SNDT Women's University, Mumbai

Master of Science (Data Science)
(M.Sc.- DS.)

as per NEP-2020

Syllabus

w.e.f.

A.Y.: 2023-24

Programme	Master of Science (Data Science) (M.Sc.- DS.)
Preamble	In pursuit of academic excellence and a comprehensive understanding of the rapidly evolving field of Data Science, the Master of Science (Data Science) program is designed to equip students with a profound knowledge base and practical skills. This program integrates a diverse range of courses that blend theoretical foundations with hands-on experiences, ensuring our graduates are well-prepared for the dynamic challenges of the data-driven era.
Programme Outcomes (POs)	Upon successful completion of the Master of Science (Data Science) program, graduates will demonstrate: • **Proficiency in Data Analysis:** Graduates will possess the skills to analyze complex datasets, employing statistical and machine learning techniques to derive meaningful insights and make informed decisions. • **Competence in Programming and Software Development:** Graduates will be proficient in programming languages such as Python and R, capable of developing and implementing data science solutions effectively. • **Mastery of Big Data Technologies:** Graduates will have a comprehensive understanding of big data technologies and tools, enabling them to handle and process large volumes of data efficiently. • **Application of Machine Learning and Deep Learning:** Graduates will be able to apply machine learning and deep learning techniques to solve real-world problems, including tasks such as classification, regression, clustering, and natural language processing. • **Business Intelligence and Decision Support:** Graduates will be equipped with the skills to leverage data for strategic decision-making, bridging the gap between data science and business intelligence. • **Ethical and Responsible Data Practices:** Graduates will demonstrate an understanding of ethical considerations in data science, adhering to responsible data practices and respecting privacy and confidentiality. • **Effective Communication and Visualization**: Graduates will be adept at communicating complex technical concepts to diverse audiences and utilizing data visualization tools to present findings in a clear and compelling manner. • **Research and Innovation in Data Science:** Graduates will have the ability to engage in research

	activities, contributing to the advancement of knowledge in the field of data science, and fostering innovation in data-driven solutions. • **Specialized Knowledge in Chosen Elective Areas:** Graduates will exhibit specialized knowledge in elective areas chosen during the program, such as cyber security, artificial intelligence, database systems, or other relevant domains. • **Practical Experience through Internships and Projects:** Graduates will have practical experience gained through internships, on-the-job training (OJT), and research projects, enhancing their ability to apply theoretical knowledge in real-world settings. • **Continuous Learning and Adaptability:** Graduates will demonstrate a commitment to continuous learning, staying abreast of emerging technologies and industry trends in the rapidly evolving field of data science • **Collaboration and Teamwork:** Graduates will be effective collaborators, able to work seamlessly within interdisciplinary teams to address complex data science challenges.
Programme Specific Outcomes (PSOs)	Programme Specific Outcomes (PSOs) for an MSC (Data Science) specify the particular skills, knowledge, and abilities that students are expected to gain upon completion of the program. • **Advanced Data Analysis Proficiency**- Graduates will be proficient in employing advanced statistical and machine learning techniques for data analysis, extracting meaningful insights and making data-driven decisions in diverse domains. • **Programming and Software Development Skills**- Graduates will demonstrate advanced programming skills, with the ability to develop and implement data science solutions using languages such as Python and R. • **Expertise in Big Data Technologies** -Graduates will exhibit expertise in utilizing and managing big data technologies and tools, demonstrating proficiency in handling and processing large-scale datasets. • **Application of Machine Learning and Deep Learning** -Graduates will showcase expertise in applying machine learning and deep learning techniques to solve complex problems, including tasks such as classification, regression, clustering, and natural language processing. • **Business Intelligence and Strategic Decision Support** -Graduates will possess the skills to integrate data science insights with business intelligence, supporting strategic decision-making processes within organizations.

	• **Ethical and Responsible Data Practices**- Graduates will adhere to ethical considerations in data science, practicing responsible data handling, and demonstrating an understanding of privacy and confidentiality issues. • **Effective Communication and Data Visualization**- Graduates will effectively communicate complex technical concepts and present data-driven findings using visualization tools, catering to diverse audiences. • **Research and Innovation in Data Science** - Graduates will engage in research activities, contributing to the advancement of knowledge in data science and fostering innovation in data-driven solutions. • **Specialized Knowledge in Elective Areas** - Graduates will apply specialized knowledge gained in elective areas, such as cyber security, artificial intelligence, database systems, or other chosen domains, to address specific data science challenges. • **Practical Experience through Internships and Projects -**Graduates will demonstrate practical experience gained through internships, on-the-job training (OJT), and research projects, showcasing their ability to apply theoretical knowledge in practical scenarios. • **Continuous Learning and Adaptability**- Graduates will exhibit a commitment to continuous learning, staying updated on emerging technologies and industry trends in the rapidly evolving field of data science. • **Collaboration and Teamwork Skills** -Graduates will demonstrate effective collaboration and teamwork skills, working seamlessly within interdisciplinary teams to address complex data science challenges.
Eligibility Criteria for Programme	A woman Graduate in any B.Sc. (Physics), B.Sc. (Mathematics), B.Sc. (Electronics), B.Sc. (Information Technology), B.Sc. (Computer Science), B.Sc. (IT) or BCA or any engineering graduate in allied subject from the recognized university with aggregate marks not less than 50% for Open Category and 45% Reserved Category.
Intake	60
Duration	4 semesters (2 years)

Master of Science (Data Science) (M.Sc.- DS.)

Year -I

Code	Subjects	Type of Course	Credits	Marks	Int.	Ext.
	Semester – I					
115611	Computer Oriented Statistical Techniques-I	Major (Core) Theory	4	100	50	50
115612	Data Structure and Analysis of Algorithms	Major(Core) Theory	4	100	50	50
115613	Python Programming	Major(Core) Theory	2	50	50	0
115624	Computer Oriented Statistical Techniques– Lab(Using R)	Major (Core) Practical	2	50	25	25
115625	Data Base Management Systems-Lab	Major (Core) Practical	2	50	25	25
125611/ 125612/ 125613/ 125614/ 125615	Elective-I	Major (Elective) Theory	4	100	50	50
135611	Research Methodology	Minor Stream (RM) Theory	4	100	50	50
			22	550	300	250
	Semester-II					
Code	**Subjects**	**Type of Course**	**Credit**	**Marks**	**Int.**	**Ext.**
215611	Data Mining with Analytics	Major (Core) Theory	4	100	50	50
215612	Applied Artificial Intelligence	Major (Core) Theory	4	100	50	50
215613	Introduction to Data Science	Major (Core) Theory	2	50	0	50
215624	Data Mining with Analytics –Lab	Major (Core) Practical	2	50	25	25
215625	Applied Artificial Intelligence–Lab	Major (Core) Practical	2	50	25	25
225611/ 225612/ 225613/ 225614/ 225615	Elective-II	Major (Elective) Theory	4	100	50	50
245641	On Job training (OJT)	OJT	4	100	50	50
			22	550	250	300

Exit option (44 credits):Post Graduate Diploma in Data Science

Year -II

Code	Subjects	Type of Course	Credit	Marks	Int.	Ext.
	Semester-III					
315611	Big Data Analytics	Major (Core) Theory	4	100	50	50
315612	Machine Learning	Major (Core) Theory	4	100	50	50
315615	Business Intelligence	Major (Core) Theory	2	50	0	50
315623	Big Data Analytics-Lab	Major (Core) Practical	2	50	25	25
315624	Machine Learning-Lab	Major (Core) Practical	2	50	25	25
325611/ 325612/ 325613/ 325614/ 325615	Elective-III	Major (Elective) Theory	4	100	50	50
355631	Research Project	RP	4	100	50	50
			22	550	250	300
	Semester-IV					
415611	Deep Learning	Major (Core) Theory	4	100	50	50
415612	Natural Language Processing	Major (Core) Theory	4	100	50	50
415623	Deep Learning-Lab	Major (Core) Practical	2	50	50	0
425611/ 425612/ 425613/ 425614/ 425615	Elective-IV/MOOC/SWAYAM	Major (Elective) Theory	4	100	50	50
445641	Internship	OJT	8	200	100	100
			22	550	300	250

Code	Elective-I	Code	Elective-II
125611	1. Cyber Security	225611	1. Ethical Hacking
125612	2. Digital Image Processing	225612	2. Project Management
125613	3. Software Engineering	225613	3. Fuzzy Logic and Neural Network
125614	4. Artificial Intelligence	225614	4. Linear Algebra
125615	5. Database Systems for Data Science	225615	5. Inferential Statistics

Code	Elective-III	Code	Elective-IV
325611	1. Block chain	425611	1. Information Security
325612	2. GIS and Remote Sensing	425612	2. Cloud Computing
325613	3. Software Testing	425613	3. Robotic Process Automation
325614	4. Data Visualization	425614	4. Social network Analysis
325615	5. Data Governance	425615	5. Agile Methodology

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester III		
315611	**BIG DATA ANALYTICS** **Major (Core) Theory**		**4**
	Course Outcomes: **Learners will be able to:** • Apply big data analytics approaches, including conceptualization, summarization, and machine learning techniques. • Analyze the characteristics of datasets to distinguish between trivial data and big data for various applications. • Evaluate solutions for problems related to batch learning, online learning, and big data characteristics like high dimensionality and scalability issues. • Design scalable solutions to manage and process dynamically growing big data.		
Module 1	**INTRODUCTION**		1
	LOs: Learners will be able to • Apply understanding of fundamental Big Data concepts to address modern data challenges effectively. • Analyze traditional and Big Data approaches to select appropriate technologies for large dataset processing and analysis. • Evaluate infrastructure requirements for efficient Big Data handling, focusing on data analytics for valuable insights extraction. • Design strategies to overcome challenges inherent in Big Data, considering volume, velocity, and variety factors for effective data management and analysis.	**Module Contents:** • Introduction to Big Data, Big Data Characteristics, Types of Big Data • Traditional Versus Big Data Approach, Technologies Available for Big Data • Infrastructure for Big Data, Use of Data Analytics, Big Data Challenges, Desired Properties of a Big Data System, Case Study of Big Data Solutions	
Module 2	**Analytical Theory and Methods**		1
	LOs: Learners will be able to • Apply clustering algorithms, association rules, and the Apriori Algorithm to recognize patterns effectively across diverse datasets. • Analyse real-world scenarios by applying	**Module Contents:** • Clustering and Associated Algorithms, Association Rules, Apriori Algorithm, Candidate Rules • Applications of Association Rules, Validation and Testing, Diagnostics, Regression, Linear Regression, Logistic Regression, Additional Regression Models	

	association rules, mastering validation, and testing for model reliability. • Evaluate linear and logistic regression models, implementing and interpreting regression analyses with proficiency. • Design additional regression models to enhance analytical capabilities for addressing diverse data-driven challenges.		
Module 3	**Hadoop**		1
	LOs: Learners will be able to • Apply foundational principles to comprehend Hadoop, including its definition, core components, and the role of operating systems in Big Data processing. • Analyse Hadoop architecture, ecosystem components, and technologies like Hive to develop proficiency and understanding. • Evaluate the limitations of Hadoop while exploring practical applications, particularly in recommendation systems. • Design practical strategies to address the limitations of Hadoop and optimize its application in real-world scenarios.	**Module Contents:** • Introduction, What is Hadoop?, Core Hadoop Components, Operating System for Big Data • Concepts, Hadoop Architecture, Hadoop Ecosystem, Hive, , Hadoop Limitations , Recommendation Systems.	
Module 4	**NoSQL**		1
	LOs: Learners will be able to • Apply comprehensive learning to grasp NoSQL, covering its definition, business drivers, case studies, and data architectural patterns for informed decision-making in data management. • Analyse the practical application of MapReduce within the new software stack, understanding its role and	**Module Contents:** • What is NoSQL?, NoSQL Business Drivers, NoSQL Case Studies, NoSQL Data Architectural Patterns • Variations of NoSQL Architectural Patterns, Using NoSQL to Manage Big Data • Map Reduce: MapReduce and The New Software Stack, MapReduce, Algorithms Using MapReduce	

	implementing algorithms for efficient big data processing. • Evaluate the effectiveness of MapReduce in handling large-scale data processing tasks, considering its scalability and performance. • Design strategies to optimize MapReduce algorithms and workflows for specific big data processing requirements, ensuring efficient and effective data processing.		
Assignments/ Activities			
	These assignments aim to apply theoretical concepts to practical application and critical thinking. **Module 1:** • Students will conduct research on clustering algorithms, association rules, and the Apriori Algorithm. **Module 2:** • They will analyze real-world applications of association rules and perform validation and testing of these algorithms. **Module 3:** • Students will delve into regression analysis, including linear regression, logistic regression, and additional regression models. **Module 4:** • They will evaluate the effectiveness and limitations of regression models in different data-driven scenarios. • Finally, students will design and implement regression models for specific analytical tasks, applying the learned theories and methods.		

