title	L	T	P	S	C
22CM201701	**COST MANAGEMENT OF ENGINEERING PROJECTS**	3	-	-	-	3

Pre-Requisite   -
Anti-Requisite   -
Co-Requisite   -

COURSE DESCRIPTION: This course will provide an understanding of the cost tools and techniques that can be used throughout a project's design and development. The students will be exposed to the methods, processes, and tools needed to conduct economic analysis, estimation of Project.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1   Understand the costing concepts and their role in decision-making.
CO2   Understand the project management concepts and their various aspects in selection.
CO3   Interpret costing concepts with project execution.
CO4   Knowledge of costing techniques in the service sector and various budgetary control techniques.
CO5   Become familiar with quantitative techniques in cost management.

CO-PO-PSO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	3	-	-	-	-	-	-
CO2	3	-	-	-	-	-	-
CO3	-	3	-	-	-	-	-
CO4	3	-	-	-	-	-	-
CO5	3	-	-	-	-	-	-
Course Correlation Mapping	**3**	**3**	**-**	**-**	**-**	**-**	**-**

Correlation Levels: *3: High;*   *2: Medium;*   *1: Low*

COURSE CONTENT

Module 1:   INTRODUCTION TO COSTING CONCEPTS   *(05 Periods)*

Objectives of a Costing System; Cost concepts in decision-making; Relevant cost, Differential cost, Incremental cost, and Opportunity cost; Creation of a Database for operational control.

Module 2:   INTRODUCTION TO PROJECT MANAGEMENT   *(10 Periods)*

Project: meaning, Different types, why to manage, cost overruns centers, various stages of project execution: conception to commissioning. Project execution as conglomeration of technical and nontechnical activities, Detailed Engineering activities, Pre-project execution main clearances and documents, Project team: Role of each member, Importance Project site: Data required with significance, Project contracts

Module 3   PROJECT EXECUTION AND COSTING CONCEPTS   *(10 Periods)*

Project execution Project cost control, Bar charts and Network diagram, Project commissioning: mechanical and process, Cost Behavior and Profit Planning Marginal Costing; Distinction between Marginal Costing and Absorption Costing; Break-even Analysis, Cost-Volume-Profit Analysis, Various decision-making problems, Pricing strategies: Pareto Analysis, Target costing, Life Cycle Costing

Module 4   COSTING OF SERVICE SECTOR AND BUDGETARY CONTROL   *(10 Periods)*

Just-in-time approach, Material Requirement Planning, Enterprise Resource Planning, ActivityBased Cost Management, Bench Marking; Balanced Score Card and Value-Chain Analysis, Budgetary Control: Flexible Budgets; Performance budgets; Zero-based budgets

Module 5   QUANTITATIVE TECHNIQUES FOR COST MANAGEMENT   *(10 Periods)*

Linear Programming, PERT/CPM, Transportation problems, Assignment problems, Learning Curve Theory.

Total Periods: 45

EXPERIENTIAL LEARNING

1   Prepare a mini-project report regarding cost control techniques in manufacturing units.
2   Prepare a report on real-life engineering project case studies, especially those that faced cost overruns or successfully managed costs
3   Conduct hands-on budgeting exercises where participants are given a project scope, and they have to create detailed budgets.

RESOURCES

TEXT BOOKS:

1.   John M. Nicholas, Herman Steyn Project Management for Engineering, Business and Technology, Taylor & Francis, 2 August 2020, ISBN: 9781000092561
2.   Albert Lester ,Project Management, Planning and Control, Elsevier/Butterworth-Heinemann, 2007, ISBN: 9780750669566, 075066956X.

REFERENCE BOOKS:

1.   Charles T. Horngren et al Cost Accounting a Managerial Emphasis, Prentice Hall of India, New Delhi, 2011.
2.   Ashish K. Bhattacharya, Principles & Practices of Cost Accounting A. H. Wheeler publisher, 1991.

