E.G.S. PILLAY ENGINEERING COLLEGE

(Autonomous)

Approved by AICTE, New Delhi | Affiliated to Anna University, Chennai

Accredited by NAAC with 'A++' Grade | Accredited by NBA (BE- CIVIL, CSE, ECE, EEE, MECH & IT)

NAGAPATTINAM – 611 002

B.TECH. ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

Full Time Curriculum and Syllabus

Second Year – Fourth Semester

B.Tech - AI&DS (R2023) – IV Semester									
COURSE CODE	**COURSE NAME**	**CATEGORY**	**L**	**T**	**P**	**C**	**MAX. MARKS**		
							CA	**ES**	**TOTAL**
Theory Courses									
2301GEX07	Environmental Sciences & Sustainability	BSC	1	0	2	2	50	50	100
2301MA401	Discrete Mathematics & Graph Theory	BSC	3	0	0	3	40	60	100
2302AS401	Software Engineering	PCC	3	0	0	3	40	60	100
2302AS402	Data Mining Concepts	PCC	3	0	0	3	40	60	100
2302AS403	Data Analytics and Visualization	PCC	3	0	2	4	50	50	100
2302AS404	Machine Learning Techniques	PCC	3	0	0	3	40	60	100
Laboratory Courses									
2302AS451	Data Mining Laboratory	PCC	0	0	2	1	60	40	100
2302AS452	Machine Learning Laboratory	PCC	0	0	2	1	60	40	100
Other Courses									
2301MC403	Mandatory Course - I - Innovation & Entrepreneurship Fundamentals	MC	3	0	0	3	100	-	100
2304GE401	Professional Development Course - II	EEC	0	0	2	1	100	-	100
2301LS401	Life Skills – IV	-	0	0	0	0	100	-	100
		TOTAL	**16**	**0**	**10**	**21**	**680**	**420**	**1100**

2301GEX07	ENVIRONMENTAL SCIENCES AND SUSTAINABILITY	L	T	P	C
	(Common to all Branches of B.E/ B.Tech)	1	0	2	2
PREREQUISITE:					
	Basic Knowledge about the valuable environment Basic Knowledge to conserve the precious environment				

COURSE OBJECTIVES:

	Realize the interdisciplinary and holistic nature of the environment
	Understand how natural resources and environment affect the quality of life and stimulate the quest for sustainable development

COURSE OUTCOMES:

At the end of the course the student will be able to

CO1	Describe the importance of ecosystem.
CO2	Describe the various environmental issues and its prevention.
CO3	Organize various natural resources and the immediate need to conserve it.
CO4	Select the various ways of conservation of biodiversity.
CO5	Investigate the different types of pollution and its effects.

COs Vs POs MAPPING:

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PSO1	PSO2
CO1	2	1	-	-	-	3	-	-	-	-	-	-	-
CO2	2	1	-	-	-	3	-	-	-	-	-	-	-
CO3	3	2	1	1	1	1	3	2	2	2	3	-	-
CO4	3	2	1	1	1	1	3	2	2	2	3	-	-
CO5	3	2	1	1	1	1	3	2	2	2	3	-	-

COURSE CONTENTS:

Module I	**ECOSYSTEM**	**8 Hours**

Concept of an ecosystem – structure and function of an ecosystem – producers, consumers and decomposers. Introduction, types, characteristic features, structure and function of the (a) forest ecosystem (b) grassland ecosystem (c) desert ecosystem (d) aquatic ecosystems (ponds, rivers, oceans)

Module II	**ENVIRONMENT AL ISSUES AND SOLUTIONS**	**7 Hours**

Current Environmental Issues: Acid rain, Ozone layer depletion, Global warming, Greenhouse effect.
Solutions: 12 principles of green chemistry-Rain water harvesting.

Mini Project Modules

Module III	**BIODIVERSITY**	**10 Hours**

Introduction to biodiversity -genetic, species and ecosystem diversity – value of biodiversity: consumptive use, productive use, social, ethical, aesthetic and option values – hot-spots of biodiversity – threats to biodiversity: habitat loss, poaching of wildlife, man-wildlife conflicts – conservation of biodiversity: In-situ and ex-situ conservation of biodiversity.

Module IV	**NATURAL RESOURCES**	**10 Hours**

Forest resources: Use and over-exploitation, deforestation- timber extraction, mining, dams and their effects on forests and tribal people – Water resources: Use and over utilization of surface and ground water, dams-benefits and problems – Mineral resources: Use and exploitation, environmental effects of extracting and using mineral resources – Food resources: World food problems, changes caused by agriculture and overgrazing, effects of modern agriculture, fertilizer-pesticide problems, water logging, salinity– Energy resources: Growing energy needs, renewable and nonrenewable energy sources, use of alternate energy sources. Energy Conversion processes Biogas – production and uses, anaerobic digestion – Land resources: Land as a resource, land degradation, man induced landslides, soil erosion and desertification – role of an individual in conservation of natural resources – Equitable use of resources for sustainable lifestyles.

