E.G.S. PILLAY ENGINEERING COLLEGE

(Autonomous)

NAGAPATTINAM – 611002

(Affiliated to Anna University, Chennai | Accredited by NAAC with 'A++' Grade|Accredited by NBA T1(B.E. – CSE, CIVIL, ECE, EEE, MECH& B.Tech – IT) | Approved by AICTE, New Delhi)

B.TECH.- ARTIFICIAL INTELLIGENCE AND DATA SCIENCE R – 2023

THIRD YEAR

CURRICULUM AND SYLLABUS FOR FIFTH SEMESTER

V Semester									
COURSE CODE	**COURSE NAME**	**CATEGORY**	**L**	**T**	**P**	**C**	**MAX. MARKS**		
							CA	**ES**	**TOTAL**
Theory Courses									
2302AS501	Operating System Concepts	PCC	3	0	0	3	40	60	100
2302AS502	Modern Computer Architecture	PCC	3	0	0	3	40	60	100
2302AS503	Big Data Frameworks	PCC	3	0	0	3	40	60	100
2302AS504	Deep Learning Techniques	PCC	3	0	2	4	50	50	100
	Elective – I	PCC	2	0	2	3	50	50	100
	Open Elective - I	OEC	3	0	0	3	40	60	100
Laboratory Courses									
2302AS551	Programming with Large Datasets Laboratory	PCC	0	0	2	1	60	40	100
2304AS551	Mini Project – I (Business Case Study)	EEC	0	0	2	1	100	-	100
Other Courses									
2304GE501	Professional Development Course – III	EEC	0	0	2	1	100	-	100
2301LS501	Life Skills – V	-	0	0	0	0	100	-	100
TOTAL			**17**	**0**	**10**	**22**	**620**	**380**	**1000**

2302AS501 OPERATING SYSTEM CONCEPTS

L	T	P	C
3	0	0	3

PREREQUISITE:

1. Understand the role of an operating system in managing hardware and software resources, including processes, threads, and system calls.
2. Explore CPU scheduling, inter-process communication, synchronization, deadlock handling, and memory management techniques like paging and segmentation.
3. Understand file and directory structures, disk storage organization, and file allocation methods.
4. Analyze security concerns, resource protection strategies, and architectural complexities in modern operating systems using real-world case studies.

COURSE OBJECTIVES:

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

COURSE OUTCOMES:

CO1: Understand and build applications using inter-process and inter-thread data sharing and communication mechanisms

CO2: Analyze memory management concepts and evaluate the importance of memory safety, space separation, and protection techniques

CO3: Interpret file and directory structures, analyse disk allocation and scheduling algorithms, and recommend strategies to optimize storage utilization and file access

CO4: Evaluate data security concerns and implement mechanisms for ensuring data safety, including access control techniques

CO5: Assess the features of contemporary operating systems through case studies

COs Vs POs MAPPING:

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

COURSE CONTENTS:

Module-I Introduction **9 Hours**

Operating System structure, Process management, Memory management, Storage management, Distributed systems, System calls, OS design and implementation, Virtual machines

Module-II Process Management **9 Hours**

Process concept – Scheduling, Inter process communication, Multithreading models, Process scheduling, Scheduling algorithms, Thread scheduling

Module-III Process Coordination **9 Hours**

Synchronization, Critical section problem, Semaphores, Atomic transactions, Deadlocks – Model, Characterization, Handling deadlocks, Prevention, Avoidance, Detection and recovery

Module-IV Memory Management **9 Hours**

Contiguous memory allocation, Paging, Segmentation, Virtual memory management

Module-V Storage Management **9 Hours**

File concept, Access methods, Structure, Implementation, Directory implementation, Free space management, Secondary storage structure – Disk structure, Scheduling, Management, Swap-space management, RAID structure

TOTAL: 45 HOURS

REFERENCES:

1. A. Silberschatz and P.B. Galvin, "Operating System Concepts", 8th Ed., Wiley, 2008.
2. W. Stalling, "Operating Systems: Internals and Design Principles", 6th Ed., Pearson Education, 2008
3. A.S. Tanenbaum, "Modern Operating System", 3rd Ed., Pearson, 2007.
4. D.M. Dhamdhere," Operating Systems-A Concept Based Approach", McGraw-Hill, 2008.
5. EktaWalia, "Operating System Concepts", Khanna Book Publishing, 2019.
6. https://nptel.ac.in/courses/106105214
7. https://onlinecourses.nptel.ac.in/noc20_cs04/preview

2302AS502 MODERN COMPUTER ARCHITECTURE

L	T	P	C
3	0	0	3

PREREQUISITE:

Nil

COURSE OBJECTIVES:

1.Explain the basic principles and components of computer systems

2.Apply various logic design techniques to solve digital circuit design problem

3.Utilize high-performance computing architectures to develop efficient computing solutions for complex tasks

COURSE OUTCOMES:

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

CO1: Understand the organization of the Control unit, Arithmetic and Logical unit, Memory unit and the I/O unit.

CO2: Analyse different computer architectures and their applications

CO3: Categorize distributed computing concepts to design and analysesynchronizationand communication in parallel systems

CO4: Understand HPC architecture and distinguish between data and task parallelism

CO5: Apply CUDA to develop parallel programs and manage GPU-TPU data exchange

COs Vs POs MAPPING:

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

COURSE CONTENTS:

Module-I Introduction **9 Hours**

Designing combinational and sequential logic, Computers registers and instructions, Timing, and control, Instructions cycle, Memory reference instruction, I/O interruption, Adder and Subtractor circuits, Booth Multiplication Algorithm

Module-II Multi-core Architecture **9 Hours**

Memory technologies, Hierarchical memory systems, Locality principle and caching, Direct-mapped caches, Associative caches, DRAM – organisation, access techniques, scheduling algorithms and signal systems. Tiled Chip Multicore Processors (TCMP), Network on Chips (NoC), NoC router – architecture, design, routing algorithms and flow control techniques

Module-III Distributed Computing Systems and Concurrency 9 Hours

Relation to Parallel Multiprocessors/multicomputer Systems, Distributed and Concurrent Programs, Message Passing vs. Shared Memory Systems, Synchronous vs. Asynchronous Executions, Design Issues and Challenges, Distributed Computing Technologies, Clocks and Synchronization, Coordination and Agreement Algorithms, Global State and Distributed Transactions

Module-IV High Performance Computing (HPC) 9 Hours

HPC Architecture, Parallel Processing, Parallel Memory Models, Data vs. Task Parallelism, High Throughput Computing, Vectorization, Multithreading

Module-V High Performance Computing with CUDA 9 Hours

CUDA programming model, Basic principles of CUDA programming, Concepts of threads and blocks, GPU,TPU and NPU data exchange

TOTAL: 45 HOURS

REFERENCES:

1. M. Morris Mano, "Computer System & Architecture", Prentice Hall of India, 2002.
2. John L. Hennessy and David A Patterson, "Computer Architecture-A quantitative approach", Morgan Kaufmann/ Elsevier, 4th Edition, 2007.
3. J. P. Hayes, "Computer Architecture and Organization", McGraw-Hill Companies, 1998
4. David Culler and J.P. Singh, "Parallel Computer Architecture: A Hardware/Software Approach" 1998.
5. https://onlinecourses.nptel.ac.in/noc20_cs41/preview
6. https://www.coursera.org/learn/introduction-high-performance-computing

2302AS503 **BIG DATA FRAMEWORKS** **L T P C**

3 0 0 3

PREREQUISITE:

2302AS302 - Database Systems

COURSE OBJECTIVES:

1. Introduce the importance of big data and role of Hadoop framework in analyzing large datasets.
2. Familiarizeinwriting queries using Pig and Hive to process big data
3. Discuss various big data frameworks and applications using Spark and Scala.
4. Understand the concept and writing applications using SparkSQL
5. Provide the concepts of NoSQL databases and study the working mechanisms of MongoDB

COURSE OUTCOMES:

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

CO1: Describe Hadoop Distributed File System and apply MapReduce architecture to solve real world problems.