3   Vohra N.D., Quantitative Techniques in Management, Tata McGraw Hill Book Co. Ltd, 2007

4   Robert S Kaplan Anthony A. Alkinson, Management & Cost Accounting, 2003

VIDEO LECTURES:

1.   https://www.youtube.com/watch?v=rck3MnC7OXA

2.   https://www.youtube.com/watch?v=QWD1LMzStI4

WEB RESOURCES:

1.   https://www.superfastcpa.com/what-are-cost-concepts-in-decision-making

2.   https://www.indeed.com/career-advice/career-development/project-cost-controls

3.   https://www.geeksforgeeks.org/difference-between-pert-and-cpm/

UNIVERSITY ELECTIVE

Course Code	Course Title	L	T	P	S	C
22CE201701	**DISASTER MANAGEMENT**	3	-	-	-	3

Pre-Requisite   -

Anti-Requisite   -

Co-Requisite   -

COURSE DESCRIPTION: This course provides a detailed discussion on disaster prone areas in India, repercussions of disasters and hazards, disaster preparedness and management, risk assessment and disaster management.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Analyze the vulnerability of an area to natural and man-made disasters/hazards as per the guidelines to solve complex problems using appropriate techniques ensuring safety, environment and sustainability.

CO2. Analyze the causes and impacts of disasters using appropriate tools and techniques and suggest mitigation measures ensuring safety, environment and sustainability besides communicating effectively in graphical form.

CO3. Suggest the preparedness measures using appropriate tools and techniques and suggest mitigation measures ensuring safety, environment and sustainability.

CO4. Analyze the Risk Assessment using appropriate tools and techniques and suggest mitigation measures ensuring safety, environment and sustainability.

CO5. Design disaster management strategies to solve pre, during and post disaster problems using appropriate tools and techniques following the relevant guidelines and latest developments ensuring safety, environment and sustainability besides communicating effectively in graphical form.

CO-PO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	3	3	-	2	2	2	2
CO2	3	3	-	2	2	2	2
CO3	3	3	-	2	2	2	2
CO4	3	3	-	3	2	2	2
CO5	3	2	3	2	2	2	1
Course Correlation Mapping	**3**	**3**	**3**	**3**	**2**	**2**	**2**

Correlation Levels: 3: High;   2: Medium;   1: Low

COURSE CONTENT

Module 1:   DISASTER PRONE AREAS IN INDIA   *(09 Periods)*
Introduction: Disaster: Definition, Factors and Significance; Difference Between Hazard and Disaster; Natural and Manmade Disasters: Difference, Nature, Types And Magnitude.
Disaster Prone Areas: Study Of Seismic Zones; Areas Prone To Floods And Droughts, Landslides And Avalanches; Areas Prone To Cyclonic And Coastal Hazards With Special Reference To Tsunami; Post-Disaster Diseases And Epidemics.

Module 2:   REPERCUSSIONS OF DISASTERS AND HAZARDS   *(09 Periods)*
Economic Damage, Loss of Human and Animal Life, Destruction of Ecosystem. Natural Disasters: Earthquakes, Volcanisms, Cyclones, Tsunamis, Floods, Droughts And Famines, Landslides And Avalanches, Man-made disaster: Nuclear Reactor Meltdown, Industrial Accidents, Oil Slicks And Spills, Outbreaks Of Disease And Epidemics, War And Conflicts.

Module 3:   DISASTER PREPAREDNESS AND MANAGEMENT   *(11 Periods)*
Preparedness: Monitoring Of Phenomena Triggering A Disaster Or Hazard; Evaluation Of Risk: Application Of Remote Sensing, Data From Meteorological And Other Agencies, Media Reports: Governmental And Community Preparedness.

Module 4:   RISK ASSESSMENT   *(08 Periods)*
Disaster Risk: Concept and Elements, Disaster Risk Reduction, Global and National Disaster Risk Situation. Techniques of Risk Assessment, Global Co-Operation In Risk Assessment And Warning, People's Participation In Risk Assessment. Strategies for Survival.