Module V	**ENVIRONMENTAL POLLUTION**	**10 Hours**

Definition – Source, causes, effects and control measures of: (a) Air pollution (b) Water pollution(c) Soil pollution - soil waste management: causes, effects and control measures of municipal solid wastes – (d) Marine pollution (e) Noise pollution –(f) Nuclear pollution (g) Thermal pollution role of an individual in prevention of pollution.

	TOTAL:	**45 HOURS**

MINI PROJECT ADDITIONAL TOPICS

Soil Science

1. Effects of climate change on soil erosion.
2. The role of land management in maintaining soil health.
3. Effects of salinity in coastal region Agricultural activity.
4. The effects of climate change on agriculture.

Urban Ecology

1. How road construction impacts biodiversity and ecosystems.
2. The effects of urbanization and city planning on water cycles.
3. Impacts of noise pollution on human health.

Pollution and Bio-remediation

1. The role of bio-remediation in removing “forever” chemicals from the environment.
2. Impacts of air pollution on human health.
3. How to improve plastic recycling processes.
4. Individual measures to reduce consumption and creation of microplastics.

General Topics

1. Impact of Urbanization on Local Biodiversity
2. Renewable Energy Options for Sustainable Living.
3. Waste Management Strategies in Urban Areas
4. Climate Change and Its Effects on Local Ecosystems
5. Air Quality Monitoring in Urban centers
6. Water Quality Assessment in Local Water Bodies
7. Green Roof Technology and Its Environmental Benefits
8. Impact of Plastic Pollution on Marine Life.
9. Eco-friendly Practices in Agriculture:
10. The Role of Community Gardens in Urban Sustainability
11. Alternate energy sources for community Development.
12. E-Waste Management.
13. Energy Audit of a building.
14. Rainwater harvesting system.
15. Population growth variation among nations.
16. Population explosion.
17. Family welfare programme.
18. Women welfare programme.
19. Child welfare programme.
20. Environmental impact analysis.
21. Role of information technology in environmental protection and human health.

REFERENCES:

1. Trivedi.R.K., "Handbook of Environmental Laws, Rules, Guidelines, Compliances and Standards", Vol. I and II, Enviro Media, 3rd edition, BPB publications, 2010.
2. Cunningham, W.P. Cooper, T.H. Gorhani, "Environmental Encyclopedia", Jaico Publ., House, Mumbai, 2001.
3. Dharmendra S. Sengar, "Environmental law", Prentice hall of India PVT LTD, New Delhi, 2007.
4. Rajagopalan, R, "Environmental Studies-From Crisis to Cure", Oxford University Press, 2005.
5. Benny Joseph, "Environmental Science and Engineering", Tata McGraw-Hill, New Delhi, 2006
6. https://en.wikipedia.org/wiki/Carbon_capture_and_storage
7. Ravikrishnan “Environmental Science and Engineering” Sri Krishna Hi-tech Publishing

2301MA401	DISCRETE MATHEMATICS AND GRAPH THEORY	L	T	P	C
		3	0	0	3

PREREQUISITE:

1. Basic concepts of functions and algebra.
2. Basic concept of Graph theory.

COURSE OBJECTIVES:

To introduce concepts of mathematical logic for analyzing propositions and proving theorems.

To work with relations and investigate their properties.

To investigate functions as relations and their properties.

To introduce basic concepts of graphs, digraphs and trees.

COURSE OUTCOMES:

On the successful completion of the course, students will be able to

CO1:	Correlate the properties of different kinds of functions and solve recurrence relations.
CO2:	Compute the validity of logical arguments and construct simple mathematical proofs.
CO3:	Determine whether given graphs are isomorphic and apply Dijkstra's algorithm to find the shortest path.
CO4:	Explain the concepts and properties of algebraic structures such as groups, rings and fields.
CO5:	Apply the concept and significance of lattice and Boolean algebra in computer science.

COs Vs POs MAPPING:

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11
CO1	3	2	1	-	-	-	-	-	-	-	-
CO2	3	2	1	-	-	-	-	-	-	-	-
CO3	3	2	1	-	-	-	-	-	-	-	-
CO4	3	2	1	-	-	-	-	-	-	-	-
CO5	3	2	1	-	-	-	-	-	-	-	-

COs Vs PSOs MAPPING

COs	PSO1	PSO2	PSO3
CO1	-	-	-
CO2	-	-	-
CO3	-	-	-
CO4	-	-	-
CO5	-	-	-

COURSE CONTENTS:

MODULE I	FUNCTIONS AND RECURRENCE RELATIONS	9 Hours

Functions –Type of functions – Injective, surjective and bijective functions –Composition of functions – Inverse functions –Permutation functions - Recurrence relations-Solving linear recurrence relations.