CO2: Build scripts using Pig over large datasets and query using Hive.

CO3: Explore Apache Spark's architecture, APIs, toolset, machine learning, and advanced analytics.

CO4: Apply Resilient Distributed Datasets (RDD) for creating applications in Spark and query using SparkSQL.

CO5: Analyze NoSQL databases and develop data models using MongoDB

COs Vs POs MAPPING:

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

COURSE CONTENTS:

Module-I **Big Data Concepts** **9 Hours**

Big Data, Importance of Big Data, Big Data use cases. **The Hadoop Distributed Files system**: The Design of HDFS, HDFS Concepts, HDFS Federation, HDFS High Availability, Basic File system Operations, Hadoop File systems, Anatomy of a File Read, Anatomy of a File Write. **Map Reduce**: What is Map reduce, Architecture of map reduce How **Map Reduce Works**: Anatomy of a Map Reduce Job Run, Failures in Map Reduce. **Map Reduce Types and Formats**: Map Reduce Types, The Default Map Reduce Job, Input Formats, Input Splits and Records, Text Input, Output Formats, Text Output, Developing a Map Reduce Application

Module-II Pig and Hive **9 Hours**

Pig: Installing and Running Pig, Generating Examples, Comparison with Databases, Pig Latin, UserDefined Functions, Data Processing Operators, Pig in Practice. **Hive**: Installing Hive, The Hive Shell, An Example, Running Hive, Comparison with Traditional Databases, Hive QL, Tables, Querying Data, User-Defined Functions, Writing a User Defined Functions, Writing a User Defined Aggregate Function

Module-III **8 Hours**

Introduction to Spark: What is Apache Spark, History of Spark, Present and Future of Spark, Running Spark, Spark's Basic Architecture, Spark Applications, Spark's Language APIs, Spark's APIs, Starting Spark, The Spark Session Data Frames, Partitions, Transformations, Lazy Evaluation, Actions, Spark UI, An End-to-End Example, Data Frames and SQL. **Spark's Toolset:** Running Production Applications, Datasets: Type-Safe Structured APIs, Structured Streaming, Machine Learning and Advanced Analytics, Lower-Level APIs, Spark R, Spark's Ecosystem and Packages

Module-IV INDUCTION MOTOR DRIVE **12 Hours**

Spark SQL: SQL, Big Data and SQL: Apache Hive, Spark SQL, Spark's Relationship to Hive, How to Run Spark SQL Queries, Catalog, Tables, Views, Databases, Select Statements, Datasets: When to Use Datasets, Creating Datasets, Actions, Transformations **Resilient Distributed Datasets**: Introduction to RDDs, Creating RDDs, Manipulating RDDs, Transformations, Actions, Saving Files, Caching, Check pointing, Pipe RDDs to System Commands

Module-V **8 Hours**

No SQL Databases: Review of traditional Databases, Need for NoSQL Databases, Columnar Databases, Failover and reliability principles, CAP Theorem, Differences between SQL and NoSQL databases, **Working mechanisms of Mongo DB:**Overview, Data Modelling, Create and Drop Database, Create and collection, Data types, Insert, Query, Update and Delete operations, Limiting and Sorting records, Indexing, Aggregation. Usecases: Aadhar database, Google BigTable, Amazon Dynamo

TOTAL: 45 HOURS

REFERENCES:

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

2. Bill Chambers, MateiZaharia, “Spark: The Definitive Guide”, 4th Edition, O'Reilly Media Inc, 2018
3. P. J. Sadalage, M. Fowler, "NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence", Addison-Wesley Professional, 2012
4. Michael Minelli, Michelle Chambers, and AmbigaDhiraj, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses", Wiley, 2013
5. V.K. Jain, Data Science and Analytics (with Python, R and SPSS Programming), Khanna Book Publishing Company
6. V.K. Jain, Big Data and Hadoop, Khanna Book Publishing Company 2020
7. Alan Gates, "Programming Pig", O'Reilly Media Inc, 2011
8. Edward Capriolo, Dean Wampler, and Jason Rutherglen, "Programming Hive", O'Reilly Media Inc, October 2012
9. http://www.planetcassandra.org/what-is-nosql
10. https://class.coursera.org/datasci-001/lecture
11. https://onlinecourses.nptel.ac.in/noc20_cs92/preview

