

## **ARMY INSTITUTE OF TECHNOLOGY**

An Autonomous Institute Affiliated to

### Savitribai Phule Pune University, Maharashtra, India

### DEPARTMENT OF COMPUTER ENGINEERING



### NEP 2020 COMPLIANT CURRICULUM FOR

### M. Tech. Data Science

Effective from: Academic Year 2025-26

www.aitpune.com





#### VISION

To become a "Globally Recognized" technical institute providing world class education and research facilities to the wards of Defence personnel.

#### MISSION

(a) Provide the right environment, to the wards of Defence personnel, for development of physical, intellectual, emotional and spiritual quotients, with a view to produce total quality engineers.

(b) Create an ecosystem which can foster the culture of research, innovation, creative thinking and higher studies.

(c) Develop an education system which creates entrepreneurs and technology leaders who are committed towards sustainable development of society and nation building.

#### CORE VALUES

Excellence, Honesty, Integrity, Team Work, Continuous Learning and Innovation





### **Department of Computer Engineering**

#### VISION

To provide best quality education and research facilities to the wards of army personnel to produce globally competent computer engineering graduates.

#### MISSION

M1: To provide state of the art infrastructure to the wards of army personnel to produce total quality computer engineer

M2: To create an environment which can promote the culture of research, innovation, creative, thinking, and higher studies

M3: To collaborate with industry fraternity to create entrepreneurs and technology leaders who are committed towards sustainable development of information society.

The Program Educational Objectives (PEOs)

**PEO1:** Graduates of the program will be employed in industry, government and entrepreneurial endeavors to have a successful professional career.

**PEO2:** Graduates of the program will pursue higher education or research.

**PEO3:** Graduates of the program utilize the acquired technical skills and knowledge for the benefit of society.





#### **Program Outcomes (POs)**

**PO1:** Ability to independently carry out research / investigation and development work to solve practical problems.

**PO2:** Ability to write and present a technical report / document.

**PO3:** Students should be able to demonstrate a degree of mastery over the area as per the specialization of the program. The mastery should be at a level higher than the requirements in the appropriate bachelor program.

#### **Program Specific Outcomes (PSOs)**

**PSO1:** Technical: Expertise in acquiring data from diverse sources, cleaning, transforming, and preparing it for analysis. Proficiency in applying statistical methods like hypothesis testing, regression analysis, time series forecasting, and dimensionality reduction. In-depth understanding of supervised and unsupervised learning algorithms and their implementation. Strong proficiency in essential data science programming tools and libraries. Familiarity with big data frameworks like Hadoop, Spark, and distributed computing concepts.

**PSO2:** Problem-Solving: Ability to effectively explore and visualize complex datasets to identify patterns and insights. Skill to choose the appropriate machine learning models based on problem requirements and evaluate their performance using relevant metrics. Expertise in creating new features from existing data to improve model accuracy. Knowledge of deploying models into production environments and continuously monitoring their performance.

**PSO3:** Communication and Critical Thinking: Ability to interpret data analysis results and translate complex technical concepts into actionable insights for decision-making. Skill to present data findings in clear, concise, and professional reports for diverse audiences. Awareness of ethical issues related to data collection, privacy, and responsible use of data science techniques.





Acronym	Title	No. of courses	Credits
PCC	Program Core Course	08	23
PEC	Program Elective Course	03	09
VSEC	Vocational and Skill Enhancement Course	04	12
HSS	Humanities and Social Science	05	09
AEC	Ability Enhancement Course	05	27
	Total	25	80

Semester	Credits		Minimum Requirement [NEP-2020]	
Ι	20	40	40	
II	20	40	40	
III	20	40	40	
IV	20	40	40	
Total	80		80	





	M. Tech Data Science [Effective from: Academic Year 2025-26]									
-		Semester-	·I							
T			Teac Sch Ho W	ching eme urs/ eek	]	Exam a	inatio nd M	n Sch arks	eme	Credits
Туре	Course Code	Course	Theory	Practical	CIE	End-Sem	ΤW	OR/ PR	Total	cicuits
PCC	MCE25611A0A	Mathematical Foundations for Data Science	03		50	50		-	100	03
PCC	MCE25612A0A	Basics of Data Science	03		50	50		-	100	03
PCC	MCE25613A0A	Data Preparation, Analysis, Modeling and Visualization	03		50	50		-	100	03
PCC	MCE25614A0A	Essential Tools for Data Science	02		50			-	50	02
	MCE25615B1A	Artificial Intelligence for Data Science	- 02							
PEC	MCE25615B2A	Business Analytics	05		50	50		-	100	03
	MCE25615B3A	Information Systems Management								
VSEC	MCE25616A0B	Laboratory Proficiency-I [Data Preparation, Analysis, Modeling and Visualization]		06			50	50	100	03
VSEC	MCE25617A0B	Laboratory Proficiency-II [Essential Tools for Data Science]		06			50	50	100	03
		Total	14	12	250	200	100	100	650	20

**Elective-I:** 

MCE25615B1A: Artificial Intelligence for Data Science

MCE25615B2A: Business Analytics

MCE25615B3A: Information Systems Management



### ARMY INSTITUTE OF TECHNOLOGY, PUNE An Autonomous Institute, Affiliated to Savitribai Phule Pune University



	Semester II									
Trues			Teac Sch Ho W	ching eme urs/ eek	]	Exam a	inatio nd M	n Sch arks	eme	
Туре	Course Code	Course	Theory	Practical	CIE	End-Sem	ML	OR/ PR	Total	Credits
PCC	MCE25621A0A	Machine Learning	03		50	50		-	100	03
PCC	MCE25622A0A	Big Data Analytics	03		50	50		-	100	03
HSS	MCE25623A0A	Research Methods for Engineers	03		50	50	-	-	100	03
	MCE25624B1A	Recommender Systems								
PEC	MCE25624B2A	Video Analytics	03		50	50		-	100	03
	MCE25624B3A	Real Time Analytics								
VSEC	MCE25625A0B	Laboratory Proficiency-III [Big Data Analytics]		06			50	50	100	03
VSEC	MCE25626A0B	Laboratory Proficiency-IV [Machine Learning]		06			50	50	100	03
AEC	MCE25627A0B	Mini Project with Seminar		02			25	25	50	01
HSS	MCE25628A0A	Value Education	01		25				25	01
		Total	13	14	225	200	125	125	675	20
Elective-II	Elective-II:									
MCE25624B1A: Recommender Systems										
MCE25624	MCE25624B2A: Video Analytics									
IVICE23024	+doa: keai time	Апациися								





	Semester-III									
			Teac Sch Hou We	ching eme rs/ ek	E	xamin	ation S Mari	Schem ks	e and	Credits
Гуре	Course Code	Course	Theory	Practical	CIE	End-Sem	ΜT	OR/ PR	Total	cicuits
PCC	MCE25631A0A	Data Storage Technologies and Networks	03		50	50		-	100	03
PCC	MCE25632A0A	Data Warehousing and Mining	03		50	50		-	100	03
	MCE25633B1A	Soft Computing								
PEC	MCE25633B2A	Deep Learning	03		50	50		-	100	03
	MCE25633B3A	Computational Linguistic Analytics								
AEC	MCE25634A0B	Industry Internship-1 OR In-house Research Project-1		06			50		50	03
AEC	MCE25635A0B	Dissertation Stage-1		14			50	50	100	07
HSS	MCE25636A0A	Personality Development through Life Enlightenment Skills.	01		25				25	01
		Total	10	20	175	150	100	50	475	20

Elective-III:

MCE25633B3A: Computational Linguistic Analytics

MCE25633B2A: Deep Learning

MCE25633B1A: Soft Computing



### ARMY INSTITUTE OF TECHNOLOGY, PUNE An Autonomous Institute, Affiliated to Savitribai Phule Pune University



	Semester-IV									
			Teac Sche Hou We	ching me rs/ ek	E	xamin	ation S Mar	Schem ks	e and	
Туре	Course Code	Course	Theory	Practical	CIE	End-Sem	MT	OR/ PR	Total	Credits
HSS	MCE25641A0A	English for Research Paper Writing	02		50				50	02
HSS	MCE25642A0A	Intellectual Property Rights	02		50				50	02
AEC	MCE25643A0B	Industry Internship-2 OR In-house Research Project-2		06			50		50	03
AEC	MCE25644A0B	Dissertation Stage-2		26			150	150	300	13
		Total	04	32	100		200	150	450	20

**Exit option to qualify for PG Diploma in Data Science:** Students who exit at the end of the first year after successful completion of all first-year courses (Semester I & II) shall be awarded a post graduate Diploma in Data Science by the institute.





Index

Subject Name	Page No.
SEM-I	
MCE25611A0A: Mathematical foundation for Data Science	13
MCE25612A0A: Basics of Data Science	14
MCE25613A0A: Data Preparation, Analysis, Modelling and Visualization	15
MCE25614A0A: Essential Tools for Data Science	17
MCE25615B1A: Artificial Intelligence and Data Science	18
MCE25615B2A: Business Analytics	19
MCE25615B3A: Information Systems Management	21
MCE25616A0B: Laboratory Proficiency-I [Data Preparation, Analysis, Modelling and	22
Visualization]	
MCE25617A0B: Laboratory Proficiency-II [Essential Tools for Data Science]	23
SEM-II	-
MCE25621A0A: Machine Learning	26
MCE25622A0A: Big Data Analytics	27
MCE25623A0A: Research Methods for Engineers	28
MCE25624B1A: Recommender Systems	30
MCE25624B2A: Video Analytics	31
MCE25624B3A : Real Time Analytics	33
MCE25625A0B: Laboratory Proficiency-III [Big Data Analytics]	34
MCE25626A0B: Laboratory Proficiency-IV [Machine Learning]	35
MCE25627A0B: Mini Project with Seminar	36
MCE25628A0A: Value Education	36
SEM-III	
MCE25631A0A: Data Storage Technologies and Networks	39
MCE25632A0A: Data Warehousing and Mining	41
MCE25633B1A: Soft Computing	42
MCE25633B2A: Deep Learning	43
MCE25633B3A: Computational Linguistic Analytics	45
MCE25634A0B: Industry Internship-1 OR Inhouse Research Project-1	46
MCE25635A0B: Dissertation Stage–I	47
MCE25636A0A: Personality Development through Life Enlightenment Skills.	48
SEM-IV	
MCE25641A0A: English for Research Paper Writing	50
MCE25642A0A: Intellectual Property Rights	51
MCE25643A0B: Industry Internship-2 OR In-house Research Project-2	52
MCE25644A0B: Dissertation Stage II	52





#### Preface

The New Education Policy (NEP) 2020 has ushered a new era of change, in India's education system to better meet the needs of the 21st century. SPPU is committed to the effective and fruitful implementation of NEP 2020 in its true spirits emphasizing holistic and multidisciplinary education as per the directives of Maharashtra government. It emphasizes a multidisciplinary approach, aiming to develop critical thinking and creativity, thereby contributing to the holistic development of individuals.

We are delighted to present the M. Tech Data Science syllabus -2025 pattern, which has been meticulously designed in alignment with the NEP 2020 with effect from academic year 2025-26. This curriculum aims to provide students with a holistic approach to engineering education, ensuring a strong foundation in Program Core Courses and Ability Enhancement Courses. This curriculum also includes components of vocational and skill enhancement courses, Value Education and Co-curricular courses to shape well-rounded engineers who can adapt to global demands. Also, this document provides information on the credit system, course contents and examination and evaluation scheme along with guidelines to make best use of the curriculum designed.

The syllabus encourages experiential learning, where theoretical concepts are supported by practical laboratory sessions. Also promotes research and innovation, encouraging students to engage in projects from the early stages of their academic journey. We believe that this syllabus, crafted with the essence of the NEP 2020, will equip our students with the necessary skills and knowledge to excel in their future endeavours. We look forward to embarking on this exciting academic journey with our students.





#### **Abbreviations**

- AEC Ability Enhancement Course
- BSC Basic Science Course
- CCC Co-Curricular Courses
- CCE Comprehensive Continuous Evaluation
- CEP Common Engineering Project
- CO Course Outcomes
- ELC Experiential Learning Courses
- ESC Engineering Science Course
- FP Field Project
- IKS Indian Knowledge System
- INT Internship
- MDM Multidisciplinary Minor
- NEP National Education Policy
- OE Open Elective
- OJT On Job Training
- PCC Program Core Course
- PEC Programme Elective Course
- PO Program Outcomes
- PR Practical
- PRJ Project
- PSO Program Specific Outcomes
- RM Research Methodology
- TH Theory
- TU Tutorials
- VEC Value Education Course
- VSE Vocational and Skill Enhancement Course



# ARMY INSTITUTE OF TECHNOLOGY, PUNE An Autonomous Institute Affiliated to Savitribai Phule Pune University, Maharashtra, India

**National Education Policy (NEP) Compliant Curriculum** 

# Semester - I



M. Tech. Data Science (2025 Pattern)

www.aitpune.com





MCE25611A0A: Mathematical foundation for Data Science						
Teaching Scheme: TH: 03 hr/week	Credits:03	Examination Scheme: CIE: 50 Marks End Semester: 50 Marks				
Prerequisite: Engineering	Mathematics					
Course Objectives:						
<ul> <li>a) To introduce the funda probability.</li> </ul>	mental concepts of discrete mathematics, includ	ing set theory, graph theory, and				
b) To explore data distribution measures of central ten	bution and statistical measures, focusing on varied dency and dispersion.	ous probability distributions and				
c) To understand statistic for decision-making.	al inference techniques such as hypothesis testing	g, ANOVA, and Bayes' Theorem				
d) To study linear algeb	ra and calculus, including matrix operations,	multivariate calculus, and their				
e) To understand regress functions, and model e	ion models, focusing on simple and multivaria valuation.	able regression techniques, cost				
Course Outcomes:						
1) Students will acquire probability theory.	knowledge of discrete mathematical concept	ts like sets, graphs, and basic				
2) Students will be able measures.	to represent and analyse data using various dis	stributions and central tendency				
3) Students will gain skill data interpretation	s in statistical inference, including hypothesis ter	sting and correlation analysis for				
<ul><li>4) Students will understar</li><li>problems in modelling</li></ul>	nd linear algebra and calculus concepts, and app	ly them in solving mathematical				
<ol> <li>Students will be profice using different regression</li> </ol>	eitent in regression analysis, including making pr	edictions and evaluating models				
using different regressi	Theory: Course Contents					
Unit I	Discrete mathematics for Data Science	(08 Hours)				
Concept of set, cardinal Principle of inclusion Ex graph traversal, travelling tree.	ity of set, finite, infinite and uncountably infi clusion, Graph: Basic terminologies, representa salesperson problem, Trees: Basic terminologies	inite sets, Basic set operations, tion of graph, path and circuit, es, search tree: Binary & M-ary				
Unit II	Data Distribution	(08 Hours)				
Data Representation, Average, Spread, Experiments, Outcomes, Events, Random Variables, Probability Distributions, Mean and Variance of a Distribution, Binomial, Poisson, and Hyper Geometric Distributions, Normal Distribution, Distributions of Several Random Variables. Measures of Central Tendency: Mean, Median, Mode, Mid- range, Measures of Dispersion: Range, Variance, Mean Deviation, Standard Deviation.						
Unit III	Statistical Inference	(08 Hours)				
Coefficient of variation: Moments, Skewness, Kurtosis, one sample hypothesis testing, hypothesis, Testing of Hypothesis, Chi-Square Tests, t-test, ANOVA. Measure of Relationship: Covariance, Pearson Correlation, Measures of Position: Percentile, Z-score, Quartiles, Bayes' Theorem, Bayes Classifier, Bayesian network.						
Unit IV	Linear Algebra and Calculus	(08 Hours)				
Linear Algebra: Matrix an Matrix factorization conce	d vector algebra, systems of linear equations usir pt/LU decomposition, Eigen values and eigenve	ng matrices, linear independence, ctors, Understanding of calculus:				





concept of function and derivative, Multivariate calculus: concept, Partial Derivatives, chain rule, the Jacobian and the Hessian

Unit V	Regression Model	(08 Hours)							
Introduction, types of regression. Simple regression- Types, Making predictions, Cost function, Gradient									
descent, Training, Model e	descent, Training, Model evaluation. Multivariable regression: Growing complexity, Normalization, making								
predictions, initialize weig	thts, Cost function, Simplifying with matrices, Bi	as term, Model evaluation							
References:									
1. Applied Statistics and	Probability for Engineers, Douglas C. Montgo	mery, George C. Runger, 2018,							
Wiley									
2. Introduction to. Mathematics. Statistics. Robert V. Hogg. Allen T. Craig, Pearson Education									
2 Probability and Statistic	of for Engineers Dichard A. Johnson Invin Mille	r John Fraund							

- 3. Probability and Statistics for Engineers. Richard A. Johnson, Irwin Miller, John Freund
- 4. Mathematical Statistics with Applications. Irwin Miller, Marylees Miller, Pearson Education
- 5. Elements of discrete mathematics.,Liu, Chung Laung. Tata McGraw-Hill Education

MCE25612A0A: Basics of Data Science								
Teaching Scheme: TH: 03 hr/week	ing Scheme: 3 hr/week Credits: 03 Examination Scheme: End Semester: 50 Marks End Semester: 50 Marks							
Prerequisites: Basic Mathe	ematics and Statistics							
Course Objectives:								
a) To introduce the basics	s of data analysis, including types of data, data dis	tribution, and outlier treatment.						
b) To explore machine le	earning algorithms, such as linear regression, K	NN, decision trees, and random						
forests.								
c) To understand the pri- effectively.	inciples and techniques of data visualization a	nd apply them to present data						
d) To develop an unders	tanding of recommendation systems, including	content-based and collaborative						
e) To explore social netw	ork analysis, focusing on graph theory and cluster	ering methods for social network						
data.								
Course Outcomes:								
1) Students will gain the calculating basic statistic	ne ability to analyze various types of data, i tical measures.	ncluding handling outliers and						
2) Students will develop classification and clust	a strong understanding of machine learning algoring tasks.	orithms and their applications in						
3) Students will be able patterns and outliers.	to apply visualization techniques to effectively	represent data and identify key						
4) Students will learn to b	build and evaluate recommendation systems using	content-based and collaborative						
filtering methods.								
5) Students will acquire	skills in social network analysis and be able	to analyze and interpret social						
network graphs and co	mmunities.							
	Theory: Course Contents							
Unit I	Data Analysis	(08 Hours)						
Types of Data: Structure	d vs. Unstructured Data, Quantitative vs. Categ	orical Data, Big Data vs. Little						
Data, Data science process, Role of Data Scientist. Introduction-Population and samples, Data Distribution,								
Outlier Treatment, Measuring Symmetry, Continuous Distribution, Kernel Density, Estimation: Sample and								
Estimated Mean, Variance	e and Standard Scores, Covariance, and Spearman	's Rank Correlation						