MODULE II	**LOGIC**	**9 Hours**
Propositions- Logical operators- Normal forms –Rules of inference-Consistency and inconsistency-Propositional logic- Proofs-Predicates- Quantifiers- Universe of discourse – Logical equivalences and implications for quantified statements-Rules of specification and generalization – Validity of arguments.		
MODULE III	**GRAPH THEORY**	**9 Hours**
Graphs- Types of graphs- Matrix representation of graphs- Graph isomorphism- Walk – Path - Cycles-Eulerian graphs -Hamiltonian graphs- Planar graphs- Euler formula- Shortest path algorithm: Dijkstra's algorithm.		
MODULE IV	**ALGEBRAIC STRUCTURES**	**9 Hours**
Algebraic systems – Semi groups and monoids - Groups – Subgroups – Homomorphism's – Normal subgroup and cosets – Lagrange's theorem – Definitions and examples of Rings and Fields.		
MODULE V	**LATTICES AND BOOLEAN ALGEBRA**	**9 Hours**
Partial ordering – Posets – Lattices as posets – Properties of lattices - Lattices as algebraic systems – Sub lattices – Direct product and homomorphism – Some special lattices – Boolean algebra.		
		TOTAL :45 HOURS

REFERENCES:

1. Rosen, K.H., "Discrete Mathematics and its Applications", 7th Edition, Tata McGraw Hill Pub. Co. Ltd., New Delhi, Special Indian Edition, 2011.
2. Tremblay, J.P. and Manohar.R, " Discrete Mathematical Structures with Applications to Computer Science", Tata McGraw Hill Pub. Co. Ltd, New Delhi, 30th Reprint, 2011.
3. Grimaldi, R.P. "Discrete and Combinatorial Mathematics: An Applied Introduction", 4th Edition, Pearson Education Asia, Delhi, 2007.
4. Lipschutz, S. and Mark Lipson., "Discrete Mathematics", Schaum's Outlines, Tata McGraw Hill Pub. Co. Ltd., New Delhi, 3rd Edition, 2010.
5. Koshy, T. "Discrete Mathematics with Applications", Elsevier Publications, 2006.
6. https://onlinecourses.nptel.ac.in/noc20_cs82/preview(Link for NPTEL/SWAYAM/MOOC Courses)

2302AS401	SOFTWARE ENGINEERING	L	T	P	C
		3	0	0	3

PREREQUISITE: NIL

COURSE OBJECTIVES:

1.	Learn the role of software, different software development models and their use in projects.
2.	Understand how to gather and document software requirements.
3.	Learn how to design software systems using UML diagrams and best design practices.
4.	Explore different ways to test software to make sure it works correctly.
5.	Learn how to manage risks and ensure software quality using industry standards like ISO-9000.

COURSE OUTCOMES:

Upon successful completion of the course, students will be able to

CO1	Infer the evolving role of software engineering**,** including its changing nature and common software myths.
CO2	Comprehend the key concepts of the software engineering process, including various process models like the Waterfall, Spiral, and Agile methodologies.
CO3	Apply the principles of software requirements engineering, distinguishing between functional and non-functional requirements, and contributing to requirements elicitation, validation, and management.
CO4	Apply design engineering concepts in developing software architectures, using UML diagrams for structural and behavioral modeling.
CO5	Examine various software testing strategies, including validation, system testing, and debugging techniques.

COURSE CONTENTS:

Module-I	**Introduction to Software Engineering**	**9 Hours**

The evolving role of software, changing nature of software, software myths. A Generic view of process: Software engineering- a layered technology, a process framework, capability maturity model integration (CMMI). Process models: The waterfall model, Spiral model and Agile methodology

Module-II	**Software Requirements**	**9 Hours**

Software Requirements: Functional and non-functional requirements, user requirements, system requirements, interface specification, software requirements document. Requirements engineering process: Feasibility studies, requirements

elicitation and analysis, requirements validation, requirements management

Module-III	**Design Engineering**	**9 Hours**

Design Engineering: Design process and design quality, design concepts, design model. Creating an architectural design: software architecture, data design, architectural styles and patterns, architectural design, conceptual model of UML, basic structural modeling, class diagrams, sequence diagrams, collaboration diagrams, use case diagrams, component diagrams.

Module-IV	**Testing Strategies**	**9 Hours**

Testing Strategies: A strategic approach to software testing, test strategies for conventional software, black-box and white-box testing, validation testing, system testing, the art of debugging. Metrics for Process and Products: Software measurement, metrics for software quality.

Module-V	**Risk and Quality management**	**9 Hours**

Risk management: Reactive Vs proactive risk strategies, software risks, risk identification, risk projection, risk refinement, RMMM. Quality Management: Quality concepts, software quality assurance, software reviews, formal technical reviews, statistical software quality assurance, software reliability, the ISO 9000 quality standards – Software process improvement – Emerging Trends in software engineering – Creating A Viable software plan.