2302AS504 DEEP LEARNING TECHNIQUES

L	T	P	C
3	0	2	4

PREREQUISITE:

2302AS404 - Machine Learning Techniques

COURSE OBJECTIVES:

1. Introduce the foundational concepts of deep learning, emphasizing its architectures, optimization methods, and ongoing research activities
2. Develop practical skills in implementing and testing learning algorithms using TensorFlow, focusing on CNNs and RNNs for image and sequential data tasks
3. Encourage the application of deep learning techniques to solve domain-specific challenges and design innovative solutions for real-world problems

COURSE OUTCOMES:

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

CO1: Comprehend the fundamentals of deep learning and the main research activities in this field

CO2: Interrelate architectures and optimization methods for deep neural network training

CO3: Implement, apply and test relevant learning algorithms in TensorFlow

CO4: Explore practical applications of CNNs and RNNs in diverse fields, assessing their effectiveness in handling image and sequential data tasks.

CO5: Apply deep learning techniques to solve domain-specific problems and create innovative applications across various industries

COs Vs POs MAPPING:

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

COURSE CONTENTS:

Module-I Introduction to Deep learning **9 Hours**

Introduction History of Deep Learning, McCulloch Pitts Neuron, Multilayer Perceptron's (MLPs), Representation Power of MLPs, Sigmoid Neurons, Feed Forward Neural Networks, Back propagation

Module-II Activation functions and parameters 9Hours

Gradient Descent (GD), Momentum Based GD, Nesterov Accelerated GD, Stochastic GD, Principal Component Analysis and its interpretations, Singular Value Decomposition, Parameters v/s Hyper-parameters

Module-III Auto-encoders & Regularization 9 Hours

Auto encoders and relation to PCA, Regularization in auto encoders, Denoising auto encoders, Sparse auto encoders, Regularization: Bias Variance Tradeoff, L2 regularization, Early stopping, Dataset augmentation, Encoder Decoder Models, Attention Mechanism, Attention over images, Batch Normalization

Module-IV Deep Learning Models 9 Hours

Introduction to CNNs, Architecture, Convolution/pooling layers, CNN Applications,AlexNet,VGGNet, GoogLeNet, ResNet. Introduction to RNNs, Back propagation through time (BPTT), Vanishing and Exploding Gradients, Truncated BPTT, GRU.

Module-V Deep Learning Applications 9 Hours

Image Processing Natural Language Understanding using LLMs, Video Analytics and Smart Monitoring Systems, Agentic AI for Autonomous Decision-Making and Task Execution.

TOTAL: 45 HOURS

Module-II

LIST OF EXPERIMENTS:

1)Implement a simple neural network with one hidden layer to classify a given dataset.	3 Hours
2)Implementbackpropagation in a two-layer network and update weights.	3 Hours
3) Implement momentum-based optimization and observe the training speed.	3 Hours
4)Apply Stochastic Gradient Descent to minimize loss in a neural network.	3Hours
5)Build a CNN from scratch for image classification using a given dataset. Explain how each convolutional layer extracts features.	4 Hours
6)Implement the VGG16 pretrained model for image classification on a given dataset	4 Hours
7)Implement an RNN to generate a sequence (e.g., text or music). Demonstrate how the model retains memory of past inputs.	3Hours
8)Mini Project: Implement an LSTM model to predict stock prices using historical data. Include data preprocessing and evaluate prediction accuracy.	6Hours

TOTAL: 15 HOURS

REFERENCES:

1. Jason Gregory, A K Peters, "Game Engine Architecture", 3rd Edition, 2019
2. Georgios N. Yannakakis and Julian Togelius, "Artificial Intelligence and Games", January 26, 2018, Springer
3. Ian Goodfellow, YoshuaBengio, Aaron Courville. "Deep Learning", MIT press, 2016
4. .https://www.google.com/url?q=https://nptel.ac.in/courses/106105215&sa=D&source=editors &ust=1737711220976334&usg=AOvVaw1Sr46B5IEAm-ojT3nHT-YG
5. https://www.coursera.org/specializations/deep-learning
6. https://onlinecourses.nptel.ac.in/noc21_cs76/preview

