

SN	Courses, Modules and Outcomes	Course Contents	Cr
	<b>Semester III</b>		
<b>315511</b>	<b>BIG DATA ANALYTICS Major (Core) Theory</b>		<b>4</b>
	<b>Course Outcomes: Learners will be able to:</b> <ul style="list-style-type: none"> <li>Understand IBM's approach to big data and analytics.</li> <li>Understand the flow of data in a Hadoop ecosystem.</li> <li>Describe the steps involved in processing and analyzing data in Hadoop.</li> <li>Explain the Hive data warehousing and SQL-like query language.</li> <li>Describe the services and architecture of Apache Hive.</li> </ul>		
<b>Module 1</b>	<b>INTRODUCTION TO BIG DATA AND HADOOP, HDFS (Hadoop Distributed File System)</b>		<b>1</b>
	<b>LOs:</b> Learners will be able to <ul style="list-style-type: none"> <li>Understanding the various types of digital data, including structured, semi-structured, and unstructured data.</li> <li>Grasping the fundamental concepts of Big Data, including the three V's (Volume, Velocity, Variety).</li> <li>Understanding the challenges and opportunities presented by Big Data.</li> <li>Developing skills in using Unix tools for data analysis.</li> </ul>	<b>Module Contents:</b> <ul style="list-style-type: none"> <li>Types of Digital Data, Introduction to BigData, BigData Analytics, History of Hadoop, Apache Hadoop, Analyzing Data with Unix tools, Analyzing Data with Hadoop, Hadoop Streaming, Hadoop Echo System, IBM Big Data Strategy, Introduction to Info sphere Big Insights and BigSheets.</li> <li>The Design of HDFS, HDFS Concepts, Command Line Interface, Hadoop file system interfaces, Data flow, Data Ingest with Flume and Scoop and Hadoop archives, Hadoop I/O: Compression, Serialization, Avro and File-Based Data Structures</li> </ul>	
<b>Module 2</b>	<b>MapReduce</b>		<b>1</b>
	<b>LOs:</b> Learners will be able to <ul style="list-style-type: none"> <li>Understanding the key components and phases involved in a MapReduce job.</li> <li>Exploring the sequence of steps from job submission to completion.</li> <li>Understanding the job</li> </ul>	<b>Module Contents:</b> <ul style="list-style-type: none"> <li>Anatomy of a Map Reduce Job Run, Failures, Job Scheduling, Shuffle and Sort, Task Execution, MapReduce Types and Formats, MapReduce Features</li> </ul>	

	<p>scheduling process in a MapReduce framework.</p> <ul style="list-style-type: none"> <li>Exploring how tasks are scheduled and allocated resources in a distributed environment.</li> </ul>		
<b>Module 3</b>	<b>Hadoop EcoSystem</b>		1
	<p><b>LOs: Learners will be able to</b></p> <ul style="list-style-type: none"> <li>Learning about the different execution modes of Pig, including local and MapReduce modes.</li> <li>Understanding the advantages and use cases for each execution mode.</li> <li>Understanding the syntax and semantics of Pig Latin, the scripting language for Pig.</li> <li>Learning to write Pig Latin scripts for data processing tasks.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>Pig: Introduction to PIG, Execution Modes of Pig, Comparison of Pig with Databases, Grunt, Pig Latin, User Defined Functions, Data Processing operators. Hive : Hive Shell, Hive Services, HiveMetastore, Comparison with Traditional Databases, HiveQL, Tables, Querying Data and User Defined Functions. Hbase: HBase Basics, Concepts, Clients, Example, Hbase Versus RDBMS. Big SQL : Introduction</li> </ul>	
<b>Module 4</b>	<b>Data Analytics with R Machine Learning:</b>		1
	<p><b>LOs: Learners will be able to</b></p> <ul style="list-style-type: none"> <li>Understanding the significance of big data in the analytics landscape.</li> <li>Exploring the challenges and opportunities posed by large-scale data.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>Introduction, Supervised Learning, Unsupervised Learning, Collaborative Filtering. Big Data Analytics with Big R.</li> </ul>	
<b>Assignments/ Activities</b>			
	<p>These assignments aim to apply theoretical concepts to practical application and critical thinking.</p> <ul style="list-style-type: none"> <li>Set up a small Hadoop cluster or use a cloud environment.</li> <li>Assign tasks where students perform common HDFS operations using the command line interface.</li> <li>Assign a case where students need to optimize the scheduling of MapReduce jobs.</li> <li>Consider factors such as resource utilization, data locality, and overall cluster efficiency.</li> <li>Provide a dataset and task students with building a supervised learning model using Big R.</li> </ul>		

### **Bibliography:**

White, T. (2012). Hadoop: The Definitive Guide (3rd ed.). O'Reilly Media.