TOTAL: 45 HOURS

COs Vs POs & PSOs MAPPING:

CO#	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PSO1	PSO2	PSO3
CO1	2	2	2	2	-	-	-	-	-	-	-	3	-	-
CO2	3	2	2	3	-	-	-	-	-	-	-	3	-	-
CO3	2	3	2	2	-	-	-	-	-	-	-	2	-	-
CO4	3	2	2	2	-	-	-	-	-	-	-	2	-	-
CO5	3	2	2	2	-	-	-	-	-	-	-	2	-	-

REFERENCES:

1. Roger S. Pressman, "Software Engineering, A practitioner's Approach", 6th edition, McGraw Hill International Edition, 2004
2. Ian Sommerville, "SoftwareEngineering", 9thedition, Pearson Education, 2011
3. Grady Booch, James Rambaugh and Ivar Jacobson, "The unified modeling language user guide", Pearson Education, 1999
4. James F. Peters and Witold Pedrycz, "Software Engineering, an Engineering approach", 1999
5. Waman S Jawadekar, "Software Engineering principles and practice" , The McGraw-Hill Companies, 2004
6. Meiler page-Jones, "Fundamentals of object-oriented design using UML", Pearson Education, 1999
7. https://onlinecourses.nptel.ac.in/noc20_cs68/preview
8. https://archive.nptel.ac.in/courses/106/101/106101061

2302AS402	DATA MINING CONCEPTS	L	T	P	C
		3	0	0	3

PREREQUISITE: 2302AS302- Database Systems

COURSE OBJECTIVES:

1.	To train the basic concepts and techniques of Data Mining.
2.	To introduce mathematical statistics foundations of the Data Mining Algorithms.
3.	To include a wide range of clustering, estimation, prediction, and classification algorithms.
4.	To experiment basic principles, concepts and applications of cluster analysis.

COURSE OUTCOMES:

Upon successful completion of the course, students will be able to

CO1	Comprehend the data preprocessing techniques and data cubes for data warehouses.
CO2	Apply association rule mining to address real-world data mining challenges.
CO3	Use classification techniques to tackle practical data mining problems.
CO4	Implement clustering approaches to resolve real-time data mining problems.
CO5	Make use of outlier detection to improve the quality and effectiveness of the data mining process.

COURSE CONTENTS:

Module I	**Introduction to Data Mining and Data Warehouse**	**9 Hours**

Introduction, Stages of the Data Mining Process, Applications and Issues, Data preprocessing: Data cleaning, Data transformation, Data reduction, Discretization and generating concept hierarchies, Data Warehouse: Basic Concepts, Data Warehouse Modeling, Data Warehouse Design and Usage, Data Warehouse Implementation.

Module II	**Mining Frequent Patterns, Association and Correlations**	**9 Hours**

Basic concepts, Market Basket Analysis, Mining Methods: Apriori algorithm, Associations and Correlations, Mining various kinds of Association Rules, Correlation Analysis, Constraint Based Association Mining.

Module III	**Classification**	**9 Hours**

Basic Concepts, Decision Tree Induction, Bayesian Classification, Rule Based Classification, Classification by Backpropagation, Support Vector Machines**,** Associative Classification, Lazy Learners, Other Classification Methods.

Module IV	**Cluster Analysis**	**9 Hours**

Overview, Partitioning Methods, Hierarchical Methods, Density Methods, Grid Methods, Evaluation of Clustering, Probabilistic Model Based Clustering, Clustering High Dimensional Data, Clustering Graph and Network Data, Clustering with Constraints

Module V	**Outlier detection**	**9 Hours**

Outlier and Outlier Analysis, Outlier Detection Techniques, Statistical Approaches, Proximity-based Approaches, Clustering/Classification based Approaches, Mining Contextual and Collective Outliers, Outlier Detection in High-Dimensional Data, Data Mining Applications, Case Study: Data Mining for Intelligence

TOTAL: 45 HOURS

COs Vs POs & PSOs MAPPING:

CO#	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PSO1	PSO2	PSO3
CO1	3	3	3	2	3	-	-	-	-	-	-	3	2	-
CO2	3	3	2	3	-	-	-	-	-	-	3	3	-	-
CO3	3	3	3	2	3	-	-	-	-	-	-	2	3	-
CO4	3	3	2	1	3	-	-	-	-	-	3	3	2	-
CO5	3	3	3	2	3	-	-	-	-	-	-	3	1	-

REFERENCES:

1. Jiawei Han, Micheline Kamber and Jian Pei, “Data Mining concepts and techniques”, 3rd edition, Morgan Kauffman publishing, 2012.
2. Ian. H.Walton and Eibe Frank, “Data Mining Practical machine learning tools and techniques", 4th edition, Morgan Kauffman publishing, 2017.
3. Gareth James, Daniela Witten, Trevor Hastie, Rob Thibshrani, “an Introduction to Statistical Learning”, Springer Publishing, 2013.
4. Bertrand Clarke, Ernest Fokoue, Hao Helen Zhang, “Principles and Theory for Data Mining and Machine Learning”, Springer Publishing, 2011.
5. Galit Shmueli, Peter C.Bruce,“Data Mining for Business Analytics: Concepts, Techniques, and Applications in R”, Wiley Publishers, 2017.
6. https://onlinecourses.nptel.ac.in/noc21_cs06/preview
7. https://www.coursera.org/specializations/data-mining
8. https://www.mygreatlearning.com/academy/learn-for-free/courses/data-mining
9. https://www.simplilearn.com/free-introduction-to-data-mining-course-skillup
10. https://www.udemy.com/topic/datamining/?srsltid=AfmBOopZk4ZW2KTUT9YGuwZhAjPk6WB6nC25jWAIqLT3KS-VbWN4jMqi

