PRESIDENCY UNIVERSITY
Presidency University Act, 2013 of the Karnataka Act No. 41 of 2013 | Established under Section 2(f) of UGC Act, 1956
Approved by AICTE, New Delhi
GAIN MORE KNOWLEDGE REACH GREATER HEIGHTS
PRESIDENCY GROUP OVER 40 YEARS OF ACADEMIC WISDOM

School of Engineering

Department of Computer Science and Engineering

CURRICULUM STRUCTURE

Based on Choice Based Credit System (CBCS) and Outcome Based Education (OBE)

Program: MASTER OF TECHNOLOGY IN DATA SCIENCE

M.TECH. [DATA SCIENCE]

M.Tech. [DSC]

2021-2023

Regulations No: PU/AC-18.8/CSE16/DSC/2021-2023

Resolution No. 8 of the 18th Meeting of the Academic Council held on 3rd August, 2022, and Ratified by the Board of Management in its 19th Meeting held on 4th August, 2022.

August 2022

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

I. Name of the Program: M. TECH. [DATA SCIENCE]

II. Program Code: M.Tech. [DSC]

III. Program Needs:

Master of Technology in Data Science is a postgraduate degree program that familiarises students with the conceptual framework which is used in Analytics, this would include statistics, computational physics, algebra, CS and all allied subjects. This course provides an opportunity to learn principles, tools, and techniques to model various real-world problems, analyze them, and discover useful information. The interdisciplinary field of data science uses key skills of a wide range of fields including machine learning, statistics, visualization etc. It enables us to identify meaning and appropriate information from huge volumes of data to make informed decisions in technology, science, business etc.

This course will make the students to understand and master CSE and also make them specialist in the field of Data Science – which is one of the emerging fields of Computer Science and Highest paid in the market. Companies like Google, Microsoft, Face Book, Twitter, Instagram, Cisco, Amazon, Oracle are the major users of Data Science which is highly driven by data.

Program Educational Objectives [PEOs]: After four years of successful completion of the program, the graduates shall be:

PEO 01: To prepare graduates who will be successful professionals in industry, government, academia, research, entrepreneurial pursuit and consulting firms.

PEO 02: To prepare graduates who will contribute to society as broadly educated, expressive, ethical and responsible citizens with proven expertise.

PEO 03: To prepare graduates who will achieve peer recognition as individuals or in a team through demonstration of good analytical, research, design and implementation skills.

PEO 04: To prepare graduates who will thrive to pursue life-long reflective learning to fulfil their goals.

IV. Program Outcome [POs]: On successful completion of the Program, the students shall be able to:

PO 1: An ability to analyze, manage and supervise engineering systems and processes with the aid of appropriate advanced tools.

PO 2: An ability to design a system and process within constraints of health, safety, security, economics, manufacturability to meet desired needs.

PO 3: An ability to carry out research in the respective discipline and publish the findings.

PO 4: An ability to effectively communicate and transfer the knowledge/ skill to stakeholders.

PO 5: An ability to realize the impact of engineering solutions in a contemporary, global, economical, environmental, and societal context for sustainable development.

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Program Specific Outcomes [PSOs]: Upon completion of the M.Tech in Data Science programme, students will be able to:

PSO 01: [Problem Analysis]: Identify, formulate, research literature, and analyze complex engineering problems related to Data science principles and practices, Programming and Computing technologies reaching substantiated conclusions using first principles of mathematics, natural sciences and engineering sciences.

PSO 02: [Design/development of Solutions]: Design solutions for complex engineering problems related to Data science principles and practices, Programming and Computing technologies and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, cultural, societal and environmental considerations.

PSO 03: [Modern Tool usage]: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modelling to complex engineering activities related to Data science principles and practices, Programming Data science Computing & analytics with an understanding of the limitations.

V. Curriculum Structure: The curriculum structure is composed of the following baskets:

1] **School Core** (Common & Compulsory to all the programs of the School consists of 32 Credits)

2] **Program Core** (Specific to the M.Tech. [DS] Program, consists of 15 credits).