- Acharya, S., & Chellappan, S. (2015). *Big Data Analytics*. Wiley.
- Berthold, M., & Hand, D. J. (2007). *Intelligent Data Analysis*. Springer.
- Liebowitz, J. (2013). *Big Data and Business Analytics*. Auerbach Publications, CRC Press.
- Plunkett, T., Hornick, M. (2013). *Using R to Unlock the Value of Big Data: Big Data Analytics with Oracle R Enterprise and Oracle R Connector for Hadoop*. McGraw-Hill/Osborne Media, Oracle Press.
- Rajaraman, A., & Ullman, J. D. (2012). *Mining of Massive Datasets*. Cambridge University Press.
- Franks, B. (2012). *Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics*. John Wiley & Sons.
- Myat, G. J. (2007). *Making Sense of Data*. John Wiley & Sons.
- Warden, P. (2011). *Big Data Glossary*. O'Reilly.
- Minelli, M., Chambers, M., & Dhiraj, A. (2013). *Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses*. Wiley Publications.
- Sathi, A. (2012). *Big Data Analytics: Disruptive Technologies for Changing the Game*. MC Press.
- Zikopoulos, P., DeRoos, D., Parasuraman, K., Deutsch, T., Giles, J., & Corigan, D. (2012). *Harness the Power of Big Data: The IBM Big Data Platform*. Tata McGraw Hill Publications.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	<b>Semester III</b>		
<b>315512</b>	<b>Machine Learning Major (Core)</b>		<b>4</b>
	<b>Course Outcomes:</b> <b>Learners will be able to:</b> <ul style="list-style-type: none"> <li>Recognize real-world examples and applications of machine learning.</li> <li>Define and comprehend the principles of unsupervised learning.</li> <li>Grasp the concepts of subset selection for dimensionality reduction.</li> <li>Handle multiclass classification using One vs One and One vs Rest approaches.</li> </ul>		
<b>Module 1</b>	<b>Introduction :</b>		<b>1</b>
	<b>LOs:</b> Learners will be able to <ul style="list-style-type: none"> <li>Understand the role of machine learning in automating tasks and making predictions.</li> <li>Understand the importance of splitting data into training and testing sets for model evaluation.</li> </ul>	<b>Module Contents:</b> <ul style="list-style-type: none"> <li><b>Introduction:</b> What is Machine Learning, Examples of Machine Learning applications, Training versus Testing, Positive and Negative Class, Cross- validation.</li> <li>Types of Learning: Supervised, Unsupervised and Semi-Supervised Learning.</li> <li>Dimensionality Reduction: Introduction to Dimensionality Reduction, Subset Selection, Introduction to Principal Component Analysis.</li> </ul>	
<b>Module 2</b>	<b>Binary and Multiclass Classification:</b>		<b>1</b>
	<b>LOs:</b> Learners will be able to <ul style="list-style-type: none"> <li>Define and interpret a confusion matrix.</li> <li>Understand how true positives, true negatives, false positives, and false negatives contribute to classification assessment.</li> </ul>	<b>Module Contents:</b> <ul style="list-style-type: none"> <li>Assessing Classification Performance, Handling more than two classes, Multiclass Classification-One vs One, One vs Rest Linear Models: Perceptron, Support Vector Machines (SVM), Soft Margin SVM, Kernel methods for non-linearity</li> </ul>	

<b>Module 3</b>	<b>Regression</b>		1
	<p><b>LOs:</b> Learners will be able to</p> <ul style="list-style-type: none"> <li>Interpret the implications of different error metrics in the context of regression problems.</li> <li>Identify factors that contribute to overfitting in regression models.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>Assessing performance of Regression – Error measures, Overfitting and Underfitting, Catalysts for Overfitting, VC Dimensions</li> <li>Linear Models: Least Square method, Univariate Regression, Multivariate Linear Regression, Regularized Regression - Ridge Regression and Lasso</li> <li>Theory of Generalization: Bias and Variance Dilemma, Training and Testing Curves Case Study of Polynomial Curve Fitting.</li> </ul>	
<b>Module 4</b>	<b>LOGIC BASED AND ALGEBRAIC MODELS, TRENDS IN MACHINE LEARNING</b>		1
	<p><b>LOs:</b> Learners will be able to</p> <ul style="list-style-type: none"> <li>Understand the decision-making process based on the proximity of instances.</li> <li>Define decision trees and their role in classification and regression.</li> <li>Understand the process of tree construction and decision-making.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>Distance Based Models: Neighbors and Examples, Nearest Neighbor Classification, Distance based clustering algorithms - K-means and K-medoids, Hierarchical clustering.</li> <li>Rule Based Models: Rule learning for subgroup discovery, Association rules mining – Apriori Algorithm, Confidence and Support parameters.</li> <li>Tree Based Models: Decision Trees, Minority Class, Impurity Measures – Gini Index and Entropy, Best Split</li> <li>Ensemble Learning: Combining Multiple</li> </ul>	

		<p>Models, Bagging, Randomization, Boosting, Stacking</p> <ul style="list-style-type: none"> <li>• Reinforcement Learning: Exploration, Exploitation, Rewards, Penalties</li> <li>• Deep Learning: The Neuron, Expressing Linear Perceptron as Neurons, Feed Forward Neural Networks, Linear Neurons and their Limitations, Sigmoid, Tanh and ReLU Neurons</li> </ul>	
<b>Assignments/ Activities towards CCE</b>			
		<ul style="list-style-type: none"> <li>• Conducting training and testing on sample datasets.</li> <li>• Implementing cross-validation techniques for model assessment.</li> <li>• Implementing supervised learning algorithms on labeled datasets.</li> <li>• Exploring unsupervised learning techniques for clustering or dimensionality reduction.</li> <li>• Discussing scenarios where semi-supervised learning is advantageous.</li> <li>• Evaluating classification performance on various datasets.</li> <li>• Discussing and analyzing confusion matrices and ROC curves.</li> </ul>	