Unit II	Machine Learning Algorithms	(08 Hours)						
Linear Regression, K-Nearest Neighbor(KNN) Algorithm, K-means clustering, Naive Bayes classifier,								
Spam Filtering using N	aive Bayes classifier, Decision Trees in Macl	hine Learning, Random Forest						
Algorithm								
Unit III	Data Visualization	(08 Hours)						
Visual Encoding, Planar V	Ariables, Retinal Variables, Mapping features to	encodings, Data Representation,						
Dot Plot, Stem and Leaf	Plot, Bar Chart, Histogram, Difference Between	Bar Chart and Histogram, Box						
and Whisker Plot, Identify	ving Outliers							
Unit IV	Recommendation Systems	(08 Hours)						
The Utility Matrix, The	Long Tail, Applications of Recommendation S	Systems, Populating the Utility						
Matrix, Content-Based R	ecommendations: Item Profiles, Discovering Fea	atures of Documents, Obtaining						
Item Features From Tags,	Representing Item Profiles, User Profiles, Record	mmending Items to Users Based						
on Content, Collaborativ	e Filtering: Measuring Similarity, Clustering U	Jsers and Items, Evaluation of						
Recommendation System								
Unit V	Social Network Analysis	(08 Hours)						
Social Networks as Graph	is, Varieties of Social Networks, Graphs With Se	veral Node Types, Clustering of						
Social-Network Graphs:	Distance Measures for Social-Network Graphs	, Applying Standard Clustering						
Methods, Betweenness, T	he Girvan-Newman Algorithm, Using Betweenne	ss to Find Communities						
References:								
1. J. Leskovek, A. Rajaran	nan and J. Ullman, Mining of Massive Datasets,	Cambridge University Press.						
2. C. O'Neil and R, Schutt, Doing Data Science, Straight Talk from the Frontline, O'Reilly.								
3. L. Igual and S. Segui,	Introduction to Data Science: A Python Approa	ch to Concepts, Techniques and						
Applications, Springer.								
L Joal Crus Data Science from Scratch: First Principles with Puthon Shroff/O'Pailly								

- 4. Joel Grus, Data Science from Scratch: First Principles with Python, Shroff/O'Reilly
- 5. Lillian Peirson, Data Science for Dummies, Wiley

MCE25613A0A: Data Preparation, Analysis, Modelling and Visualization						
Teaching Scheme: TH: 03 hr/weekCredits: 03Examination Scheme: CIE: 50 Marks End Semester: 50 Marks						
Prerequisites: Basics of F	Probability theory, statistics, Python/R Progra	mming				
Course Objectives:						
<ul> <li>a) To introduce students to data mining, knowledge discovery, and various data pre-processing methods.</li> <li>b) To teach students how to perform statistical analysis and evaluation of data mining techniques.</li> <li>c) To provide students with skills in data preparation, including cleaning, normalization, and transformation.</li> <li>d) To develop students' abilities to visualize data effectively and understand the principles of data arrangement.</li> <li>a) To have to develop and the principles of data arrangement.</li> </ul>						
multidimensional data.						
1) Students will unders	tand the basic concepts of data mining t	pre-processing and different learning				
<ul> <li>2) Students will understand the basic concepts of data mining, pre-processing, and different learning paradigms.</li> <li>2) Students will be able to apply statistical tests for comparing data mining methods and evaluate their performance.</li> </ul>						





- 3) Students will gain the ability to clean, normalize, and transform data for machine learning models.
- 4) Students will be proficient in visualizing data, understanding its structure, and effectively representing data with various visualization techniques.
- 5) Students will be able to analyse and interpret patterns in static and moving data using advanced perceptual theories.

Theory: Course Contents				
Unit I	Introduction to Data Preprocessing	(08 Hours)		
Data mining and knowle	edge discovery, data mining methods, supervised lea	arning, unsupervised learning,		
other learning paradigms	, introduction to data preprocessing- data preparation,	data reduction.		
Unit II	Data Sets and Proper Statistical Analysis of Data Mining Techniques.	(08 Hours)		
Data sets and partitions methods- conditions for algorithms, non-parametri tests for multiple compar	- data set partitioning, performance measures, usin the safe use of parametric tests, normality test over ric tests for comparing two algorithms in multiple dat isons among more than two algorithms	g statistical tests to compare er the group of data sets and ta set analysis, non-parametric		
Unit III	Data Preparation Basic Models	(08 Hours)		
Data preparation basic models, data integration, data cleaning, data normalization, transform and enrich data: data transformation, linear transformations, quadratic transformations, non-polynomial approximations of transformations, rank transformations, box-cox transformations, spreading the histogram, nominal to binary transformation.				
Unit IV	Data Visualization, Arrange Tables, Networks and Trees	(08 Hours)		
Need for data visualization, types of data, stages of data visualization, human visual perception and cognition. why arrange? arrange by keys and values, express: quantitative values- scatterplots, separate, order, and align: categorical regions- list alignment: one key, matrix alignment: two keys, volumetric grid: three keys, recursive subdivision: multiple keys, spatial axis orientation, spatial layout density, arrange networks and trees- connection: link marks, matrix views, costs and benefits: connection versus matrix, containment: hierarchy marks				
Unit V	Static and Moving Data	(08 Hours)		
Gestalt laws, texture theory and data mapping, perception of transparency; overlapping data, perceiving patterns in multidimensional discrete data, patterns in motion				
<ol> <li>References:         <ol> <li>Salvador García, Julián Luengo, Francisco Herrera, "Data Preprocessing in Data Mining</li> <li>Colin Ware, Information visualization perception for design, MK publication</li> <li>Tamara Munzner, Visualization analysis and design.</li> <li>Mark Gardner, "Beginning R: The Statistical Programming Language, Wrox Publication,</li> <li>David Dietrich, Barry Hiller, Data Science and Big Data Analytics, EMC education services, Wiley Publications</li> </ol> </li> </ol>				





MCE25614A0A: Essential Tools for Data Science			
Teaching Scheme: TH:02 hr/week	Credits:02	Examination Scheme: CIE: 50 Marks	
Prerequisites: Knowledge	of programming		
<ul> <li>Course Objectives:</li> <li>a) To introduce students to the Python programming language, IDEs, and essential programming concepts like OOP, file handling, and web scraping.</li> <li>b) To provide a solid understanding of data pre-processing techniques, including feature engineering, scaling, and visualization using Python libraries.</li> <li>c) To familiarize students with Weka, R programming, and TensorFlow for data pre-processing, model optimization, and machine learning model development.</li> <li>d) To teach students advanced Excel functions, data visualization, and analytics using Power BI and Tableau for better data insights.</li> <li>e) To expose students to Qlik Sense, SAS, and Google Data Studio for data loading, visualization, and reporting.</li> </ul>			
<ol> <li>Course Outcomes:</li> <li>Students will be able handling, and NumPy</li> <li>Students will apply da Matplotlib and Seaborn</li> <li>Students will use Wel machine learning mode</li> <li>Students will demons visualization, and repo</li> <li>Students will utilize Q report generation in real</li> </ol>	to write Python code using basic and advance for data manipulation. ta pre-processing techniques and create visualizan. xa, R, and TensorFlow for data manipulation, p els. strate proficiency in using Excel, Power BI, rting. lik Sense, SAS, and Google Data Studio for data al-world projects.	ed features like OOP, exception ations using Python libraries like re-processing, and implementing and Tableau for data analysis, a manipulation, visualization, and	
	Theory: Course Contents		
Unit I	Basics of Python	(06 Hours)	
Introduction to Python and IDEs The basics of the Python programming language and how you can use various IDEs for python development like Jupyter Notebook, PyCharm, Variables, Data Types, Loops, Conditional Statements, functions, decorators, lambda functions, file handling, exception handling, Object-Oriented Programming: An Introduction to OOP concepts like classes, objects, inheritance, abstraction, polymorphism, encapsulation, Web scraping and interaction with APIs, NumPy Arrays, CRUD Operations, etc. Linear Algebra: Matrix Multiplication, CRUD Operations, Inverse, Transpose, Rank, Determinant of a matrix, Scalars, Vectors, Matrices, Loading the data, dataframes, series, splitting the data, etc.			
Unit II Dat	a rre-processing and visualization using Pyth	on (06 Hours)	
Value Imputations, Outlier Analysis and Handling, VIF, Bias-variance Trade-Off, Cross Validation Techniques, train-test split, Bar charts, scatter plots, count plots, line plots, pie charts, donut charts, etc. with the Python Matplotlib Library Regression plots, categorical plots, area plots, etc., with the Python Seaborn Library, Scikitlearn library.			
	Weka, K programming and TensorFlow	(06 Hours)	
R Programming: Introduction	ction to R, Understanding R data structure, imp	porting data, Manipulating Data,	





Using functions in R, R Programming, Charts and Plots				
rensorFlow: Inti-	oduction to Tensorriow, Tensorriow Fundamentals: Tens	sor operations (addition,		
TensorFlow	snapling, shering), Loading Data for MiL Woders, Droadea	isting and data types in		
	Advanced Excel, Power RI and Tableau	(06 Hours)		
Advanced Excel:	Excel Introduction, Customizing Excel, Basic Functions, J	Formatting and Proofing.		
Mathematical Fun	ctions, Text Functions, Lookup Functions, Pivot Tables, VBA N	Aacro,		
Introduction and	overview of Tableau and Power BI, connecting to Data, Data	Preparation and Cleaning,		
Building Visualiza	tions, Calculations and Analytics in Tableau and Power BI.			
Unit V	Qlik Sense, SAS and Google Data studio	(06 Hours)		
Introduction to Q	ik Sense, Qlik Sense vs. Qlik View, Loading Data. Creating Ch	nart, Editing Visualization,		
Uses of Variables	Dynamic Loading, Script Files & Sub-routines, Advanced Script	ipt Functions, Introduction		
to SAS, SAS Fran	nework, SAS Programming Concepts, SAS Applications, data vi	isualization with SAS,		
Introduction to Go	ogle Data studio, Data Preparation and Management, Building a	and Customizing Report.		
References:				
1. R. Thareja, "Py	1. R. Thareja, "Python Programming Using Problem Solving Approach", Oxford University Press			
2. S. Arora and L	2. S. Arora and L. Malik, Data Science and Analytics with Python, Universities Press			
3. Y. Jahnavi, Classification and knowledge analysis using Weka: A Data mining approach, BlueRoseONE				
V. Batula and A. V. Krishna Prasad, Mastering Data Science Using R Programming: Unleashing Insights				
with R, BS Publications				
5. I. Hale, Practical TensorFlow: Hands-On Applications in Data Science, independently published				
N. K. Arora, Mastering Advanced Excel, BPB Publications				
/. I. Carter, Data	. T. Carter, Data Analysis with Excel and Power BI: A Beginner's Guide to Data Analysis Using Excel and			
Power BI, Ama	Power BI, Amazon Asia-Pacific Holdings			
A. Unaturvedi and Praveen Malik, Mastering Data Visualization with Tableau, BPB Publications				

9. E. Fouda, Learn Data Science Using SAS Studio:A Quick-Start Guide, Apress S. Garg, Qlik Sense for Business Intelligence, BPB Publications

	MCE25615B1A: Artificial Intelligence and Data Science			
Teaching Scheme: TH:03 hr/week		Credits:03	Examination Scheme: CIE: 50 Marks End Semester: 50 Marks	
Pr	erequisites: Data structure,	Algorithms		
Co	ourse Objectives:			
a) b) c) d) e)	<ul> <li>a) To provide a foundational understanding of artificial intelligence and intelligent agents.</li> <li>b) To explore problem-solving techniques using various search strategies and optimization problems.</li> <li>c) To examine knowledge representation, reasoning, and classical planning in AI systems.</li> <li>d) To study uncertainty and probabilistic reasoning in decision-making processes.</li> <li>e) To understand learning methods in AI and explore its applications across various industries.</li> </ul>			
Co	ourse Outcomes:			
1) 2) 3) 4) 5)	<ol> <li>Students will understand the basics of intelligent agents and their environments.</li> <li>Students will gain proficiency in problem-solving techniques using search strategies and optimization.</li> <li>Students will be able to apply knowledge representation and reasoning techniques in AI systems.</li> <li>Students will be equipped to handle uncertain knowledge and make decisions using probabilistic reasoning.</li> <li>Students will apply AI learning methods to real-world applications in various fields like healthcare and finance.</li> </ol>			





Theory: Course Contents				
Unit I	Introduction and Intelligent Agents	(08 Hours)		
History of Artificial Intellige	ence, The State of the Art Intelligent Agent	s: Agents and Environments, Good		
Behavior: The Concept of Ra	tionality, The Nature of Environments, and	The Structure of Agents.		
Unit II	Problem-solving	(08 Hours)		
Solving Problems by Sear	ching: Problem-Solving Agents, Uninfor	med Search Strategies, Informed		
(Heuristic) Search Strategies,	Heuristic Functions, Beyond Classical Sear	rch		
Local Search Algorithms and	d Optimization Problems, Local Search in (	Continuous Spaces, Searching with		
Nondeterministic Actions, S	Searching with Partial Observations, Onli	ine Search Agents and Unknown		
Environments.				
Unit III	Knowledge, Reasoning, and Planning	(08 Hours)		
Knowledge based Agents, Fi	irst-Order Logic and Its Inference, Classica	l Planning, Planning and Acting in		
the Real World, Knowledge I	Representation			
Unit IV	Uncertain Knowledge and Reasoning	(08 Hours)		
Quantifying Uncertainty, Pr	obabilistic Reasoning, Probabilistic Reaso	oning over Time, Making Simple		
Decisions, Making Complex	Decisions			
Unit V	Learning and Applications	(08 Hours)		
Learning from Examples, Knowledge in Learning, Learning Probabilistic Models, Reinforcement Learning,				
AI Applications in various fields in marketing, healthcare, banking, finance, etc.				
References:				
1. Russell S. and Norvig P., Artificial Intelligence: A Modern Approach. Prentice-Hall.				
2. Elaine Rich, Kevin Knight and Nair, Artificial Intelligence, Tata McGraw-Hill.				
3. George F. Luger Artificial Intelligence: Structures and Strategies for Solving Complex Problems,				
Pearson Education.				
4. Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning, MIT Press.				

Melanie Mitchell, Artificial Intelligence: A Guide for Thinking Humans, Farrar, Straus and Giroux

	MCE25615B2A: Business Analytics			
Te Tl	eaching Scheme: H:03 hr/week	Credits:03	Examination Scheme: CIE: 50 Marks End Semester: 50 Marks	
Pr	erequisites: Basic Statis	tics, Basic Mathematics, Basic Ma	nagement and Basics of Data Mining	
Co	ourse Objectives:			
a)	To provide an overview	v of Business Analytics, its scope,	processes, and role in organizations.	
b)	To explore descriptive	analytics techniques, including sta	tistical tools and data visualization for business	
	decision-making.			
c)	To study predictive and	lytics methods such as regression	analysis, forecasting, and trend analysis.	
d)	d) To understand prescriptive analytics, focusing on optimization models and their applications in various			
	business scenarios.			
e)	e) To learn about simulation techniques, risk analysis, and decision analysis for strategic business planning			
	and forecasting.			
Co	Course Outcomes:			
1)	) Students will understand the role of Business Analytics in organizations and how it drives competitive			
	advantages.			
2)	2) Students will apply statistical tools and methods for descriptive analytics and data exploration.			
3)	Students will model rel	ationships and trends in data using	regression analysis and forecasting techniques.	
4)	Students will solve o	ptimization problems and apply	prescriptive analytics to real-world business	