Module 5:   DISASTER MANAGEMENT   *(08 Periods)*
Disaster management organization and methodology, Disaster management cycle, Disaster management in India – Typical cases and Cost–benefit analysis, Disaster management programs implemented by NGOs and Government of India, Usage of GIS andRemote sensing techniques in disaster management, Leadership and Coordination in
Disaster management, Emerging trends in disaster management.
                                                        Total Periods: 45

EXPERIENTIAL LEARNING

1. Perform hazard assessment and vulnerability analysis for any nearby town/city and prepare a detailed report of possible impacts of various disasters on environment, infrastructure and development.

2. Prepare a detailed report on the causes and effects of Tsunami that was occurred in the year 2004. Also discuss various advancements in Tsunami warning systems.

3. Identify the major causes of urban floods in cities like Chennai, Hyderabad & Mumbai. Also list various mitigation strategies to reduce the impact of floods.

4. Prepare a detailed report on how various man-made activities are directly/indirectly related to the occurrence of landslides that occurred in recent days in India.

5. Visit AP State Disaster Response and Fire Services Department and record about various methods used by them in mitigating disasters and their management.

RESOURCES

TEXT BOOKS:

1. Sharma V. K., *Disaster Management, Medtech Publishing, 2nd Edition, 2013*.

2. Anand S. Arya, Anup Karanth, and Ankush Agarwal, *Hazards, Disasters and Your Community: A Primer for Parliamentarians*, GOI–UNDP Disaster Risk Management Programme, Government of India, National Disaster Management Division, Ministry of Home Affairs, New Delhi, Version 1.0, 2005

REFERENCE BOOKS:

1. Donald Hyndman and David Hyndman, *Natural Hazards and Disasters*, Cengage Learning, USA, 5th Edition, 2015.

2. *Disaster Management in India,* A Status Report, Ministry of Home Affairs, Govt. of India, May 2011.

3. Rajendra Kumar Bhandari, *Disaster Education and Management: A Joyride for Students, Teachers, and Disaster Managers*, Springer India, 2014.

4. Singh R. B., *Natural Hazards and Disaster Management*, Rawat Publications, 2009.

5. R. Nishith, Singh AK, *Disaster Management in India: Perspectives, issues and strategies,* New Royal book Company.

6. Sahni, PardeepEt.Al. (Eds.), *Disaster Mitigation Experiences And Reflections*, Prentice Hall of India, New Delhi.

7. Goel S. L. , *Disaster Administration And Management Text And Case Studies*, Deep &Deep Publication Pvt. Ltd., New Delhi

VIDEO LECTURES:

1. https://nptel.ac.in/courses/105104183

2. https://www.digimat.in/nptel/courses/video/124107010/L01.html

WEB RESOURCES:

1. https://egyankosh.ac.in/handle/123456789/25093

2. https://www.egyankosh.ac.in/handle/123456789/25912

3. https://www.nios.ac.in/media/documents/333courseE/12.pdf

4. https://ndmindia.mha.gov.in/images/public-awareness/Primer%20for%20Parliamentarians.pdf

UNIVERSITY ELECTIVE

Course Code	Course Title	L	T	P	S	C
22SS201701	**VALUE EDUCATION**	3	-	-	-	3

Pre-Requisite -
Anti-Requisite -
Co-Requisite -

COURSE DESCRIPTION: This course deals with understanding the value of education and self-development, Imbibe good values in students, and making them know about the importance of character.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Demonstrate the knowledge of values and self-development
CO2. Analyze the importance of the cultivation of values.
CO3. Learn suitable aspects of personality and behavioral development
CO4. Function as a member and leader in multi-disciplinary teams by avoiding faulty thinking.
CO5. Develop character and competence for effective studies.