**Bibliography:**

Kevin Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012

Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning, Springer 2009

David Barber, Bayesian Reasoning and Machine Learning (Cambridge University Press). Online version available

Tom Mitchell. Machine Learning (McGraw Hill) , First Edition

Richard O. Duda, Peter E. Hart, David G. Stork. Pattern Classification (John Wiley & Sons)

Ethern Alpaydin, Introduction to Machine Learning, MIT Press, 3rd Edition.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	<b>Semester III</b>		
<b>315513</b>	<b>DATA SCIENCE Major (Core) Theory</b>		<b>2</b>
	<b>Course Outcomes:</b> <b>Learners will be able to:</b> <ul style="list-style-type: none"> <li>• Understanding of Data Science Concepts.</li> <li>• Acquire skills in collecting and handling data through various methods, including APIs.</li> <li>• Develop a solid understanding of statistical concepts and terminologies.</li> <li>• Learn the principles and applications of Naive Bayes classification algorithms.</li> </ul>		
<b>Module 1</b>	<b>Introduction to core concepts and technologies, Data collection and management, Data analysis:</b>		<b>1</b>
	<b>LOs:</b> Learners will be able to <ul style="list-style-type: none"> <li>• Understand the fundamental concepts and terminologies related to data science.</li> <li>• Understand practical applications of data science across different industries, showcasing real-world scenarios where data science is applied.</li> <li>• Acquire knowledge on collecting data through various methods, with a focus on understanding and using APIs.</li> <li>• Gain an understanding of the Naive Bayes algorithm and its use in probabilistic classification.</li> </ul>	<b>Module Contents:</b> <ul style="list-style-type: none"> <li>• Introduction, Terminology, data science process, data science toolkit, Types of data, Example applications.</li> <li>• Introduction, Sources of data, Data collection and APIs, Exploring and fixing data, Data storage and management, Using Multiple data sources.</li> <li>• Introduction, Terminology and concepts, Introduction to statistics, Central tendencies and distributions, Variance, Distribution properties and arithmetic, Samples/CLT.</li> <li>• Basic machine learning algorithms, Line a regression, SVM, Naive Bayes.</li> </ul>	
<b>Module 2</b>	<b>Data visualization and applications:</b>		<b>1</b>
	<b>LOs:</b> Learners will be able to <ul style="list-style-type: none"> <li>• Identify and categorize different types of data visualizations.</li> <li>• Understand the technologies and tools used for effective data visualization.</li> <li>• Understand the methods and tools used in developing applications for data science purposes.</li> </ul>	<b>Module Contents:</b> <ul style="list-style-type: none"> <li>• Introduction, Types of data visualization, Data for visualization: Data types, Data encodings, Retinal variables, Mapping variables to encodings, Visual encodings.</li> <li>• Applications of Data Science, Technologies for visualization.</li> <li>• Recent trends in various data collection and analysis techniques, various visualization techniques, application development methods of used in data science.</li> </ul>	

<b>Assignments/ Activities</b>	
	<p>These assignments aim to apply theoretical concepts to practical application and critical thinking.</p> <ul style="list-style-type: none"> <li>• Creating a conceptual framework that illustrates the key components of data science, including data collection, processing, analysis, and visualization.</li> <li>• Creating a conceptual framework that illustrates the key components of data science, including data collection, processing, analysis, and visualization.</li> <li>• Organize a workshop where students practice collecting data from different sources, including APIs.</li> <li>• Require them to write a report explaining their findings, including measures of central tendency, variance, and distribution properties.</li> <li>• Discuss the scenarios in which Naive Bayes might be more suitable and encourage critical thinking about algorithm selection.</li> </ul>

**Bibliography:**

O'Neil, C., & Schutt, R. (2013). Doing Data Science: Straight Talk From The Frontline. O'Reilly Media.

Leskovec, J., Rajaraman, A., & Ullman, J. (2014). Mining of Massive Datasets (2nd ed.). Cambridge University Press.



SN	Courses, Modules and Outcomes	Course Contents	Cr	
	<b>Semester III</b>			
<b>315524</b>	<b>Big Data Analytics Lab: Practical Major (Core)</b>		<b>2</b>	
	<p><b>Course Outcomes:</b> <b>Learners will be able to:</b></p> <ul style="list-style-type: none"> <li>• Optimize business decisions and create competitive advantage with Big Data analytics .</li> <li>• Imparting the architectural concepts of Hadoop and introducing map reduce paradigm.</li> <li>• Derive business benefit from unstructured data.</li> <li>• Introduce programming tools PIG &amp; HIVE in Hadoop echo system.</li> <li>• Developing Big Data applications for streaming data using Apache Spark.</li> </ul>			
<b>Module 1</b>	<b>Exploring Big Data with Hadoop</b>		<b>1</b>	
	<p><b>LOs:</b> Learners will be able to</p> <ul style="list-style-type: none"> <li>• Utilize and navigate big data tools and platforms (e.g., Hadoop etc.) for data storage, retrieval, and processing.</li> <li>• Implement distributed computing techniques to handle and analyze massive datasets efficiently.</li> <li>• Proficiency in Apache Hadoop cluster and MapReduce.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>• Perform setting up and Installing Hadoop in its two operating modes: <ul style="list-style-type: none"> <li>1)Pseudo distributed,</li> <li>2)Fully distributed.</li> </ul> </li> <li>• Use web based tools to monitor your Hadoop setup. Implement the following file management tasks in Hadoop: <ul style="list-style-type: none"> <li>1)Adding files and directories</li> <li>2)Retrieving files</li> <li>3)Deleting files</li> </ul> </li> <li>• Benchmark and stress test an Apache Hadoop cluster Run a basic Word Count Map Reduce program to understand Map Reduce Paradigm. <ul style="list-style-type: none"> <li>1)Find the number of occurrences of each word appearing in the input file(s)</li> <li>2)Performing a MapReduce Job for word search count (look for specific keywords in a file)</li> </ul> </li> <li>• Stop word elimination problem: <ul style="list-style-type: none"> <li>1)Input: A large textual file containing one sentence per line. A small file containing a set of stop words (One stop word per line)</li> <li>2)Output: A textual file containing the same sentences of the large input file without the words appearing in the small file.</li> </ul> </li> </ul>		