References:

1. Shankarmani, R. (2016). Big Data Analytics. Wiley.
2. Chellappan, S., & Acharya, S. (2015). Big Data and Analytics. Wiley.
3. Prajapati, V. (2013). Big Data Analytics with R and Hadoop. Pack Publishing.
4. Dasgupta, N. (2018). Practical Big Data Analytics. Pack Publishing.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester III		
315612	**Machine Learning** **Major (Core) Theory**		**4**
	Course Outcomes: **Learners will be able to:** • Apply knowledge of human learning aspects to enhance computer-based learning processes. • Analyze primitives in the learning process by computers to understand their foundational elements. • Evaluate the nature of problems solved with machine learning to identify their scope and applicability. • Design effective strategies for integrating human learning aspects into computer-based learning systems and addressing various problems using machine learning techniques.		
Module 1	**Introduction**		1
	LOs: Learners will be able to • Apply profound knowledge of machine learning principles to differentiate between learning and designing, and to understand the characteristics of different machine learning tasks. • Analyze diverse machine learning models, including geometric, logical, and probabilistic models, to comprehend their strengths and limitations. • Evaluate feature engineering techniques, covering feature types, construction, transformation, and selection, for enhancing model development. • Design effective machine learning models by integrating various techniques and selecting appropriate models based on task requirements.	**Module Contents:** • Machine learning, Examples of Machine Learning Problems, Structure of Learning, learning versus Designing, Training versus Testing, Characteristics of Machine learning tasks, Predictive and descriptive tasks, • **Machine learning Models**: Geometric Models, Logical Models, Probabilistic Models. • **Features:** Feature types, Feature Construction and Transformation, Feature Selection.	
Module 2	**Classification and Regression**		1
	LOs: Learners will be able to •Apply proficiency in assessing binary and multiclass classification	**Module Contents:** • Binary Classification-Assessing Classification performance, Class probability	

	performance, including accurate class probability estimation. •Analyze regression performance assessment, exploring error measures and identifying factors contributing to overfitting. •Evaluate the theory of hypothesis in regression modeling to understand its implications and limitations. •Design robust evaluation strategies for both classification and regression tasks, considering various performance metrics and addressing potential challenges like overfitting.	Estimation Assessing class probability Estimates, Multiclass Classification. • **Regression:** Assessing performance of Regression-Error measures, Overfitting-Catalysts for Overfitting, Polynomial Regression. • Theory of hypothesis.	
Module 3	**Linear and Tree based Models**		1
	LOs: Learners will be able to •Apply various linear models such as the Least Squares method, Multivariate Linear Regression, Regularized Regression, Logistic Regression, and Support Vector Machines (SVM) to diverse datasets. •Analyse the principles and applications of tree-based models, including Decision Trees, Regression Trees, and Clustering Trees, for effective pattern recognition and data analysis. •Evaluate the strengths and weaknesses of different linear and tree-based models in different scenarios, considering factors like model complexity and interpretability. • Design and implement appropriate linear and tree-based models based on the characteristics of the dataset and the objectives of the analysis.	**Module Contents:** • **Linear Models:** Least Squares method, Multivariate Linear Regression, Regularized Regression, Bias/Variance Trade-off, Dimension Reduction Logistic Regression, Gradient Descent, Perceptron, Support Vector Machines SVM, Soft Margin SVM, Time Series Analysis, Forecasting. • **Tree Based Models:** Decision Trees, Regression trees, Clustering Trees.	

Module 4	Logic and Rule based models		1
	LOs: Learners will be able to •Apply logic-based, algebraic, distance-based, rule-based models, and ensemble learning techniques such as bagging, boosting, online learning, deep learning, and reinforcement learning to various data analysis tasks. •Analyse the principles and methodologies of each model and ensemble learning technique to identify their strengths and weaknesses in different contexts. •Evaluate the performance of different models and ensemble techniques using appropriate metrics to assess their effectiveness in solving real-world problems. •Design integrated approaches combining logic-based, algebraic, distance-based, and rule-based models with ensemble learning techniques to enhance predictive accuracy and robustness in data analysis tasks.	**Module Contents:** • **Logic Based and Algebraic Model:** Distance Based Models: Neighbours and Examples, Nearest Neighbours Classification, Distance based clustering K-means Algorithm, Hierarchical clustering, • **Rule Based Models**: Association rule mining. • **Ensemble Learning:** Introduction to Ensemble Learning, Bagging and Boosting, Online learning and Sequence Prediction, Deep Learning, Reinforcement Learning.	
Assignments/ Activities towards CCE			
	Module 1: •Problem Identification: Choose a real-world problem suitable for machine learning analysis. Identify the type of problem (predictive or descriptive) and define the learning task (classification or regression). **Module 2:** •Data Preparation: Collect or generate data relevant to the chosen problem. Perform feature construction, transformation, and selection to prepare the dataset for analysis. **Module 3:** •Model Selection: Select appropriate machine learning models based on the problem type and task. Choose from geometric, logical, and probabilistic models discussed in the module. •Model Implementation: Implement the selected models using Python or R programming languages. Train the models on the prepared dataset and evaluate their performance using relevant		

	metrics. **Module 4:** • Analysis and Interpretation: Analyse the results obtained from the models and interpret their performance. Discuss the strengths and limitations of each model in addressing the chosen problem. • Documentation: Prepare a comprehensive report documenting the entire process, including problem statement, data description, methodology, results, analysis, and conclusions.	

References:

1. Flach, P. (2012). Machine Learning: The Art and Science of Algorithms that Make Sense of Data. Cambridge University Press.
2. Murphy, K. (2012). Machine Learning: A Probabilistic Perspective. MIT Press.
3. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning. Springer.
4. Barber, D. (2012). Bayesian Reasoning and Machine Learning [Online version]. Cambridge University Press.
5. Mitchell, T. (1997). Machine Learning (1st ed.). McGraw Hill.
6. Duda, R. O., Hart, P. E., & Stork, D. G. (2007). Pattern Classification. John Wiley & Sons.
7. Alpaydin, E. (2015). Introduction to Machine Learning (3rd ed.). MIT Press.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester III		
315615	**BUSINESS INTELLIGENCE** **Major (Core) Theory**		**2**
	Course Outcomes: **Learners will be able to:** • Gain an overview of the importance of Business Intelligence (BI) in modern contexts, emphasizing its relevance in decision-making and organizational efficiency. • Examine the evolution of data analytics methods, comparing traditional approaches with modern techniques, and identify their respective strengths and limitations in BI applications. • Develop an understanding of Decision Support Systems (DSS), Artificial Intelligence (AI), Expert Systems, and Knowledge Management Systems (KMS) within the context of BI, designing strategies to leverage these technologies for enhancing organizational decision-making and business performance. • Assess various mathematical models, classification, and clustering methods utilized in Data Analytics for BI, evaluating their effectiveness in extracting insights from large datasets.		
Module 1	**Introduction to Business Intelligence and Decision Making**		1
	LOs: Learners will be able to • Gain proficiency in Business Intelligence (BI) and Decision Support Systems (DSS) through understanding their fundamentals and the importance of timely decisions. • Explore the role of mathematical models and data preparation techniques in BI and DSS, assessing their contribution to decision-making processes. • Assess the effectiveness of BI and DSS in facilitating timely decisions, considering their impact on organizational decision-making. • Design decision support systems integrating mathematical models and data preparation techniques to enhance decision-making efficiency and effectiveness.	**Module Contents:** • **Business Intelligence:** Introduction to Business Intelligence, Significance of Effective and timely decisions in Business, The role of mathematical models, Business Intelligence architectures, Ethics and Business Intelligence • **Data Preparation:** Representation of input data, Data validation, Data transformation, Data reduction and data mining process, Analysis methodologies • **Mathematical models for decision making**: Structure of mathematical models, Development of a model, Classes of models • **Decision Support System:** Definition of system, Representation of the decision-making process, Evolution of information systems, Definition of decision support system, Development of a decision support system	
Module 2	**Business Intelligence Model and Knowledge Management**		1

	LOs: Learners will be able to • Apply advanced proficiency in business intelligence applications, including marketing and logistic models. • Analyze efficiency using data envelopment techniques. • Evaluate knowledge management skills in organizational learning, IT integration, and system implementation. • Design comprehensive knowledge management solutions integrating organizational learning, IT systems, and implementation.	**Module Contents:** • **Business intelligence applications:** **Marketing models:** Relational marketing, Sales force management, **Logistic and production models:** Supply chain optimization, Optimization models for logistics planning, Revenue management systems • **Data envelopment analysis:** Efficiency measures, Efficient frontier, The CCR model, Identification of good operating practices • **Knowledge Management:** Introduction to Knowledge Management, Organizational Learning and Transformation, Knowledge Management Activities, Approaches to Knowledge Management, Information Technology (IT) In Knowledge Management, Knowledge Management Systems Implementation, Roles of People in Knowledge Management.	
Assignments/ Activities			
	These assignments aim to apply theoretical concepts to practical application and critical thinking. **Module 1:** • Activity: Case study analysis of a company using business intelligence applications in marketing. Focus on relational marketing and sales force management. • Assignment: Write a report on logistic and production models, highlighting supply chain optimization and optimization models for logistics planning. **Module 2:** • Activity: Conduct an efficiency analysis of a provided dataset using the CCR model and identify good operating practices. • Assignment: Create a presentation explaining the concept of the efficient frontier and how efficiency measures are applied in business contexts.		

References:

1. Vercellis, C. (2009). Business intelligence: Data mining and optimization for decision making. Wiley Publications.
2. Turban, E., Sharda, R., & Delen, D. (2011). Decision support and business intelligence systems (9th ed.). Pearson Publications.
3. Grossmann, W., & Rinderle-Ma, S. (2015). Fundamentals of business intelligence. Springer Publications.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester III		
315623	**Big Data Analytics Lab: Practical** **Major (Core)**		**2**
	Course Outcomes: **Learners will be able to:** • Apply hands-on expertise in big data analytics by mastering the installation and practical use of PySpark for linear and logistic regression. • Analyze graphical data processing, Hive database management, window functions, and time series analysis using PySpark. • Evaluate practical applications of PySpark in real-world scenarios, focusing on advanced skills in data processing. • Design solutions for complex data challenges using PySpark, including linear and logistic regression models.		
Module 1	**Introduction to Hadoop and Pyspark**		1
	LOs: Learners will be able to • Apply hands-on expertise in Hadoop and PySpark by defining installation steps. • Analyze and perform linear and logistic regression using PySpark. • Evaluate practical skills through MapReduce programming for word count problems using PySpark. • Design and implement solutions using PySpark for complex data processing tasks.	**Module Contents:** • Define steps to install hadoop and pyspark • Practical to perform linear regression using pyspark • Practical to perform logistic regression using Pyspark • Practical to perform map reduce program for word count problem.	
Module 2	**Implementation using Pyspark**		1
	LOs: Learners will be able to • Apply advanced data processing skills with PySpark, focusing on creating and accessing graphical data. • Analyze structured database management using Hive and the implementation of window functions. • Evaluate practical applications in Time Series Analysis using PySpark. • Design solutions incorporating aggregate functions for complex data scenarios with	**Module Contents:** • Create graphical data and access the graphical data using spark • Practical to use hive to create and store structured databases • Practical to perform window function using Pyspark. • Practical to perform Times Series Analysis using Pyspark • Practical to perform Aggregate function using Pyspark.	