<ul> <li>problems.</li> <li>5) Students will build simulation models, assess risks, and make informed decisions using decision trees and risk analysis techniques.</li> </ul>					
	Theory: Course Contents				
Unit I	Overview of Business analytics	(08 Hours)			
Business analytics: Overv relationship of business a Organization structures o policy, outsourcing, ensur Database analytics	view of business analytics, scope of business analytics process and organization, competitive a f business analytics, team management, management, management, data quality, measuring contribution of busin	lytics, business analytics process, advantages of business analytics. ent issues, designing information less analytics, managing changes.			
Unit II	Descriptive Analytics	(06 Hours)			
Descriptive Analytics: s probability distribution resources, business analyt and exploring data, busine	atatistical tools: statistical notation, descriptive and data modeling, sampling and estimation tics personnel, data and models for business analy ess analytics technology.	statistical methods, review of methods overview. Important tics, problem solving, visualizing			
Unit III	Predictive Analytics	(08 Hours)			
Trend lines and Regression Analysis Learning Objectivess: Modeling Relationships and Trends in Data, Simple Linear Regression, Residual Analysis and Regression Assumptions, Multiple Linear Regression, Building Good Regression Models, Regression with Categorical Independent Variables, Regression Models with Nonlinear Terms. Forecasting Methods: Forecasting Techniques: Qualitative and Judgmental Forecasting, Statistical Forecasting Models, Forecasting Models for Stationary Time Series, Forecasting Models for Time Series with a Linear Trend, Forecasting Time Series with Seasonality, Regression					
Unit IV	Prescriptive Analytics	(08 Hours)			
Linear Optimization : Optimization Models, Linear Optimization in Bank Financial Planning, Analytics in Practice: Using Optimization Models for Sales Planning at NBC, Developing Linear Optimization Models, Identifying Decision Variables, the Objectives, and Constraints, Developing a Mathematical Model, Implementing Linear Optimization Models, Solving Linear Optimization Models, Graphical Interpretation of Linear Optimization with Two Variables, Applications of Linear Optimization, Integer Linear Optimization Models, Models with Binary Variables, Nonlinear Optimization Models					
Unit V	Simulation, Risk Analysis and Risk Analysis	(10 Hours)			
Model-Building Strategies: Building Models Using Logic and Business Principles, Building Models Using Influence Diagrams, Building Models Using Historical Data, Model Assumptions, Complexity, and Realis. Analyzing uncertainty. Introduction of simulation and Risk Analysis, Types of simulations, Risk Management, Risk Assessment, Impact Analysis, Monte Carlo Simulation Using Analytic Solver Platform, New-Product Development Model, Newsvendor Model, Overbooking Model, Cash Budget, Decision Analysis: Formulating Decision Problems, Decision Strategies with the without Outcomes Probabilities, Decision Trees, The Value of Information, Utility and Decision Making. Recent Trends in: Embedded and collaborative business intelligence, Visual data recovery, Data Storytelling and Data journalism, Decision Tree and Risk					
References:					
<ol> <li>James Evans, Business Analytics: Methods, Models and Decisions, Pearson</li> <li>Marc J. Schniederjans, Dara G. Schniederjans, Christopher M. Starkey, Business analytics Principles, Concepts, and Applications, Pearson FT Press.</li> <li>U Dinesh Kumar, Business Analytics: The Science of Data - Driven Decision Making, Wiley</li> <li>Carlo Vercellis, Business Intelligence - Data Mining and Optimization for Decision Making, Wiley</li> </ol>					
<ul><li>Publications,</li><li>5. Thomas H. Davenport and Jeanne G. Harris, Competing on Analytics: The New Science of Winning, Harvard Business Review Press</li></ul>					





MCE25615B3A: Information Systems Management				
Teaching Scheme: Current Curre				
TH:03 hr/week	Credits: 03	CIE:50 Marks End Semester: 50 Marks		
Prerequisites: Information	Systems and Engineering Economics			
Course Objectives:				
a) To understand the fund	damentals of Management Information Systems	(MIS) and their role in business		
b) To explore Business	Intelligence (BI) concepts, architecture, and a	pplications in modern business		
c) To study decision supp	ort systems and their role in business decision-ma	aking processes		
d) To learn about data v	isualization techniques and their significance in	n managing information system		
projects.				
implementation.	a applications of Business Intelligence and	strategies for successful BI		
Course Outcomes:	understending of here MIC surgests husiness an			
systems.	understanding of now MIS supports business pro	ocesses and integrates enterprise		
<ul><li>2) Students will be able challenges</li></ul>	e to design and implement Business Intelligen	ice systems and understand its		
<ol> <li>Students will apply dec</li> </ol>	cision support tools and systems to improve busin	ess decision-making.		
4) Students will demons	strate proficiency in data visualization techniq	ues and managing information		
systems projects.	and implement Dusiness Intelligence strategies t	a anhanaa husinaaa nanfammanaa		
in various case studies.	and implement business intemgence strategies of	o enhance business performance		
	Theory: Course Contents			
Unit I	Management Information System (MIS)	(08 Hours)		
Managing Information S Choices, Information Sy	ystems, Ethical and Social Issues, Information ystems Security and Control, Managing Data	Technology Infrastructure and Resources, Business Process		
Integration and Enterprise	Business Intelligence	(10 Hours)		
Business Intelligence an I	introduction: Introduction, Definition, History and	d Evolution. Difference between		
Information and Intelliger	nce, Factors of Business Intelligence System - Bu	usiness Intelligence Architecture,		
Real time Business Inte	lligence, Business Intelligence Applications, B	Business Intelligence Essentials:		
Introduction, Creating B	usiness Intelligence Environment, Business Int	telligence Landscape, Types of		
Business Intelligence, Business Intelligence Platform, Dynamic roles in Business Intelligence, Roles of				
Business Intelligence in Modern Business- Challenges of BI Business Intelligence User Model:				
Content Management System End User Segmentation Basic Reporting and Ouerving Online Analytical				
Processing, OLAP Techni	ques. OLAP Applications. Applying the OLAP to	o Data Warehousing, Benefits of		
using OLAP, Dashboard, A	Advanced/Emerging BI Technologies, Future of E	Business Intelligence		
Unit III	Building Advanced Information Systems	(06 Hours)		
Decision Support in Busi	ness, Decision Support Trends, Decision Support	Systems, Management,		
Information Systems, Online Analytical Processing, Using Decision Support Systems, Executive				
Information Systems, Enter	erprise Portais and Decision Support, Knowledge	Management Systems		
Unit IV	Systems Projects	(00 110015)		





Data visualisation: Introduction, Types of data visualisation, Data for visualisation: Data types, Data encodings, Retinal variables, Mapping variables to encodings, Visual encodings, The importance of project management, Selecting projects, Establishing the business value of Information Systems, Managing project risk

Applications of Business Intelligence (08 Hours)

Business Intelligence Strategy and Road Map: Introduction, planning to implement a Business Intelligence Solution, Understand Limitations of Business Intelligence, Business Intelligence Usage, How to make the best use of Business Intelligence? Implementing Business Intelligence: Implementation Strategy, Fundamental decisions, Business Intelligence Case Studies: Improving Operational Efficiency –Audi AG, Maximizing Profitability- The Frank Russell Company

#### **References:**

Unit V

- 1. Rahul De, -MIS: Management Information Systems in Business, Government and Societyl, Wiley India,
- 2. Chan S. Park, Fundamentals of Engineering Economics, 3rd Edition, Pearson Education
- 3. Kenneth C. Laudon, Jane P. Laudon, Management Information Systems
- 4. Managing the Digital Firm, Prentice Hall
- 5. James A. O'Brien, George M. Marakas, Introduction to Information Systems, McGraw-Hill

MCE25616A0B: Laboratory Proficiency-I [Data Preparation, Analysis, Modelling and Visualization]		
Teaching Scheme: TW/PR: 06 hr/week	Credits: 03	Examination Scheme: TW: 50 Marks PR: 50 Marks
Prerequisites: Python Pre	ogramming, Data Visualization.	
	List of Assignments	
<ul> <li>Assignment 1: Choose a dataset and perform the following using Python.</li> <li>a) Compute and display summary statistics for each feature available in the dataset. (eg. minimum, maximum, mean, range, standard deviation, variance and percentiles). Use a bar-graph to demonstrate the results.</li> <li>b) Data Visualization-Create a histogram for each feature in the dataset to illustrate the feature distributions.</li> <li>c) Create a boxplot for each feature in the dataset. All the boxplots should be combined into a single plot. Compare distributions and identify outliers.</li> </ul>		
<b>Assignment 2</b> : Take any dataset and perform regression analysis using Python on it. Demonstrate the results using appropriate visualization techniques for numerical and categorical features (e.g. histogram, scatter plot, heat map, box plot).		
<b>Assignment 3</b> : Implement Naive Bayes algorithm, using Python to classify a dataset. (Do not use built-in function for Naive Bayes). Compare the performance of implementation with the built-in function for Naive Bayes. Present the Confusion matrix for each classifier. For measuring performance, use at least five metrics such as accuracy, precision, recall, F-measure etc.		
Assignment 4: Take a sample dataset. Plot the data using appropriate graphs (e.g. scatter diagram). Perform normality and symmetry tests on it using at least one graph method and at least one statistical test in Python. Analyse the results, then evaluate Spearman's Rank Correlation for this data. Assignment 5: Implement the Box-Cox Transformation in Python on a chosen dataset to achieve variance		
stabilization and normality improvement.		





**Assignment 6**: Perform Data Exploration (Descriptive Statistics, Skewness and Kurtosis of each column, Correlation between variables) through Summary Statistics using Python.

Assignment 7: Perform visualization of non-numerical data using Python for finding possible patterns in the data.

- a) Show the frequency distribution of each category and compare frequencies of categorical data.
- b) Show the proportion of each category in the dataset.
- c) Visualize the distribution of categories and how they are split by another categorical variable

**Assignment 8**: To implement and analyze the arrangement of networks and trees using connection techniques such as link marks and matrix views, visualize these networks using Python. provide a detailed report of findings

**Assignment 9**: To implement and analyze the visualization of quantitative values using scatterplots, with a focus on separating, ordering, and aligning categorical regions, using Python. provide a detailed report of your findings.

**Assignment 10**: Implement and analyze visualizations using Python that incorporate Gestalt laws, texture theory, and data mapping to enhance the perception of transparency, overlapping data, multidimensional discrete data patterns, and patterns in motion.

MCE25617A0B: Laboratory Proficiency-II [Essential Tools for Data Science]			
Teaching Scheme: PR: 06 hr/week	Credits:03		Examination Scheme: TW: 50 Marks PR: 50 Marks
Prerequisite: Basic Progra	mming Skills		
	List of As	signments	
Assignment 1: a) Create and manipulate b) Write a program in Py	e Tuple, List, Dictionary usin thon for String Manipulatio	ng Python. ns.	
Assignment 2: a) Implement CRUD operations using Python and Pandas on any dataset. b) Implement Logistic Regression in ScikitLearn			
<ul> <li>Assignment 3:</li> <li>a) Analyze any data set; display the data and provide brief description of the data using Numpy and Pandas.</li> <li>b) Plot histogram to visualize distribution of any 2 features using Matplotlib</li> </ul>			
c) Plot a pairwise scatter plot for the features and identify core relation between them using Seaborn.			
<ul> <li>a) Create a Relation with the help of data mining tool WEKA, then add an attribute, remove an attribute and normalize the relation.</li> <li>b) Perform matrix multiplication in R programming.</li> <li>c) Classify images using TensorFlow.</li> </ul>			
<ul> <li>Assignment 5:</li> <li>a) Perform the following advanced Excel operations to analyze the dataset: Data Cleaning &amp; Preprocessing, Exploratory Data Analysis (EDA), Correlation &amp; Regression Analysis, Automation with Visual Basic for Applications. Create Pivot Tables &amp; Pivot Charts.</li> <li>b) Use Power BI to create interactive reports and dashboards that can provide actionable insights to key decision-makers.</li> </ul>			





#### Assignment 6:

- a) Create a dashboard using Tableau for any application system.
- b) Identify trends, anomalies, and actionable insights for better decision-making using Tableau.

Assignment 7:

a) Perform data Analysis with Qlik Sense for any application.

Assignment 8:

a) Perform comprehensive analysis on any dataset and develop a data analysis dashboard and extract final report in PDF or HTML format using SAS Output Delivery System (ODS) for easy sharing.

Assignment 9:

a) Develop a comprehensive dashboard for any business application and monitor performance in real time by aggregating data from various departments using Google Data Studio.

Assignment 10:

a) Analyze and visualize key performance metrics across multiple data sources to identify trends, patterns and insights that supports decision making using Google Data studio.





# **ARMY INSTITUTE OF TECHNOLOGY, PUNE**

# An Autonomous Institute Affiliated to Savitribai Phule Pune University,Maharashtra, India

### **National Education Policy (NEP) Compliant Curriculum**

### Semester - II



M. Tech. Data Science (2025 Pattern)

www.aitpune.com





MCE25621A0A: Machine Learning			
<b>Teaching Scheme:</b>	C	Examination Scheme:	
TH: 03 hr/week	Credits:03	CIE: 50 Marks End Somostor: 50 Marks	
Prerequisites: Basics of D	ata Science	End Semester: 50 Warks	
Course Objectives:			
a) Introduce students to applications.	the foundational concepts of machine lear	rning, including types, workflows, and	
<ul> <li>b) Equip students with the regression tasks.</li> </ul>	ne knowledge to implement supervised lea	arning algorithms for classification and	
c) Teach students the fu dimensionality reducti	ndamentals and applications of unsupervion techniques.	ised learning, including clustering and	
d) Provide an understand	ding of ensemble learning methods and t	their applications in improving model	
<ul><li>e) Familiarize students dynamic environments</li></ul>	with reinforcement learning concepts an	nd algorithms for decision-making in	
<ul> <li>Course Outcomes:</li> <li>1) Students will be able to differentiate between traditional and machine learning approaches and apply machine learning workflows.</li> <li>2) Students will implement supervised learning algorithms for real-world classification and regression problems.</li> <li>3) Students will apply unsupervised learning techniques to analyse and extract patterns from unlabelled data.</li> <li>4) Students will optimize and evaluate ensemble learning methods to enhance model accuracy.</li> <li>5) Students will apply reinforcement learning techniques to solve decision-making problems with dynamic environments.</li> </ul>			
Introduction: Learning, designing a Learning System, Traditional Learning vs Machine Learning, Various types of Machine Learning: Supervised Learning, Unsupervised Learning, Reinforcement Learning, Machine Learning workflow, Machine Learning issues and challenges, Machine Learning Applications in real world problems. Predictive Models: Regression, Multivariate Regression, Types of Regression Models – Estimation of Regression coefficients, issues and challenges, applications			
Unit II	Supervised Learning	(08 Hours)	
Classification Models: Introduction, Different types of classifiers: Perceptron, Naive Bayes, Decision Tree, Logistic Regression, K-Nearest Neighbor, Artificial Neural Networks, Support Vector Machine, Evolution metrics for supervised learning, Issues and challenges, applications.			
Unit III	Unsupervised Learning	(08 Hours)	
Partitioning based clustering, Hierarchical based clustering, Density based clustering, Grid based clustering, Mixture Models and EM Algorithm, Fuzzy k-Means Algorithm, Evolution metrics for clustering models - Dimensional Reduction Techniques: Need, Various types: PCA, ICA, FA, t-SNE - Case studies.			
Unit IV	Ensemble Learning	(08 Hours)	
Boosting, AdaBoost Algorithm, Bagging, Random Forest, NoFree-Lunch Theorem, XGBoost Algorithm, Stacking, Voting, Ensemble Diversity, Error Decomposition, Diversity Measures, Evaluating Ensembles of Classifiers, Case studies			
Unit V	Reinforcement Learning	(08 Hours)	





Introduction to Reinforcement Learning, single state case, elements of reinforcement learning, model-based learning, Temporal difference learning: exploration strategies, deterministic rewards and actions, non-deterministic rewards and actions, eligibility traces generalization, partial observable states, the tiger problem.

- 1. Andriy Burkov, The Hundred Page Machine Learning Book, Ayalaux Publications.
- 2. Ethem Alpaydin, Introduction to Machine Learning, MIT Press.
- 3. Peter Flach, Machine Learning: The Art and Science of Algorithms that make sense of data, Cambridge University Press
- 4. Kevin Murphy, Machine Learning: A Probabilistic Approach, MIT Press
- 5. Andrew Ng, Machine Learning Yearning.

MCE25622A0A: Big Data Analytics			
Teaching Scheme: TH: 03 hr/week	Credits:03	Examination Scheme: CIE: 50 Marks End Semester: 50 Marks	
Prerequisites: DBMS, Bas	ics of Data Science		
<ul> <li>Course Objectives:</li> <li>a) Understand Big Data Concepts</li> <li>b) Explore Data Analysis and Pre-processing Techniques</li> <li>c) Master Hadoop Ecosystem and Tools</li> <li>d) Learn Big Data Analytics with Spark</li> <li>e) Leverage Cloud Computing and AWS for Big Data</li> </ul>			
<ul> <li>Course Outcomes:</li> <li>Students will be able to define big data, describe its characteristics, types, and sources, and identify business problems suitable for big data analytics.</li> <li>Students will be able to apply the data analysis lifecycle and preprocessing techniques to clean and prepare data for analysis, ensuring readiness for big data tools.</li> <li>Students will gain the skills to use Hadoop components like HDFS, MapReduce, YARN, and HBase for distributed big data processing and storage, and utilize tools like Hive and Pig for data analysis.</li> <li>Students will be able to use Apache Spark and its core concepts, including RDD optimization and Spark SQL, and apply PySpark and RHadoop for big data analytics tasks such as text mining and data mining.</li> <li>Students will demonstrate the ability to use AWS services like EC2, S3, and database services for big data storage, computation, monitoring, and automation, supporting cloud-based big data solutions.</li> </ul>			
	Theory: Course Contents		
Unit IBig Data Overview(08 Hours)Big data: characteristics, types, sources, architectures, Data analysis process, Data analytics lifecycle, Pre- processing data, Market and Business Drivers for Big Data Analytics, Business Problems Suited to Big Data Analytics			
Unit II	Exploratory Data Analytics	(08 Hours)	
Distributed and Parallel Computing for Big Data, Cloud Computing and Big Data, In-Memory Computing Technology for Big Data, Introduction to Hadoop, HDFS, MapReduce, YARN, HBase, Combining HDFS and HBase			
Unit III	Hadoop Ecosystem for Big Data	(08 Hours)	





	Analytics		
Hadoop ecosystem: Sqoop, Impala, Apache Flume, Pig, Hive, Data transformation and analysis using Pig, Data analysis using Hive and Impala, Mahout, Oozie, Zookeeper etc.			
Unit IV	Big Data Analytics with Apache Spark using PySpark and R Programming	(08 Hours)	
Apache Spark, Spark core, Interactive data analysis with spark shell, Writing a spark application, Spark RDD Optimization Techniques, Spark Algorithm, Spark SQL, Big data analytics with PySpark: Python and Apache Spark Big data analytics with RHadoop: R and Hadoop, Text mining in RHadoop, Data mining in Hive, Data Analysis MapReduce techniques using RHadoop.			
Unit V	Amazon Web Services (AWS)	(08 Hours)	
Introduction to Cloud Computing & AWS, Elastic Compute and Storage Volumes, Autoscaling and DNS, Simple Storage Service (S3), Database Services, Access Management and Monitoring Services, Automation and Configuration management, AWS Migration			
References:			
1. DT Editorial Services	, Big Data, Black Book: Covers Hadoop 2,	MapReduce, Hive, YARN, Pig, R and	
<ol> <li>Data Visualization, Dreamtech Press</li> <li>Holden Karau, Andy Konwinski, Patrick Wendell, and Matei Zaharia , Learning Spark: Lightning-Fast Big Data Analysis, O'Reilly Media.</li> </ol>			
3. David Dietrich, Barry Hiller, Data Science and Big Data Analytics, EMC education services, Wiley Publications			
<ol> <li>Mohammed Guller, Big Data Analytics with Spark: A Practitioner's Guide to Using Spark for Large Scale Data Analysis</li> </ol>			
5. Viktor Mayer-Schönberger and Kenneth Cukier, Big Data: A Revolution That Will Transform How We Live, Work, and Think, Houghton Mifflin Harcourt			
MCE25623A0A: Research Methods for Engineers			

MCL25625AUA: Research Methods for Engineers			
Teaching Scheme: TH:03 hr/week	Credits:03	Examination Scheme: CIE: 50 Marks End Semester: 50 Marks	
Prerequisites: Critical Thinking and Problem-Solving Skills.			
Course Objectives:			
a) To introduce the fundamentals of research methodology, its significance, and types of research in engineering.			
b) To provide knowledge on conducting literature searches, reviewing publications, and developing a research plan.			
c) To familiarize student	c) To familiarize students with survey research methods, ethical considerations, and statistical analysis.		
d) To develop skills for presenting and reporting research findings, including writing research papers and thesis reports.			
e) To instill an understanding of ethical practices in research, intellectual property rights, and patent processes in engineering.			
Course Outcomes:			
1) Students will gain an understanding of the research process research types and the importance of			

- erstanding of the research process, research ιyμ g engineering ethics in research.
- Students will be able to conduct literature reviews, develop research plans, and design research proposals 2) effectively.