2303AS002 **ROBOTIC PROCESS AUTOMATION (PROFESSIONAL ELECTIVE – I)**

L	T	P	C
2	0	2	3

PREREQUISITE:

NIL

COURSE OBJECTIVES:

1.This syllabus is designed to provide a comprehensive learning path for developers who want to build and deploy robotic process automation (RPA) solutions using UiPath Studio. It covers the essential topics required for both the Automation Developer Associate and Automation Developer Professional certifications.
2.The syllabus is divided into five units, each focusing on a specific area of UiPath Studio development

COURSE OUTCOMES:

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

- **CO1:** **Explain** the fundamental principles of Robotic Process Automation (RPA) and **describe** its diverse applications across various industries.
- **CO2:** **Apply** industry-standard RPA tools to **automate** specified business processes with demonstrable efficiency.
- **CO3:** **Create** well-structured and efficient RPA workflows, incorporating best practices for robustness and maintainability.
- **CO4:** **Execute** the implementation and deployment of RPA solutions, adhering to security protocols and compliance requirements.
- **CO5:** **Investigate** advanced RPA concepts, including cognitive RPA and AI-driven automation techniques, and **discuss** their potential impact

COs Vs POs MAPPING:

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	3	2	1	2	2	-	-	-	-	-	-	1	1	-	3
CO2	3	3	3	2	3	-	-	-	-	-	-	2	2	-	3
CO3	3	3	3	2	3	-	-	-	-	-	2	2	2	-	3
CO4	3	3	3	2	3	3	-	-	-	-	2	2	1	-	3
CO5	3	3	2	2	2	2	-	-	-	2	2	2	2	-	3

COURSE CONTENTS:

Module-I **Introduction to RPA** **6 Hours**

Definition and scope of RPA, Benefits and challenges of RPA, RPA vs. Automation vs. Artificial Intelligence, RPA architecture and components, Process discovery and analysis, RPA use cases in various industries

Module-II RPA Tools and Technologies 6 Hours

Overview of popular RPA tools (UiPath, Automation Anywhere, Blue Prism), Tool selection criteria, Installation and configuration of RPA tools, Basic components of RPA tools (bots, workflows, activities), Recording and playback of processes, Object identification and selectors

Module-III RPA Development and Design 6 Hours

RPA development lifecycle, Process mapping and flowcharting, Designing RPA workflows, Handling exceptions and errors, Control flow statements (conditions, loops), Data manipulation and transformation, Integration with other systems (databases, APIs), Best practices for RPA development

Module-IV RPA Implementation and Deployment 6 Hours

RPA deployment strategies (attended, unattended, hybrid), Testing and validation of RPA processes, RPA bot maintenance and monitoring, Scaling RPA solutions, Security and compliance considerations, Return on investment (ROI) analysis

Module-V Advanced RPA Topics 6 Hours

Artificial intelligence and machine learning in RPA, Cognitive RPA, RPA and process mining, RPA and business process management, Future trends in RPA

TOTAL: 30 HOURS

LIST OF EXPERIMENTS:

1.**Understanding RPA Basics and Use Cases:** 3 Hours

- Research and document three different real-world RPA use cases from different industries (e.g., finance, healthcare, manufacturing). For each use case, describe:
  - The manual process being automated.
  - The benefits of implementing RPA.
  - Potential challenges in automating the process.
- Create a simple UiPath sequence using basic activities (like Write Line, Message Box) to simulate a very basic automated task, illustrating the concept of a bot performing steps. For example, a bot that greets the user with a message box

2.**UiPath Studio Interface and Basic Activities:** 3 Hours

- Explore the different panels in UiPath Studio (Project, Activities, Properties, Outline, Locals).
- Create a UiPath sequence to:
  - Take user input using the Input Dialog activity.