2302AS403	DATA ANALYTICS AND VISUALIZATION	L	T	P	C
		3	0	2	4

PREREQUISITE: 2302AS101 - Fundamentals of Data Science
2302AS302 - Database Systems

COURSE OBJECTIVES:

1.	To provide a broad overview of data analysis and visualization techniques.
2.	To build descriptive and predictive models, and validating their models against the actual outcomes.
3.	To perform data wrangling, cleaning, and sampling to get a suitable data set; exploratory data analysis; generating hypotheses and building intuition; prediction or statistical learning; communication – summarizing results through various visualization techniques and providing interpretable summaries.

COURSE OUTCOMES:

Upon successful completion of the course, students will be able to

CO1	Define and explain key data analytics and business intelligence concepts
CO2	Examine, Collect, clean, and prepare data for analysis
CO3	Perform exploratory data analysis (EDA) and descriptive statistics
CO4	Design and create effective data visualizations using industry-standard tools
CO5	Communicate data-driven insights to stakeholders in a clear and concise manner

COURSE CONTENTS:

Module I	**Introduction to Data Analytics and Business Intelligence**	**09 Hours**

Introduction, Data and its importance in business: Types of data, Data lifecycle, Data quality, Business intelligence and its applications: BI concepts and terminology, BI architecture, BI tools and technologies, Data visualization and its role in BI: Visualization principles and best practices, Popular data visualization tools.

Module II	**Data Collection and Preparation**	**09 Hours**

Data sources and collection methods: Internal and external data sources, Data collection strategies and techniques, Data cleaning and pre-processing: Data wrangling techniques, Handling missing values, Data transformation.

Module III	**Exploratory Data Analysis and Descriptive Statistics**	**09 Hours**

Exploratory data analysis (EDA): Identifying trends and patterns, Discovering relationships between variables, Outlier detection, Descriptive statistics: Measures of central tendency, Measures of variability, Data distribution analysis.

Module IV	**Data Visualization**	**09 Hours**

Data visualization principles and best practices: Choosing the right chart type for the data (Quantitative data & Qualitative data), Effective use of color, fonts, and layout, Creating interactive dashboards, Design principles for effective data visualization, Data visualization tools: Tableau, Power BI, Python libraries (matplotlib, seaborn), Other visualization tools, Best practices for presenting data visualizations.

Module V	**Communicating Data Insights**	**09 Hours**

Effective storytelling with data: Structuring a compelling narrative, Using data to support key points, Avoiding common pitfalls, Presentation skills: Delivering clear and concise presentations, Engaging your audience, Answering questions effectively.

TOTAL: 45 HOURS

LIST OF EXPERIMENTS:

1. Explore Data sources and collection methods	4 Hours
2. Explore Data cleaning and pre-processing	4 Hours
3. Perform Exploratory data analysis (EDA)	4 Hours
4. Perform Descriptive statistics	4 Hours
5. Choosing the right chart type for the data (Quantitative data & Qualitative data)	4 Hours
6. Creating interactive dashboards	4 Hours
7. Communicating Data Insights	3 Hours
8. Capstone Project	3 Hours
TOTAL: 30 HOURS	

COs Vs POs & PSOs MAPPING:

CO#	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PSO1	PSO2	PSO3
CO1	2	2	-	-	-	-	-	-	-	-	2	1	2	-
CO2	3	3	3	2	-	-	-	-	-	-	2	-	3	-
CO3	3	3	3	3	2	-	-	-	-	-	1	2	3	-
CO4	3	3	3	3	3	-	-	-	2	2	2	2	3	2
CO5	3	3	2	2	1	-	-	-	2	3	-	1	3	2

REFERENCES:

1. Skiena, Steven S, "The Data Science Design Manual", CRC press, 2017.
2. Pang-Ning Tan, Michael Steinbach, Anuj Karpatne and Vipin Kumar, "Introduction to Data Mining", Second Edition, 2016.
3. V.K. Jain, "Data Science and Analytics with Python, R and SPSS Programming", Khanna Book Publishing Company, 2025.
4. V.K. Jain, "Big Data and Hadoop", Khanna Book Publishing Company, 2022.
5. Tamara Munzner, “Visualization Analysis and Design”, A K Peters/CRC Press; 1st edition, 2014.
6. Michael Minelli, Michelle Chambers, and Ambiga Dhiraj, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses", Wiley, 2013.
7. Matthew O. Ward, Georges Grinstein and Daniel Keim, “Interactive Data Visualization: Foundations, Techniques, and Applications”, 2nd Edition, CRC press, 2015.
8. Gururajan Govindan, Shubhangi Hora and Konstantin Palagachev, “The Data Analysis Workshop”, Publisher(s): Packt Publishing, July 2020.
9. https://www.oreilly.com/library/view/the-data-analysis/9781839211386
10. https://github.com/Harvard-IACS/2019-CS109A.git
11. https://learn.microsoft.com/en-us/power-bi/learning-catalog/learning-catalog-data-analyst