- 3) Students will be proficient in conducting surveys, analyzing data, and adhering to ethical research practices.
- 4) Students will be able to present research findings clearly through well-structured papers and presentations.
- 5) Students will have knowledge of ethical codes in research, intellectual property laws, and how to handle patents and copyrights.

Theory: Course Contents		
Unit I Introduction (8 Hours)		
Introduction: Evolution of Research Methodology: Meaning, nature, scope, and significance of research; Research paradigm; The purpose and Products of Research; Reasons for doing research, Objectivess of research, Motivation for research; Postulates underlying scientific investigations; Types of research; Research process and work flow. Engineering Research-Why? Research Questions, Engineering Ethics, conclusive proof-what constitutes, A research project-Why take on? Engineering dictionary, Shodhganga, The Library of Congress, Research gate, Google Scholar, Bibliometrics, Citations, Impact Factor, h-index, I index, plagiarism, copyright infringement. Collect data for overbooking decision for demand and revenue management of flights.		
Unit II Literature Search & Review, Developing Research Plan (08 Hours)		
Research Plan: Archival Literature, why should engineers be ethical? Types of publications- Journal papers, conference papers, books, standards, patents, theses, trade magazine, newspaper article, infomercials, advertisement, Wikipedia & websites, Measures of research impact, Literature review, publication cost. Developing Research Plan: Research Proposals, finding suitable research questions, The elements of research proposals-title, details, budget, Design for Outcomess-1D data, 2D data, 3D data, N-D data, The research tools-Experimental measurements, numerical modelling, theoretical derivations & Calculations, curve matching.		
Unit III Survey Research Methods (06 Hours)		
Survey Research Methods: Why undertake a survey, Ergonomics and human factors, Ethics approval, General survey guidelines, Survey statements, Survey delivery, Respondent selection, Survey timelines, Statistical analysis Reporting		
Unit IV Research Presentation (08 Hours)		
Research presentation: Introduction, Standard terms, Standard research methods and experimental techniques, Paper title and keywords, Writing an abstract, Paper presentation and review, Conference presentations, Poster presentations. Reporting Research: Thesis, Structure and Style for writing thesis, Dissemination of research findings; Reporting and interpretation of results; cautions in interpretations, Type of reports, Typical report outlines. The path forward: Publication trends, Getting started in research, Quality assurance (QA) Occupational health		
Unit V Code of Ethics and Patents (10 Hours)		
Code of Ethics, IEEE Code of Ethics, ACM Software Engineering Code of Ethics and Professional Practice, Code of Ethics especially covering Engineering discipline, various aspects- environment, sustainable Outcomess, employer, general public, & Nation, Engineering Disasters, GNU PSPP Tool, SOFA, NOST- Dataplot, Intellectual Property India- services, InPASS - Indian Patent Advanced Search System, US patent, IPR, Copyright, Patents, IEEE / ACM Paper templates Patent act, 1970 and Patent Rules 1972 (with amendments), Qualitative Analysis Tools- AQUAD, CAT. IP related laws in India, Google Optimization Tools, OpenMDAO		
References:		
<ol> <li>David V Thiel, Research Methods- for Engineers, Cambridge University Press.</li> <li>Kothari C.R., Research Methodology. New Age International.</li> <li>Caroline Whitbeck, Ethics in Engineering Practice and Research, Cambridge University Press</li> <li>Gordana DODIG-CRNKOVIC, Scientific Methods in Computer Science, Department of Computer Science Malardalen University, Vasteas, Sweden</li> </ol>		





MCE25624B1A: Recommender Systems			
Teaching Scheme: Examination Scheme:			
TH-03 hr/wook	Credits:03	CIE: 50 Marks	
111.05 III/WEEK		End Semester: 50 Marks	
Prerequisites: Understand	ing User Behavior		
Course Objectives:			
a) To introduce the funda	mentals of recommender systems, their f	unctions, and applications.	
b) To explore collaborative	e filtering techniques, including user-bas	sed and item-based approaches.	
c) To examine content-ba	ised and knowledge-based recommendat	tion methods, including the creation of item	
d) To study hybrid rec	commender systems understanding v	arious strategies for combining different	
approaches.	ommender systems, understanding ve		
e) To understand the eval	uation of recommender systems, includir	ng error metrics and user-centered evaluation	
methods.	•	-	
Course Outcomes:			
1) Students will be able t	o describe the basic functions and applic	ations of recommender systems and identify	
common challenges.			
2) Students will gain the	ability to implement collaborative filtering	ng techniques and understand their strengths	
and weaknesses.	able of analyting content based filterin	a matheda and building Imaviladas based	
s) Students will be capa recommendation system	able of applying content-based intering	g methods and building knowledge-based	
4) Students will learn how	v to design and apply hybrid recommend	ation approaches using various hybridization	
strategies.	v to design and apply hybrid recommend	ation approaches asing various nyonaization	
5) Students will unders	tand and apply different evaluation	techniques to assess the performance of	
recommender systems.		· ·	
	Theory: Course Conte	nt	
Unit I	Introduction to Recommender syst	em (06 Hours)	
Recommender system fur	ctions, Understanding ratings, Application	ons of recommendation systems, Issues with	
recommender system.			
Unit II	Collaborative Filtering	(08 Hours)	
User-based nearest neigh	approaches Attacks on collaborative reco	est neighbor recommendation, Model based	
	Content & knowledge based recommen	valation (10 Hours)	
High level architecture of	content-based systems. Advantages and	drawbacks of content-based filtering Item	
profiles. Discovering feat	ures of documents. Obtaining item feat	ures from tags. Representing item, profiles.	
Methods for learning use	er profiles, Similarity based retrieval, C	Classification algorithms, Knowledge based	
recommendation: Knowle	recommendation: Knowledge representation and reasoning, Constraint based recommenders, Case based		
recommenders.			
Unit IV	Hybrid approaches	(08Hours)	
Opportunities for hybridization, Monolithic hybridization design: Feature combination, Feature, augmentation,			
Parallelized hybridization design: Weighted, Switching, Mixed, Pipelined hybridization design: Cascade Meta-			
level, Limitations of hybridization strategies			
Unit V	Evaluating Recommender System	n decigne Evaluation on historical detects	
Fror metrics Decision	upport metrics. User-Centred metrics	in designs, Evaluation on historical datasets,	
References	apport metrics, Oser-Centred metrics.		
1. F. Ricci, L. Rokach, a	nd B. Shapira, eds. Recommender Syster	ns Handbook.	





- 2. D. Jannach, M. Zanker and F. Fering, Recommender Systems: An Introduction, Cambridge University Press.
- 3. C. C. Aggarwal, Recommender Systems: The Textbook, Springer.
- 4. N. Manouselis, H. Drachsler, K. Verbert, E. Duval, Recommender Systems for Learning, Springer.
- 5. Manish Kumar, Building Data Streaming Applications with Apache Kafka: Design, Develop, and Stream Real-time Data Pipelines, Packt Publishing

MCE25624B2A: Video Analytics			
Teaching Scheme: TH:03 hr/week	Credits:03	Examination Scheme: CIE: 50 Marks End Semester: 50 Marks	
Prerequisites: Linear Alg	gebra/Probability Review/Matrix theory		
Course Objectives:			
a) To introduce the fur vision tasks.	a) To introduce the fundamentals of computer vision, including image preprocessing, segmentation, and 3D vision tasks.		
b) To provide a deep un	derstanding of motion analysis techniques, including of	otical flow and video tracking.	
c) To explore object i	ecognition methods, focusing on statistical, neural	network, and syntactic pattern	
recognition.			
d) To develop skills in r	eal-time video analytics, including activity recognition	and behavior analysis.	
e) To examine the state-	of-the-art video analytics applications, with emphasis	on deep learning and future trends	
in computer vision.		1 5	
1			
Course Outcomes:			
1) Students will underst	and the core principles of computer vision, including in	nage enhancement, segmentation,	
and 3D vision tasks.			
2) Students will be abl	e to apply motion analysis techniques like optical flo	w and tracking in various video	
applications.			
3) Students will gain	proficiency in object recognition using various p	attern recognition methods and	
optimization techniqu	les.		
4) Students will acquir	e knowledge to implement intelligent video analytic	s for real-time applications and	
behavior analysis.			
5) Students will explore	e and analyze cutting-edge video analytics technologie	es and their future applications in	
various fields.			
TT VIT	Theory: Course Content	(00.11	
	Introduction to Computer Vision	(U8 Hours)	
Motivation, Relationshi	ps to other fields, Image pre-processing, Image En	hancement, Image segmentation,	
Feature Extraction: Shape representation and description: Contour-based shape representation and description,			
region-based shape representation and description, statistical and syntactic texture description methods Camera			
Nodels: Cameras: Finnole cameras, cameras with lenses, the Human eye, Sensing, 5D cameras. 5D vision tasks,			
from Y			
	Motion Analysis	(09 Hours)	
Differential Motion And	Insis methods Change detection Segmentation using a	motion Image flow sogmentation	
Uniciential Motion Analysis methods, Change detection, Segmentation using motion, image now, segmentation using Moving comerce Optical flow. Analysis based on correspondence of interest points, detection of specific			
motion natterns, video tracking motion models to aid tracking			
Unit III	Object Recognition	(08 Hours)	





Knowledge representation, Statistical Pattern Recognition, Neural Nets, Syntactic pattern recognition, Recognition as graph matching, Optimization techniques in recognition, fuzzy systems, texture recognition methods		
Unit IV	Intelligent Video Analytics	(08 Hours)
Real-time video analyti	cs and video mining, temporal and spatial event reco	gnition, Vision-based activity
recognition, Behaviour A	analysis, Content-Based Analysis of Digital Video	
Unit V	Video Analytics State-of-the-art and the Future	(08 Hours)
Video Analytics: state o	f the art applications with reference to computer vision	applications, Deep learning in
video analytics, Human motion recognition and its applications, Video Analytics for Business Intelligence,		
Virtual reality/Augmented reality applications, and Healthcare applications.		
References:		
1. Sonka, Hlavac, Boyle, Digital Image Processing and Computer Vision- Cengage Learning, Indian Edition		
2. Ramesh Jain, Kasturi, Schunck, Machine Vision, McGraw-Hill		
3. Milan Sonka, Vaclav Hlavac, Roger Boyle, Image Processing, Analysis, and Machine Vision, Thomson		
Learning		
4. David Forsyth, Jean Ponce, Computer Vision, Pearson Education		

5. Jan Eril Solem, Programming Computer Vision with Python, O'Reilly





MCE25624B3A: Real Time Analytics		
Teaching Scheme: TH:03 hr/week	Credits:03	Examination Scheme: CIE: 50 Marks End Semester: 50 Marks
Prerequisites: Time serie	s and Forecasting	
<ul> <li>Course Objectives:</li> <li>a) To provide a comprehensive understanding of data analytics, including various types and steps such as data pre-processing, cleaning, and visualization.</li> <li>b) To introduce the fundamentals of Python for data analytics, covering libraries like NumPy, Pandas, SciPy, and Matplotlib.</li> <li>c) To explore the application of the Box-Jenkins methodology and ARIMA models for time series analysis.</li> <li>d) To understand the architecture and processes involved in real-time streaming data analysis and management.</li> <li>e) To learn about market basket analysis, recommender systems, and real-time object detection using YOLO in modern e-commerce applications.</li> </ul>		
<ul> <li>Course Outcomes:</li> <li>1) Students will be able to apply various data pre-processing, cleaning, and transformation techniques in data analytics projects.</li> <li>2) Students will gain proficiency in using Python for data analysis, leveraging libraries like Pandas, SciPy, and Matplotlib for advanced analytics.</li> <li>3) Students will be capable of implementing ARIMA models for time series forecasting and examining correlation and stationarity in data.</li> <li>4) Students will understand the architecture and techniques involved in processing and storing streaming data for real-time analysis.</li> <li>5) Students will be able to implement market basket analysis and recommender systems, as well as apply YOLO for real-time object detection in e-commerce.</li> </ul>		
	Theory: Course Contents	
Unit IFundamentals of Data Analytics(08 Hours)Data Analytics Basics, Data Types, Analytics Types, Data Analytics Steps: Data Pre-Processing, Data Imputation, Data Cleaning, Data Transformation, Data Visualization, and Data Engineering. Descriptive, Predictive, and Prescriptive Analytics		
Unit II         Data Analytics with Python         (08 Hours)           Data Analytics using Python, Statistical Procedures, Web Scraping in Python, Advanced analytics, NumPy, Pandas, SciPy, Matplotlib.         NumPy, Pandas, SciPy, Matplotlib.		
Unit IIIAdversarial Search and Games(08 Hours)Box-Jenkins Methodology for ARIMA models: Examining correlation and stationarity of time series data, ARIMA models for time series data (An Auto-regressive model of order one and a Moving Average Model of order one).		
Unit IV	Streaming Data Analysis	(08 Hours)
Streaming Analytics Architecture: Designing Real-Time Streaming Architectures, Service Configuration and Coordination, Data-Flow Management in Streaming Analysis		
Unit V	Market Basket Analysis and Recommender system	(08 Hours)
Processing Streaming Data, Storing Streaming Data, Today's ecommerce system, apriori algorithm. YOLO: real time object Detection		
References: 1. Anil Maheshwari, Data Analytics made accessible, Amazon Digital Publication.		





- 2. Thomas H. Davenport, Jeanne G. Harris and Robert Morison, Analytics at a) Work: Smarter Decisions, Better Results, Harvard Business Press.
- 3. Byron Ellis, Real–Time Analytics: Techniques to Analyze and Visualize Streaming Data, WILEY Publication.
- 4. Michael Minelli, Michelle Chambers, and Ambiga Dhiraj, Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses, Wiley Publication.
- 5. Tyler Akidau, Slava Chernyak, and Reuven Lax, Streaming Systems: The What, Where, When, and How of Large-Scale Data Processing, O'Reilly Media.

MCE25625A0B: Laboratory Proficiency-III [Big Data Analytics]		
Teaching Scheme: TW/PR- 6hr/week	Credits:03	Examination Scheme: TW: 50 Marks PR: 50 Marks
Prerequisite: Basic Knowl	edge of Data analytics, DBMS	
	List of Assign	nents
Assignment 1: Take any text or image dat Python.	taset and perform sentiment Analys	is or image classification analysis on it using
Assignment 2: Implement a Python code insights from historical v models for future weather	e for analysing and forecasting we veather data to understand seasons trends.	ather data. pre-process, analyse, extract actionable al patterns, detect anomalies, and build predictive
Assignment 3: Implement an efficient data processing pipeline that can handle large-scale sales data and calculate the total sales, average sales, sales by region/location, sales growth, sales per customer, top products/services, profitability analysis, customer segmentation analysis, sales forecasting, sales distribution over time while ensuring scalability performance and minimal resource consumption using Python		
Assignment 4: Create a large dataset and perform a comparative analysis of In-Memory Computing vs. Traditional Disk-Based Methods using Python.		
Assignment 5: Perform data transformation(Load, Filter, projection, renaming columns, grouping data, aggregating data, joining data, sorting data, flattens nested data ) and analysis(calculating employee tenure, calculating salary growth, salary distribution across the organisation, employee age analysis, department performance analysis, employee retention and departmental insights) tasks using Pig on employee dataset		
Assignment 6: Perform data transformation(Load, Filter, Aggregation and Grouping) and analysis(Treatment Plan Analysis, Trend Analysis Over Time, Patient Age Group Analysis, Top Medication and Treatment Plan Analysis, Data Quality and Missing Value Analysis, Hospital Performance Analysis) tasks in Healthcare dataset Using Hive and Impala.		
Assignment 7: Demonstrate application of Apache spark to analyse sentiment streaming data from social media. (Installation of multi-node Hadoop as well as Spark is required.)		
Assignment 8: Develop a Spark application to process(Creating New Columns, Filter, Aggregations, Data Sorting, Join with Another Dataset, Window Functions for Advanced Analysis, Handling Time Series Data, Data Saving) a large dataset containing transaction details from an e-commerce website.		
Assignment 9: Create business intelligence dashboards with Amazon QuickSight.		





#### Assignment 10:

Extract data from different data sources, apply suitable transformations(Data Cleaning, Filtering and Selecting Columns, Type Casting/Converting Columns, Joining Data, Aggregations and Grouping, String Manipulations, Date Manipulations, Data Loading into Destination Tables) and load into destination tables using an AWSglue.