CO-PO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	3	-	-	-	-	-	-
CO2	2	3	-	-	2	-	-
CO3	2	-	-	-	2	-	-
CO4	2	-	-	-	-	-	-
CO5	2	2	-	-	-	-	-
Course Correlation Mapping	**2**	**3**	**-**	**-**	**2**	**-**	**-**

Correlation Levels: 3: High;   2: Medium;   1: Low

COURSE CONTENT

Module 1: VALUES AND SELF-DEVELOPMENT *(09 Periods)*

Values and self-development –Social values and individual attitudes. Work ethics, Indian vision of humanism. Moral and non-moral valuation. Standards and principles. Value judgements- Case studies

Module 2: IMPORTANCE OF CULTIVATION OF VALUES. *(09 Periods)*

Importance of cultivation of values. Sense of duty. Devotion, Self-reliance. Confidence, Concentration. Truthfulness, Cleanliness. Honesty, Humanity. Power of faith, NationalUnity. Patriotism. Love for nature, Discipline- Case studies

Module 3: PERSONALITY AND BEHAVIOR DEVELOPMENT *(09 Periods)*

Personality and Behavior Development - Soul and Scientific attitude. Positive Thinking. Integrity and discipline, Punctuality, Love and Kindness - Case studies

Module 4: AVOID FAULTY THINKING. *(09 Periods)*

Avoid fault Thinking. Free from anger, Dignity of labour. Universal brotherhood andreligious tolerance. True friendship. Happiness Vs suffering, love for truth. Aware of self- destructive habits. Association and Cooperation. Doing best for saving nature - Case studies

Module 5: CHARACTER AND COMPETENCE *(09 Periods)*

Character and Competence –Holy books vs Blind faith. Self-management and Good health. Science of reincarnation, Equality, Nonviolence, Humility, Role of Women. All religions and the same message. Mind your Mind, Self-control. Honesty, Studying effectively- Case studies

Total Periods: 45

EXPERIENTIAL LEARNING

1. Demonstrate orally using your experiences of what values are naturally acceptable in a relationship to nurture or exploit others.

2. Prepare a report by identifying and analyzing the importance of cultivation of values.

3. Present a poster on different attitudes and behaviours.

4. Students give a PowerPoint presentation on doing best for nature.

5. Students are encouraged to bring a daily newspaper to class or to access any news related to the need for human values and note down the points.

6. Prepare a case study on how to maintain harmony with different religious people through character and competence.

(It's an indicative one. The Course Instructor may change the activities and the same shall be reflected in the Course Handout)

RESOURCES

TEXTBOOKS:

1. R. Subramanaian, *Professional Ethics*, Oxford Higher Education, 2013.
2. Mike W. Martin and Roland Schinzinger, *Ethics in Engineering*, Tata McGraw-Hill, 3rd Edition, 2007.
3. Chakravarthy, S.K.: Values and ethics for Organizations: Theory and Practice, Oxford University Press, NewDelhi, 1999.

REFERENCE BOOKS:

1. M.G. Chitakra: Education and Human Values, A.P.H. Publishing Corporation, New Delhi, 2003
2. Awakening Indians to India, Chinmayananda Mission, 2003
3. Satchidananda, M.K.: Ethics, Education, Indian Unity and Culture, Ajantha Publications, Delhi, 1991

VIDEO LECTURES:

1. https://www.youtube.com/watch?v=90VQPZURN5c
2. https://www.youtube.com/watch?v=6ofPcK0uDaA
3. https://www.youtube.com/watch?v=5_f-7zCi79A
4. https://www.youtube.com/watch?v=2ve49BWAJRE
5. https://www.youtube.com/watch?v=kCOIfnxxQ5U

WEB RESOURCES:

1. https://www.livingvalues.net/
2. https://livingvalues.net/materials-for-schools/
3. https://www.edb.gov.hk/en/curriculum-development/4-key-tasks/moral-civic/index.html

UNIVERSITY ELECTIVE

Course Code	Course Title	L	T	P	S	C
22SS201702	**PEDAGOGY STUDIES**	3	-	-	-	3

Pre-Requisite -

Anti-Requisite -

Co-Requisite -

COURSE DESCRIPTION: This course deals with understanding pedagogical practices thatare being used by teachers in formal and informal classrooms, the effectiveness of pedagogical practices, teacher education (curriculum and practicum), and the school curriculum and guidance materials that can best support effective pedagogy.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1 Demonstrate knowledge of pedagogical methodology

CO2 Analyze the functional knowledge in Pedagogical practices, Curriculum, and Teacher Education

CO3 Learn effective pedagogical practices and apply strategies.