	PySpark.		
Assignments/ Activities towards Comprehensive Continuous Evaluation			
	Module 1: Assignment: Develop a step-by-step installation guide for Hadoop and PySpark, then perform a linear regression analysis on a given dataset using PySpark. **Module 2:** Assignment: Create and visualize a graphical dataset using Spark, and store the structured database using Hive. Then, perform a window function analysis on the dataset using PySpark.		

References:

1. Marz, N., & Warren, J. (2015). Big Data: Principles and Best Practices of Scalable Realtime Data Systems. Manning Publications.
2. White, T. (2015). Hadoop: The Definitive Guide. O'Reilly Media.
3. Guller, M. (2015). Big Data Analytics with Spark: A Practitioner's Guide to Using Spark for Large Scale Data Analysis. Apress.
4. Srinivasan, S. (2018). Big Data Analytics: Methods and Applications. CRC Press.
5. Gates, A., Thusoo, A., & et al. (2015). Hive: The Definitive Guide. O'Reilly Media.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester III		
315624	**Machine Learning Lab: Practical** **Major (Core)**		**2**
	Course Outcomes: **Learners will be able to:** •Apply proficiency in handling and analyzing diverse datasets using GitHub. •Analyze data through Exploratory Data Analysis (EDA) and Data Pre-processing techniques. •Evaluate the implementation of machine learning algorithms, including Linear and Logistic Regression, Decision Trees, Support Vector Machines (SVM), K-Nearest Neighbours (KNN), Time Series Forecasting, and either Recommendation Systems or Random Forest. •Design versatile applications in statistical and machine learning domains.		
Module 1	**EDA**		1
	LOs: Learners will be able to • Apply collaborative data work through a GitHub account, loading diverse data formats for statistical summarization. • Analyze data using Exploratory Data Analysis (EDA) and Data Pre-processing techniques. • Evaluate regression modeling techniques, including Linear and Logistic Regression, demonstrating applied skills in data analytics. • Design and implement practical approaches for data analysis, showcasing proficiency in regression modeling.	**Module Contents:** • Setup Github Account, loading data from different source files formats (csv, excel) and summarizing data with statistics. • Practical to implement Exploratory Data Analysis (EDA)& Data Pre-processing (Outlier Detection, Handling Missing Data, Encoding Categorical Data) • Practical to implement Linear Regression (Single/Multiple) • Practical to implement Logistic Regression	
Module 2	**Implementation of ML algorithm**		1
	LOs: Learners will be able to • Apply Decision Tree, Support Vector Machine (SVM), and K-Nearest Neighbours (KNN) algorithms for classification and clustering tasks, demonstrating practical competence. • Analyze Time Series	**Module Contents:** • Practical to implement Decision Tree Algorithm • Practical to implement Support Vector Machine (SVM) Algorithm • Practical to implement K-Nearest Neighbours KNN Algorithm • Practical to implement Time Series Forecasting	

	Forecasting techniques, showcasing proficiency in predictive modeling for sequential data. • Design and implement either Recommendation Systems or the Random Forest Algorithm, showcasing practical skills and a comprehensive understanding of diverse machine learning applications. • Showcase versatility in machine learning applications through practical competence in classification, clustering, predictive modeling, and recommendation systems or random forest algorithms.	• Practical to implement Recommendation Systems or Practical to implement Random Forest Algorithm	
Assignments/ Activities towards Comprehensive Continuous Evaluation			
	Module 1: • Assignment: Using your GitHub account, load data from different source file formats (e.g., csv, excel) and perform statistical summarization. Then, implement Exploratory Data Analysis (EDA) and Data Pre-processing techniques (Outlier Detection, Handling Missing Data, Encoding Categorical Data), followed by practical implementation of Linear Regression (Single/Multiple) and Logistic Regression. **Module 2:** • Assignment: Practically implement the Decision Tree Algorithm, Support Vector Machine (SVM) Algorithm, and K-Nearest Neighbours (KNN) Algorithm. Additionally, implement Time Series Forecasting and choose either Recommendation Systems or the Random Forest Algorithm for practical implementation.		

References:

1. Flach, P. (2012). Machine learning: The art and science of algorithms that make sense of data. Cambridge University Press.
2. Murphy, K. P. (2012). Machine learning: A probabilistic perspective. MIT Press.
3. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning. Springer.
4. Barber, D. (2012). Bayesian reasoning and machine learning. Cambridge University Press. [Online version available]
5. Mitchell, T. M. (2017). Machine learning (1st ed.). McGraw-Hill.
6. Duda, R. O., Hart, P. E., & Stork, D. G. (2007). Pattern classification (2nd ed.). John Wiley & Sons.
7. Alpaydin, E. (2014). Introduction to machine learning (3rd ed.). MIT Press.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester III		
325611	**BLOCK CHAIN** **Major (Elective) Theory**		**4**
	Course Outcomes: **Learners will be able to:** • Apply understanding of Blockchain Concepts, including its decentralized nature, distributed ledger, and cryptographic security features. • Analyze the historical context and evolution of blockchain technology, including the development of the first blockchain. • Evaluate the Nakamoto consensus and different consensus algorithms. • Design exploration of concepts such as interoperability, portability in Hyperledger Fabric, and the concept of sharding in blockchain.		
Module 1	**Fundamentals of Blockchain**		1
	LOs: Learners will be able to • Apply principles of distributed databases, including architecture, advantages, and challenges, to design and manage data across multiple network nodes. • Analyze complexities of consensus in distributed systems, considering conflicting or malicious information, and understand the significance of Byzantine fault tolerance. • Evaluate the concept of ASIC resistance in crypto-currencies, exploring motivations and implications of designing systems to resist mining centralization through specialized hardware. • Design a comprehensive understanding of cryptography principles, including confidentiality, integrity, and authenticity, learning applications and functions of hash functions, digital signatures (specifically ECDSA), memory-hard algorithms, and zero knowledge proofs.	**Module Contents:** • Distributed Database, Two General Problem, Byzantine General problem and Fault Tolerance, Hadoop Distributed File System, Distributed Hash Table, ASIC resistance, Turing Complete. • Cryptography: Hash function, Digital Signature - ECDSA, Memory Hard Algorithm, Zero Knowledge Proof.	

Module 2	**Blockchain , Distributed Consensus:**		1
	LOs: Learners will be able to • Apply understanding of fundamental blockchain concepts, including its decentralized nature, distributed ledger, and cryptographic security features. • Analyze the structure and operation of a blockchain network, including nodes, peers, and the peer-to-peer communication model. • Evaluate the differences between private and public blockchains, considering use cases, access control, and levels of decentralization. • Design an exploration of the Nakamoto consensus and different consensus algorithms such as Proof of Work, Proof of Stake, and Proof of Burn, understanding their strengths and weaknesses.	**Module Contents:** • Introduction, Advantage over conventional distributed database, Blockchain Network, Mining Mechanism, Distributed Consensus, Merkle Patricia Tree, Gas Limit, Transactions and Fee, Anonymity, Reward, Chain Policy, Life of Blockchain application, Soft & Hard Fork, Private and Public blockchain. • Nakamoto consensus, Proof of Work, Proof of Stake, Proof of Burn, Difficulty Level, Sybil Attack, Energy utilization and alternate.	
Module 3	**Cryptocurrency , Cryptocurrency Regulation:**		1
	LOs: Learners will be able to • Apply understanding of the historical context and evolution of blockchain technology, including the development of the first blockchain in the context of Bitcoin. • Analyze the construction of the Ethereum blockchain and its components. • Evaluate the concept and implementation of smart contracts. • Design exploration of the Decentralized Autonomous Organization (DAO) within the Ethereum ecosystem.	**Module Contents:** • History, Distributed Ledger, Bitcoin protocols - Mining strategy and rewards, Ethereum - Construction, DAO, Smart Contract, GHOST, Vulnerability, Attacks, Sidechain, Namecoin. • Stakeholders, Roots of Bit coin, Legal Aspects-Crypto currency Exchange, Black Market and Global Economy. Applications: Internet of Things, Medical Record Management System, Domain Name Service and future of Blockchain.	
Module 4	**Hyperledger , Scalability and other challenges :**		1
	LOs: Learners will be able to • Apply understanding of Hyperledger as a blockchain protocol, including its reference	**Module Contents:** • Hyperledger as a protocol :The reference architecture Requirements and design goals of	

	architecture, design goals, and modular approach. • Analyze the features of Hyperledger Fabric, such as its modular architecture, privacy and confidentiality mechanisms, scalability considerations, deterministic transactions, and identity management. • Evaluate the scalability challenges in blockchain networks and understand how Hyperledger Fabric addresses these concerns in the network, consensus, storage, and view planes. • Design solutions for interoperability and portability in Hyperledger Fabric, and explore the concept of sharding to improve scalability through parallel processing of transactions.	Hyperledger Fabric: The modular approach Privacy and confidentiality, Scalability, Deterministic transactions Identity, Auditability Interoperability Portability Rich data queries Fabric Hyperledger Fabric Membership services Blockchain services Consensus services Distributed ledger ,The peer to peer protocol Ledger storage Chaincode services ,Components of the fabric • Scalability and Other Challenges: Scalability Network plane ,Consensus plane, Storage plane View plane ,Block size increase ,Block interval reduction Invertible Bloom, Lookup Tables Sharding State channels Private blockchain, Proof of Stake Sidechains Subchains Tree chains (trees) Block propagation Bitcoin-NG, Plasma ,Privacy Indistinguishability Obfuscation Homomorphic encryption ,Zero-Knowledge Proofs State channels Secure multiparty computation Usage of hardware to provide confidentiality Coin Join Confidential transactions, Mimble Wimble Security Smart contract security Formal verification and analysis Oyente tool	
Assignments/ Activities			
	These assignments aim to apply theoretical concepts to practical application and critical thinking. **Module 1:** • Assignment: Develop a comprehensive report explaining the Byzantine General problem and Fault Tolerance. Additionally, implement a simple blockchain model that demonstrates the principles of a distributed hash table and ASIC resistance. Include a practical exercise to create and verify a digital signature using ECDSA. **Module 2:** • Assignment: Create a detailed presentation comparing Proof of Work, Proof of Stake, and Proof of Burn consensus mechanisms. Explain their strengths, weaknesses, and susceptibility to Sybil attacks. Include a practical component where you simulate a mining mechanism and calculate transaction fees in a mock blockchain network. **Module 3:** • Assignment: Write an essay on the history and evolution of Bitcoin and Ethereum, focusing on their protocols, mining strategies, and the role of smart contracts. Analyze a case study on the DAO attack, detailing its impact on the Ethereum network and subsequent regulatory		

	implications. Include a section on the legal aspects of cryptocurrency exchanges and their influence on the global economy. **Module 4:** • Assignment: Design and document a Hyperledger Fabric network architecture, detailing the requirements and design goals. Implement a small-scale Hyperledger Fabric network to demonstrate modularity, privacy, confidentiality, and identity management. Address scalability challenges and propose solutions, such as sharding and state channels, to improve network efficiency.	