MCE25626A0B: Laboratory Proficiency-IV [Machine Learning]		
Teaching Scheme: PR – 6hr/week	Credits: 03	Examination Scheme: TW: 25 Marks PR: 25 Marks
Prerequisite: Knowledge	of any programming language, Python bas	sics.
	List of Assignments	
<ul><li>Assignment 1: Use appropriate dataset from UCI machine learning repository and implement two different decision tree algorithms using Python.</li><li>Assignment 2: Extract confusion matrix from the test results of Assignment 1 using Python. Compare the</li></ul>		
Assignment 3: Classify th performance of your imple	e data using various decision tree algorithms nentations with the results from the Weka tool	from the Weka tool and compare the
<b>Assignment 4:</b> Implement k-NN classifier to classify a standard dataset (from UCI machine learning repository). Use Python for implementation. Test the performance for various values of k. Now, classify the same dataset using distance-weighted k-NN and locally weighted averaging methods. Compare the performance on at least six standard performance measures.		
Assignment 5: Implement	Naïve-bayes algorithm for text data classificat	ion in Python.
Assignment 6: Implement fuzzy k-means algorithm for Image Segmentation in Python.		
Assignment 7: Implement XGBoost algorithm for anomaly detection in Python.		
Assignment 8: Implement Random Forest algorithm for Time Series Forecasting in Python.		
<ul> <li>Assignment 9:</li> <li>(a) Implement the Q-Learning algorithm in Python, which updates the Q-table based on the agent's experience (state, action, reward, and next state).</li> <li>(b) Train the agent using multiple episodes and observe its learning process.</li> </ul>		

Decision Making in Python.



MCE25627AOR. Mini Project with Somin



		WICE25027A0D. WIIII I TOJECT WITH	Semmar
To T	eaching Scheme: W/PR: 02 hr/week	Credits:01	Examination Scheme: OR: 25 Marks TW: 25 Marks
C a) b) c) d)	ourse Objectives: To identify the domain To learn to communic To categorize the reser To work in profession	of research ate in a scientific language through collaboration arch material confined to the domain of choice al environment	on with a guide.
C( 1) 2) 3) 4)	ourse Outcomes: Conduct thorough lite Develop presentation Furnish the report of t Analyse the findings a	rature survey confined to the domain of choice skills to deliver the technical contents ne technical research domain nd work of various authors confined to the cho	sen domain
M be hc cc pr th	lini Project should inclu- eginning of the second buse Research Project – eeds to be allocated bas bllege has to monitor a rogress through regular e progress needs to be o	<b>Conduction guidelines</b> ide partial/complete project implementation. S semester itself and same guide should be com I. The preferences/choices of the domain should ed on the preference/choices. In case of Indus and evaluate the progress of the student. The reporting and presentations and proper docum locumented unambiguously.	Student should be allocated guide in the atinued for the: Industry Internship-I/ In Id be taken from the students. The guide stry Internship, the assigned guide from e student has to exhibit the continuous mentation. The continuous assessment of
		MCF25628A0A · Value Educat	ion
T T	eaching Scheme: H: 01 hr/week	Credits:01	Examination Scheme: CIE: 25 Marks
<ul> <li>Course Objectives:</li> <li>a) To understand the importance of value education and its role in achieving happiness and prosperity.</li> <li>b) To explore harmony in human relationships, society, and nature, fostering holistic well-being.</li> <li>c) To cultivate personal values, self-discipline, and moral development for individual growth.</li> <li>d) To develop positive thinking, integrity, and self-awareness for improved personality and behavior.</li> <li>e) To emphasize character building, self-management, and the development of competence for holistic living.</li> </ul>			
C(1) 2)	ourse Outcomes: Students will gain ins happiness and prosper Students will understa	ights into value education and apply the princ ity. nd and promote harmony in human relationshi	ciples of self-exploration for continuous
3)	<ol> <li>Students will understand and promote namony in numan relationsmps, society, and nature, fostering initial fulfilment.</li> <li>Students will cultivate essential values such as honesty, discipline, and social responsibility for personal and societal growth.</li> </ol>		
(4)	development.	a positive minuser, self-discipline, and em	otional interligence for better personal

5) Students will enhance their character and competence, integrating values like humility, nonviolence, and self-control for effective life management.

Theory: Course Contents

Unit IValue Education and Harmony in Human being(3 hours)





Understanding Value Education, Self-exploration as the Process for Value Education, Continuous Happiness and Prosperity – the Basic Human Aspirations, Right Understanding, Relationship and Physical Facility, Happiness and Prosperity – Current Scenario, Method to fulfil the Basic Human Aspirations Understanding Human being as the Co-existence of the Self and the Body, distinguishing between the Needs of the Self and the Body, The Body as an Instrument of the Self, Understanding Harmony in the Self, Harmony of the Self with the Body, Program to ensure self-regulation and Health			
Unit II         Harmony in the Family. Society and Nature         (3 hours)			
The Basic Unit of Human Interaction, Values in Human–to–Human Relationship, Nine universal values in relationships viz. Trust, Respect, Affection, Care, Guidance, Reverence, Glory, Gratitude, Love. Understanding Harmony in Society, Vision for the Universal Human Order, Human Order Five Dimension. Understanding Harmony in Nature, self–regulation & mutual fulfilment among the Four orders of Nature, Realizing Existence as coexistence at all levels holistic perception of harmony in existence			
Unit IIIValues and self-development(3 hours)			
Social values and individual attitudes. Work ethics, Indian vision of humanism, Moral and non- moral valuation, Standards and principles, Value judgements. Importance of cultivation of values, Sense of duty. Devotion, Self-reliance. Confidence, Concentration, Truthfulness, Cleanliness, Honesty, Humanity. Power of faith, National Unity, Patriotism. Love for nature, Discipline.			
Unit IV Personality and Behaviour Development (2 hours)			
Soul and Scientific attitude. Positive Thinking, Integrity and discipline, Punctuality, Love and Kindness, Avoid fault Thinking. Free from anger, Dignity of labour, Universal brotherhood and religious tolerance, True friendship, Happiness Vs suffering, love for truth, Aware of self-destructive habits, Association and Cooperation, Doing best for saving nature.			
Soul and Scientific attitude. Positive Thinking, Integrity and discipline, Punctuality, Love and Kindness, Avoid fault Thinking. Free from anger, Dignity of labour, Universal brotherhood and religious tolerance, True friendship, Happiness Vs suffering, love for truth, Aware of self-destructive habits, Association and Cooperation, Doing best for saving nature.			
Soul and Scientific attitude. Positive Thinking, Integrity and discipline, Punctuality, Love and Kindness, Avoid fault Thinking.Free from anger, Dignity of labour, Universal brotherhood and religious tolerance, True friendship, Happiness Vs suffering, love for truth, Aware of self-destructive habits, Association and Cooperation, Doing best for saving nature.Unit VCharacter and Competence(2 hours)			
Soul and Scientific attitude. Positive Thinking, Integrity and discipline, Punctuality, Love and Kindness, Avoid fault Thinking.         Free from anger, Dignity of labour, Universal brotherhood and religious tolerance, True friendship, Happiness Vs suffering, love for truth, Aware of self-destructive habits, Association and Cooperation, Doing best for saving nature.         Unit V       Character and Competence       (2 hours)         Holy books vs Blind faith, Self-management and Good health, Science of reincarnation, Equality, Nonviolence, Humility, Role of Women, All religions and same message, Mind your Mind, Self-control, Honesty, Studying effectively.			





# ARMY INSTITUTE OF TECHNOLOGY, PUNE

## An Autonomous Institute Affiliated to Savitribai Phule Pune University,Maharashtra, India

### **National Education Policy (NEP) Compliant Curriculum**

### **Semester - III**



M. Tech. Data Science (2025 Pattern)

www.aitpune.com





MCE25631A0A: Data Storage Technologies and Networks			
<b>Teaching Scheme:</b>	Feaching Scheme:     Examination Scheme:       CIE: 50 Marks		
TH: 03 hr/week		End Semester: 50 Marks	
Prerequisites: Basic know	Prerequisites: Basic knowledge of Computer Architecture, Operating Systems, and Computer Networking.		
Course Objectives:			
a) To understand the f key components of a	undamentals of information storage, data types, a data center.	storage architecture evolution, and the	
b) To introduce intell traditional storage m	igent storage systems, storage provisioning t nethods.	echniques, and compare virtual and	
c) To explore storage data access and tran	networking technologies, including Fibre Chann sfer.	el SAN, iSCSI, and FCIP for efficient	
<ul> <li>d) To provide knowled access and retrieval</li> </ul>	lge on NAS, object-based storage, and unified methods.	storage solutions with a focus on data	
e) To understand the domains, and securi	principles of securing storage infrastructure, ty implementations in cloud and virtualized envir	including risk management, security conments.	
Course Outcomes:			
1) Students will be a infrastructure.	able to explain the components of informati	on storage systems and data center	
2) Students will gain provisioning, and th	proficiency in understanding intelligent storage eir applications.	systems, traditional vs. virtual storage	
<ul><li>3) Students will be ca</li><li>Channel SAN iSCS</li></ul>	pable of understanding and implementing stora	ge networking technologies like Fibre	
<ul><li>4) Students will be ab</li></ul>	le to work with NAS, object-based storage, and	d unified storage solutions to optimize	
<ul><li>5) Students will acqui implementing securi</li></ul>	agement. hire skills in securing storage infrastructure ity measures in various storage environments.	, addressing security concerns, and	
	Theory: Course Contents		
Unit I	Introduction to Storage System	(08 Hours)	
Information to Information Storage - Information Storage, Data, Types of Data, Big Data, Information, Storage, Evolution of Storage Architecture, Data Centre Infrastructure- Core Elements of a Data Centre, Key Characteristics of a Data Centre, Managing a Data Centre, Data Centre Environment – Application, Database Management System (DBMS), Host (Compute), Connectivity, Storage, Disk Drive Components, Disk Drive Performance, Host Access to Data, Direct-Attached Storage, Storage Design Based on Application, Disk Native Command Queuing, Introduction to Flash Drives, RAID Implementation Methods, RAID Array Components, RAID Techniques, RAID Levels, RAID, Impact on Disk Performance, RAID Comparison, Components of an Intelligent Storage System- Front end, Cache, Back End,			
Unit II         Intelligent Storage Systems and Virtualization         (08 Hours)			
Comparison between Virtual and Traditional Storage Provisioning, Types of Intelligent Storage Systems- High- End Storage Systems, Midrange Storage Systems, Server and Storage I/O Fundamentals- Server and I/O Architectures, Storage Hierarchy, From Bits to Bytes, Disk Storage Fundamentals, Initiators and Targets, How to write and read from a Storage Device, Storage Sharing vs. Data Sharing Different Types of Storage, I/O Connectivity and Networking Fundamentals, IT Clouds, Virtualization: Servers, Storage, and Networking, Virtualization and Storage Services, Data and Storage Access			
	Storage Networking rechnologies - SAN, 18		





Fibre Channel Storage Area Networks - Fibre Channel: Overview, The SAN and Its Evolution, Components of FC SAN, FC Connectivity, Switched Fabric Ports, Fibre Channel Architecture- Fibre Channel Protocol Stack, Fibre Channel Addressing, World Wide Names, FC Frame, Structure and Organization of FC Data, Flow Control, Classes of Service, Zoning, FC SAN Topologies, IP SAN – iSCSI- Components of iSCSI, iSCSI Host Connectivity, iSCSI Topologies, iSCSI Protocol Stack, FCIP - FCIP Protocol Stack, FCoE - I/O Consolidation Using FCoE, Components of an FCoE Network

Unit IV	Storage Networking Technologies - NAS, Object-Based and Unified storage	(08 Hours)
---------	--	------------

Introduction to NAS, Benefits of NAS, File Systems and Network File Sharing- Accessing a File System, Network File Sharing, Components of NAS, NAS I/O Operation, NAS Implementations- Unified NAS, Unified NAS Connectivity, Gateway NAS, Gateway NAS Connectivity, Scale-Out NAS, Scale-Out NAS Connectivity, NAS File-Sharing Protocols – NFS, CIFS, Object-Based and Unified Storage – Object-Based Storage Devices - Object-Based Storage Architecture, Components of OSD, Object Storage and Retrieval in OSD, Benefits of Object-Based Storage, Common Use Cases for Object-Based Storage, Content-Addressed Storage

Unit V Secu

Securing the Storage Infrastructure

(08 Hours)

Information Security Framework, Risk Triad, Storage Security Domains – Securing application access domain, securing management access domain, Security Implementations in Storage Networking- FC SAN, NAS, IP SAN, Securing Storage Infrastructure in Virtualized and Cloud Environments – Security concerns, Security measures

- 1. EMC Education Services, Information storage and management, Wiley Publication.
- 2. Greg Schulz, Cloud and Virtual Data Storage Networking, CRC Press.
- 3. Storage Networks: The Complete Reference, Robert Spalding", Publisher: McGraw-Hill Osborne Media
- 4. Storage area network essentials, Richard Barker, Paul Massiglia, Wiley Publication
- 5. K.L. JAMES, Data Storage Technologies Kindle Edition.