2302AS404	MACHINE LEARNING TECHNIQUES	L	T	P	C
		3	0	0	3

PREREQUISITE: Nil

COURSE OBJECTIVES:

1.	To introduce the foundational concepts of machine learning, including learning tasks, decision tree learning, and hypothesis space search.
2.	To equip students with the knowledge of supervised learning techniques and unsupervised learning algorithms.
3.	To explore reinforcement learning concepts, including Q-learning, temporal-difference learning, and dynamic programming applications.
4.	To enable students to design and optimize neural networks using advanced methods like backpropagation, regularization, and hyper parameter tuning.

COURSE OUTCOMES:

Upon successful completion of the course, students will be able to

CO1	Infer the fundamentals of machine learning, including learning problems, decision tree learning, and hypothesis space exploration.
CO2	Apply regression and classification techniques in supervised learning to solve real-world problems and evaluate performance.
CO3	Implement unsupervised learning algorithms and ensemble methods for clustering and data analysis.
CO4	Analyze various dimensionality reduction techniques and the fundamentals of Neural Networks.
CO5	Analyze reinforcement learning concepts, including Q-learning, temporal-difference learning, and dynamic programming

COURSE CONTENTS:

Module-I	**Basics of Machine Learning**	**9 Hours**

Learning Problems Perspectives and Issues -Concept Learning Task -Version Spaces and Candidate Eliminations – Inductive bias – Decision Tree learning – Representation – Algorithm –Hypothesis Space Search.

Module-II	**Supervised Learning**	**9 Hours**

Supervised learning: Regression models: Simple Linear Regression, Multiple linear Regression, Cost Function, Gradient Descent, Performance Metrics. Classification models: Decision Trees-ID3, CART, K-Nearest-Neighbours (KNN), Multinomial Logistic Regression - Support vector machine, Decision Tree, Random Forests.

Module-III	**Unsupervised Learning**	**9 Hours**

Combining multiple learners: Model combination schemes, Voting, Ensemble Learning - Bagging, boosting, stacking, Unsupervised learning: K-means, Instance Based Learning: KNN, Gaussian mixture models and Expectation maximization.

Module-IV	**Dimensionality Reduction & Neural Networks**	**9 Hours**

Dimensionality reduction techniques: PCA, LDA, ICA. Neural Networks - Perceptron - Activation functions, Introduction to Deep Learning, Feature Representation Learning, GPUs in Deep Learning Networks.

Module-V	**Reinforcement Learning**	**9 Hours**

Basics of RL - Learning Task - Q-Learning - Non deterministic Rewards and actions - Temporal-difference learning - Relationship to Dynamic Programming - Active reinforcement learning - Generalization in reinforcement learning. Case studies: Autonomous Driving by Tesla and Waymo, Robotics - OpenAI's Dactyl.

TOTAL: 45 HOURS

COs Vs POs & PSOs MAPPING:

CO#	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PSO1	PSO2	PSO3
CO1	3	3	2	2	2	-	-	-	-	-	-	2	2	-
CO2	3	3	3	3	3	-	-	-	-	-	-	3	2	-
CO3	3	3	3	3	3	-	-	-	-	-	2	3	2	-
CO4	3	3	3	3	3	2	-	-	-	2	2	3	3	-
CO5	3	3	3	3	3	2	-	-	-	3	3	3	3	-

REFERENCES:

1. Tom M. Mitchell, “Machine Learning”, McGraw-Hill, 2010
2. EthemAlpaydin, “Introduction to Machine Learning (Adaptive Computation and Machine Learning)”, The MIT Press, 2004
3. Kevin Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012.
4. Bishop, Christopher. Neural Networks for Pattern Recognition. New York, NY: Oxford University Press, 1995
5. https://www.udacity.com/course/intro-to-machine-learning--ud120
6. https://www.coursera.org/learn/machine-learning-duke
7. https://onlinecourses.nptel.ac.in/noc23_cs18/preview
8. https://onlinecourses.nptel.ac.in/noc20_cs49/preview

2302AS451	DATA MINING LABORATORY	L	T	P	C
		0	0	2	1

PREREQUISITE: 1902AS302–Database Management Systems

COURSE OBJECTIVES:

1.	To learn to perform data mining tasks using a data mining toolkit
2.	To demonstrate the working of algorithms for data mining tasks such association rule mining, classification, clustering and regression.
3.	To obtain Practical Experience Working with all real datasets.