3] **Discipline Electives** (Specific to the M.Tech. [DS] Program and categorized under various specialized groups, the minimum number of credits to be earned in this basket is 15)

4] **Open Electives** (Consists of courses from various schools to provide an opportunity for multi-disciplinary learning and the minimum number of credits to be earned from this basket is 06)

Table 1: Summary of Minimum Credit Contribution from various Baskets

Baskets	Credit Contribution
SCHOOL CORE	32
PROGRAM CORE	15
DISCIPLINE ELECTIVE	15
OPEN ELECTIVE	06
TOTAL CREDITS	**Min. 68**

The curriculum structure is designed as per the AI and incorporating OBE Principles. The students are provided with at most flexibility in selection of the courses of their choice. The students are free to prepare their own Course grid for every semester from the Basket of courses subject to satisfying the pre-requisites for the courses selected and adhering to the Minimum and Maximum Credit requirement as per the Program Regulations.

The students have an option to decide the pace of his/ her learning [The number of semesters to complete the program]. The slot time table system provides the opportunity to the students to decide the time slot of the course and to select the faculty member among those who are offering the course.

The curriculum provides an opportunity to the students to obtain a specific specialization in the basic degree of **MASTER OF TECHNOLOGY IN M. Tech. (DATA SCIENCE) – DSC.**

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

VI. Basket wise courses:

A] School Core: Minimum Credits to be earned from this basket = 32 Credits

Sl. No.	Course Code	Course Name	L	P	Credits
SCHOOL CORE - Minimum Credits to be earned from this basket					**32**
1	MAT6001	Advanced Engineering Mathematics	3	0	3
2	ENG5001	English for Employability	2	2	3
3	SEM5001	Seminar - I	-	-	1
4	SEM5002	Seminar - II	-	-	1
5	PIP6001	DISSERTATION/ INTERNSHIP - I	-	-	10
6	PIP6002	DISSERTATION/ INTERNSHIP - II	-	-	14

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

B] Program Core: Minimum Credits to be earned from this basket 15 Credits

PROGRAM CORE - Minimum Credits to be earned from this basket					15
1	CSE5008	Programming in Data Science	2	2	3
2	CSE5007	Machine Learning Algorithms	2	2	3
3	CSE5009	Data Analytics and Visualization	2	2	3
4	CSE6001	Deep Learning	2	2	3
5	CSE6003	Big Data Analytics Tools and Techniques	2	2	3

C] Discipline Electives: Minimum Credits to be earned from this basket 15 Credits

1	CSE5011	Data Science with Cloud Computing	3	0	3
2	CSE5015	Data Security and Access Control	3	0	3
3	CSE5013	Soft Computing Techniques	3	0	3
4	CSE6002	Natural Language Processing Techniques	2	2	3
5	CSE6004	Time Series Analysis and Forecasting	3	0	3
6	CSE6005	Intelligent Information Retrival	3	0	3

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

7	CSE6007	IOT Data Analytics	3	0	3
8	CSE6008	Probabilistic Graph Models	3	0	3
9	CSE6009	Artificial Neural Networks	3	0	3
10	CSE6010	Social Network Analysis	3	0	3
11	CSE5016	Essentials for Machine Learning	3	0	3
12	CSE6011	Application of Probability theory in Computer Science	3	0	3
13	CSE5017	NoSQL Databases	2	2	3

A] Open Electives: Minimum Credits to be earned from this basket 6 Credits

1	CIV5001	Sustainable Smart Cities	3	0	3	Civil Engineering Basket
2	CIV5002	Systems Design for Sustainability	3	0	3	
3	CIV5003	Self-Sustainable Buildings	3	0	3	
4	CIV5004	Energy and Buildings	3	0	3	
1	LAW5001	International Trade Law	3	0	3	Law Basket
2	LAW5002	Law relating to Business Establishment	3	0	3	
3	LAW5003	Data Protection Law	3	0	3	