Module 2	MapReduce Implementation	1
	<p><b>LOs:</b> Learners will be able to</p> <ul style="list-style-type: none"> <li>• Develop MapReduce programs to process and analyze semi-structured, record-oriented data.</li> <li>• Implement algorithms to calculate average, maximum, and minimum temperatures for each year in a large dataset using MapReduce.</li> <li>• Utilize Pig Latin and Hive to perform various data operations, including sorting, grouping, joining, projecting, and filtering on diverse datasets.</li> <li>• Create, alter, and drop databases, tables, views, functions, and indexes in Hive, demonstrating proficiency in managing data structures in a distributed environment.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>• Write a Map Reduce program that mines weather data. Weather sensors collecting data every hour at many locations across the globe gather large volume of log data, which is a good candidate for analysis with MapReduce, since it is semi structured and record-oriented. Data available at: <a href="https://github.com/tomwhite/hadoop-book/tree/master/input/ncdc/all">https://github.com/tomwhite/hadoop-book/tree/master/input/ncdc/all</a>.             <ol style="list-style-type: none"> <li>1) Find average, max and min temperature for each year in NCDC data set?</li> <li>2) Filter the readings of a set based on value of the measurement, Output the line of input files associated with a temperature value greater than 30.0 and store it in a separate file.</li> </ol> </li> <li>• Purchases.txt Dataset             <ol style="list-style-type: none"> <li>1) Instead of breaking the sales down by store, give us a sales breakdown by product category across all of our stores What is the value of total sales for the following categories? Toys, Consumer Electronics</li> <li>2) Find the monetary value for the highest individual sale for each separate store What are the values for the following stores? Reno, Toledo, Chandler</li> <li>3) Find the total sales value Across all the stores, and the total number of sales.</li> </ol> </li> <li>• Install and Run Pig then write Pig Latin scripts to sort, group, join, project, and filter your data.</li> <li>• Write a Pig Latin scripts for finding TF-IDF value for book dataset (A corpus of eBooks available at: Project Gutenberg)</li> <li>• Install and Run Hive then use Hive to create, alter, and drop databases, tables, views, functions,</li> </ul>

		<p>and indexes.</p> <ul style="list-style-type: none"> <li>• Install, Deploy &amp; configure Apache Spark Cluster. Run apache spark applications using Scala.</li> <li>• Data analytics using Apache Spark on Amazon food dataset, find all the pairs of items frequently reviewed together.</li> <li>• Write a single Spark application that: <ul style="list-style-type: none"> <li>▪ Transposes the original Amazon food dataset, obtaining a PairRDD of the type:</li> <li>▪ &lt;user_id&gt; → &lt;list of the product_ids reviewed by user_id&gt;</li> <li>▪ Counts the frequencies of all the pairs of products reviewed together;</li> </ul> </li> <li>• Writes on the output folder all the pairs of products that appear more than once and their frequencies. The pairs of products must be sorted by frequency.</li> </ul>	
<b>Assignments/ Activities towards Comprehensive Continuous Evaluation</b>			
	<ul style="list-style-type: none"> <li>• Assign MapReduce programs to analyze provided weather data. Ask them to calculate average, maximum, and minimum temperatures for each year in the dataset.</li> <li>• Analyse filter specific readings based on temperature values and store the filtered data into separate files using MapReduce.</li> <li>• Illustrate dataset similar to the Purchases.txt dataset. Allocate to write scripts (using tools like Hadoop or Spark) to generate sales breakdowns by product categories across multiple stores.</li> <li>• Assign tasks to calculate total sales values for specified product categories and determine the highest individual sale values for distinct stores.</li> <li>• Create exercises where students use Pig to perform sorting, grouping, joining, projecting, and filtering operations on sample datasets.</li> <li>• Assign exercises where students create, alter, and drop databases, tables, views, functions, and indexes using Hive.</li> </ul>		

### Bibliography

Marz, N., & Warren, J. (2015). Big Data: Principles and Best Practices of Scalable Realtime Data Systems. Manning Publications.

White, T. (2015). Hadoop: The Definitive Guide. O'Reilly Media.

Guller, M. (2015). Big Data Analytics with Spark: A Practitioner's Guide to Using Spark for Large Scale Data Analysis. Apress.

Srinivasan, S. (2018). Big Data Analytics: Methods and Applications. CRC Press.