- Display a personalized message using the Message Box activity, incorporating the user's input.
- Write the user's input and the personalized message to the Output panel using the Write Line activity.

3.**Recording and Playback of a Simple Process:** 3 Hours

- Use the Basic Recording feature to automate the steps of opening a simple application (like Notepad), typing some text, and saving the file with a specific name.
- Run the recorded sequence and observe the bot's actions.
- Analyze the generated activities and selectors.

4.**Object Identification and Selectors:** 3 Hours

- Automate interacting with a simple webpage (e.g., a login page with username and password fields and a button) using the Web Recording feature.
- Inspect the selectors generated for the different UI elements.
- Modify one of the selectors (e.g., by changing an attribute) and observe how it affects the bot's ability to identify the element. Correct the selector to make the automation work again.

5.**Designing RPA Workflows with Control Flow:** 3 Hours

- Create a UiPath flowchart or sequence that asks the user for a number using Input Dialog.
- Use an If activity to check if the number is even or odd and display a corresponding message box.
- Use a While loop to repeatedly ask the user for input until they enter a specific keyword (e.g., "exit").

6.**Handling Exceptions and Errors:** 3 Hours

- Create a workflow that attempts to open a file at a specific path using the Open Application activity.
- Use a Try-Catch block to handle the scenario where the file does not exist. In the Catch block, display a user-friendly error message using Message Box and log the error using Log Message.
- Simulate different error scenarios (e.g., incorrect file path) to test the exception handling.

7.**Data Manipulation and Transformation:** 3 Hours

- Read data from an Excel file (with at least two columns, e.g., Name and Age) using the Read Range activity.
- Use activities like For Each Row and Assign to:
  - Create a new DataTable or a new column in the existing DataTable.
  - Perform a simple data transformation (e.g., concatenate "Hello" with each name).
  - Filter the data based on a condition (e.g., only include rows where Age is greater than 25).
- Write the transformed data to a new Excel file using the Write Range

activity.

8.**Integration with Databases:** 3 Hours

- Connect to a sample database (e.g., a local SQLite database or a simple Access database) using the Database activities (you might need to install the relevant database connector package).
- Use the Connect activity to establish a connection.
- Execute a Run Query activity to retrieve data from a table.
- Use a For Each Row activity to process the retrieved data (e.g., display each record using Write Line).
- Use an Insert activity to add a new record to the table.
- Use the Disconnect activity to close the database connection

9.**Attended vs. Unattended Automation:** 3 Hours

- Attended Automation: Create a simple automation that requires user interaction (e.g., a bot that guides the user through filling out a form by displaying instructions in message boxes and waiting for user input). Run this bot from UiPath Assistant.
- Unattended Automation (Simulation): Create an automation that runs without user intervention (e.g., a bot that reads data from a predefined Excel file and writes it to another file). While you might not have a full Orchestrator setup in a basic lab, discuss how this bot could be scheduled and run unattended via Orchestrator. Document the steps involved in deploying and scheduling such a bot.

10. **Exploring Cognitive RPA Concepts (Basic)::** 3 Hours

- Use a UiPath activity that incorporates basic cognitive capabilities (you might need to install relevant packages like UiPath.IntelligentOCR.Activities or explore activities related to Document Understanding if available in your UiPath version).
- For example:
  - Use the Read PDF Text activity to extract text from a PDF document.
  - Explore basic OCR capabilities to extract text from an image within a document.
  - Discuss how more advanced cognitive services (like AI Computer Vision, Language Processing APIs) could be integrated with UiPath to handle more complex unstructured data or decision-making processes. This part might involve research and a conceptual workflow design rather than full implementation depending on the complexity and available resources

TOTAL: 30 HOURS

REFERENCES:

1. Richard Murphy, "Learning RPA: Robotic Process Automation".
2. William Murphy, "RPA: The Robotic Process Automation Handbook".
3. Adir Moshe, “UiPathStudioX Robotic Process Automation: A Beginner's Guide”.
4. AdeelJaved, AnumSundrani, Naida Malik and Sidney Madison Prescott, “Robotic Process Automation with UiPath Studio: Design, Build, and Deploy Automated Solutions”, Apress,

2021.