COURSE OUTCOMES:

Upon successful completion of the course, students will be able to

CO1	Comprehend the data preprocessing techniques and data cubes for data warehouses.
CO2	Apply association rule mining to address real-world data mining challenges.
CO3	Use classification techniques to tackle practical data mining problems.
CO4	Implement clustering approaches to resolve real-time data mining problems.
CO5	Make use of outlier detection to improve the quality and effectiveness of the data mining process.

LIST OF EXPERIMENTS:

1.Import an excel file in Rapid miner	3 Hours
2.Reduce the size of a data set by 50% using macros	3 Hours
3.Pre-processing Techniques on Data Set	3 Hours
4.Generate Association rules using Aprori algorithm	3 Hours
5.Generate decision tree for given dataset using Rapid miner	3 Hours
6.Implement Naive bayes classifier on a given data set	3 Hours
7.Exploring SVM classification on a given data set using Rapid miner	3 Hours
8.Implement K-Means Clustering for given dataset using Rapid miner	3 Hours
9.Detect outlier on a given data set using Rapid miner	3 Hours
10.Implementation of KNN- Cross validation	3 Hours
TOTAL: 30 HOURS	

COs Vs POs & PSOs MAPPING:

CO#	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PSO1	PSO2	PSO3
CO1	3	3	3	2	3	-	-	-	-	-	-	3	2	-
CO2	3	3	2	3	-	-	-	-	-	-	2	3	-	-
CO3	3	3	3	2	3	-	-	-	-	-	-	2	3	-
CO4	3	3	2	1	3	-	-	-	-	-	3	3	2	-
CO5	3	3	3	2	3	-	-	-	-	-	-	3	1	-

REFERENCES:

1. Jiawei Han, Micheline Kamber and Jian Pei, "Data Mining concepts and techniques", 3rd edition, Morgan Kauffman publishing, 2012
2. Ian. H.Walton and Eibe Frank, "Data Mining", Practical machine learning tools and techniques, 3rd edition, Morgan Kauffman publishing , 2017
3. Gareth James, Daniela Witten, Trevor Hastie and Rob Thibshrani, "an Introduction to Statistical Learning", Springer Publishing, 2013
4. Bertrand Clarke, Ernest Fokoue and Hao Helen Zhang, "Principles and Theory for Data Mining and Machine Learning", Springer Publishing, 2011
5. Galit Shmueli and Peter C.Bruce, "Data Mining for Business Analytics: Concepts, Techniques, and Applications in R", Wiley Publishers, 2017.
6. Peter C.Bruce, " Data Mining for Business Analytics: Concepts, Techniques, and Applications with XLMiner", 3rd edition, Wiley, 2016
7. https://academy.rapidminer.com/learning-paths/get-started-with-rapidminer-and-machine-learning

2302AS452	MACHINE LEARNING LABORATORY	L	T	P	C
		0	0	2	1

PREREQUISITE: Nil

COURSE OBJECTIVES:

1.	To understand pattern classification algorithms to classify multivariate data
2.	To understand the implementation of genetic algorithms
3.	To gain knowledge about Q-Learning
4.	To create new machine learning techniques.

COURSE OUTCOMES:

Upon successful completion of the course, students will be able to

CO1	Recognize the characteristics of Machine Learning techniques that enable to solve real world problems
CO2	Apply supervised and unsupervised algorithms, probabilistic and evolutionary approaches for the given problems
CO3	Design neural network to solve classification and function approximation problems
CO4	Analyze optimal classifiers using genetic algorithms
CO5	Perform Evaluation of Machine Learning algorithms and Model Selection.

LIST OF EXPERIMENTS [SUGGESSTED]

1. Implement ML models using SVM	3 Hours
2. Implement ML models using KNN	3 Hours
3. Implement ML models using K-Means	3 Hours
4. Implement ML models using Logistic Regression	3 Hours
5. Implement ML models using Linear Regression	3 Hours
6. Implementation of Bayesian networks	3 Hours
7. Implement Naïve Bayes theorem to classify the English text	3 Hours
8. Implement an algorithm to demonstrate the significance of genetic algorithm	3 Hours
9. Implement the finite words classification system using Backpropagation algorithm	3 Hours
10. Train a neural network using TensorFlow or PyTorch.	3 Hours
	Total Hours:30

Mode of Assessment: Continuous Assessment, PAT, ESP

COs Vs POs & PSOs MAPPING:

CO#	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PSO1	PSO2	PSO3
CO1	3	3	2	2	2	-	-	-	-	-	-	2	2	-
CO2	3	3	3	3	3	-	-	-	-	-	-	3	2	-
CO3	3	3	3	3	3	-	-	-	-	-	2	3	2	-
CO4	3	3	3	3	3	2	-	-	-	-	2	3	33	-

CO5	3	3	3	3	3	2	-	-	-	-	3	3	33	-

REFERENCES:

1. Tom M. Mitchell, “Machine Learning”, McGraw-Hill, 2010
2. Bishop, Christopher, "Neural Networks for Pattern Recognition", New York, NY: Oxford University Press, 1995
3. EthemAlpaydin, “Introduction to Machine Learning (Adaptive Computation and Machine Learning)”, The MIT Press, 2004
4. Tastie, R. Tibshirani and J. H. Friedman, “The Elements of Statistical Learning”, Springer(2nded.), 2009
5. https://www.udacity.com/course/intro-to-machine-learning--ud120
6. .https://www.coursera.org/learn/machine-learning-duke
7. https://onlinecourses.nptel.ac.in/noc23_cs18/preview
8. https://onlinecourses.nptel.ac.in/noc20_cs49/preview

2301MC403	INNOVATION & ENTREPRENEURSHIP FUNDAMENTALS	L	T	P	C
		3^	0	0	3^

PREREQUISITE:

The course assumes no prior skill or background in design, art or engineering. It is open to all undergraduates and graduate students with an interest in learning design thinking, and is especially recommended for those students planning social-venture and other kinds of design interventions

COURSE OBJECTIVES:

1.	Cultivate the mindset and skills of successful entrepreneurs
2.	Lead innovative teams
3.	Develop and refine your strategy in today's fast-changing, dynamic markets
4.	Grow your customer base through inbound and outbound marketing

COURSE OUTCOMES:

Upon successful completion of the course, students will be able to

CO1	Explain the basics of Entrepreneurship & Innovation
CO2	Analyze Leadership Styles and compare them
CO3	Choose business models based on the requirement and justify with cases
CO4	Develop a method or mechanism for Innovation marketing and sustainability
CO5	Develop a Business Model and Strategy framework and demonstrate through presentation

COURSE CONTENTS:

Module-I	**Entrepreneurship Basics**	**9 Hours**

Entrepreneurship Basics – Skills, Mindset, Myth vs Fact, Entrepreneurial Leadership: Navigating Uncertainty, Critical lessons in entrepreneurial leadership, innovation, teamwork, communications, and problem-solving & Risk management. Business Opportunity Identification, Idea Validation, Case Study : Entrepreneurs Story

Module-II	**Innovation & Creativity**	**9 Hours**

Analyzing the Current Business Scenario, Innovation and Creativity- An Introduction, Innovation in Current Environment, Types of Innovation, School of Innovation. Challenges of Innovation, Steps of Innovation Management, Idea Management System, Divergent V/s Convergent Thinking, Design Thinking and Entrepreneurship

Module-III	**Business Models & Strategies for Innovation**	**9 Hours**

Experimentation in Innovation Management, Idea Championship, Participation for Innovation, Co-creation for Innovation, Proto typing to Incubation. What is a Business Model, who is an Entrepreneur, Social Entrepreneurship, Blue Ocean Strategy-I, Blue Ocean Strategy-II

Module-IV	**Marketing & Sustainability of Innovation**	**9 Hours**

Marketing of Innovation, Technology Innovation Process, Technological Innovation Management Planning, Technological Innovation Management Strategies, Technology Forecasting. Sustainability Innovation and Entrepreneurship, Innovation Sustainable Conditions, Innovation: Context and Pattern, SME'S strategic involvement in sustainable development, Exploration of business models for material efficiency services

Module-V	**Managing Innovation: IPR**	**9 Hours**

Management of Innovation, creation of IPR, Management of Innovation, creation of IPR, Types of IPR, Patents and Copyrights, Patents in India, Business Models and value proposition, Business Model Failure: Reasons and Remedies, Incubators: Business Vs Technology, Managing Investor for Innovation, Future markets and Innovation needs for India.

TOTAL: 45 HOURS

COs Vs POs & PSOs MAPPING:

CO#	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO 10	PO 11	PSO 1	PSO 2	PSO 3
CO1	2	-	-	-	-	-	-	-	-	-	3	-	-	-

CO2	**2**	**-**	**-**	**-**	**-**	**2**	**2**	**2**	**2**	**-**	**3**	**-**	**-**	**-**
CO3	**2**	**-**	**-**	**-**	**-**	**2**	**2**	**2**	**2**	**-**	**3**	**-**	**-**	**-**
CO4	**2**	**-**	**-**	**-**	**-**	**2**	**2**	**2**	**2**	**-**	**3**	**-**	**-**	**-**
CO5	**2**	**-**	**-**	**-**	**-**	**2**	**2**	**2**	**2**	**-**	**3**	**-**	**-**	**-**

REFERENCES:

1. 8 Steps toInnovation: Going from Jugaad to Excellence- Book by Rishikesha T. Krishnan and Vinay Dabholkar
2. Innovation and Entrepreneurship Book by Peter Drucker
3. HBS series on Innovation and Entrepreneurship
4. The Lean Startup Book by Eric Ries, 2013
5. Zero to One Book by Blake Masters and Peter Thiel, 2014
6. Founders at Work: Stories of Startups' Early Days Book by Jessica Livingston, 2001
7. Crossing the Chasm Book by Geoffrey Moore, 1991
8. Hooked: How to Build Habit-Forming Products Book by Nir Eyal, 2013
9. Rework Book by David Heinemeier Hansson and Jason Fried, 2010
10. https://nptel.ac.in/courses/127/105/127105007/