Gates, A., Thusoo, A., & et al. (2015). Hive: The Definitive Guide. O'Reilly Media.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	<b>Semester III</b>		
<b>325511</b>	<b>BLOCK CHAIN Major (Elective) Theory</b>		<b>3</b>
	<p><b>Course Outcomes: Learners will be able to:</b></p> <ul style="list-style-type: none"> <li>• Understanding of Block Chain Concepts.</li> <li>• Including its decentralized nature, distributed ledger, and cryptographic security features.</li> <li>• Understand the historical context and evolution of blockchain technology, including the development of the first blockchain</li> <li>• Explore the Nakamoto consensus and different consensus algorithms.</li> <li>• Explore the concepts of interoperability and portability in Hyperledger Fabric.</li> <li>• Understand the concept of sharding in blockchain.</li> </ul>		
<b>Module 1</b>	<b>Fundamentals of Blockchain</b>		<b>1</b>
	<p><b>LOs:</b> Learners will be able to</p> <ul style="list-style-type: none"> <li>• Understand the principles of distributed databases, including their architecture, advantages, and challenges. Learn how to design and manage data across multiple nodes in a network.</li> <li>• Explore the complexities of consensus in distributed systems when some nodes may provide conflicting or malicious information. Understand the significance of Byzantine fault tolerance in distributed systems.</li> <li>• Understand the concept of ASIC resistance in the context of crypto-currencies. Explore the motivations and implications of designing systems to resist mining centralization through specialized hardware.</li> <li>• Gain a comprehensive understanding of cryptography principles, including confidentiality, integrity, and authenticity. Learn the applications and functions of hash functions, digital signatures (specifically ECDSA), memory-hard algorithms, and zero</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>• 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.</li> </ul>	

	knowledge proofs.		
<b>Module 2</b>	<b>Blockchain , Distributed Consensus:</b>		1
	<p><b>LOs:</b> Learners will be able to</p> <ul style="list-style-type: none"> <li>• Understand the fundamental concepts of blockchain, including its decentralized nature, distributed ledger, and cryptographic security features.</li> <li>• Gain an understanding of the structure and operation of a blockchain network, including nodes, peers, and the peer-to-peer communication model.</li> <li>• Differentiate between private and public blockchains, understanding their use cases, access control, and levels of decentralization.</li> <li>• Explore the Nakamoto consensus and different consensus algorithms such as Proof of Work, Proof of Stake, and Proof of Burn, understanding their strengths and weaknesses.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>• 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 &amp; Hard Fork, Private and Public blockchain.</li> <li>• Nakamoto consensus, Proof of Work, Proof of Stake, Proof of Burn, Difficulty Level, Sybil Attack, Energy utilization and alternate.</li> </ul>	
<b>Module 3</b>	<b>Cryptocurrency , Cryptocurrency Regulation:</b>		1
	<p><b>LOs:</b> Learners will be able to</p> <ul style="list-style-type: none"> <li>• Understand the historical context and evolution of blockchain technology, including the development of the first blockchain in the context of Bitcoin.</li> <li>• Explore the construction of the Ethereum blockchain, learn about the Decentralized Autonomous Organization (DAO), and understand the concept and implementation of smart contracts.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>• History, Distributed Ledger, Bitcoin protocols - Mining strategy and rewards, Ethereum - Construction, DAO, Smart Contract, GHOST, Vulnerability, Attacks, Sidechain, Namecoin.</li> <li>• 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.</li> </ul>	
<b>Module 4</b>	<b>Hyperledger , Scalability and other challenges :</b>		1
	<p><b>LOs:</b> Learners will be able to</p> <ul style="list-style-type: none"> <li>• Understand Hyperledger as a blockchain protocol, including its reference architecture, design goals, and the modular approach it adopts.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>• Hyperledger as a protocol :The reference architecture Requirements and design goals of Hyperledger Fabric: The modular approach</li> </ul>	

	<ul style="list-style-type: none"> <li>• Explore the features of Hyperledger Fabric, such as its modular architecture, privacy and confidentiality mechanisms, scalability considerations, deterministic transactions, and identity management.</li> <li>• Explore the scalability challenges in blockchain networks and understand how Hyperledger Fabric addresses scalability concerns in the network, consensus, storage, and view planes.</li> <li>• Explore the concepts of interoperability and portability in Hyperledger Fabric, understanding how it facilitates integration with other systems and ensures the portability of applications.</li> <li>• Understand the concept of sharding in blockchain networks, and how it enables the parallel processing of transactions to improve scalability.</li> </ul>	<p>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</p> <ul style="list-style-type: none"> <li>• 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</li> </ul>	
<b>Assignments/ Activities</b>			
	<p>These assignments aim to apply theoretical concepts to practical application and critical thinking.</p> <ul style="list-style-type: none"> <li>• Research and write a comprehensive essay or create a presentation that explains the fundamental concepts of blockchain, highlighting its decentralized nature, the role of a distributed ledger, and the cryptographic security features that contribute to its integrity.</li> <li>• Create a timeline or infographic that visually represents the historical evolution of blockchain technology. Include key milestones, developments, and influential figures in the field.</li> <li>• Write a research paper or prepare a presentation that delves into the concepts of interoperability and portability within Hyperledger Fabric.</li> <li>• Design a workshop or hands-on exercise where participants simulate the sharding process in a blockchain network. Document the findings and insights.</li> </ul>		

## **Bibliography:**

Antonopoulos, Andreas M. (2014). *Mastering Bitcoin: Unlocking Digital Cryptocurrencies*. O'Reilly Media.