CO4 Function effectively as an individual and as a member of the Professional development.

CO5 Understand research Gaps and provide future Directions.

CO-PO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	2	1	-	-	-	-	-
CO2	2	3	-	-	3	-	-
CO3	2	2	-	-	3	-	-
CO4	1	1	-	-	-	-	-
CO5	-	-	-	-	-	-	-
Course Correlation Mapping	2	2	-	-	3	-	-

Correlation Levels: 3: High;   2: Medium;   1: Low

COURSE CONTENT

Module 1:   **INTRODUCTION AND METHODOLOGY**   *(09 Periods)*

Aims and rationale, Policy background, Conceptual framework and terminology Theories of learning, Curriculum, Teacher education. Conceptual framework, Research questions**.** Overview of Methodology and Searching- Case studies

Module 2:   **THEMATIC OVERVIEW**   *(09 Periods)*

Pedagogical practices are being used by teachers in formal and informal classrooms in developing countries. Curriculum, Teacher Education- Case studies

Module 3   **EFFECTIVENESS OF PEDAGOGICAL PRACTICES**   *(09 Periods)*

Evidence on the effectiveness of pedagogical practices, Methodology for the in-depth stage: quality assessment of included studies, teacher education (curriculum and practicum) and the school curriculum and guidance materials best support effective pedagogy, Theory of change, Strength and nature of the body of evidence for effective pedagogical practices. Pedagogic theory and pedagogical approaches. Teachers' Attitudes and beliefs and Pedagogic strategies- Case studies

Module 4   **PROFESSIONAL DEVELOPMENT**   *(09 Periods)*

alignment with classroom practices and follow-up support, Peer support, and Support from the head teacher and the community. Curriculum and assessment, Barriers to learning: limited resources and large class sizes- Case studies

Module 5   **RESEARCH GAPS AND FUTURE DIRECTIONS**   *(09 Periods)*

Research design**,** Contexts**,** Pedagogy**,** Teacher Education**,** Curriculum and Assessment**,** Dissemination and research impact- Case studies

Total Periods: 45

EXPERIENTIAL LEARNING

1. List out the self-improvement in you after going through pedagogical methodologies.
2. Discuss different practices that you would like to adopt in the curriculum.
3. Describe in your own words how can you bring effectiveness to the curriculum.
4. Imagine you are a head teacher and illustrate different barriers to learning.
5. Assume you are a teacher and Interpret different directions that you would bring for the assessment of the students.

(It's an indicative one. The Course Instructor may change the activities and the same shall be reflected in the Course Handout)

RESOURCES

TEXTBOOK:

1. Ackers J, Hardman F (2001) Classroom interaction in Kenyan primary schools, Compare, 31 (2): 245-261.
2. Alexander RJ (2001) Culture and pedagogy: International comparisons in primary education.

REFERENCES:

1. Akyeampong K (2003) Teacher training in Ghana - does it count? Multi-site teacher education
2. Agrawal M (2004) Curricular reform in schools: The importance of evaluation, Journal of Curriculum Studies, 36 (3): 361-379.Oxford and Boston: Blackwell.
3. Akyeampong K, Lussier K, Pryor J, Westbrook J (2013) Improving teaching and learning of basic maths and reading in Africa: Does teacher preparation count? International Journal Educational Development, 33 (3): 272–282.
4. Chavan M (2003) Read India: A mass scale, rapid, 'learning to read' campaign.