References:

1. Antonopoulos, A. M. (2014). Mastering Bitcoin: Unlocking digital cryptocurrencies. O'Reilly Media.
2. Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. [Link to the Bitcoin Whitepaper].
3. Wood, G. (2014). ETHEREUM: A secure decentralized transaction ledger. Yellow paper. [Link to the Ethereum Yellow Paper].
4. Atzei, N., Bartoletti, M., & Cimoli, T. (2017). A survey of attacks on Ethereum smart contracts.
5. Bashir, I. (2018). Mastering blockchain. Wiley.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester III		
325612	**GIS AND REMOTE SENSING** **Major(Elective) Theory**		**4**
	Course Outcomes: **Learners will be able to:** • Apply fundamentals and technical skills in data acquisition and management. • Analyze spatial data through spatial analysis, mapping, and visualization techniques. • Evaluate remote sensing applications and the integration of various technologies. • Design effective problem-solving strategies and enhance communication skills.		
Module 1	**Fundamentals of GIS**		1
	LOs: Learners will be able to • Apply the principles of GIS, including its components and spatial data characteristics, to real-world scenarios. • Analyze spatial data through maps and spatial data modeling techniques. • Evaluate attribute data management using database data models in GIS applications. • Design GIS solutions and assess developments in database technologies for improved spatial data management.	**Module Contents:** • Defining GIS, components of GIS, spatial data, spatial data-maps, characteristics, spatial data modeling, attribute data management-database data model, GIS applications and developments in database.	
Module 2	**Input-Output and Data Analysis in GIS**		1
	LOs: Learners will be able to • Apply various data input and editing methods to integrate and refine spatial data. • Analyze spatial data using techniques such as measurements, queries, reclassification, buffering, map overlay, interpolation, and network analysis. • Evaluate analytical modeling in GIS for physical, environmental, and human processes to support decision-making. • Design outputs from GIS, including maps, non-cartographic outputs, spatial multimedia, and	**Module Contents:** • Data input and editing– methods, editing, integration, Data analysis-measurements, queries, reclassification, buffering, map overlay, interpolation, analysis of surfaces, network analysis, spatial analysis, Analytical modeling in GIS-physical, environment and human processes, output from GIS –maps, non-cartographic output, spatial multimedia, decision support.	

	decision support systems.		
Module 3	**Issues in GIS:**		1
	LOs: Learners will be able to • Apply computer methods and address human and organizational issues in GIS. • Analyze GIS data quality and perform error analysis. • Design GIS projects, focusing on project implementation and evaluation. • Evaluate the future of GIS, leveraging internet resources and enhancing communication skills.	**Module Contents:** • Development of computer methods for spatial data, Issues in GIS- data quality and errors, sources of errors, human and organizational issues, GIS project design and management–problem identification, designing a data model, project management, Implementation, evaluation, the future of GIS, Internet resources of GIS.	
Module 4	Remote Sensing, Global Positioning Systems (GPS)		1
	LOs: Learners will be able to • Apply principles of remote sensing and classify remote sensing systems based on imaging characteristics. • Analyze methods for extracting information from remote sensing images and integrate remote sensing with GIS. • Evaluate the accuracy and applications of GPS, including differential GPS techniques. • Design solutions that integrate GIS and GPS for comprehensive spatial data analysis.	**Module Contents:** • Principles of remote sensing, remote sensing system-classification, Imaging, characteristics, extraction of information from images–metric and thematic, Integration of RS and GIS. • Introduction to GPS, Accuracy of GPS, Differential GPS, Applications of GPS, Integration of GIS and GPS.	
Assignments/ Activities			
	These assignments aim to apply theoretical concepts to practical application and critical thinking. **Module 1:** • Assignment: Develop a presentation defining GIS and its components, including spatial data and its characteristics. Explore spatial data modeling and attribute data management using database models. Analyze GIS applications and advancements in database technology. **Module 2:** • Assignment: Create a practical exercise demonstrating various data input and editing methods in GIS. Perform data analysis tasks such as measurements, queries, buffering, and spatial analysis. Design analytical models for physical, environmental, and human processes in GIS. Produce outputs including maps, non-cartographic outputs, and decision support systems. **Module 3:** • Assignment: Write a research paper on the development of computer methods for spatial data analysis in GIS. Investigate issues related to data quality and errors in GIS, including sources of errors and organizational challenges. Design a GIS project, from problem identification to project		

	management, and evaluate its implementation and future prospects using internet resources. **Module 4:** • Assignment: Create a comprehensive report on the principles of remote sensing, including classification and imaging characteristics of remote sensing systems. Explore methods for extracting information from remote sensing images and integrating remote sensing with GIS. Additionally, analyze the principles of GPS, its accuracy, applications, and integration with GIS.	

References:

1. Heywood, I., Cornelius, S., & Carver, S. (2000). An introduction to geographical information systems. Pearson Education Asia.
2. Lo, C. P., & Yeung, A. (2016). Concepts and techniques of geographic information systems. PHI, New Delhi.
3. Demers, M. N. (1999). Fundamentals of geographic information systems (2nd ed.). John Wiley & Sons (Asia) Pte Ltd.
4. Razvi, M. (2002). ArcGIS developer's guide for Visual Basic applications. Onword Press.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester III		
325613	**Software Testing** **Major (Elective) Theory**		**4**
	Course Outcomes: **Learners will be able to:** • Analyse fundamental principles and concepts of software testing, its purpose, objectives, and role in software development. • Proficient in test design techniques, including equivalence partitioning, boundary value analysis, decision tables, and state transition testing, to create effective test cases. • Apply testing methods such as functional, non-functional, unit, integration, system, regression, and acceptance testing to diverse software systems. • Evaluate quality assurance principles and best practices, emphasizing testing's importance in the software development lifecycle, along with ethical and professional responsibilities associated with software testing.		
Module 1	**Overview of Software Testing**		1
	LOs: Learners will be able to • Apply software testing terminologies, methodologies, and life cycles proficiently. • Analyze the economic aspects of testing and its influence on organizational structures. • Develop skills in policy creation, test strategies, and risk management to meet customer needs effectively. • Evaluate the advantages of structured testing processes and their associated cost implications, demonstrating proficiency in the seven-step software testing process.	**Module Contents:** • Software Testing Terminology and Methodology Software Testing Terminology, Software Testing Life Cycle, Writing a Policy for Software Testing, Economics of Testing, Testing – An organizational Issue, Management Support for Software Testing, Fig. of Software Testing Methodology, Risk associated with not meeting customer needs, Developing Test Strategy • Overview of Software Testing Process Advantages of Following a Process, The Cost of Computer Testing, The Seven-Step Software Testing Process • Verification and Validation Verification and Validation (V&V) Activities, Verification, Verification of Requirements, Verification of High –level Design, Verification of Low –level Design, How to Verify Code? ,Validation • Static Testing Inspections, Structured Walkthroughs, Technical Reviews.	
Module 2	**Validation and Regression Testing**		1
	LOs: Learners will be able to • Execute validation activities,	**Module Contents:** • Validation Activities Unit	

	including unit, integration, function, system, and acceptance testing, to assess software functionality and ensure compliance with requirements. • Differentiate between progressive and regressive testing, understanding the significance of regression testing in maintaining software quality. • Apply regression testing techniques to identify potential issues from software changes, ensuring stability and reliability. • Identify regression testing objectives, determine suitable instances for conducting tests, define regression test problems, and select appropriate types and strategies to minimize risks.	Validation Testing, Integration Testing, Function Testing, System Testing , Acceptance Testing • Regression Testing Progressive vs. Regressive Testing, Regression Testing Produces Quality Software, Regression Testability, Objectives of Regression Testing, When is Regression Testing Done? , Regression Testing Types, Defining Regression Test Problem, Regression Testing Techniques.	
Module 3	**Testing Management and Metrics**		1
	LOs: • Establish a comprehensive understanding of test management structures, facilitating effective test planning and detailed design. • Recognize the necessity of software metrics and demonstrate the capability to define, classify, and apply them within the software development life cycle. • Evaluate entities to be measured in software projects, with a focus on size metrics and their implications for software management. • Formulate testing-specific measurement objectives, identifying attributes and relevant metrics for monitoring and controlling the testing process.	**Module Contents:** • Test Management Test Organization, Structure of Testing Group, Test Planning, Detailed Test Design and Test Specifications • Software Metrics Need for Software Management, Definition of Software Metrics, Classification of Software Metrics, Entities to be Measured, Size Metrics • Testing Metrics for Monitoring and Controlling the Testing Process Measurement Objectives for Testing, Attributes and Corresponding Metrics in Software Testing, Attributes, Estimation Models for Estimating Testing Efforts (include only topic Halstead Metrics), Test Point Analysis (TPA) – introduction only.	
Module 4	**Automation Testing Tool**		1
	LOs: Learners will be able to • Evaluate the necessity and	**Module Contents:** • Testing Process Maturity Models	

	significance of test process maturity, measuring, assessing, and improving processes using established models. • Identify the rationale for automation in testing, categorize testing tools, and apply selection criteria considering associated costs. • Analyze guidelines for automated testing and gain an overview of commercial tools to implement automation effectively. • Apply agile methodologies to enhance testing, recognizing agility's importance, overcoming inhibitors, and implementing solutions within an agile framework.	Need for Test Process Maturity, Measurement and Improvement of a Test Process, Test Process Maturity Models • Automation and Testing Tools Need for Automation, Categorization of Testing Tools, Selection of Testing Tools, Cost Incurred in Testing Tools, Guidelines for Automated Testing, Overview of Some Commercial Testing Tools Testing Object Oriented Software Object-Oriented Testing • Using Agile Methods to Improve Software Testing The importance of Agility, Building an Agile Testing Process, Agility Inhibitors, Is Improvement Necessary, Compressing Time, Challenges, Solutions , Measuring Readiness , The Seven-Step Process 4.5 Test Plan.	
Assignments/ Activities towards CCE			
	Module 1: • To develop a comprehensive software testing policy to ensure efficient testing practices within an organization. **Module 2:** • Begin by researching and understanding the software testing terminology, methodologies, and the software testing life cycle. • Evaluate the economics of testing and its organizational impact, including management support for software testing. • Develop an understanding of the seven-step software testing process and its advantages. **Module 3:** • Identify the verification and validation activities, including verification of requirements, high-level design, low-level design, and code validation. • Explore static testing techniques such as inspections, structured walkthroughs, and technical reviews. **Module 4:** • Based on your understanding of the above concepts, draft a policy document outlining the software testing process. • Your policy document should include sections on software testing terminology, the software testing life cycle, test strategy development, verification and validation activities,		

References:

1. Chauhan, N. (2016). Software testing principles and practices. Oxford University Press.

2. Perry, W. E. (2006). Effective methods of software testing (3rd ed.). Wiley, India.
3. Desikan, S., & Ramesh, G. (2005). Software testing principles and practices. Pearson Education.
4. Patton, R. (2005). Software testing (2nd ed.). Pearson Education.
5. Dustin, E. (2003). Effective software testing: 50 specific ways to improve your testing. Pearson Education.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester III		
325614	**Data Visualization** **Major(Elective) Theory**		**4**
	Course Outcomes: **Learners will be able to:** • Gain understanding of basic data science concepts. • Learn to detect and diagnose common data issues like missing values, outliers, and inconsistencies. • Explore various machine learning techniques for data prediction. • Understand the importance of data quality and the impact of data issues on analysis and decision-making.		
Module 1	**Introduction to Data &Data data transformation**		1
	LOs: Learners will be able to • Define and categorize diverse data types proficiently. • Conduct Exploratory Data Analysis (EDA) within the Data Science lifecycle. • Master data collection and extraction techniques. • Apply data transformations like dimension reduction and feature extraction for effective analysis and modeling.	**Module Contents:** • What is Data? Different kinds of data, Data Sources, Different types of data sources, • Exploratory Data Analysis (EDA), Data Science lifecycl**e**, Data Collection • Data Extraction, Data Analysis & Modelling • Data transformations :Dimension reduction, Feature extraction, Smoothing and aggregating	
Module 2	**Python concepts used in data Science**		1
	LOs: Learners will be able to • Proficiently manipulate arrays using NumPy, performing mathematical operations and manipulating shapes. • Master the pandas library for efficient data structure handling, data insertion, and export. • Acquire skills in data cleansing, including checking and filling missing data. • Perform advanced data operations such as aggregations and joins.	**Module Contents:** • The World of arrays with Numpy : creating an array, Mathematical operations, Indexing and slicing, Shape manipulation. • Empowering Data analysis with pandas :the data structure of pandas, Inserting and exporting data • Data Cleansing**:** checking missing data, filling missing data, merging operations • Data Operations**:** Aggregation operations, Joins	
Module 3	**Inferential Statistics & Data Visualization**		1