Nakamoto, Satoshi. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. [Link to the Bitcoin Whitepaper](#)

Wood, Gavin. (2014). *ETHEREUM: A Secure Decentralized Transaction Ledger*. Yellow paper. [Link to the Ethereum Yellow Paper](#)

Atzei, Nicola; Bartoletti, Massimo; Cimoli, Tiziana. (2017). *A survey of attacks on Ethereum smart contracts*.

Bashir, Imran. (2018). *Mastering Blockchain*. Wiley.



SN	Courses, Modules and Outcomes	Course Contents	Cr
	<b>Semester III</b>		
<b>325512</b>	<b>GIS AND REMOTE SENSING Major(Elective) Theory</b>		<b>4</b>
	<b>Course Outcomes:</b> <b>Learners will be able to:</b> <ul style="list-style-type: none"> <li>Understanding Fundamentals, Technical Skills, Data Acquisition and Management.</li> <li>Spatial Analysis, Mapping and Visualization, Remote Sensing Applications, Integration of Technologies</li> <li>Problem Solving, Communication Skills.</li> </ul>		
<b>Module 1</b>	<b>Fundamentals of GIS</b>		<b>1</b>
	<b>LOs:</b> Learners will be able to <ul style="list-style-type: none"> <li>Define GIS (Geographic Information System), Understand Components of GIS, Comprehend Spatial Data, Understand Characteristics of Spatial Data</li> <li>Analyze Spatial Data Maps, Attribute Data Management - Database Data Model, GIS Applications, Developments in Database for GIS.</li> </ul>	<b>Module Contents:</b> <ul style="list-style-type: none"> <li>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.</li> </ul>	
<b>Module 2</b>	<b>Input-Output and Data Analysis in GIS</b>		<b>1</b>
	<b>LOs:</b> Learners will be able to <ul style="list-style-type: none"> <li>Data Input and Editing, Data Analysis</li> <li>Analytical Modeling in GIS, Output from GIS.</li> </ul>	<b>Module Contents:</b> <ul style="list-style-type: none"> <li>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.</li> </ul>	
<b>Module 3</b>	<b>Issues in GIS:</b>		<b>1</b>
	<b>LOs:</b> Learners will be able to <ul style="list-style-type: none"> <li>Development of Computer Methods, Human and Organizational Issues in GIS, GIS Data Quality and Error Analysis, GIS Project Design</li> </ul>	<b>Module Contents:</b> <ul style="list-style-type: none"> <li>Development of computer methods for spatial data, Issues in GIS- data quality and errors, sources of errors, human and organizational</li> </ul>	

	<p>and Management</p> <ul style="list-style-type: none"> <li>Project Implementation and Evaluation, Understanding the Future of GIS, Internet Resources for GIS, Communication Skills.</li> </ul>	<p>issues, GIS project design and management–problem identification, designing a data model, project management, Implementation, evaluation, the future of GIS, Internet resources of GIS.</p>	
<b>Module 4</b>	Remote Sensing, Global Positioning Systems (GPS)		1
	<p><b>LOs:</b> Learners will be able to</p> <ul style="list-style-type: none"> <li>Principles of Remote Sensing, Remote Sensing System Classification, Imaging Characteristics, Extraction of Information from Images, Integration of Remote Sensing and GIS</li> <li>Introduction to GPS, Accuracy of GPS, Differential GPS, Applications of GPS, Integration of GIS and GPS.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>Principles of remote sensing, remote sensing system-classification, Imaging, characteristics, extraction of information from images–metric and thematic, Integration of RS and GIS.</li> <li>Introduction to GPS, Accuracy of GPS, Differential GPS, Applications of GPS, Integration of GIS and GPS.</li> </ul>	
<b>Assignments/ Activities</b>			
	<p>These assignments aim to apply theoretical concepts to practical application and critical thinking.</p> <ul style="list-style-type: none"> <li>Discuss the importance of metadata in data warehousing, to create a metadata management plan for a given data warehouse, outlining how metadata will be collected, stored, and utilized.</li> <li>From a dataset extract relevant information, transform it according to a predefined business rule, and load it into a data warehouse.</li> <li>Create a set of business queries related to a hypothetical business problem</li> <li>Find a dataset suitable for clustering analysis. then use clustering algorithms to identify natural groupings within the data and interpret the results.</li> <li>web mining project (perform web scraping, and apply web mining techniques)</li> </ul>		

**Bibliography:**

Heywood, I., Cornelius, S., & Carver, S. (2000). *An Introduction to Geographical Information Systems*. Pearson Education Asia.

Lo, C. P., & Yeung, A. (n.d.). *Concepts and Techniques of Geographic Information Systems*. PHI, New Delhi.

Demers, M. N. (n.d.). *Fundamentals of Geographic Information Systems, 2nd Edition*. John Wiley & Sons (Asia) Pte Ltd.

Razvi, M. (2002). *ArcGIS Developer's Guide for Visual Basic Applications*. Onword Press.