Teaching Scheme: TH: 03 hr/week         Credits:03         Examination Scheme: CIE: 50 Marks End Semester: 50 Marks           Prerequisites: Database Management Systems, Data Mining Course Objectives:         In understand the need for data warehousing, its evolution, and its significance in business intelligence.           10         To understand the need for data warehousing, its evolution, and its significance in business intelligence.           10         To captore the architecture, components, and building blocks of data warehousing, including dimensional modelling techniques.           c)         To captore the architecture, components, and building blocks of data warehousing, including dimensional modelling techniques.           c)         To captore the architecture, components, and delivery mechanisms, including OLAP and web-enabled data warehouse delivery.           of the principles of data mining, including its applications, techniques like classification and clustering, and current trends.           Course Outcomes:         1)           1)         Students will be proficient in designing and understanding data warehouse architecture and dimensional modelling.           3)         Students will understand how information is accessed and delivered from a data warehouse domains.           5)         Students will understand how information is accessed and delivered from a data warehouse domains.           5)         Students will understand how information is accessed and delivered from a data warehouse domains.           5)         Students will equipped	MCE25632A0A: Data Warehousing and Mining			
TH: 03 hr/week         Credits:03         CIE: 50 Marks End Somester: 50 Marks           Prerequisites: Database Management Systems, Data Mining         Course Objectives:           a) To understand the need for data warehousing, its evolution, and its significance in business intelligence.         b) To explore the architecture, components, and building blocks of data warehousing, including dimensional modelling techniques.           c) To learn the processes of data extraction, transformation, and loading (ETL) for integrating data into a warehouse.         d) To gain knowledge of information access and delivery mechanisms, including OLAP and web-enabled data warehouse.           d) To gain knowledge of information access and delivery mechanisms, including OLAP and web-enabled data warehouse.         To comprehend the principles of data mining, including its applications, techniques like classification and clustering, and current trends.           Course Outcomes:         1) Students will be able to articulate the need and significance of data warehouse architecture and dimensional modelling.           1) Students will acquire hands-on knowledge of the ETL process for data integration and transformation.           4) Students will understand how information is accessed and delivered from a data warehouse using OLAP and web-based tools.           5) Students will be equipped with skills in data mining techniques and their applications in various domains.           Entroduction to data warehousing and (08 Hours)           Need for Data warehousing: Escalating need for strategic information, failure of past system, operational versus decision support systems, data warehous	<b>Teaching Scheme:</b>		Examination Scheme:	
End Semester: 50 Marks           Prerequisites: Database Management Systems, Data Mining           Course Objectives:           a) To understand the need for data warehousing, its evolution, and its significance in business intelligence.           b) To explore the architecture, components, and building blocks of data warehousing, including dimensional modelling techniques.           c) To learn the processes of data extraction, transformation, and loading (ETL) for integrating data into a warehouse.           d) To gain knowledge of information access and delivery mechanisms, including OLAP and web-enabled data warehouse delivery.           e) To comprehend the principles of data mining, including its applications, techniques like classification and clustering, and current trends.           Course Outcomes:           1) Students will be able to articulate the need and significance of data warehousing in business intelligence.           2) Students will acquire hands-on knowledge of the ETL process for data integration and transformation.           4) Students will be quipped with skills in data mining techniques and their applications in various domains.           Theory: Course Coutents           10         Introdection to data warehouse movement, Evolution of Business intelligence.           5) Students will be quipped with skills in data mining techniques and their applications in various domains.           Course Outourse         Introdection to data warehouse movement, Evolution of Business Intelligence.           5) Students will be quipped with skills in	TH: 03 hr/week	Credits:03	CIE: 50 Marks	
Prerequisites: Database Management Systems, Data Mining Course Objectives: a) To understand the need for data warehousing, its evolution, and its significance in business intelligence. b) To explore the architecture, components, and building blocks of data warehousing, including dimensional modelling techniques. c) To learn the processes of data extraction, transformation, and loading (ETL) for integrating data into a warehouse. d) To gain knowledge of information access and delivery mechanisms, including OLAP and web-enabled data warehouse delivery. e) To comprehend the principles of data mining, including its applications, techniques like classification and clustering, and current trends. Course Outcomes: 1) Students will be able to articulate the need and significance of data warehouse architecture and dimensional modelling. Students will be proficient in designing and understanding data warehouse architecture and dimensional modelling. Students will acquire hands-on knowledge of the ETL process for data integration and transformation. 4) Students will understand how information is accessed and delivered from a data warehouse using OLAP and web-based tools. 5) Students will be due to averhousing: Escalating need for strategic information, failure of past system, operational versus decision support systems, data warehouse and bata Marts, architecture, components, metadata Trends in warehousing: Continue growth in data warehouse and bata Marts, architecture, components, metadata Trends in warehousing: Continue growth in data warehouse and bales, sincellaneous dimensions, snowflake schema, Advanced dimension modelling: Dimensional modelling basics, star schema, advantages, examples. Advanced dimension modelling: Dimensional modelling basics, star schema, advantages, examples. Advanced dimension modelling: Dimensional modelling basics, star schema, advantages, examples. Advanced dimension modelling: Dimensional modelling basics, star schema, advantages, examples. Advanced dimension modelling: Dimensional modelli			End Semester: 50 Marks	
Course Objectives: a) To understand the need for data warehousing, its evolution, and its significance in business intelligence. b) To explore the architecture, components, and building blocks of data warehousing, including dimensional modelling techniques. c) To learn the processes of data extraction, transformation, and loading (ETL) for integrating data into a warehouse. d) To gain knowledge of information access and delivery mechanisms, including OLAP and web-enabled data warehouse. d) To gain knowledge of information access and delivery mechanisms, including OLAP and web-enabled data warehouse. d) To comprehend the principles of data mining, including its applications, techniques like classification and clustering, and current trends. Course Outcomes: 1) Students will be able to articulate the need and significance of data warehouse architecture and dimensional modelling. Students will acquire hands-on knowledge of the ETL process for data integration and transformation. 4) Students will equipe hands-on knowledge of the ETL process for data integration and transformation. 4) Students will acquire hands-on knowledge of the ETL process for data integration and transformation. 4) Students will acquire hands-on knowledge of the ETL process for data integration and transformation. 5) Students will be equipped with skills in data mining techniques and their applications in various domains. 5) Students will be equipped with skills in data warehousing ging and contexpts 10 Introduction to data warehousing data warehouse movement, Evolution of Business Intelligence. 5) Bata warehousing: Escalating need for strategic information, failure of past system, operational versus decision support systems, data warehouse defined, data warehouse movement, Evolution of Business Intelligence. 5) Internation modelling blocks: Data warehouse and Data Marts, architecture, components, metadata Trends in warehousing blocks: Data warehousing, significant trends 5) Intelligence 6) Internation from the data warehouse, sindicent reds 5)	Prerequisites: Database	Management Systems, Data Mining		
a) To understand the need for data warehousing, its evolution, and its significance in business intelligence.         b) To explore the architecture, components, and building blocks of data warehousing, including dimensional modelling techniques.         c) To learn the processes of data extraction, transformation, and loading (ETL) for integrating data into a warehouse.         d) To gain knowledge of information access and delivery mechanisms, including OLAP and web-enabled data warehouse delivery.         e) To comprehend the principles of data mining, including its applications, techniques like classification and clustering, and current trends.         Course Outcomes:         1) Students will be able to articulate the need and significance of data warehouse in business intelligence.         2) Students will be proficient in designing and understanding data warehouse architecture and dimensional modelling.         3) Students will acquire hands-on knowledge of the ETL process for data integration and transformation.         4) Students will be equipped with skills in data mining techniques and their applications in various domains. <b>Unit 1</b> Introduction to data warehousing and users system, operational versus decision support systems, data warehouse ad for data warehousing using significant rends.         Need for Data warehousing: Escalating need for strategic information, failure of past system, operational versus decision is upport systems, data warehouse ad Data Marts, architecture, components, metadata Trends in warehousing: Continue growth in data warehousing, significant trends         Unit 1       Architecture and laferistroteure </td <td>Course Objectives:</td> <td></td> <td></td>	Course Objectives:			
b) 10 explore the architecture, components, and building blocks of data warehousing, including dimensional modelling techniques.         c) To learn the processes of data extraction, transformation, and loading (ETL) for integrating data into a warehouse.         d) To gain knowledge of information access and delivery mechanisms, including OLAP and web-enabled data warehouse delivery.         e) To comprehend the principles of data mining, including its applications, techniques like classification and clustering, and current trends.         Course Outcomes:         1) Students will be able to articulate the need and significance of data warehousing in business intelligence.         2) Students will be proficient in designing and understanding data warehouse architecture and dimensional modelling.         3) Students will acquire hands-on knowledge of the ETL process for data integration and transformation.         4) Students will understand how information is accessed and delivered from a data warehouse using OLAP and web-based tools.         5) Students will be equipped with skills in data mining techniques and their applications in various domains.         Theory: Course Contents         Unit 1         Introduction to data warehousing and colspan="2">(08 Hours)         versus decision support systems, data warehouse defined, data warehouse movement, Evolution of Business Intelligence.         Students will be proficient in designing and understanding and understanding and understanding and understanding and ureneousing significant trends	a) To understand the ne	ed for data warehousing, its evolution, and its	significance in business intelligence.	
c) To learn the processes of data extraction, transformation, and loading (ETL) for integrating data into a warehouse. d) To gain knowledge of information access and delivery mechanisms, including OLAP and web-enabled data warehouse delivery. e) To comprehend the principles of data mining, including its applications, techniques like classification and clustering, and current trends. Course Outcomes: 1) Students will be able to articulate the need and significance of data warehouse architecture and dimensional modelling. 3) Students will acquire hands-on knowledge of the ETL process for data integration and transformation. 4) Students will acquire hands-on knowledge of the ETL process for data integration and transformation. 4) Students will understand how information is accessed and delivered from a data warehouse using OLAP and web-based tools. 5) Students will be equipped with skills in data mining techniques and their applications in various domains. Theory: Course Contents Unit 1 Introduction to data warehousing and (08 Hours) Courses doring and concepts Unit 1 Introduction to data warehouse movement, Evolution of Business Intelligence. Data warehousing: Escalating need for strategic information, failure of past system, operational versus decision support systems, data warehouse defined, data warehouse movement, Evolution of Business Intelligence. Data ware housing building blocks: Data ware house and Data Marts, architecture, components, metadata Trends in warehousing: Continue growth in data warehousing, significant trends. Unit 1 Acchitecture and Infrastructures (08 Hunrs) Principles of dimension modelling: Unders to dimensional modelling basics, star schema, advantages, examples. Advanced dimension modelling: other integration approaches Unit H Acchitecture and Liferstructures (08 Hunrs) Unit H Information, Data Loading,	b) To explore the archit modelling techniques	ecture, components, and building blocks of s.	data warehousing, including dimensional	
<ul> <li>d) To gain knowledge of information access and delivery mechanisms, including OLAP and web-enabled data warehouse delivery.</li> <li>e) To comprehend the principles of data mining, including its applications, techniques like classification and clustering, and current trends.</li> <li>Course Outcomes:         <ol> <li>Students will be able to articulate the need and significance of data warehousing in business intelligence.</li> <li>Students will be proficient in designing and understanding data warehouse architecture and dimensional modelling.</li> <li>Students will acquire hands-on knowledge of the ETL process for data integration and transformation.</li> <li>Students will understand how information is accessed and delivered from a data warehouse using OLAP and web-based tools.</li> <li>Students will be equipped with skills in data mining techniques and their applications in various domains.</li> </ol> </li> <li> <b>Theory: Course Contents</b> </li> <li> <b>Unit   Introduction to data warehousing and</b>         (08 Hours)          <b>Course Contents Unit   Introduction to data warehousing and</b>         (08 Hours)         <b>Reed</b> for Data warehousing: Escalating need for strategic information, failure of past system, operational versus decision support systems, data warehouse and Data Marts, architecture, components, metadata Trends in warehousing: Continue growth in data warehousing, significant trends         <b>Unit   Architecture and Infristructure   OB Hours) Principles of dimension modelling:</b> Dimensional modelling basics, star schema, advantages, examples. Advanced dimension modelling: updates to dimensional tables, miscellaneous dimensions, snowflake schema, Aggregate Fact tables, families of Stars.         <b>Data Extraction, Transformat</b></li></ul>	c) To learn the process warehouse.	es of data extraction, transformation, and le	oading (ETL) for integrating data into a	
In comprehend the principles of data mining, including its applications, techniques like classification and clustering, and current trends.           Course Outcomes:         1) Students will be able to articulate the need and significance of data warehousing in business intelligence.           2) Students will be proficient in designing and understanding data warehouse architecture and dimensional modelling.           3) Students will acquire hands-on knowledge of the ETL process for data integration and transformation.           4) Students will understand how information is accessed and delivered from a data warehouse using OLAP and web-based tools.           5) Students will be equipped with skills in data mining techniques and their applications in various domains.           Theory: Course Contents           Unit I           Introduction to data warehouse movement, Evolution of Business Intelligence.           Data warehousing: Escalating need for strategic information, failure of past system, operational versus decision support systems, data warehouse defined, data warehouse movement, Evolution of Business Intelligence.           Data warehousing: Continue growth in data warehouse information, failure of past system, operational versus docision support systems, data warehouse and Data Marts, architecture, components, metadata Trends in warehousing: Continue growth in data warehousing isgnificant trends           Out III Architecture and Interstructure           (ON Hours)           Principles of dimension modelling: Dimensional modelling basics, star schema, advantages, exampl	d) To gain knowledge o	f information access and delivery mechanism	ns, including OLAP and web-enabled data	
Course Outcomes:         1) Students will be able to articulate the need and significance of data warehousing in business intelligence.         2) Students will be proficient in designing and understanding data warehouse architecture and dimensional modelling.         3) Students will acquire hands-on knowledge of the ETL process for data integration and transformation.         4) Students will understand how information is accessed and delivered from a data warehouse using OLAP and web-based tools.         5) Students will be equipped with skills in data mining techniques and their applications in various domains.         Theory: Course Contents         Unit 1       Introduction to data warehousing and concepts         (08 Hours)         Need for Data warehousing: Escalating need for strategic information, failure of past system, operational versus decision support systems, data warehouse and Data Marts, architecture, components, metadata Trends in warehousing: Continue growth in data warehousing, significant trends         Unit I         Vertice (08 Hours)         Principles of dimension modelling: Dimensional modelling basics, star schema, advantages, examples. Advanced dimension, Transformation, and Loading: ETL overview, requirement, Data extraction, Data Transformation, Tansformation from the data warehouse: Demand for OLAP, major features and functions, OLAP models, OLAP implementation considerations         Data ware house and Delivery         Unit II         Arechitecture an	<ul> <li>e) To comprehend the p clustering, and currer</li> </ul>	principles of data mining, including its applient trends.	cations, techniques like classification and	
<ul> <li>Students will be able to articulate the need and significance of data warehousing in business intelligence.</li> <li>Students will be proficient in designing and understanding data warehouse architecture and dimensional modelling.</li> <li>Students will acquire hands-on knowledge of the ETL process for data integration and transformation.</li> <li>Students will understand how information is accessed and delivered from a data warehouse using OLAP and web-based tools.</li> <li>Students will be equipped with skills in data mining techniques and their applications in various domains.</li> <li>Theory: Course Contents         <ul> <li>Unit 1</li> <li>Introduction to data warehousing and concepts</li> <li>Concepts</li> <li>Unit 1</li> <li>Introduction to data warehouse and concepts</li> <li>Introduction to data warehouse defined, data warehouse movement, Evolution of Business Intelligence.</li> </ul> </li> <li>Data warehousing: Escalating need for strategic information, failure of past system, operational versus decision support systems, data warehouse and Data Marts, architecture, components, metadata Trends in warehousing: Continue growth in data warehousing, significant trends</li> <li>Unit 11</li> <li>Architecture and Infrastructure</li> <li>Unit 11</li> <li>Architecture and Infrastructure</li> <li>Unit 11</li> <li>Architecture and Data Marts, architecture, bate extraction, Data Transformation, and Loading: ETL overview, requirement, Data extraction, Data Transformation, and Loading: ETL overview, requirement, Data extraction, Data Transformation.</li> <li>Unit 11</li> <li>Information Access and Dolivery</li> <li>(08 Hours)</li> <li>Users of Information: Information from the data warehouse: Demand for OLAP, major features and functions, OLAP models, OLAP implementation considerations</li> <li>Data Warehouse and the web: Web enables data warehouse: Dema</li></ul>	Course Outcomes:			
modelling.         3) Students will acquire hands-on knowledge of the ETL process for data integration and transformation.         4) Students will understand how information is accessed and delivered from a data warehouse using OLAP and web-based tools.         5) Students will be equipped with skills in data mining techniques and their applications in various domains.         Theory: Course Contents <b>Unit 1 Theory: Course Contents Unit 11</b> <td colspan<="" td=""><td><ol> <li>Students will be able</li> <li>Students will be pro</li> </ol></td><td>to articulate the need and significance of data ficient in designing and understanding data</td><td>a warehousing in business intelligence. warehouse architecture and dimensional</td></td>	<td><ol> <li>Students will be able</li> <li>Students will be pro</li> </ol></td> <td>to articulate the need and significance of data ficient in designing and understanding data</td> <td>a warehousing in business intelligence. warehouse architecture and dimensional</td>	<ol> <li>Students will be able</li> <li>Students will be pro</li> </ol>	to articulate the need and significance of data ficient in designing and understanding data	a warehousing in business intelligence. warehouse architecture and dimensional
<ul> <li>4) Students will understand how information is accessed and delivered from a data warehouse using OLAP and web-based tools.</li> <li>5) Students will be equipped with skills in data mining techniques and their applications in various domains.</li> <li>5) Students will be equipped with skills in data mining techniques and their applications in various domains.</li> <li>5) Students will be equipped with skills in data mining techniques and their applications in various domains.</li> <li>608 Hours)</li> <li>Concepts</li> <li>Concept</li></ul>	modelling. 3) Students will acquire	hands-on knowledge of the ETL process for	data integration and transformation.	
Students will be equipped with skills in data mining techniques and their applications in various domains.         Theory: Course Contents         Unit I         Introduction to data warchousing and concepts         (08 Hours)         Need for Data warehousing: Escalating need for strategic information, failure of past system, operational versus decision support systems, data warehouse defined, data warehouse movement, Evolution of Business Intelligence.         Data ware housing building blocks: Data ware house and Data Marts, architecture, components, metadata Trends in warehousing: Continue growth in data warehousing, significant trends         Unit II         Architecture and Infrastructure       (08 Hours)         Principles of dimension modelling: Dimensional modelling basics, star schema, advantages, examples. Advanced dimension modelling: updates to dimensional tables, miscellaneous dimensions, snowflake schema, Aggregate Fact tables, families of Stars.         Data Extraction, Transformation, and Loading: ETL overview, requirement, Data extraction, Data Transformation, Data Loading, other integration approaches         Unit II       Information Access and Delivery         Users of Information: Information from the data warehouse; whe will use information delivery, delivery tools, OLAP models, OLAP implementation considerations         Data ware house and the web: Web enables data warehouse; web-based information delivery, OLAP and the WEB, building a web-enabled data warehouse         Unit IV	4) Students will unders	tand how information is accessed and delive	ered from a data warehouse using OLAP	
Theory: Course ContentsUnit IIntroduction to data warehousing and concepts(08 Hours)Need for Data warehousing: Escalating need for strategic information, failure of past system, operational versus decision support systems, data warehouse defined, data warehouse movement, Evolution of Business Intelligence.Data ware housing building blocks: Data ware house and Data Marts, architecture, components, metadata Trends in warehousing: Continue growth in data warehousing, significant trendsUnit IIArchitecture and Infrastructure(08 Hours)Principles of dimension modelling: Dimensional modelling basics, star schema, advantages, examples. Advanced dimension modelling: updates to dimensional tables, miscellaneous dimensions, snowflake schema, Aggregate Fact tables, families of Stars. Data Extraction, Transformation, and Loading: ETL overview, requirement, Data extraction, Data Transformation, Data Loading, other integration approachesUnit IIIInformation Access and Delivery(08 Hours)Users of Information: Information from the data warehouse, who will use information, information delivery, delivery tools, OnLine Analytical Processing in the data warehouse. Demand for OLAP, major features and functions, OLAP models, OLAP implementation considerations Data ware house and the web: Web enables data warehouse, web-based information delivery, OLAP and the WEB, building a web-enabled data warehouseUnit IVData Mining(08 Hours)Why data mining, what is Data Mining, Data mining as a process of Knowledge Discovery, Major issues in Data Mining, Mining Frequent Pattern, Classification: basic concept and methods, Clustering: basic concept	<ul><li>5) Students will be equip</li></ul>	pped with skills in data mining techniques an	d their applications in various domains.	
Unit IIntroduction to data warehousing and concepts(08 Hours)Need for Data warehousing: Escalating need for strategic information, failure of past system, operational versus decision support systems, data warehouse defined, data warehouse movement, Evolution of Business Intelligence. Data ware housing building blocks: Data ware house and Data Marts, architecture, components, metadata Trends in warehousing: Continue growth in data warehousing, significant trendsUnit IIArchitecture and Infrastructure(08 Hours)Principles of dimension modelling: Dimensional modelling basics, star schema, advantages, examples. Advanced dimension modelling: updates to dimensional tables, miscellaneous dimensions, snowflake schema, Aggregate Fact tables, families of Stars. Data Extraction, Transformation, and Loading: ETL overview, requirement, Data extraction, Data Transformation, Data Loading, other integration approaches(08 Hours)Unit IIInformation Access and Delivery(08 Hours)Users of Information: Information from the data ware house, who will use information, information delivery, delivery tools, OLAP models, OLAP implementation considerations Data ware house and the web: Web enables data warehouse, web-based information delivery, OLAP and the WEB, building a web-enabled data warehouse(08 Hours)Why data mining, what is Data Mining, Data mining as a process of Knowledge Discovery, Major issues in Data Mining, Mining Frequent Pattern, Classification: basic concept and methods, Clustering: basic concept		Theory: Course Contents	1	
Need for Data warehousing: Escalating need for strategic information, failure of past system, operational versus decision support systems, data warehouse defined, data warehouse movement, Evolution of Business Intelligence.         Data ware housing building blocks: Data ware house and Data Marts, architecture, components, metadata Trends in warehousing: Continue growth in data warehousing, significant trends         Unit II       Architecture and Infrastructure       (08 Hours)         Principles of dimension modelling: Dimensional modelling basics, star schema, advantages, examples. Advanced dimension modelling: updates to dimensional tables, miscellaneous dimensions, snowflake schema, Aggregate Fact tables, families of Stars.         Data Extraction, Transformation, and Loading: ETL overview, requirement, Data extraction, Data Loading, other integration approaches       (08 Hours)         Users of Information: Information from the data ware house, who will use information, information delivery, delivery tools, OnLine Analytical Processing in the data warehouse: Demand for OLAP, major features and functions, OLAP models, OLAP implementation considerations       Data ware house and the web: Web enables data warehouse, web-based information delivery, OLAP and the WEB, building a web-enabled data warehouse         Unit IV       Data Mining       (08 Hours)         Why data mining, What is Data Mining, Data mining as a process of Knowledge Discovery, Major issues in Data Mining, Mining Frequent Pattern, Classification: basic concept and methods, Clustering: basic concept	Unit I	Introduction to data warehousing and concepts	(08 Hours)	
Interligence.Data ware housing building blocks: Data ware house and Data Marts, architecture, components, metadata Trends in warehousing: Continue growth in data warehousing, significant trendsUnit IIArchitecture and Infrastructure(08 Hours)Principles of dimension modelling: Dimensional modelling basics, star schema, advantages, examples. Advanced dimension modelling: updates to dimensional tables, miscellaneous dimensions, snowflake schema, Aggregate Fact tables, families of Stars. 	Need for Data warehou versus decision support	using: Escalating need for strategic information systems, data warehouse defined, data warehouse defin	ation, failure of past system, operational ehouse movement, Evolution of Business	
Unit IIArchitecture and Infrastructure(08 Hours)Principles of dimension modelling: Dimensional modelling basics, star schema, advantages, examples. Advanced dimension modelling: updates to dimensional tables, miscellaneous dimensions, snowflake schema, Aggregate Fact tables, families of Stars. Data Extraction, Transformation, and Loading: ETL overview, requirement, Data extraction, Data 	Data ware housing build	ing blocks: Data ware house and Data Marts, Continue growth in data warehousing signifi	architecture, components, metadata	
Principles of dimension modelling: Dimensional modelling basics, star schema, advantages, examples.Advanced dimension modelling: updates to dimensional tables, miscellaneous dimensions, snowflake schema, Aggregate Fact tables, families of Stars.Data Extraction, Transformation, and Loading: ETL overview, requirement, Data extraction, Data Transformation, Data Loading, other integration approachesUnit IIIInformation Access and Delivery(08 Hours)Users of Information: Information from the data ware house, who will use information, information delivery, delivery tools, OnLine Analytical Processing in the data warehouse: Demand for OLAP, major features and 	Unit II	Architecture and Infrastructure	(08 Hours)	
Aggregate Fact tables, families of Stars.DataExtraction, Transformation, and Loading: ETL overview, requirement, Data extraction, DataTransformation, Data Loading, other integration approachesUnit IIIInformation Access and DeliveryUsers of Information: Information from the data ware house, who will use information, information delivery, delivery tools, OnLine Analytical Processing in the data warehouse: Demand for OLAP, major features and functions, OLAP models, OLAP implementation considerationsData ware house and the web: Web enables data warehouse, web-based information delivery, OLAP and the WEB, building a web-enabled data warehouseUnit IVData MiningWhy data mining, what is Data Mining, Data mining as a process of Knowledge Discovery, Major issues in Data Mining, Mining Frequent Pattern, Classification: basic concept and methods, Clustering: basic concept	Principles of dimension Advanced dimension me	n modelling: Dimensional modelling basic odelling: updates to dimensional tables, misc	cs, star schema, advantages, examples. ellaneous dimensions, snowflake schema,	
Transformation, Data Loading, other integration approachesUnit IIIInformation Access and Delivery(08 Hours)Users of Information: Information from the data ware house, who will use information, information delivery, delivery tools, OnLine Analytical Processing in the data warehouse: Demand for OLAP, major features and functions, OLAP models, OLAP implementation considerations Data ware house and the web: Web enables data warehouse, web-based information delivery, OLAP and the WEB, building a web-enabled data warehouseUnit IVData Mining(08 Hours)Why data mining, what is Data Mining, Data mining as a process of Knowledge Discovery, Major issues in Data Mining, Mining Frequent Pattern, Classification: basic concept and methods, Clustering: basic concept	Aggregate Fact tables, families of Stars. Data Extraction, Transformation, and Loading: ETL overview, requirement, Data extraction, Data			
Unit IIIInformation Access and Delivery(08 Hours)Users of Information: Information from the data ware house, who will use information, information delivery, delivery tools, OnLine Analytical Processing in the data warehouse: Demand for OLAP, major features and functions, OLAP models, OLAP implementation considerations Data ware house and the web: Web enables data warehouse, web-based information delivery, OLAP and the 	Transformation, Data Loading, other integration approaches			
Osers of information: information from the data ware house, who will use information, information derivery, delivery tools, OnLine Analytical Processing in the data warehouse: Demand for OLAP, major features and functions, OLAP models, OLAP implementation considerations         Data ware house and the web: Web enables data warehouse, web-based information delivery, OLAP and the WEB, building a web-enabled data warehouse         Unit IV       Data Mining         Why data mining, what is Data Mining, Data mining as a process of Knowledge Discovery, Major issues in Data Mining, Mining Frequent Pattern, Classification: basic concept and methods, Clustering: basic concept		Information Access and Delivery	(08 Hours)	
Intervery tools, OnEnic Analytical Processing in the data watchouse. Demand for OEAF, major relatives and functions, OLAP models, OLAP implementation considerations         Data ware house and the web: Web enables data warehouse, web-based information delivery, OLAP and the WEB, building a web-enabled data warehouse         Unit IV       Data Mining         Why data mining, what is Data Mining, Data mining as a process of Knowledge Discovery, Major issues in Data Mining, Mining Frequent Pattern, Classification: basic concept and methods, Clustering: basic concept	Users of information: information from the data ware house, who will use information, information delivery,			
Interesting of the interest	functions. OLAP models. OLAP implementation considerations			
WEB, building a web-enabled data warehouse         Unit IV       Data Mining       (08 Hours)         Why data mining, what is Data Mining, Data mining as a process of Knowledge Discovery, Major issues in Data Mining, Mining Frequent Pattern, Classification: basic concept and methods, Clustering: basic concept	Data ware house and the web: Web enables data warehouse, web-based information delivery, OLAP and the			
Unit IVData Mining(08 Hours)Why data mining, what is Data Mining, Data mining as a process of Knowledge Discovery, Major issues in Data Mining, Mining Frequent Pattern, Classification: basic concept and methods, Clustering: basic concept	WEB, building a web-enabled data warehouse			
Why data mining, what is Data Mining, Data mining as a process of Knowledge Discovery, Major issues in Data Mining, Mining Frequent Pattern, Classification: basic concept and methods, Clustering: basic concept	Unit IV	Data Mining	(08 Hours)	
	Why data mining, what Data Mining, Mining Fi	is Data Mining, Data mining as a process or requent Pattern, Classification: basic conception	of Knowledge Discovery, Major issues in t and methods, Clustering: basic concept	