5. Adir Sharon, “Learning UiPath Robotic Process Automation - Building Software Robots for Business Automation”.
6. Mohammed Al-Buraiki, “UiPathStudioX Robotics Process Automation: A Beginner's Guide to Building Software Robots”.
7. Bernd Hinz, "RPA Implementation: A Guide to Successful Robotic Process Automation".
8. Ramya Devi Vaddi , "UiPath Advanced Development with Computer Vision and Machine Learning".
9. https://www.uipath.com/rpa/academy
10. Richard Murphy, "Learning RPA: Robotic Process Automation".

2303AS051 **ROBOTIC PROCESS AUTOMATION (OPEN ELECTIVE - I)**

L	T	P	C
3	0	0	3

PREREQUISITE:

NIL

COURSE OBJECTIVES:

This syllabus is designed to provide a comprehensive learning path for developers who want to build and deploy robotic process automation (RPA) solutions using UiPath Studio. It covers the essential topics required for both the Automation Developer Associate and Automation Developer Professional certifications.

The syllabus is divided into five units, each focusing on a specific area of UiPath Studio development.

COURSE OUTCOMES:

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

CO1: Explain the fundamental principles of Robotic Process Automation (RPA) and describe its diverse applications across various industries

CO2: Apply industry-standard RPA tools to automate specified business processes with demonstrable efficiency

CO3: Create well-structured and efficient RPA workflows, incorporating best practices for robustness and maintainability

CO4: Execute the implementation and deployment of RPA solutions, adhering to security protocols and compliance requirements

CO5: Investigate advanced RPA concepts, including cognitive RPA and AI-driven automation techniques, and discuss their potential impact

COs Vs POs MAPPING:

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

COURSE CONTENTS:

Module-I **Introduction to RPA** **9 Hours**

Definition and scope of RPA, Benefits and challenges of RPA, RPA vs. Automation vs. Artificial Intelligence, RPA architecture and components, Process discovery and analysis, RPA use cases in various industries

Module-II RPA Tools and Technologies 9 Hours

Overview of popular RPA tools (UiPath, Automation Anywhere, Blue Prism), Tool selection criteria, Installation and configuration of RPA tools, Basic components of RPA tools (bots, workflows, activities), Recording and playback of processes, Object identification and selectors

Module-III RPA Development and Design 9 Hours

RPA development lifecycle, Process mapping and flowcharting, Designing RPA workflows, Handling exceptions and errors, Control flow statements (conditions, loops), Data manipulation and transformation, Integration with other systems (databases, APIs), Best practices for RPA development

Module-IV RPA Implementation and Deployment 9 Hours

RPA deployment strategies (attended, unattended, hybrid), Testing and validation of RPA processes, RPA bot maintenance and monitoring, Scaling RPA solutions, Security and compliance considerations, Return on investment (ROI) analysis

Module-V Advanced RPA Topics 9 Hours

Artificial intelligence and machine learning in RPA, Cognitive RPA, RPA and process mining, RPA and business process management, Future trends in RPA

TOTAL: 45 HOURS

REFERENCES:

1. Richard Murphy, "Learning RPA: Robotic Process Automation".
2. William Murphy, "RPA: The Robotic Process Automation Handbook".
3. Adir Moshe, “UiPathStudioX Robotic Process Automation: A Beginner's Guide”.
4. AdeelJaved, AnumSundrani, Naida Malik and Sidney Madison Prescott, “Robotic Process Automation with UiPath Studio: Design, Build, and Deploy Automated Solutions”, Apress, 2021.
5. Adir Sharon, “Learning UiPath Robotic Process Automation - Building Software Robots for Business Automation”.
6. Mohammed Al-Buraiki, “UiPathStudioX Robotics Process Automation: A Beginner's Guide to Building Software Robots”.
7. Bernd Hinz, "RPA Implementation: A Guide to Successful Robotic Process Automation".
8. Ramya Devi Vaddi , "UiPath Advanced Development with Computer Vision and Machine Learning".
9. https://www.uipath.com/rpa/academy

2302AS551	PROGRAMMING WITH LARGE DATASETS LABORATORY	L	T	P	C
		0	0	2	1

PREREQUISITE:

2302AS302 - Database Systems

COURSE OBJECTIVES:

The objective of this course is to provide hands-on training in writingprograms to analyze, model and visualize large datasets. In particular, the students willlearn programming using Map-reduce, Python, R etc. to solve. Students will learn how tosolve common data analysis problems using datasets from a variety of domains: websearch, e-commerce, social-networking, machine learning etc.