SN	Courses, Modules and Outcomes	Course Contents	Cr	
	<b>Semester III</b>			
325513	<b>Software Testing Major (Elective) Theory</b>		<b>4</b>	
	<p><b>Course Outcomes:</b> <b>Learners will be able to:</b></p> <ul style="list-style-type: none"> <li>• Understand the fundamental principles and concepts of software testing, including its purpose, objectives, and the role it plays in software development.</li> <li>• Proficient in various test design techniques, including equivalence partitioning, boundary value analysis, decision tables, and state transition testing, enabling them to create effective test cases.</li> <li>• Apply various testing methods such as functional, non-functional, unit, integration, system, regression, and acceptance testing to diverse software systems.</li> <li>• Understand quality assurance principles and best practices, emphasizing the importance of testing in the software development lifecycle.</li> <li>• Understand the ethical and professional responsibilities associated with software testing.</li> </ul>			
<b>Module 1</b>	<b>Overview of Software Testing</b>		<b>1</b>	
	<p><b>LOs:</b> Learners will be able to</p> <ul style="list-style-type: none"> <li>• Gain proficiency in software testing terminologies, methodologies, and life cycles.</li> <li>• Comprehend the economic aspects of testing and its impact on organizational structures.</li> <li>• Develop skills in creating policies, test strategies, and risk management to ensure meeting customer needs.</li> <li>• Analyze the advantages of structured testing processes and their cost implications.</li> <li>• Demonstrate proficiency in the seven-step software testing process.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>• 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</li> <li>• Overview of Software Testing Process Advantages of Following a Process, The Cost of Computer Testing, The Seven-Step Software Testing Process</li> <li>• Verification and Validation Verification and Validation (V&amp;V) Activities, Verification, Verification of Requirements, Verification of High –level Design, Verification of Low – level Design, How to Verify Code? ,Validation</li> </ul>		

		<ul style="list-style-type: none"> <li>• Static Testing Inspections, Structured Walkthroughs, Technical Reviews.</li> </ul>	
<b>Module 2</b>	<b>Validation and Regression Testing</b>		1
	<p><b>LOs:</b> Learners will be able to</p> <ul style="list-style-type: none"> <li>• Execute various validation activities including unit, integration, function, system, and acceptance testing to assess different aspects of software functionality and ensure its compliance with requirements.</li> <li>• Differentiate between progressive and regressive testing, comprehending the significance of regression testing in maintaining software quality.</li> <li>• Apply regression testing techniques to identify potential issues arising from software changes or updates, thereby ensuring the stability and reliability of the software product.</li> <li>• Identify the objectives of regression testing and determine appropriate instances for conducting regression tests in the software development life cycle.</li> <li>• Define regression test problems, select suitable types of regression testing, and utilize effective regression testing strategies to address software changes and minimize the risk of introducing new defects into the system.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>• Validation Activities Unit Validation Testing, Integration Testing, Function Testing, System Testing , Acceptance Testing</li> <li>• 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.</li> </ul>	
<b>Module 3</b>	<b>Testing Management and Metrics</b>		1
	<p><b>LOs:</b></p> <ul style="list-style-type: none"> <li>• Establish a comprehensive understanding of test management structures, including the organization and composition of testing groups, enabling effective test planning and detailed</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>• Test Management Test Organization, Structure of Testing Group, Test Planning, Detailed Test Design and Test Specifications</li> <li>• Software Metrics Need for Software Management, Definition of Software Metrics, Classification</li> </ul>	

	<p>test design.</p> <ul style="list-style-type: none"> <li>Recognize the need for software metrics and demonstrate the ability to define, classify, and apply various metrics within the software development life cycle.</li> <li>Evaluate entities to be measured within software projects, particularly focusing on size metrics and their implications on software management.</li> <li>Formulate measurement objectives specific to testing, identifying attributes and corresponding metrics relevant to monitoring and controlling the testing process.</li> </ul>	<p>of Software Metrics, Entities to be Measured, Size Metrics</p> <ul style="list-style-type: none"> <li>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.</li> </ul>	
<b>Module 4</b>	<b>Automation Testing Tool</b>		1
	<p><b>LOs:</b> Learners will be able to</p> <ul style="list-style-type: none"> <li>Evaluate the necessity and significance of test process maturity, demonstrating the ability to measure, assess, and improve test processes within an organization using established maturity models.</li> <li>Identify the rationale behind automation in testing, categorize various testing tools, and apply criteria for selecting appropriate tools while considering associated costs.</li> <li>Analyze guidelines for automated testing and gain an overview of commercial testing tools, fostering the skills required for implementing automated testing effectively.</li> <li>Apply agile methodologies to enhance software testing, recognizing the importance of agility, overcoming inhibitors, and implementing solutions to improve testing processes within an agile framework.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>Testing Process Maturity Models Need for Test Process Maturity, Measurement and Improvement of a Test Process, Test Process Maturity Models</li> <li>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</li> <li>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.</li> </ul>	

<b>Assignments/ Activities towards CCE</b>	
	<ul style="list-style-type: none"> <li>• Provide a case study of an organization's testing process and have students analyze and propose improvements based on maturity model principles.</li> <li>• Organize a session where students demonstrate how a specific testing tool works, highlighting its features, benefits, and practical applications.</li> <li>• Conduct a simulation where students participate in an agile testing environment, taking on roles (developer, tester, product owner) to experience iterative development, testing, and feedback cycles.</li> <li>• Provide a software scenario and ask students to create a comprehensive test plan, including test objectives, strategies, resource allocation, and metrics for evaluating test progress.</li> <li>• Divide students into teams, assigning each team a different validation testing type (e.g., unit, integration). They role-play scenarios to validate a software component or system.</li> </ul>

### **Bibliography:**

Chauhan, Naresh. *Software Testing Principles and Practices*. Oxford University Press.