	LOs: Learners will be able to • Possess a solid understanding of inferential statistics, including distributions, z-scores, p-values, and confidence intervals. • Master data interpretation through visualization, including chart customization and creating diverse plots. • Effectively use visualization tools like boxplots, heatmaps, scatter plots with histograms, and bubble charts. • Develop skills in interpreting data correlations, Chi-square distribution, and ANOVA results through visualization techniques.	**Module Contents:** • Inferential Statistics**:** Various forms of distribution, z-score, p-value, Type 1 and Type 2 errors, Confidence Interval, Correlation, Chi-square distribution, ANOVA • Making Sense of Data Through Visualization: Controlling the line properties of a chart, creating multiple plots, styling your plots, Boxplots, Heatmaps, Scatter plots with histogram, Bubble charts	
Module 4	**Machine Learning basics & Generating Recommendation systems**		1
	LOs: Learners will be able to • Gain comprehensive understanding of machine learning, including linear regression, logistic regression, decision trees, and clustering algorithms. • Develop proficiency in generating recommendation systems through user-based collaborative filtering, item-based collaborative filtering, and context-based filtering. • Demonstrate practical implementation skills in a case study analyzing unstructured data using text mining techniques. • Master various machine learning techniques for data analysis and recommendation system development.	**Module Contents:** • Uncovering Machine Learning: Different types of Machine Learning, Linear Regression, Logistic Regression, Decision Tree, K-means Clustering, Hierarchical Clustering • Generating Recommendations Systems:User Based collaborative filtering, Item Based collaborative filtering, Context Based filtering • Case Study Theory:Analyzing Unstructured Data using Text mining techniques. (Case Study Practical Implementation to be performed in lab as part of Practical's)	
Assignments/ Activities			
	These assignments aim to apply theoretical concepts to practical application and critical thinking. **Module 1:** • Form small teams and propose a data science project. Clearly outline the problem statement, the dataset to be used, and the goals of the project. • Specify the types of data involved, potential sources, and the relevance of the project to real-world applications. • Use EDA techniques to explore the dataset. Document the findings and insights gained from the exploration. • Apply data extraction methods, focusing on the use of NumPy and Pandas for handling arrays and data structures. **Module 2:** • Perform data cleansing operations, including checking for missing data, filling		

	gaps, and merging datasets. • Apply inferential statistics concepts to analyze the dataset. Utilize statistical tests and visualization tools to draw meaningful conclusions. **Module 3:** • Create visualizations showcasing distributions, correlations, and other relevant statistical insights using Matplotlib and Seaborn. • Implement machine learning algorithms such as linear regression, logistic regression, decision trees, and clustering techniques using scikit-learn. • Evaluate the performance of the models and document the results. • Implement recommendation systems, incorporating collaborative filtering and contextual filtering techniques. **Module 4:** • Present the generated recommendations and assess the effectiveness of the system. • Each team presents their project, covering the entire data science lifecycle from problem formulation to machine learning and recommendation system implementation. • Discuss challenges faced, solutions implemented, and lessons learned.	

References:

1. Madhavan, S. (2015). Mastering Python for Data Science: Explore the world of data science through Python and learn how to make sense of data. Packt Publishing.

2. Vander Plas, J. (2016). Python Data Science Handbook: Essential Tools for Working with Data. O'Reilly.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester III		
325615	**Data Governance** **Major(Elective) Theory**		**4**
	Course Outcomes: **Learners will be able to:** • Gain understanding of basic Data Governance concepts. • Learn various Data Governance strategies and their implementation. • Understand Data Governance within Organizational Culture. • Familiarize with Data Governance Policies and Procedures.		
Module 1	**Introduction to Data Governance**		1
	LOs: Learners will be able to • Gain understanding of data governance, its components, and significance in the modern landscape. • Comprehend practical examples and the added value data governance brings to businesses. • Understand the essential ingredients of data governance, including tools and the enterprise dictionary. • Recognize the symbiotic relationship between people and processes within the governance framework.	**Module Contents:** • **What Is Data Governance?:**What Data Governance Involves, Why Data Governance Is Becoming More Important, Examples of Data Governance in Action, The Business Value of Data Governance, Why Data Governance Is Easier in the Public Cloud. • **Ingredients of Data Governance: Tools** The Enterprise Dictionary. **Ingredients of Data Governance: People and Processes:** The People, The Process, People and Process Together	
Module 2	**Data Governance Strategies**		1
	LOs: Learners will be able to • Comprehend and articulate intricacies of data governance throughout the data life cycle. • Gain profound understanding of data quality and its significance within data governance programs. • Integrate data quality techniques to enhance and ensure data quality. • Recognize the phases, management, and operationalization of data governance.	**Module Contents:** • **Data Governance over a Data Life Cycle:** What Is a Data Life Cycle?, Phases of a Data Life Cycle, Data Life Cycle Management, Applying Governance over the Data Life Cycle, Operationalizing Data Governance. • **Improving Data Quality:** What Is Data Quality?, Why Is Data Quality Important?, Why Is Data Quality a Part of a Data Governance Program?, Techniques for Data Quality	
Module 3	**Data Governance Policies and Procedures**		1

	LOs: Learners will be able to • Develop skills to govern data in transit, including data transformations and lineage tracking. • Gain expertise in policy management, simulation, monitoring, and change management. • Acquire comprehensive knowledge of data protection planning strategies and cloud-specific considerations. • Implement physical security measures, prevention of data exfiltration, and identity and access management for agile data protection.	**Module Contents:** • **Governance of Data in Flight:** Data Transformations, Lineage, Policy Management, Simulation, Monitoring, Change Management. • **Data Protection:** Planning Protection, Data Protection in the Cloud, Physical Security, Data Exfiltration, Identity and Access Management, Keeping Data Protection Agile, Data Protection Best Practices.	
Module 4	**Data Governance and Organizational Culture**		1
	LOs: Learners will be able to • Implement effective monitoring systems and define monitoring criteria. • Acquire skills in fostering a culture of data privacy and security. • Understand the importance of leadership commitment, intention, and effective communication. • Proficient in incident handling procedures and ensuring transparency in managing data-related incidents.	**Module Contents:** • **Monitoring:** What Is Monitoring?, Why Perform Monitoring?, What Should You Monitor?, What Is a Monitoring System?, Monitoring Criteria, • Important Reminders for Monitoring • **Building a Culture of Data Privacy and Security:** Data Culture: What It Is and Why It's Important, Starting at the Top—Benefits of Data Governance to the Business, Intention, Training, and Communications, Beyond Data Literacy, Maintaining Agility, Interplay with Legal and Security, • Incident Handling, Importance of Transparency	
Assignments/ Activities			
	These assignments aim to apply theoretical concepts to practical application and critical thinking. **Module 1:** • Students analyze real-world examples of data governance implementation, identifying key elements, challenges, and business benefits. • Explore the impact of adopting data governance in public cloud environments. **Module 2:** • Groups collaborate to design a comprehensive data governance framework, considering tools, people, and processes discussed in the modules. • Emphasize the integration of an enterprise dictionary and strategies for effective data governance. **Module 3:** • Students develop a data quality improvement plan, incorporating techniques discussed in the module and understanding the importance of data quality in governance. **Module 4:** • Simulate data protection and security scenarios, focusing on planning, cloud		

	considerations, physical security, access management, and incident handling. • Discuss best practices to keep data protection agile. • Each group presents their monitoring plan, highlighting the criteria, system, and important reminders discussed in the module. • Discuss building a culture of data privacy and security, emphasizing the interplay with legal and security aspects, incident handling, transparency, and the role of organizational culture.	

Bibliography:

1. Eryurek, E., Gilad, U., Lakshmanan, V., Kibunguchy-Grant, A., & Ashdown, J. (2021). Data Governance: The Definitive Guide People, Processes, and Tools to Operationalize Data Trustworthiness. O'Reilly.
2. Ladley, J. (2012). Data Governance: How to Design, Deploy and Sustain an Effective Data Governance Program. Morgan Kaufmann.

<table>
<tr><th>SN</th><th>Courses, Modules and Outcomes</th><th>Course Contents</th><th>Cr</th></tr>
<tr><td></td><td colspan="2">Semester IV</td><td></td></tr>
<tr><td>415611</td><td colspan="2">Deep Learning
Major(Core) Theory</td><td>4</td></tr>
<tr><td></td><td colspan="2">Course Outcomes:
Learners will be able to:
• Implement perceptron with understanding of input features, weights, bias, and activation function.
• Explore deep learning concepts like activation functions and forward propagation.
• Implement convolution and pooling layers in TensorFlow, understanding convolution operations.
• Train different RNN architectures including one-to-one, one-to-many, many-to-one, and many-to-many.</td><td></td></tr>
<tr><td>Module 1</td><td colspan="2">Introduction to Deep Learning</td><td>1</td></tr>
<tr><td></td><td>LOs: Learners will be able to

• Define input features, weights, bias, and activation function for building a perceptron.
• Recognize limitations of single-layer perceptron, particularly in learning non-linear relationships.
• Describe artificial neural network structure, comprising input, hidden, and output layers.
• Outline activation functions, essential for introducing non-linearity in neural networks to learn complex patterns.</td><td>Module Contents:

• Perceptron: What is a Perceptron? Implementing perceptron, Introducing & Implementing Weights & Bias, Multilayer Perceptron, Limitations of perceptron.
• Introduction to Deep Learning: What is deep learning? Biological and artificial neurons, ANN and its layers, Input layer, Hidden layer, Output layer, exploring activation functions, the sigmoid function, the tanh function, The Rectified Linear Unit function, The leaky ReLU function, The Swish function, The softmax function, Forward propagation in ANN, How does ANN learn?</td><td></td></tr>
<tr><td>Module 2</td><td colspan="2">Convolutional Neural Networks:</td><td>1</td></tr>
<tr><td></td><td>LOs: Learners will be able to

• Examine TensorFlow's representation of computations through directed acyclic graphs (DAGs).
• Understand sessions in TensorFlow for executing operations within a computational graph.
• Analyze the general architecture of Convolutional Neural Networks (CNNs), involving convolutional, pooling, and fully connected layers.
• Implement Convolutional and Pooling layers within the CNN architecture.</td><td>Module Contents:

• Getting to Know TensorFlow
• What is TensorFlow? Understanding computational graphs and sessions, Sessions, Variables, constants, and placeholders, Introducing Tensor Board, Creating a name scope.
• Back propagation Algorithm, Neural Network Training,
• Convolutional Neural Networks:
• Overall Architecture, The Convolution Layer, Issues with the Fully Connected Layer, Convolution Operations, Padding, Stride, Batch</td><td></td></tr>
</table>

		Processing, The Pooling Layer, Implementing a Convolution Layer, Implementing a Pooling Layer, Implementing a CNN, Visualizing a CNN.	
Module 3	**Optimizers in DL**		1
	LOs: Learners will be able to • Comprehend gradient descent as an optimization method for minimizing loss during training. • Introduce adaptive learning rates based on historical parameter gradients. • Understand the challenges of training Recurrent Neural Networks (RNNs) and managing sequential dependencies. • Explain backpropagation through time, the algorithm used for training RNNs by unfolding them into computational graphs over time, and explore various RNN architectures.	**Module Contents:** • Optimizers in DL: Gradient Descent, Stochastic Gradient Descent, Mini-Batch Gradient Descent, SGD with Momentum, AdaGrad (Adaptive Gradient Descent), RMS-Prop (Root Mean Square Propagation), AdaDelta, Adam (Adaptive Moment Estimation). • Introducing RNNs: • RNN implementation and training, Backpropagation through time, Vanishing & exploding gradients, long short-term memory LSTM, • Different types of RNN architectures: • One-to-one architecture • One-to-many architecture • Many-to-one architecture • Many-to-many architecture.	
Module 4	**Deep Unsupervised Learning**		1
	LOs: Learners will be able to • Grasp autoencoders as neural network architectures for unsupervised learning by encoding and decoding input data. • Explore Generative Adversarial Networks (GANs) as a framework for training generative models via adversarial training. • Understand various scenarios for different model implementations.	**Module Contents:** • Deep Unsupervised Learning: Auto encoders (standard, sparse, denoising, contractive, etc), Variational Auto encoders, • Deep Generative Models GANS.	
Assignments/ Activities			
	These assignments aim to apply theoretical concepts to practical application and critical thinking. **Module 1:** • Task students to build a simple neural network from scratch using Python or a framework like TensorFlow/Keras. They should train it on a dataset and analyze its performance. **Module 2:** • Provide pre-trained neural network models and have students visualize the learned features and activations at different layers to understand how information is processed.		