and methods, Data Mining Trends and Research Frontiers : Mining Complex Data Types, other Methodologies of Data Mining, Data Mining Applications, Data Mining trends.

Unit VImplementation and Maintenance(08 Hours)Physical design process: Physical design steps, considerations, physical storage, indexing, performance<br/>enhancement techniques. Data warehouse development: Data warehouse testing, major deployment activity,<br/>security, backup and recovery.Growth and Maintenance: Monitoring the data warehouse, user training and support, managing the data<br/>warehouse.

- 1. Paulraj Ponniah, Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals. John Wiley & Sons, Inc.
- 2. Marakas, George M. Modern data warehousing, mining, and visualization: core concepts. Upper Saddle River, NJ: Prentice Hall.
- 3. Parteek Bhatia, Data Mining and Data Warehousing: Principles and Practical Techniques, Cambridge.
- 4. Matthew A. Russell, Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More, Shroff Publishers
- 5. G.K. Gupta, Introduction to Data Mining with Case Studies, PHI Learning Private Limited

MCE25633B1A: Soft Computing			
Teaching Scheme: TH:03 hr/week	Credits:03	Examination Scheme: CIE: 50 Marks End Semester: 50 Marks	
Prerequisites: Mathematic	s, Algorithms and analysis, Programming Langua	age Python/ Java/C++	
<ul> <li>Course Objectives:</li> <li>a) To introduce the basic concepts of soft computing, fuzzy logic, and their real-world applications.</li> <li>b) To explore fuzzy systems, including membership functions, fuzzification, defuzzification, and fuzzy inference systems.</li> <li>c) To understand the working principles of genetic algorithms and their applications in optimization problems.</li> <li>d) To study neural networks, including supervised and unsupervised learning, and their applications in associative memory.</li> <li>e) To examine hybrid systems, such as neuro-fuzzy and genetic-fuzzy systems, and their role in solving</li> </ul>			
complex real-world pro	oblems.		
<ol> <li>Students will understand the fundamental concepts and components of soft computing and fuzzy logic.</li> <li>Students will gain knowledge of fuzzy systems, including fuzzy reasoning, inference systems, and fuzzy control.</li> <li>Students will develop the ability to apply genetic algorithms for optimization and problem-solving.</li> <li>Students will be able to design and implement neural networks and understand their learning principles.</li> </ol>			
5) Students will learn to integrate soft computing techniques in hybrid systems to solve complex applications like robotics and optimization problems.			
Theory: Course Contents			
Unit I	Introduction to Soft Computing and Fuzzy logic	(08 Hours)	
Introduction to soft computing, Paradigms soft computing, Features, Components, Techniques, Applications, Neural Networks, Fuzzy logic, Genetic Algorithms, Hybrid systems, Introduction to Fuzzy logic: Classical and Fuzzy sets, operations, properties, Fuzzy Relations.			
Unit II	Fuzzy Systems	(08 Hours)	





Membership Functions, F	uzzification and Methods, Defuzzification and N	Alethods, Fuzzy Logic, Fuzzy Rules	
and Fuzzy Reasoning, F	uzzy Inference Systems, Fuzzy Expert System	s, Fuzzy Decision Making. Fuzzy	
Control Systems, Fuzzy C	lassification.		
Unit III	Genetic Algorithms	(08 Hours)	
Introduction to Genetic A	Algorithms (GA), Search space, Working Princi	ple, Simple GA, Operators, Fitness	
function, Multi-level Optim	mization		
Unit IV	Introduction to Neural Networks And Advanced Neural Networks	(08 Hours)	
Naural Naturals Model	Auvaliced Neural Networks	ing miles and regions activation	
Ineural Network, Model	s, Terminologies, Supervised Learning, Learning	ing fules and various activation	
Iunctions, Single layer Per	Managina Hanfield Networks, Architectu	ire of Backpropagation(BP), Neural	
Networks as Associative	Memories - Hopfield Networks, Bidirection	al Associative Memory Networks	
Unsupervised Learning.			
Unit V	Hybrid Systems and Soft Computing	(08 Hours)	
	Applications	(00 110 110)	
Kohonen Self Organizing	g Maps and Counter Propogation Networks Networks	eural Network Classification, Deep	
learning, Neuro-Fuzzy Hy	brid Systems, Genetic Neuro Hybrid Systems, I	Fuzzy-Genetic Hybrid Systems. GA	
based Back propagation N	etworks, Optimization of TSP using GA, Fuzzy c	controllers for Robot.	
References:			
1. S.N. Sivanandam & S	N.Deepa, Principles of Soft computing, John Wi	ley & Sons	
2. S.Rajasekaran, G. A.	Vijayalakshami, Neural Networks, Fuzzy Logic a	nd Genetic Algorithms: Synthesis &	
Applications, PHI.			
3. David E. Goldberg., Genetic Algorithms: in Search and Optimization, PHI			
5. Durid D. Goldovig., C	enetic Algorithms: in Search and Optimization,P	HI	
4. Jyh:Shing Roger Jang	enetic Algorithms: in Search and Optimization,P , Chuen:Tsai Sun, EijiMizutani, Neuro:Fuzzy an	HI Id Soft Computing, Prentice:Hall of	
<ol> <li>Juria E. Goldorigi, C.</li> <li>Jyh:Shing Roger Jang India.</li> </ol>	enetic Algorithms: in Search and Optimization,P , Chuen:Tsai Sun, EijiMizutani, Neuro:Fuzzy an	HI Id Soft Computing, Prentice:Hall of	
<ol> <li>Juria E. Goldergi, C.</li> <li>Jyh:Shing Roger Jang India.</li> <li>Pratihar , Soft Computation</li> </ol>	enetic Algorithms: in Search and Optimization,P , Chuen:Tsai Sun, EijiMizutani, Neuro:Fuzzy an ting: Fundamentals and Applications, Narosa Pub	HI Id Soft Computing, Prentice:Hall of Ishing House Pvt. Ltd.	
<ol> <li>Jyh:Shing Roger Jang India.</li> <li>Pratihar , Soft Computer</li> </ol>	enetic Algorithms: in Search and Optimization,P , Chuen:Tsai Sun, EijiMizutani, Neuro:Fuzzy an ting: Fundamentals and Applications, Narosa Pub	HI Id Soft Computing, Prentice:Hall of Ilishing House Pvt. Ltd.	

MCE25633B2A: Deep Learning				
Te Tl	Teaching Scheme: TH:03 hr/weekCredits:03Examination Scheme: CIE: 50 Marks End Semester: 50 Marks		Examination Scheme: CIE: 50 Marks End Semester: 50 Marks	
Pr	erequisites: Mathemat	ical Foundation of Data Science, Mac	hine Learning	3
Co	ourse Objectives:			
a)	To introduce the fo networks, and backpr	undational concepts of deep learn opagation.	ing, including	g perceptrons, feedforward neural
b)	To explore optimiza challenges like the va	tion techniques in deep learning, s nishing gradient problem.	such as gradi	ent descent methods, and address
c)	c) To understand regularization techniques and convolutional neural networks (CNNs), focusing on architecture and applications in computer vision.			
d)	d) To delve into deep unsupervised learning techniques, including autoencoders, and the role of CNNs in feature extraction and image classification.			
e)	e) To study sequence models like RNNs, LSTMs, and GRUs, and their applications in natural language processing, image captioning, and video processing.			
Course Outcomes:				
1)	) Students will be able to explain the basic principles of deep learning, including neural networks and their training algorithms.			
2)	) Students will understand optimization techniques in deep learning and address common challenges in training neural networks.			





3) Students will learn t	3) Students will learn to implement regularization strategies and understand the use of CNNs in real-world		
applications like image recognition.			
4) Students will be cap	able of applying deep unsupervised learning techr	iques like autoencoders for feature	
extraction and classif	ication tasks.		
5) Students will acquir	e knowledge in sequence models and apply the	m in natural language processing,	
machine translation,	and other sequential tasks.		
<b></b>	Theory: Course Contents		
	Foundations of Deep learning		
Perceptron's, Perceptro	n Learning Algorithm, Sigmoid Neuron, Shallo	ow neural networks, Deep neural	
Deep learning	Neural networks, Gradient descent and the backpro	opagation algorithm Applications of	
Deep learning.	Ontimization and Deep Learning	(09 Hours)	
	Optimization and Deep Learning	(US HOURS)	
Learning Parameters of Roll Houristics for an	a feedforward neural network, the vanishing gradie	and problem, and ways to mitigate it,	
Stochastia CD AdaCre	olding dat local minima, neuristics for laster the	ainal Component Analysis and its	
interpretations Singular	Value Decomposition	cipal Component Analysis and its	
interpretations, Singular	Permission Techniques and		
Unit III	Convolutional Neural Networks.	(08 Hours)	
Bias Variance Trade o	ff, L2 regularization, Early stopping, Dataset au	gmentation, Parameter sharing and	
tying, Injecting noise at	input, Ensemble methods, Dropout, Convolutional	Neural Networks, LeNet, AlexNet,	
ZF-Net, VGGNet, O	GoogLeNet, ResNet, Visualizing Convolutio	nal Neural Networks, Guided	
Backpropagation, Deep	Dream, Deep Art, Fooling Convolutional Neural N	etworks,	
Unit IV	Deep Unsupervised Learning	(08 Hours)	
Convolutional Neural	Networks, LeNet, AlexNet, ZF-Net, VGGNet,	GoogLeNet, ResNet, Visualizing	
Convolutional Neural Networks, Guided Backpropagation, Deep Dream, Deep Art, Fooling Convolutional			
Neural Networks, Feature Extraction using autoencoders. Image classifier for identifying cat vs dogs using			
CNN.			
Unit V	Sequence Models	(08 Hours)	
RNN, LSTM, GRU models. Application to NLP, language models, machine translation, image, captioning,			
video processing, visual question answering, video processing, learning from descriptions, Attention			
Mechanism, Attention over images, Sentiment Analysis using Recurrent Neural Networks.			
References:			
1. Neural Networks: A Systematic Introduction, Raúl Rojas.			
2. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press.			
3. Deep Learning Part-I, Swayam Prof. Mitesh M. Khapra			
4. Andrew Ng, Neural Networks and Deep Learning, Coursera.			
5. Prof. Vineeth N Balasubramanian, Deep Learning for Computer Vision.			