Perry, William E. *Effective Methods of Software Testing* (3rd Edition). Wiley, India.

Desikan, Srinivasan and Ramesh, Gopaldaswamy. *Software Testing Principles and Practices*. Pearson Education.

Patton, Ron. *Software Testing* (2nd Edition). Pearson Education.

Dustin, Elfriede. *Effective Software Testing: 50 Specific Ways to Improve Your Testing*. Pearson Education.

SN	Courses, Modules and Outcomes	Course Contents	Cr
	<b>Semester III</b>		
<b>325514</b>	<b>Robotic Process Automation Major(Elective) Theory</b>		<b>4</b>
	<p><b>Course Outcomes:</b>  <b>Learners will be able to:</b></p> <ul style="list-style-type: none"> <li>Define and explain the fundamental concepts and principles of Robotic Process Automation.</li> <li>Gain proficiency in using popular RPA tools such as UiPath, Automation Anywhere, or Blue Prism.</li> <li>Develop RPA bots to automate specific tasks and processes.</li> <li>Diagnose and troubleshoot common issues encountered during RPA implementation.</li> </ul>		
<b>Module 1</b>	<b>Robotic Process Automation Foundations, UiPath, Automation Anywhere</b>		<b>1</b>
	<p><b>LOs:</b> Learners will be able to</p> <ul style="list-style-type: none"> <li>Understanding RPA Fundamentals, RPA Capabilities and Components, Benefits and Downsides of RPA, Comparison with Other Business Technologies.</li> <li>Comparison with Other Business Technologies, Automation Anywhere Overview.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>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</li> <li>What is UiPath, UiPath Studio, UiPath Robot, UiPath Orchestrator, UiPath – an integrated view</li> <li>What is Automation Anywhere, Enterprise Control Room, IQ Bot.</li> </ul>	
<b>Module 2</b>	<b>Downloading and Installing UiPath Studio and Data Manipulation</b>		<b>1</b>
	<p><b>LOs:</b>Learners will be able to</p> <ul style="list-style-type: none"> <li>UiPath Studio and Task Recorder Proficiency, Workflow Sequencing and Control Flow.</li> <li>Variables, Scope, and Collections, Arguments, Clipboard Management.</li> <li>Data Table Usage, File Operations, CSV/Excel Handling.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>Learning UiPath Studio, Task Recorder, Step by step examples using the recorder</li> <li>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.</li> <li>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</li> </ul>	

		table and vice versa.	
<b>Module 3</b>	<b>Taking Control of the Controls, Exception Handling and Debugging</b>		1
	<p><b>LOs: Learners will be able to</b></p> <ul style="list-style-type: none"> <li>• UI Automation Basics, Advanced Automation Plugins</li> <li>• Assistant Bots and Triggers.</li> <li>• Error Handling and Debugging</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>• 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</li> <li>• 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</li> <li>• What are assistant bots, Monitoring system event triggers, Monitoring image and element triggers, Launching an assistant bot on a keyboard event.</li> <li>• Exception handling, Common exceptions and ways to handle them, Logging and taking screenshots, Debugging techniques, Collecting crash dumps, Error reporting</li> </ul>	
<b>Module 4</b>	<b>Managing and Maintaining the Code.</b>		1
	<p><b>LOs: Learners will be able to</b></p> <ul style="list-style-type: none"> <li>• Updates, Project Organization, Reusability of Workflows, State Machine.</li> <li>• Using Config Files and Examples of a Config File, Using Orchestration Server to Control Bots, Publishing and Managing.</li> </ul>	<p><b>Module Contents:</b></p> <ul style="list-style-type: none"> <li>• 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</li> <li>• Publishing using publish utility, Overview of Orchestration Server, Using Orchestration Server to Control bots, Using</li> </ul>	



		Orchestration server to deploy bots, License management, Publishing and managing	
<b>Assignments/ Activities</b>			
	<p>These assignments aim to apply theoretical concepts to practical application and critical thinking.</p> <ul style="list-style-type: none"> <li>• Discuss the importance of metadata in data warehousing, to create a metadata management plan for a given data warehouse, outlining how metadata will be collected, stored, and utilized.</li> <li>• From a dataset extract relevant information, transform it according to a predefined business rule, and load it into a data warehouse.</li> <li>• Create a set of business queries related to a hypothetical business problem</li> <li>• Find a dataset suitable for clustering analysis. then use clustering algorithms to identify natural groupings within the data and interpret the results.</li> <li>• web mining project (perform web scraping, and apply web mining techniques)</li> </ul>		

**Bibliography:**

Tripathi, Alok Mani. "Learning Robotic Process Automation." Packt Publishing, 2018.

Taulli, Tom. "The Robotic Process Automation Handbook – A Guide to Implementing RPA Systems." Apress, 2020.

Sireci, Jonathan. "The Practitioner’s Guide to RPA." Farchair Solutions, 2020.  
 Bornet, Pascal; Barkin, Ian; Wirtz, Jochen. "Intelligent Automation – Welcome to the World of HYPERAUTOMATION." 2021.