	Module 3: • Assign students to create a CNN model for image classification using a dataset like CIFAR-10 or MNIST. They should experiment with different architectures and hyper parameters. **Module 4:** • Challenge students to create a GAN model capable of generating realistic images from a given dataset (e.g., faces, digits). They should evaluate the quality of generated images	

References:

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning (Adaptive Computation and Machine Learning series). The MIT Press.
2. Chollet, F. (2018). Deep Learning with Python. Manning.
3. Buduma, N., & Locascio, N. (2017). Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms. O'Reilly Media.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester IV		
415612	**NATURAL LANGUAGE PROCESSING** **Major (Core) Theory**		**4**
	Course Outcomes: **Learners will be able to:** • Provide an understanding of text processing for Natural Language Understanding and Natural Language Generation. • Cover various techniques available for natural language processing. • Introduce different approaches and algorithms for carrying out NLP tasks. • Explore the know-hows, issues, and challenges in NLP applications and their relevance in classical and modern contexts.		
Module 1	**INTRODUCTION TO NLP**		1
	LOs: Learners will be able to • Foster an understanding of Natural Language Understanding and Natural Language Generation. • Equip learners with practical skills in NLTK and spaCy, focusing on computing fundamentals with languages. • Master various text processing techniques, including Unicode, regular expressions, tokenization, stemming, lemmatization, segmentation, and formatting.	**Module Contents:** • **Natural Language Processing:** What is Natural Language Understanding and Natural Language Generation, Introduction to NLTK, spaCy, Computing with Languages – Text and Words, Searching Text counting vocabulary, List , Strings , Variable, Computing frequency Distribution • **Accessing Text Corpora, Lexical Resources and Processing Raw Text:** Introduction to Corpora, Conditional Frequency Distribution, Lexical Resources, Accessing text from web, Text Processing using Strings : Unicode, Regular Expressions Normalizing Text :Tokenizing Text, Stemming, Lemmatization, Segmentation, Formatting	
Module 2	**TAG AND TEXT**		1
	LOs: Learners will be able to • Develop expertise in categorizing and tagging words through taggers and Python dictionaries. • Introduce the application of machine learning algorithms like Decision Trees and Naïve Bayes Classifier for text classification. • Provide an understanding of supervised classification processes and skills to	**Module Contents:** • **Categorizing and Tagging Words:** Using a Taggers, Tagged Corpora Mapping words to properties using Python Dictionaries ,Tagging, How to determine category of a word • **Learning to Classify Text:** Using Machine Learning Algorithms to create classifiers, Supervised Classification, Decision Tree, Naïve Bayes Classifier, and Evaluation of the Classifier.	

	evaluate classifier performance.		
Module 3	**INFORMATIONAND SENTENCE ANALYSIS**		1
	LOs: Learners will be able to • Cultivate expertise in extracting information from text through Information Extraction, Chunking, Named Entity Recognition, and Relation Extraction. • Explore the complexities of analyzing sentence structures, including grammatical nuances, syntax utilization, context-free grammar, parsing, and dependency grammar. • Provide practical skills in building and assessing chunkers and managing recursion in linguistic structure.	**Module Contents:** • **Extracting Information from the Text: Information Extraction :**Information Extraction, Chunking, Developing and Evaluating Chunkers, Recursion in Linguistic Structure, Named Entity Recognition, Relation Extraction • **Analysing, Sentence Structure:** Grammatical Dilemmas, What's the use of syntax? Context free Grammar, Parsing with Context free Grammar, Dependency and Dependency Grammar	
Module 4	**Building feature based Grammar**		1
	LOs: Learners will be able to • Enhance proficiency in constructing feature-based grammar, encompassing grammatical features and processing feature structures. • Investigate sentence semantics, including Natural Language Understanding, Propositional Logic, First Order Logic (Predicate Logic), and Discourse Semantics. • Foster comprehension of organizational learning and transformation, while extending feature-based grammar to augment linguistic analysis capabilities.	**Module Contents:** • **Building feature based Grammar:** Grammatical Features, Processing Feature Structures Organizational Learning and Transformation, Extending a Feature-Based Grammar • **Analysing the Meaning of Sentences:** Natural Language Understanding, Propositional Logic, First Order Logic (Predicate Logic)The Semantics of English Sentences, Discourse Semantics	
Assignments/ Activities			
	These assignments aim to apply theoretical concepts to practical application and critical thinking: **Module 1:** • Apply advanced Natural Language Processing (NLP) techniques to analyze and extract information from a given dataset.		

	Module 2: • Build a feature-based grammar for a specific language or domain of interest. **Module 3:** • Analyze the meaning of sentences using propositional logic. • Investigate and apply discourse semantics to a set of interconnected sentences. **Module 4:** • Develop an NLU system for a specific application (e.g., sentiment analysis, information extraction).	

References:

1. Indurkhya, N., & Damerau, F. J. (2010). Handbook of Natural Language Processing (2nd ed.). CRC Press Taylor and Francis Group.

2. Manning, C., & Schutze, H. (2009). Natural Language Processing With Python. Wiley Publications.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester IV		
415623	**Deep Learning Lab: Practical** **Major (Core)**		**2**
	Course Outcomes: **Learners will be able to:** • Master essential deep learning concepts such as CNNs, RNNs, LSTMs, autoencoders, and GANs using TensorFlow. • Apply deep learning techniques to tasks like image and text processing, demonstrating practical skills in number prediction, text classification, and sentiment analysis. • Demonstrate proficiency in unsupervised learning and dimensionality reduction using autoencoders, and understand the applications of GANs in generating synthetic data.		
Module 1	**Implement using TensorFlow**		1
	LOs: Learners will be able to • Gain practical expertise in performing Eigen Values and Eigen Vectors calculations using TensorFlow. • Demonstrate hands-on skills in implementing Neural Networks for XOR operations and binary classification tasks. • Apply Neural Networks to real-world scenarios by performing Breast Cancer Classification, showcasing practical applications in medical data analysis.	**Module Contents:** • Practical to perform Eigen Values and Eigen Vectors using TensorFlow. • Practical to perform XOR Using Neural Networks. • Practical to perform Binary Classification Using Neural Networks. • Practical to perform Breast Cancer Classification Using Neural Networks	
Module 2	**Algorithm Implementation**		1
	LOs: Learners will be able to • Master the implementation of Number Prediction using Convolutional Neural Networks (CNN), showcasing image classification skills. • Demonstrate expertise in Text Classification using Recurrent Neural Networks (RNN), emphasizing sequential data processing. • Implement Movie Review Text Classification using Bi-	**Module Contents:** • Implement Number prediction using CNN • Implement Text Classification using RNN • Implement Movie Review Text Classification using Bi-Directional LSTM • Practical to implement Autoencoders. • Implement GANS algorithm using TensorFlow	

	Directional Long Short-Term Memory (LSTM) networks, showcasing advanced natural language processing.		
Assignments/ Activities towards Comprehensive Continuous Evaluation			
	These assignments aim to apply theoretical concepts to practical application and critical thinking: **Module 1:** • Implement a CNN for number prediction on a dataset like MNIST. • Develop an RNN for text classification on a dataset such as sentiment analysis. • Implement a Bi-Directional LSTM for movie review sentiment analysis. **Module 2:** • Implement autoencoders for dimensionality reduction or data reconstruction on a chosen dataset. • Implement a GAN for generating synthetic data in a chosen domain (e.g., images, text).		

References:

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning (Adaptive Computation and Machine Learning series). The MIT Press.

2. Chollet, F. (2018). Deep Learning with Python. Manning.

3. Buduma, N., & Locascio, N. (2017). Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms. O'Reilly Media.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester IV		
425611	**Information Security** **Major(Elective) Theory**		**4**
	Course Outcomes: **Learners will be able to:** • Apply knowledge of symmetric ciphers and their mechanisms. • Analyse the OSI Security Architecture's relevance to information security. • Evaluate classical encryption techniques within the symmetric cipher model. • Design principles of public key cryptography, digital signatures, authentication applications like Kerberos and X.500 Authentication Service, and countermeasures against malicious software threats.		
Module 1	**Symmetric Ciphers**		1
	LOs: Learners will be able to • Apply knowledge of fundamental services provided by symmetric ciphers. • Analyse common attacks on symmetric ciphers and defend against them. • Evaluate the OSI Security Architecture's role in network security. • Design secure encryption algorithms using block cipher design principles and explain different modes of operation for secure communication.	**Module Contents:** • **Overview** – Services, Mechanism and Attacks, The OSI Security. Architecture, A model for network security Classical Encryption techniques – Symmetric Cipher model, Substitution. Techniques, Transposition techniques, Rotor Machines, Steganography. Block Cipher and Data Encryption Standard – Simplified DES, Block. Chiper principles, The Data Encryption Standard, The strength of DES, Differential and Linear Cryptanalysis, Block Cipher design principles, Block Cipher mode of Operation	
Module 2	**Asymmetric Ciphers**		1
	LOs: Learners will be able to • Apply principles of public key cryptography and its applications. • Analyse the RSA algorithm, including key management practices. • Evaluate different public key cryptosystems, assessing their strengths and weaknesses. • Design authentication protocols and discuss their role in information security, including message authentication codes and secure hash functions.	**Module Contents:** • Public Key Cryptography and RSA – Principles of Public Key Cryptosystems, The RSA Algorithm Key management ; Other public key cryptosystemsKey Management, Diffe-Hellman Key Exchange, Elliptical Curve Arithmetic, Elliptical curve Cryptography Message Authentication and HASH Functions – Authentication requirements, Authentication Functions, Message Authentication	

		Codes, Hash Functions, security of Hash Functions and MACS Digital Signatures and Authentication Protocols – Digital Signatures, Authentication Protocols, Digital Signature Standard	
Module 3	**Network Security practice**		1
	LOs: Learners will be able to • Apply authentication protocols such as Kerberos and X.500. • Analyse secure email communication using PGP and S/MIME. • Evaluate IPSec architecture and components. • Design SSL/TLS protocols for securing web communication, considering Secure Electronic Transaction (SET) principles for e-commerce.	**Module Contents:** • **Network Security practice** : Authentication Applications – Kerberos, X.500 Authentication Service Electronic Mail Security – Pretty Good Privacy, S/MIME IP Security – IP Security Overview, IP Security Architecture, Authentication Header, Encapsulating security payload, Combining Security Associations, Key Management WEB Security – Web Security Considerations, Secure Socket Layer and Transport Layer Security, Secure Electronic Transaction	
Module 4	**System Security**		1
	LOs: Learners will be able to • Apply intrusion detection systems for threat identification. • Analyse password management policies. • Evaluate countermeasures against viruses. • Design and configure firewalls based on security needs.	**Module Contents:** • System Security : Intruders – Intruders, Intruder detection, Password Management, Malicious Software – Viruses and Related Threats, Virus Countermeasures, Firewall design principles, Trusted system.	
Assignments/ Activities			
	These assignments aim to apply theoretical concepts to practical application and critical thinking. **Module 1:** • Intruder Detection System Design: Design and configure an intrusion detection system (IDS) to detect and respond to unauthorized access attempts and suspicious activities on a network. **Module 2:** • Password Management Policy: Develop and document an effective password management policy that includes guidelines for creating strong passwords, regular password updates, and secure storage practices. **Module 3:** • Malicious Software Countermeasures: Research and propose countermeasures against viruses and related threats, considering techniques such as antivirus software deployment, malware scanning, and user education.		

	Module 4: • Firewall Configuration: Design and configure firewall rules based on specific security requirements, considering factors such as network topology, traffic patterns, and permitted services. • Trusted System Implementation: Identify and implement mechanisms to establish and maintain trust in computing environments, including integrity verification, secure boot processes, and software validation.	