COURSE OUTCOMES:

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

CO1: Interpret the results of statistical computations to gain insights into the dataset

CO2: Illustrate, visualize box-plots and histograms for numerical variables in large multivariatedatasets, and compute correlation metrics

CO3: Apply MapReduceprograms to perform clustering and classification of a large multi-variate dataset

CO4: Construct program solutions using classes, objects, virtual environments and packages

CO5: Build spark programs to compute box-plots and histograms of all the numerical variables in a largedataset

CO6: Analyze a spark program to perform classification and regression in a large dataset

COs Vs POs MAPPING:

COs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PSO1	PSO2	PSO3
CO1	3	2	2	2	2	-	-	-	-	-	-	2	1	-
CO2	3	2	1	2	2	-	-	-	-	-	-	2	1	-
CO3	3	2	2	2	2	-	-	-	-	-	-	2	1	-
CO4	3	2	2	2	2	-	-	-	-	-	-	2	1	-
CO5	3	2	2	2	2	-	-	-	-	-	-	2	2	-
CO6	3	2	2	2	2	-	-	-	-	-	-	2	2	-

LIST OF EXPERIMENTS:

1. Write a map reduce program to compute descriptive statistics such as mean, median, mode, standard deviation from a large dataset.
2. Write a map-reduce program to compute box-plots and histograms of all the numerical variables in a large multi-variate dataset.
3. Write a map-reduce program to compute correlation metrics between pairs of all the numerical variables in a large multi-variate dataset.
4. Write a map-reduce program to perform clustering of a large multi-variate dataset. Measure the runtime and study its scaling behaviour as more nodes are added to the cluster.
5. Write a map-reduce program to perform classification of a large multi-variate dataset into two or more classes.
6. Write a spark program to compute box-plots and histograms of all the numerical variables in a large dataset.
7. Write a spark program to perform classification in a large dataset. Measure the runtime and study its scaling behaviour as more nodes are added to the cluster.
8. Write a spark program to perform regression in a large dataset. Measure the runtime and study its scaling behaviour as more nodes are added to the cluster.

TOTAL: 30 HOURS

REFERENCES:

1. Tom White, "Hadoop: The Definitive Guide", 4th Edition, O'Reilly Media Inc, 2015
2. Bill Chambers, MateiZaharia, “Spark: The Definitive Guide”, 4th Edition, O'Reilly Media Inc, 2018
3. P. J. Sadalage, M. Fowler, "NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence", Addison-Wesley Professional, 2012
4. Michael Minelli, Michelle Chambers, and AmbigaDhiraj, "Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses", Wiley, 2013
5. V.K. Jain, Data Science and Analytics (with Python, R and SPSS Programming), Khanna Book Publishing Company
6. V.K. Jain, Big Data and Hadoop, Khanna Book Publishing Company 2020
7. Alan Gates, "Programming Pig", O'Reilly Media Inc, 2011
8. Edward Capriolo, Dean Wampler, and Jason Rutherglen, "Programming Hive", O'Reilly Media Inc, October 2012
12. http://www.planetcassandra.org/what-is-nosql
13. https://class.coursera.org/datasci-001/lecture
14. https://onlinecourses.nptel.ac.in/noc20_cs92/preview
9. Tom White, "Hadoop: The Definitive Guide", 4th Edition, O'Reilly Media Inc, 2015
10. Bill Chambers, MateiZaharia, “Spark: The Definitive Guide”, 4th Edition, O'Reilly Media Inc, 2018
11. P. J. Sadalage, M. Fowler, "NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence", Addison-Wesley Professional, 2012