VIDEO LECTURES:

1. https://www.youtube.com/watch?v=WL40UeySag4
2. https://www.youtube.com/watch?v=MMXaXDIHFJ8
3. https://www.youtube.com/watch?v=7uJL1R6M4Iw

WEB RESOURCES:

1. https://acrl.ala.org/IS/instruction-tools-resources-2/pedagogy/a-selected-list-of-journals-on-teaching-learning/

2. https://guides.douglascollege.ca/TLonline/resourcesforonlinepedagogy

3. https://www.refseek.com/directory/teacher_resources.html

UNIVERSITY ELECTIVE

Course Code	Course Title	L	T	P	S	C
22LG201701	**PERSONALITY DEVELOPMENT**	3	-	-	-	3

Pre-Requisite   -

Anti-Requisite   -

Co-Requisite   -

COURSE DESCRIPTION: This course gives awareness to students about the various dynamics of personality development.

COURSE OUTCOMES: After successful completion of the course, students will be able to:

CO1. Demonstrate knowledge in Self-Management and Planning Career

CO2. Analyze the functional knowledge in attitudes and thinking strategies

CO3. Learn and apply soft skills for professional success.

CO4. Function effectively as an individual and as a member in diverse teams

CO5. Communicate effectively in public speaking in formal and informal situations.

CO-PO Mapping Table:

Course Outcomes	Program Outcomes						
	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**
CO1	2	1	-	-	-	-	-
CO2	2	3	-	-	-	-	-
CO3	2	2	-	-	3	-	-
CO4	1	1	-	-	-	-	-
CO5	-	-	-	-	-	-	-
Course Correlation Mapping	**2**	**2**	**3**	**-**	**3**	**-**	**-**

Correlation Levels: 3: High;   2: Medium;   1: Low

COURSE CONTENT

Module 1:   SELF-ESTEEM & SELF-IMPROVEMENT   *(09 Periods)*

Know Yourself – Accept Yourself; Self-Improvement: Plan to Improve - Actively Working to Improve Yourself- Exercises- case studies

Module 2:   DEVELOPING POSITIVE ATTITUDES   *(09 Periods)*

How Attitudes Develop – Attitudes are Catching – Improve Your Attitudes – Exercises- case studies

Module 3   SELF-MOTIVATION & SELF-MANAGEMENT   *(09 Periods)*

Show Initiative – Be Responsible Self-Management; Efficient Work Habits – Stress Management – Employers Want People Who can Think – Thinking Strategies- Exercises- case studies

Module 4   GETTING ALONG WITH THE SUPERVISOR   *(09 Periods)*

Know your Supervisor – Communicating with your Supervisor – Special Communication with your Supervisor – What Should you Expect of Your Supervisor? – What your Supervisor expects of you - Moving Ahead Getting Along with your Supervisor- Exercises- case studies

Module 5   WORKPLACE SUCCESS   *(09 Periods)*

First Day on the Job – Keeping Your Job – Planning Your Career – Moving Ahead- Exercises- case studies

                                                                  Total Periods: 45

EXPERIENTIAL LEARNING

1. List out the self-improvements in you on the charts and explain in detail.
2. Discuss different famous personalities and their attitudes.
3. Describe different personalities with respect to self-motivation and self-management.
4. Imagine you are a supervisor and illustrate different special communications.
5. Assume and Interpret different experiences on the first day of your job.

RESOURCES

TEXTBOOK:

1   Harold R. Wallace and L. Ann Masters, *Personal Development for Life and Work*, Cengage Learning, Delhi, 10th edition Indian Reprint, 2011. (6th Indian Reprint 2015)
2   Barun K. Mitra, *Personality Development and Soft Skills,* Oxford University Press, 2011.

REFERENCE BOOKS:

1 K. Alex, *Soft Skills*, S. Chand & Company Ltd, New Delhi, 2nd Revised Edition, 2011.
2 Stephen P. Robbins and Timothy A. Judge*, Organizational Behaviour*, Prentice Hall, Delhi, 16th edition, 2014

VIDEO LECTURES:

1.   https://www.youtube.com/watch?v=6Y5VWBLi1es
2.   https://www.youtube.com/watch?v=H9qA3inVMrA

Web Resources:

1.   https://www.universalclass.com/.../the-process-of-perso...
2.   https://www.ncbi.nlm.nih.gov/pubmed/25545842
3.   https://www.youtube.com/watch?v=Tuw8hxrFBH8