References:

1. Stallings, W. (2016). Network Security Essentials. Pearson.

2. Anderson, R. J. (2020). Security Engineering: A Guide to Building Dependable Distributed Systems. Wiley.

3. Pfleeger, C. P., Pfleeger, S. L., & Margulies, J. (2015). Security in Computing.Pearson.

4. Schneier, B. (1995). Applied Cryptography: Protocols, Algorithms, and Source Code in C. Wiley.

5. Murdoch, D., & Lee, R. (2014). Blue Team Handbook: Incident Response Edition. CreateSpace Independent Publishing Platform.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester IV		
425612	**Cloud Computing** **Major(Elective) Theory**		**4**
	Course Outcomes: **Learners will be able to:** • Identify security aspects of each cloud model. • Develop a risk-management strategy for migrating to the Cloud. • Implement a public cloud instance with a public cloud service provider. • Apply a trust-based security model to different layers.		
Module 1	**Introduction to Cloud Computing:**		1
	LOs: Learners will be able to • Define cloud computing and its key characteristics, service models, and deployment models. • Provide an overview of the historical development of cloud computing and its evolution from traditional models. • Compare major cloud service	**Module Contents:** • Introduction to Cloud Computing • Online Social Networks and Applications • Cloud introduction and overview • Different clouds, Risks, Novel	

	providers like AWS, Azure, and GCP. • Identify and analyze potential security risks and challenges in cloud computing.	applications of cloud computing	
Module 2	**Cloud Computing Architecture, Cloud Deployment Models**		1
	LOs: Learners will be able to: • Define the requirements driving the emergence of cloud computing and explain CPU virtualization's role. • Provide an overview of basic cloud computing principles, discuss hypervisors, and explain the SPI framework. • Identify key drivers motivating cloud adoption and assess the impact on end-users and businesses. • Explore best practices for establishing effective governance structures in cloud environments.	**Module Contents:** • **Cloud Computing Architecture:** Requirements, Introduction Cloud computing architecture, On Demand Computing Virtualization at the infrastructure level, Security in Cloud computing environments, CPU Virtualization, A discussion on Hypervisors Storage Virtualization Cloud Computing Defined, The SPI Framework for Cloud Computing, The Traditional Software Model, The Cloud Services Delivery Model • **Cloud Deployment Models:** Key Drivers to Adopting the Cloud, The Impact of Cloud Computing on Users, Governance in the Cloud, Barriers to Cloud Computing Adoption in the Enterprise	
Module 3	**Security Issues in Cloud Computing and Access management**		1
	LOs: Learners will be able to • Apply knowledge of infrastructure security in cloud computing to identify key components involved. • Analyse network-level security measures and protocols relevant to cloud environments. • Evaluate application-level security practices and challenges specific to cloud-based applications. • Design strategies to ensure data security and storage in cloud computing environments.	**Module Contents:** • **Security Issues in Cloud Computing**: Infrastructure Security, Infrastructure Security: The Network Level, The Host Level, The Application Level, Data Security and Storage, Aspects of Data Security, Data Security Mitigation Provider Data and Its Security. • **Identity and Access Management:** Trust Boundaries and IAM, IAM Challenges, Relevant IAM Standards and Protocols for Cloud Services, IAM Practices in the Cloud, Cloud Authorization Management.	
Module 4	**Security Management in the Cloud, Privacy Issues**		1

	LOs: Learners will be able to • Apply security management standards relevant to cloud computing to ensure robust security measures. • Analyse availability management practices for SaaS, PaaS, and IaaS to optimize service availability. • Evaluate risk assessments specific to cloud security and propose effective mitigation strategies. • Design and implement tailored incident response plans for cloud computing scenarios to ensure swift and effective responses.	**Module Contents:** • **Security Management in the Cloud:** Security Management Standards, Security Management in the Cloud, Availability Management: SaaS, PaaS, IaaS. • **Privacy Issues:** Privacy Issues, Data Life Cycle, Key Privacy Concerns in the Cloud, Protecting Privacy, Changes to Privacy Risk Management and Compliance in Relation to Cloud Computing, Legal and Regulatory Implications, U.S. Laws and Regulations, International Laws and Regulations.	
Assignments/ Activities			
	These assignments aim to apply theoretical concepts to practical application and critical thinking. **Module 1:** • Research and analyze the evolution of cloud computing and its impact on modern technological landscapes. Identify and explore various online social networks and applications that leverage cloud computing technologies. Develop a comprehensive overview of cloud computing, highlighting different types of clouds, associated risks, and novel applications in diverse domains. **Module 2:** • Investigate the architecture of cloud computing systems, focusing on requirements and essential components. Explore virtualization at the infrastructure level and its role in cloud computing. Evaluate security measures implemented in cloud environments. Examine various cloud deployment models and discuss key drivers influencing cloud adoption. **Module 3:** • Conduct a comprehensive assessment of security issues inherent in cloud computing, covering infrastructure, network, host, and application levels. Explore data security and storage considerations, including mitigation strategies. Investigate identity and access management (IAM) challenges and relevant standards and protocols for cloud services. **Module 4:** • Examine security management standards and practices applicable to cloud computing environments. Analyze availability management for different cloud service models (SaaS, PaaS, IaaS). Investigate privacy concerns related to cloud computing, including data life cycle, key privacy concerns, and compliance with legal and regulatory frameworks.		

References:

1. Erl, T., Mahmood, Z., &Puttini, R. (2013). Cloud Computing: Concepts, Technology & Architecture. Prentice Hall.

2. Reese, G. (2009). Cloud Application Architectures: Building Applications and Infrastructure in the Cloud. O'Reilly Media.

3. Mather, T., Kumaraswamy, S., & Latif, S. (2009). Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance. O'Reilly Media.

4. Bahga, A., & Madisetti, V. (2014). Cloud Computing: A Hands-On Approach.CreateSpace Independent Publishing Platform.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester IV		
425613	**Robotic Process Automation** **Major(Elective) Theory**		**4**
	Course Outcomes: **Learners will be able to:** • Apply fundamental concepts and principles of Robotic Process Automation. • Analyze and gain proficiency in using popular RPA tools like UiPath, Automation Anywhere, and Blue Prism. • Design and develop RPA bots to automate specific tasks and processes. • Evaluate and troubleshoot common issues during RPA implementation.		
Module 1	**Robotic Process Automation Foundations, UiPath, Automation Anywhere**		1
	LOs: Learners will be able to • Apply understanding of RPA fundamentals, capabilities, and components. • Analyze the benefits and downsides of RPA, and compare it with other business technologies. • Evaluate and compare RPA with other business technologies, focusing on Automation Anywhere. • Design automation solutions using RPA tools, particularly Automation Anywhere.	**Module Contents:** • What is RPA, Flavors of RPA, History of RPA, What can RPA do, Components of RPA, The Benefits of RPA, The Downsides of RPA, RPA Compared to BPO, BPM, BPA, What is the Difference Between AI and RPA, RPA Tools and Platforms, Consumer Willingness for Automation, The Workforce of the Future • What is UiPath, UiPath Studio, UiPath Robot, UiPath Orchestrator, UiPath – an integrated view • What is Automation Anywhere, Enterprise Control Room, IQ Bot.	
Module 2	**Downloading and Installing UiPath Studio and Data Manipulation**		1
	LOs: Learners will be able to • Apply UiPath Studio to create automation workflows using the task recorder and step-by-step examples. • Analyze and implement control flow activities, including loops and decision-making, using sequences and flowcharts. • Evaluate the use of variables, collections, and arguments to manage data within automation projects. • Design and execute data table operations, including file management and CSV/Excel integrations, to enhance workflow efficiency.	**Module Contents:** • Learning UiPath Studio, Task Recorder, Step by step examples using the recorder • Sequencing the workflow, Activities, Control flow, various types of loops, and decision making, Step by step example using Sequence, Flowchart and Control Flow, Log Message. • Variables and scope, Collections, Arguments – purpose and use, Data table usage with examples, Clipboard management, File	

		operation with step-by-step example, CSV/Excel to data table and vice versa.	
Module 3	**Taking Control of the Controls, Exception Handling and Debugging**		1
	LOs: Learners will be able to • Apply techniques for finding and interacting with UI controls, utilizing UiExplorer, screen scraping, and OCR to avoid failure points. • Analyze the use of various plugins and extensions for automating tasks across different platforms such as SAP, Java, Citrix, and web applications. • Evaluate the creation and monitoring of assistant bots, including system event and image triggers, for efficient task automation. • Design robust exception handling strategies, implement logging, debugging techniques, and error reporting to ensure reliable automation workflows.	**Module Contents:** • Finding and attaching windows, Finding the control, Techniques for waiting for a control, Act on controls – mouse and keyboard activities, Working with UiExplorer, Handling events, Revisit recorder, Screen Scraping, When to use OCR, Types of OCR available, Avoiding typical failure points • Terminal plugin, SAP automation, Java plugin, Citrix automation, Mail plugin, PDF plugin, Web integration, Excel and Word plugins, Credential management, Extensions – Java, Chrome, Firefox and Silverlight • What are assistant bots, Monitoring system event triggers, Monitoring image and element triggers, Launching an assistant bot on a keyboard event. • Exception handling, Common exceptions and ways to handle them, Logging and taking screenshots, Debugging techniques, Collecting crash dumps, Error reporting	
Module 4	**Managing and Maintaining the Code.**		1
	LOs: Learners will be able to • Apply project organization techniques and updates effectively. • Analyze and implement reusability of workflows using state machines. • Evaluate the use of configuration files and orchestration servers for controlling bots. • Design, publish, and manage automation projects efficiently.	**Module Contents:** • Updates Project organization, Nesting workflows, Reusability of workflows, Commenting techniques, State Machine, When to use Flowcharts, State Machines or Sequences, Using config files and examples of a config file, Integrating a TFS Server • Publishing using publish utility, Overview of Orchestration Server, Using Orchestration Server to Control bots, Using Orchestration server to deploy bots, License management, Publishing and managing	

Assignments/ Activities		
	These assignments aim to apply theoretical concepts to practical application and critical thinking. **Module 1:** Assignment: Compare and Contrast RPA Tools • Research the history, components, benefits, and downsides of RPA. • Compare RPA to BPO, BPM, BPA, and AI. • Identify the key features of UiPath and Automation Anywhere. • Prepare a report summarizing the comparison between UiPath and Automation Anywhere in terms of features, ease of use, and integration capabilities. • Downloading and Installing UiPath Studio and Data Manipulation **Module 2:** Assignment: Build a Simple Automation Workflow • Download and install UiPath Studio. • Create a simple automation workflow using the Task Recorder. • Use sequences and flowcharts to structure the workflow. • Incorporate control flows, loops, decision making, and log messages. • Demonstrate data manipulation using variables, collections, arguments, and data tables. • Perform file operations and demonstrate CSV/Excel data handling. • Taking Control of the Controls, Exception Handling and Debugging **Module 3:** Assignment: Develop a Comprehensive UI Automation • Utilize UiExplorer to find and attach windows and controls. • Implement techniques for waiting for a control, and perform mouse and keyboard activities. • Create a workflow that incorporates screen scraping and OCR. • Use plugins (e.g., Terminal, SAP, Java, Citrix, Mail, PDF, Web, Excel, and Word) to enhance automation. • Develop assistant bots triggered by system events or keyboard events. • Implement exception handling and debugging techniques. • Create a detailed report on the common exceptions and methods to handle them, including logging, taking screenshots, and error reporting. • Managing and Maintaining the Code **Module 4:** Assignment: Project Organization and Deployment • Organize a project using updates, nesting workflows, and commenting techniques. • Demonstrate the reusability of workflows using state machines, flowcharts, and sequences. • Create and use configuration files within a project. • Integrate a TFS server for version control. • Publish the project using the publish utility and manage it using the Orchestration Server. • Deploy bots via the Orchestration Server and manage licenses. • Prepare a documentation report detailing the project organization, deployment process, and best practices for maintaining the code.	