MCE25633B3A: Computational Linguistic Analytics			
<b>Teaching Scheme:</b>		Examination Scheme:	
TH:03 hr/week	Credits:03	CIE: 50 Marks	
Prerequisites: Theory of	Computation	End Semester: 50 Marks	
Course Objectives:	Computation		
a) To introduce the fund	lamentals of natural language processing (NLP) and	text classification methods like	
Naive Bayes, logistic	regression, and regularization.	i text elussification methods fike	
b) To explore advance	ed machine learning techniques, including nonli	inear classification and neural	
networks, applied to 1	inguistic tasks such as sentiment analysis and word	sense disambiguation.	
c) To understand comp	utational morphology, language models, and the	application of Hidden Markov	
Models and sequence	labeling in machine translation.		
d) To delve into advanc	ed computational semantics, focusing on sequence	labeling, part-of-speech tagging,	
named entity recogni	tion, and dialogue systems.		
e) To study statistical N	LP and sentiment analysis techniques, including for	ormal language theory, semantic	
parsing, and neural w	ford embeddings.		
Course Outcomes:	and the same concents of NLD and linear text of	aggification techniques for text	
analysis	and the core concepts of NLF and linear text cr	assilication techniques for text	
2) Students will gain expe	ertise in advanced machine learning models like ne	ural networks and apply them to	
linguistic tasks and class	ssifiers.	and networks and apply them to	
3)Students will learn to	apply computational morphology techniques, la	anguage models, and sequence	
labeling in machine tra	nslation systems.		
4)Students will be able	to apply advanced sequence labeling methods and	1 NLP techniques for tasks like	
named entity recognition	on and dialogue systems.		
5)Students will develop	skills in statistical NLP, sentiment analysis, and	semantic parsing for effective	
natural language under	standing.		
Init I	Learning	(06 Hours)	
Natural language proce	Examing resing and its neighbors. Themes in natural lat	(00 Hours)	
classification: Naive B	aves Discriminative learning Loss functions a	nd large-margin classification	
Logistic regression, Opti	mization, Feature selection by regularization		
Unit II	Advanced Machine Learning	(08 Hours)	
Nonlinear classification	, Feedforward neural networks, Designing neu	ral networks, Learning neural	
networks, Convolutional	l neural networks, Linguistic applications of classi	fication: Sentiment and opinion	
analysis, Wordsense di	isambiguation, Design decisions for text classif	fication, Evaluating classifiers,	
Building datasets, Learn	ing without supervision, Semi-supervised learning,	Domain adaptation	
Unit III	Computational Morphology and Machine	(08 Hours)	
	Translation		
Language models: N-g	gram language models, Smoothing and discount	ing, Recurrent neural network	
language models, Evaluating language models, Sequence labeling as classification, Sequence labeling as			
Unit IV	Advanced Computational Semantics	(08 Hours)	
Discriminative sequence	labeling with features. Neural sequence labeling	Insupervised sequence labeling	
Applications of sequen	ce labeling: part-of-speech tagging Morphosyn	tactic Attributes, named entity	
recognition. Named Enti	ty Recognition, Tokenization, Code switching, Dial	ogue acts	
Unit V	Statistical NLP and Sentiment Analysis	(10 Hours)	
Formal language theory	: Regular languages, Context-free languages, Mil	dly context-sensitive languages,	





Context-free parsing, Weighted Context-Free Grammars, Learning weighted context-free grammars, Grammar refinement, Beyond context-free parsing, Dependency parsing, Logical semantics: logical representations of meaning, Semantic parsing and the lambda calculus, learning semantic parsers, Predicate-argument semantics, semantic role labeling, Distributional and distributed Semantics, design decisions for word representations, Neural word embedding.

- 1. Dan Jurafsky and James H. Martin. Speech and Language Processing, Prentice-Hall, 2000
- 2. Igor A. Bolshakov and Alexander Gelbukh, Computational Linguistics : Models, Resources, Applications
- 3. Patrick Blackburn and Kristina Striegnitz (BS), Natural Language Processing Techniques in Prolog
- 4. Patrick Blackburn and Johan Bos (BB1), Representation and Inference for Natural Language A First Course in Computational Semantics
- 5. Dipanjan Sarkar, Text Analytics with Python: A Practical Real-World Approach to Gaining Actionable Insights from your Data

MCE25634A0B: Industry Internship-1 OR In-house Research Project-1		
Teaching Scheme: PR: 06 hr/week	Credits:03	Examination Scheme: TW: 50 Marks
<ul> <li>Course Objectives:</li> <li>a) To identify the domain of research</li> <li>b) To learn to communicate in a scientific language through collaboration with a guide.</li> <li>c) To categorize the research material confined to the domain of choice</li> </ul>		
<ul> <li>Course Outcomes:</li> <li>1) Conduct thorough literature survey confined to the domain of choice.</li> <li>2) Develop presentation skills to deliver the technical contents</li> <li>3) Furnish the report of the technical research domain</li> <li>4) Analyze the findings and work of various authors confined to the chosen domain</li> </ul>		
<b>Conduction guidelines</b> The preferences/choices of the domain will be taken from the students. The guide needs to be allocated based on the preference/choices. In case of Industry Internship-I, the assigned guide from college has to monitor and evaluate the progress of the student. The student has to exhibit the continuous progress through regular reporting and presentations and proper documentation. The continuous assessment of the progress needs to be documented unambiguously.		





MCE25635A0B: Dissertation Stage–I			
Teaching Scheme: TW/PR: 14 hr/week	Credits:07	Examination Scheme: OR: 50 Marks TW: 50 Marks	
Course Objectives:			
a) To identify the doma	ain of research		
b) To learn to commun	icate in a scientific language through collab	oration with a guide.	
c) To categorize the res	search material confined to the domain of ch	noice	
Course Outcomes:	tomations approved and to the domain of al		
<ol> <li>Conduct thorough if</li> <li>Develop presentation</li> </ol>	n skills to deliver the technical contents	loice.	
3) Furnish the report of	f the technical research domain		
<ul><li>4) Analyse the findings</li></ul>	s and work of various authors confined to th	e chosen domain	
Dissertation Stage–I:	is an integral part of the Dissertation	work. In this, the student shall	
complete the partial w	ork of the Dissertation which will consist	of problem statement, literature	
review, design, scheme	e of implementation (Mathematical Model	/SRS/UML/ERD/block diagram/	
PERT chart,) and Lay	yout & Design of the Set-up. The stude	nt is expected to complete the	
dissertation at least up	to the design phase. As a part of the prog	ress report of Dissertation work	
Stage-I, the candidate	shall deliver a presentation on the advancer	nent in Technology pertaining to	
the selected dissertatio	n topic. The student shall submit the duly	approved and certified progress	
report of Dissertation	Stage-I in standard format for satisfactory	completion of the work by the	
concerned guide and he	ead of the Department. The student will be	assessed by a panel of examiners	
of which one is nece	ssarily an external examiner. The assess	ment will be broadly based on	
literature study work	undergone content delivery presentation s	kills documentation and report	
The students are expec	ted to validate their study undertaken by pu	blishing it at standard platforms	
The investigations an	d findings need to be validated appropri	riately at standard platforms	
and investigations and interings need to be variated appropriately at standard platforms -			
contribute and/or peer reviewed journal. The student has to exhibit the continuous progress			
unough regular reporting and presentations and proper documentation of the frequency of the			
activities is at the sole	activities is at the sole discretion of the PG coordinator. The continuous assessment of the progress		
needs to be documented unambiguously. For standardization and documentation, it is			
recommended to Iollo	w the formats and guidelines circulated /	as in the dissertation workbook	
approved by the Board of Studies. Follow guidelines and formats as mentioned in Dissertation			
Workbook.			





MCE25636A0A: Personality Development through Life Enlightenment Skills.					
Teaching Scheme: TH:01 hr/week	Credits:01	Examination Scheme: CIE: 25 Marks			
Course Objectives:					
a) To introduce the cond	cept of personality development and explore differer	t personality theories.			
b) To develop an unders	tanding of SWOT analysis and its role in personal s	uccess and failure.			
c) To provide insights in	nto attitude formation, motivation, and self-esteem, a	and their impact on personality.			
d) To enhance skills in b	oody language, problem-solving, conflict manageme	nt, and emotional intelligence.			
e) To foster leadership,	teamwork, time management, and decision-makin	ig skills for effective personality			
growin.					
1) Students will gain kn	owledge of personality theories and the key dimensi	ons of personality development			
2) Students will be able	to conduct a SWOT analysis and use it to assess stro	engths and weaknesses.			
3) Students will underst	and the role of attitude, motivation, and self-esteem	in shaping one's personality.			
4) Students will improv	e their interpersonal skills, including emotional inter-	elligence, leadership, and conflict			
resolution.					
5) Students will acquire	essential skills for time management, decision-mak	ing, and effective teamwork.			
<b>TT •/ T</b>	Theory: Course Contents				
Unit I	Introduction	(07 Hours)			
Theories of Personality of	development (Freud & Erickson).	Dimensions of Personality,			
Unit II	SWOT Analysis	(06 Hours)			
The concept of Success	s and Failure, Factors responsible for Success, H	urdles in achieving Success and			
Threats) analysis	auses of failure, Conducting SwO1 (Strengths,	weaknesses, Opportunities and			
Unit III	Attitude, Motivation and Self-esteem	(06 Hours)			
Attitude, Motivation and	Self-esteem: Conceptual overview of Attitude, Typ	es of Attitudes,			
Attitude Formation, A	dvantages/Disadvantages of Positive/Negative A	Attitude - Ways to Develop			
Init IV	Motivation	(06 Hours)			
Concept of motivation: I	Definition and Nature of Motivation/Motive. Interna	1 and external motives. Theories			
of Motivation, Important	ce of self- motivation- Factors leading to de- motiva	tion. Self-esteem - Definition			
and Nature of self-esteer	n, Do's and Don'ts				
to develop positive s	elf- esteem, Low self-esteem.				
Unit V	Personality Development	(06 Hours)			
Other Aspects of Pers	onality Development: Body language, Problem-s	olving, Conflict Management			
and Negotiation- Decision-making skills, Leadership and qualities of a successful leader, Character building -					
Performances:					
1) Hurlock F. P. Dersonality Development 28th Penrint New Delhi: Tete McGrew Hill					
2) Gopinath.Rashtriva Sanskrit Sansthanam P. Bhartrihari''s ThreeSatakam. Niti- sringar-vairagva. New					
Delhi, 2010					
3) Swami Swarupananda, Srimad Bhagavad Gita, Advaita Ashram, Publication Department, Kolkata					
4) Dr. Ashish Kumar Ra	i, A Guide to Personality Development, Bigfoot Pul	plications.			
5) Dr. Neha Garg, Life S	kills and Personality Development, Notion Press	5) Dr. Neha Garg, Life Skills and Personality Development, Notion Press			



# ARMY INSTITUTE OF TECHNOLOGY, PUNE An Autonomous Institute Affiliated to Savitribai Phule Pune University, Maharashtra, India

**National Education Policy (NEP) Compliant Curriculum** 

# **Semester - IV**



M. Tech. Data Science (2025 Pattern)

www.aitpune.com





MCE25641A0A · English for Research Paper Writing		
Teaching Scheme: TH:02 hr/week	Credits:02	Examination Scheme: CIE: 50 Marks
Prerequisites: Basic Kno	wledge of English Language	
Course Objectives:		
a) To understand the key	y principles of effective planning and preparation in	writing.
b) To develop skills in p	paraphrasing, conciseness, and clarity in academic w	riting.
c) To recognize and avo	id plagiarism while effectively integrating literature	into writing.
d) To improve writing s	kills, focusing on structuring reviews, methods, and	results.
e) To master the use of	phrases and refine writing through thorough final ch	ecks.
Course Outcomes:		
1) Students will enhance	e their ability to plan, structure, and organize their w	riting for clarity and coherence.
2) Students will apply p 3) Students will demon	strate knowledge of plagingism avoid redundancy, among	athical integration of literature
into their work	strate knowledge of plagfalish avoidance and the	ethical integration of interature
4) Students will effecti	vely write and structure sections like literature re	views methods and results in
academic papers.	very write and structure sections like incluture re	views, methods, and results m
5) Students will refine the	heir writing with proper phrasing, ensuring quality a	nd thorough final revisions.
	Theory: Course Contents	
Unit I	Planning and Preparation	(06 Hours)
Planning and Preparation	n, Word Order, Breaking up Long Sentences, Structu	ring Paragraphs and Sentences,
Unit II	Paraphrasing	(06 Hours)
Being Concise and Rem	oving Redundancy, Avoiding Ambiguity and Vaguer	ness, Clarifying Who Did What,
Highlighting Your Findi	ngs.	
Unit III	Plagiarism and Literature	(06 Hours)
Hedging and Criticising,	Paraphrasing and Plagiarism, Titles, Abstracts.	
Unit IV	Writing Skills	(06 Hours)
Introduction, Writing a Review of the Literature, Methods, Results.		
Unit V	Phrases and The Final Check	(06 Hours)
Discussion, Conclusions	, Useful Phrases, The Final Check.	
References:		
I. Adrian Wallwork, I	English for Writing Research Papers, Springer No	ew York Dordrecht Heidelberg
London, 2 Coldbort P Writing	for Science, Vale University Press (available on God	ade Books)

- 2. Goldbort R Writing for Science, Yale University Press (available on Google Books)
- 3. Day R How to Write and Publish a Scientific Paper, Cambridge University Press.
- 4. Highman N, Handbook of Writing for the Mathematical Sciences, SIAM, Highman's book.
- 5. Kate L. Turabian, Wayne C. Booth, Gregory G. Colomb, Joseph M. Williams , Joseph Bizup A Manual for Writers of Research Papers, Theses, and Dissertations, University of Chicago Press.





MCE25642A0A: Intellectual Property Rights			
Teaching Scheme: Th: 02 hr/week	Credits:02	Examination Scheme: CIE: 50 Marks	
<ul> <li>Course Objectives:</li> <li>a) To introduce the fundamentals and evolution of Intellectual Property Rights (IPR).</li> <li>b) To explore various types of Intellectual Property Rights and their scope in India and globally.</li> <li>c) To understand the process and practical aspects of registering different forms of IPRs.</li> <li>d) To analyze international agreements, conventions, and Indian laws governing IPRs.</li> <li>e) To examine digital innovations, IP laws, and enforcement measures in the context of modern challenges.</li> </ul>			
<ul> <li>Course Outcomes:</li> <li>1) Students will gain a foundational understanding of IPRs and their global significance.</li> <li>2) Students will identify and differentiate between various types of Intellectual Property Rights.</li> <li>3) Students will comprehend the registration process for IPRs in India and abroad.</li> <li>4) Students will critically assess international treaties and Indian legislation related to IPR.</li> <li>5) Students will evaluate digital innovations and their relation to IP laws, including enforcement and emerging issues.</li> </ul>			
Unit I	Introduction to IPR	(06 Hours)	
Introduction to IPRs, IPR in India and Abroad, Genesis and Development, the way from WTO to WIPO ,TRIPS, Nature of Intellectual Property, Industrial Property, technological Research, Inventions and Innovations Important examples of IPR			
Unit II	Details of IP	(06 Hours)	
Understanding the types of Intellectual Property Rights: -Patents-Indian Patent Office and its Administration, Administration of Patent System, Patenting under Indian Patent Act, Patent Rights and its Scope, Patent information and database, Integrated Circuits, Industrial Designs, Trademarks (Registered and unregistered trademarks), Copyrights, Geographical Indications, Trade Secrets, Geographical Indications.			
Unit III	Registration OF IPRs	(06 Hours)	
Meaning and practical Indications, Trade Secre	aspects of registration of Copy Rights, Trad ts and Industrial Design registration in India and Ab	emarks, Patents, Geographical road.	
Unit IV	Agreements and Legislations	(06 Hours)	
International Treaties and Conventions on IPRs, TRIPS Agreement, PCT Agreement, Patent Act of India, Patent Amendment Act, Design Act, Trademark Act, Geographical Indication Act.			
Unit V	Digital Products, Laws and Enforcement OF IPRs	(06 Hours)	
Digital Innovations and	Developments as Knowledge Assets, IP Laws, G	Cyber Law and Digital Content	
Protection, Unfair Competition, Meaning and Relationship between Unfair Competition and IP Laws,			
Case Studies, Infringement of IPRs, Enforcement Measures, Emerging issues, Case Studies.			
References:			
1. Gangun Prabudona, intellectual Property RightsUnleashing the Knowledge Economy, Tata McGrawHill.			
2. Umakant Dinkar Butkar, Intellectual Property Rights Indian Patent act 1970, Namya Press.			





- 3. Siva Vaidhyanathan, Intellectual Property: A Very Short Introduction, OXFORD.
- 4. Peter Goodhart, Intellectual Property: Law & the Information Society
- 5. V K Ahuja, Intellectual Property Rights in India, Lexis Nexis

MCE25643A0B: Industry Internship-2 OR In-house Research Project-2				
Teaching Scheme: PR: 06 hr/week	Credits:03	Examination Scheme: TW: 150 Marks PR: 150 Marks		
Course Objectives: a) To identify the domain of research				
<ul> <li>b) To learn to communicate in a scientific language through collaboration with a guide.</li> <li>c) To categorize the research material confined to the domain of choice</li> </ul>				
Course Outcomes:				
1) Conduct thorough literature survey confined to the domain of choice.				
2) Develop presentation skills to deliver the technical contents				
3) Furnish the report of the technical research domain				
4) Analyze the findings and work of various authors confined to the chosen domain				
Conduction guidelines				
The preferences/choices of the domain will be taken from the students. The guide needs to be allocated				
based on the preference/choices. In case of Industry Internship, the assigned guide from college has to				
monitor and evaluate the progress of the student. The student has to exhibit the continuous progress through				
regular reporting and presentations and proper documentation. The continuous assessment of the progress				
needs to be documented unambiguously.				

MCE25644A0B: Dissertation Stage II					
Tea TW	ching Scheme: //PR: 26 hr/week	Credits:13	Examination Scheme: TW: 150 Marks PR: 150 Marks		
Course Objectives:					
a)	a) To identify the domain of research				
b) To learn to communicate in a scientific language through collaboration with a guide.					
c) To categorize the research material confined to the domain of choice					
d)	d) To work in professional environment				
Course Outcomes:					
1)	1) Conduct thorough literature survey confined to the domain of choice				
2)	2) Develop presentation skills to deliver the technical contents				
3)	3) Furnish the report of the technical research domain				
4)	Analyse the findings and work of various authors confined to the chosen domain				

#### Guidelines:

In Dissertation Work Stage-II, the student shall consolidate and complete the remaining part of the dissertation which will consist of Selection of Technology, Installations, UML implementations, testing, results, measuring performance, discussions using data tables per parameter considered for the improvement





with existing/known algorithms/systems, comparative analysis, validation of results and conclusions. The student shall prepare the duly certified final report of Dissertation in the standard format for satisfactory completion of the work by the concerned guide and the head of the Department. The students are expected to validate their study undertaken by publishing it at standard platforms. The investigations and findings need to be validated appropriately at standard platforms, conference and/or peer reviewed journal.

The student has to exhibit continuous progress through regular reporting and presentations and proper documentation of the frequency of the activities in the sole discretion of the PG coordinator. The continuous assessment of the progress needs to be documented unambiguously. It is recommended to continue with guidelines and formats as mentioned in the Dissertation Workbook approved by the Board of Studies.