References:

1. Tripathi, A. M. (2018). Learning Robotic Process Automation. Packt Publishing.
2. Taulli, T. (2020). The Robotic Process Automation Handbook: A Guide to Implementing RPA Systems. Apress.
3. Sireci, J. (2020). The Practitioner's Guide to RPA. Farchair Solutions.
4. Bornet, P., Barkin, I., & Wirtz, J. (2021). Intelligent Automation: Welcome to the World of Hyperautomation.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester IV		
425614	**Social Network Analysis** **Major(Elective) Theory**		**4**
	Course Outcomes: **Learners will be able to:** • Distinguish between various network types and relationships, including binary, valued, symmetric, and asymmetric relationships. • Utilize graph theory for social network analysis, employing methods like adjacency matrices, edge-lists, and graph traversals. • Investigate the role of ontology in the Semantic Web and its relevance in network data representation. • Analyze network structures, centrality, and centralization in Social Network Analysis (SNA), encompassing measures like density, reachability, and centrality.		
Module 1	**Introduction to social network analysis (SNA)**		1
	LOs: Learners will be able to • Analyze network relationships, discerning between various types such as binary, valued, symmetric, and asymmetric. • Utilize graph theory techniques like adjacency matrices and edge-lists for social network analysis. • Apply clustering methods in social networks, connecting theoretical concepts with real-world electronic sources. • Develop skills in understanding and utilizing ego-centric and socio-centric density measures within networks.	**Module Contents:** • **Introduction to networks and relations**- analyzing relationships to understand people and groups, binary and valued relationships, symmetric and asymmetric relationships, multimode relationships, • **Using graph theory for social networks analysis-** adjacency matrices, edge-lists, adjacency lists, graph traversals and distances, social networks vs. link analysis, ego-centric and socio-centric density , clustering. • **Social Network analysis:** Development of Social Network Analysis, Electronic sources for network analysis, Blogs and online communities.	
Module 2	**Networks, Centrality, centralization and Ontology**		1
	LOs: Learners will be able to • Understand the role of ontology in the Semantic Web and its application in knowledge representation. • Gain expertise in Social Network Analysis (SNA) techniques for	**Module Contents:** • **Ontology and their role in the Semantic Web:** Semantic Web , Ontology, Ontology based knowledge Representation , Resource Description Framework –	

	analyzing network characteristics. • Analyze network features like density, reachability, and centrality measures such as degree and closeness. • Emphasize interpreting and visualizing network structures using centrality algorithms like PageRank.	Web Ontology ,State-of-the-art in network data representation ,Ontological representation of social individuals ,Ontological representation of social relationships. • **Networks, Centrality and centralization in SNA Understanding networks**-density, reachability, connectivity, reciprocity, group-external and group-internal ties in networks, ego networks, extracting and visualizing ego networks, structural holes, Centrality-degree of centrality, closeness and betweenness centrality, local and global centrality, centralization and graph centers, notion of importance within network, Google pagerank algorithm.	
Module 3	**Extraction and mining communities in web social networks**		1
	LOs: Learners will be able to • Detect and evaluate communities within web social networks with proficiency. • Utilize community detection methods and mining algorithms, including tools like Girvan Newman. • Grasp measures of similarity and structural equivalence in SNA, exploring various approaches. • Understand clustering techniques and diverse similarity metrics for network analysis.	**Module Contents:** • **Communities in Web Social Network:** Detecting communities in social networks, Definition of community, Evaluating communities, Methods for community detection and mining,Applications of community mining algorithms, Tools for detecting communities social network infrastructures and communities , Girvan Newman algorithm ,Decentralized online social networks , Multi-Relational characterization of dynamic social network communities. • **Measures of similarity and structural equivalence in SNA**: Approaches to network positions and social roles-defining equivalence or similarity, structural equivalence, automorphic equivalence, finding equivalence sets, brute force and Tabu search, regular equivalence, equivalence of distances: Maxsim, regular equivalence • **Understanding clustering:** agglomerative and divisive clusters, Euclidean,	

		Manhattan, and squared distances, binary relations, matches , exact, Jaccard, Hamming	
Module 4	**Two-mode networks for SNA**:		1
	LOs: Learners will be able to • Master two-mode networks, including bipartite data structures and quantitative analyses like SVD. • Excel in qualitative analysis techniques, such as core-periphery and factions analysis. • Explore intricacies of affiliation and attribute networks within two-mode structures. • Conduct comprehensive analysis and visualization of two-mode data sets.	**Module Contents:** • **Understanding mode networks-** Bipartite data structures, visualizing two-mode data, quantitative analysis using two-mode Singular value decomposition (SVD) analysis, two-mode factor analysis, two-mode correspondence analysis, qualitative analysis using two mode core-periphery analysis, two-mode factions analysis, affiliation and attribute networks	
Assignments/ Activities			
	These assignments aim to apply theoretical concepts to practical application and critical thinking. **Module 1:** • Write a program to compute the following for a given a network: (i) number of edges, (ii) number of nodes; (iii) degree of node; (iv) node with lowest degree; (v) the adjacency list; (vi) matrix of the graph. **Module 2:** • Perform following tasks: (i) View data collection forms and/or import one mode/two-mode datasets; (ii) Basic Networks matrices transformations • Compute the following node level measures: (i) Density; (ii) Degree; (iii) Reciprocity; (iv) Transitivity; (v) Centralization; (vi) Clustering. **Module 3:** • For a given network find the following: (i) Length of the shortest path from a given node to another node; (ii) the density of the graph; (iii) Draw egocentric network of node G with chosen configuration parameters. **Module 4:** • Write a program to distinguish between a network as a matrix, a network as an edge list, and a network as a sociogram (or "network graph") using 3 distinct networks representatives of each.		

References:

1. Mika, P. (2007). Social Networks and the Semantic Web (1st ed.). Springer.
2. Furht, B. (2010). Handbook of Social Network Technologies and Applications (1st ed.). Springer.
3. Hanneman, R. A., & Riddle, M. (2005). Introduction to Social Network Methods. University of California. [Published in digital form and available at http://faculty.ucr.edu/~hanneman/nettext/index.html].
4. Tsvetovat, M., & Kouznetsov, A. (2011). Social Network Analysis for Startups: Finding Connections on the Social Web. O'Reilly Media.
5. Scott, J. (2012). Social Network Analysis (3rd ed.). SAGE Publications.

6. Xu, G., Zhang, Y., & Li, L. (2011). Web Mining and Social Networking – Techniques and Applications (1st ed.). Springer.
7. Goh, D., & Foo, S. (2008). Social Information Retrieval Systems: Emerging Technologies and Applications for Searching the Web Effectively. IGI Global Snippet.
8. Chevalier, M., Julien, C., & Soulé-Dupuy, C. (2009). Collaborative and Social Information Retrieval and Access: Techniques for Improved User Modelling. IGI Global Snippet.
9. Breslin, J. G., Passant, A., & Decker, S. (2009). The Social Semantic Web. Springer.
10. De Nooy, W., Mrvar, A., & Batagelj, V. (2011). Exploratory Social Network Analysis with Pajek (2nd ed.). Cambridge University Press.
11. Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2013). Analyzing Social Networks. SAGE Publications.
12. Kolaczyk, E. D., & Csárdi, G. (2014). Statistical Analysis of Network Data with R. Springer.
13. Brandes, U., & Erlebach, T. (Eds.). (2005). Network Analysis: Methodological Foundations. Springer.
14. Carrington, P. J., Scott, J., & Wasserman, S. (Eds.). (2005). Models and Methods in Social Network Analysis. Cambridge University Press.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	Semester IV		
425615	**Agile Methodology** **Major (Elective) Theory**		**4**
	Course Outcomes: **Learners will be able to:** • Apply Agile requirement techniques to streamline software development processes. • Analyse various Agile software methodologies to identify the most suitable approach for the project. • Evaluate different Agile estimation techniques to accurately plan project timelines and resource allocation. • Design an Agile testing approach to ensure the quality and functionality of software products throughout the development lifecycle.		
Module 1	**Introduction to Agile Methodologies**		1
	LOs: Learners will be able to • Apply traditional software development methodologies to understand their approach and processes. • Analyse the limitations and challenges of traditional software development methodologies to identify areas for improvement. • Evaluate the concept of Agile methodology as a more flexible and iterative approach to software development. • Design collaborative requirements analysis using the Class Responsibility Collaborator (CRC) method to enhance communication and understanding among stakeholders.	**Module Contents:** • Traditional approach of Software Development Methodology, Need of Agile software Development, Defining Agile, Agile Manifesto Principles of Agile , Values of Agile ,Business Benefits of Agile Software Development • Traditional Requirements Development , Principle of Agile Requirements Development ,Agile Requirements : Epics and User stories ,Difference between Epics and User stories ,Backlog Management, Class Responsibility Collaborator.	
Module 2	**Scrum and Kanban Methodologies**		1
	LOs: Learners will be able to • Apply Scrum framework concepts to understand its role and relevance in Agile software development. • Analyse the advantages and benefits of adopting the Scrum framework to determine its suitability for project needs. • Evaluate the underlying principles of the Scrum framework to ensure alignment with project objectives and values. • Design key artifacts in Scrum, such as the Product Backlog, Sprint Backlog,	**Module Contents:** • Introduction to Scrum framework, Advantages of Scrum Framework. Phases of Scrum, Principles of Scrum, Roles: Product owner, team members and scrum master, Scrum Ceremonies :Sprint, sprint planning, daily scrum, sprint review, and sprint retrospective, Artifacts: Product backlog, sprint backlog and increments. • Introduction to Kanban framework, Workflow, Limit the amount of work in	

	and Increments, to effectively manage project requirements and deliverables.	progress, pulling work from column to column, Kanban board, Adding policies to the board, Cards and their optimization.Kanban Practices , Kanban Flow practices.Work Item Age.Kanban vs Scrum.	
Module 3	**Extreme Programming and Agile Estimation Techniques**		1
	LOs: Learners will be able to • Apply the foundational values and principles of Extreme Programming (XP) to guide software development practices. • Analyse and implement the twelve practices of XP, such as pair programming and test-driven development, to enhance software quality and productivity. • Evaluate the life cycle stages of an XP project, from planning to release, to ensure effective project management and delivery. • Design Agile estimation techniques like Planning Poker and Shirt Sizes to optimize planning processes and enhance project estimation accuracy.	**Module Contents:** • Basic values and principles, Roles, Twelve practices of XP, Pair programming, XP team, Life cycle and tools for XP.,Good practices need to be practiced in extreme programming,Advantages of Extreme Programming • Agile Maturity Model and Agile Estimation Techniques - Planning Poker-Shirt Sizes. Dot Voting, Bucket System.	
Module 4	**Agile Testing**		1
	LOs: Learners will be able to • Apply the Agile Testing Quadrants model to classify testing activities into distinct categories. • Analyse the iterative nature of the Agile Testing Life Cycle within Agile development to ensure continuous improvement. • Evaluate the principles and practices of Behavior Driven Development (BDD) as an effective Agile testing technique. • Design Agile test metrics to measure and improve the testing process, utilizing them effectively to assess project progress and identify areas for improvement.	**Module Contents:** • Agile Testing Life Cycle, Agile Testing Quadrants, Agile Testing Techniques: Behavior Driven Development, Test Driven Development Acceptance Test Driven Development Testing.Role of Agile Tester.User stories approach in Acceptance Test Driven Development Testing.Other Techniques - Exploratory Testing , Session Based testing. • Agile Test Metrics.	
Assignments/ Activities			
	These assignments aim to apply theoretical concepts to practical application and critical thinking. **Module 1:** • Assignment: Agile Manifesto Analysis • Students will analyze the Agile Manifesto and its principles. They will identify the values and principles of Agile software development and discuss the need for Agile methodologies in contrast to traditional software development approaches. The assignment will require students to critically evaluate the business benefits of Agile software development based on the principles		

	outlined in the manifesto. **Module 2:** • Assignment: Scrum vs. Kanban Analysis • Students will compare and contrast the Scrum and Kanban frameworks. They will analyze the advantages of each framework, including their principles, roles, ceremonies, and artifacts. Through this assignment, students will design a comparative analysis highlighting the differences between Scrum and Kanban, including their workflows, work-in-progress limits, and practices. **Module 3:** • Assignment: Agile Practices Implementation Plan • Students will design an implementation plan for adopting Extreme Programming (XP) practices within a hypothetical software development team. They will apply Agile estimation techniques such as Planning Poker, Shirt Sizes, Dot Voting, and the Bucket System to plan and execute the adoption of XP practices. The assignment will require students to evaluate the advantages of XP and assess its suitability for different project scenarios. **Module 4:** • Assignment: Agile Testing Strategies Proposal • Students will propose Agile testing strategies based on the Agile Testing Quadrants and techniques such as Behavior Driven Development (BDD) and Test Driven Development (TDD). They will design a testing approach for a given software project, considering user stories and acceptance criteria. Additionally, students will evaluate the role of Agile testers and propose Agile test metrics for measuring and improving the testing process.	

References:

1. Stellman, A., & Hart, J. A. (2015). Learning Agile. O'Reilly.
2. Crispin, L., & Gregory, J. (2008). Agile Testing: A Practical Guide for Testers and Agile Teams. Addison Wesley.
3. Schwaber, K., & Beedle, M. (2002). Agile Software Development with Scrum. Pearson.
4. Martin, R. C. (2002). Agile Software Development, Principles, Patterns and Practices. Pearson.