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

4	LAW5004	Law Relating to Consumer Protection	3	0	3	
5	LAW5005	Law Relating to Infrastructure Projects	3	0	3	
1	CSE5001	Programming Methodologies using Java	3	0	3	Computer Science Basket
2	CSE5002	Human Computer Interaction	3	0	3	
3	CSE5003	IOT Applications	3	0	3	
4	CSE5004	Programming Essentials in Python	3	0	3	
1	ECE5001	Wearable Computing	3	0	3	Electronics and Communication Engineering Basket
2	ECE5002	MEMS and Nanotechnology	3	0	3	
3	ECE5003	Advanced Computer Networks	3	0	3	
4	ECE5004	Pervasive Computing	3	0	3	
1	MEC5001	Optimization Techniques	3	0	3	Mechanical Engineering Basket
2	MEC5002	Industry 4.0	3	0	3	
3	MEC5003	Six Sigma for Engineers	3	0	3	
4	MEC5004	Design for Internet of Things	3	0	3	
1	MBA3026	Essentials of Leadership	3	0	3	Management Basket
2	MBA3037	Fundamentals of Accounting	3	0	3	
3	MBA3038	Sales Techniques	3	0	3	

REGISTRAR

Registrar

4	MBA3039	Principles of Management	3	0	3	
1	RES5001	Research Methodology	3	0	3	
Research Project (Students are required to carry out research work under the guidance of a faculty member/ research scholar and the same shall be evaluated and credit will be granted as per the academic regulations)						
1	URE6001	Research Experience	-	-	3	
		Total Credits			68	

Type of Skill
F - Foundation
S - Skill Development
EM – Employability
EN – Entrepreneurship

Course Caters to
GS - Gender Sensitization
ES - Environment and sustainability
HP - Human values and Professional Ethics

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

SEMESTER GRID:2021-2023					
SEM – 1					
S. No.	**COURSE CODE**	**COURSE NAME**	**CREDIT STRUCTURE**		
			L	**P**	**CREDITS**
1	MAT6001	Advanced Engineering Mathematics	3	0	3
2	ENG5001	English for Employability	2	2	3
3	CSE5005	Artificial Intelligence	2	2	3
4	CSE5006	Knowledge Engineering and Expert Systems	3	0	3
5	CSE5007	Machine Learning Algorithms	2	2	3
6	CSEXXXX	Discipline Elective - I	3	0	3
7	CSEXXXX	Discipline Elective - II	3	0	3
8	SEM5001	Seminar – I	-	-	1
TOTAL			**18**	**6**	**22**
SEM – 2					
S. No.	**COURSE CODE**	**COURSE NAME**	**CREDIT STRUCTURE**		
			L	**P**	**CREDITS**

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

1	CSE6001	Deep Learning	2	2	3
2	CSE6002	Natural Language Processing Techniques	2	2	3
3	CSEXXXX	Discipline Elective - III	3	0	3
4	CSEXXXX	Discipline Elective - IV	3	0	3
5	CSEXXXX	Discipline Elective - V	3	0	3
6	XXXXXXX	Open Elective – I	3	0	3
7	XXXXXXX	Open Elective - II	3	0	3
8	SEM 5002	Seminar – II	-	-	1
TOTAL			**19**	**4**	**22**

SEM – 3

S. No.	**COURSE CODE**	**COURSE NAME**	**CREDIT STRUCTURE**		
			L	**P**	**CREDITS**
1	PIP6001	Dissertation/ Internship - I	-	-	10

SEM – 4

S. No.	**COURSE CODE**	**COURSE NAME**	**CREDIT STRUCTURE**		
			L	**P**	**CREDITS**
1	PIP6002	Dissertation/ Internship - II	-	-	14

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

VII. Course Catalogues:

Each course shall have a course catalogue with the following details:

i) Pre –Requisites of the course
ii) Course Description
iii) Course Outcome
iv) Course Content
iv) Reference Resources.

The Course Catalogues for the Courses offered in each basket are attached below:

Course Code: **CSE 5001**	**Course Title: Programming Methodologies using Java** **Type of Course: Open Elective** **Theory Only**	**L-P-C**	3	0	3
Version No.	1.0				
Course Pre-requisites	Basic knowledge of any structured programming: Data types, variables, constants, operators, conditional & control structures, Loops, arrays & function.				
Anti-requisites	Object Oriented Programming, Java				
Course Description	This course introduces the engineering of computer applications emphasizing modern software engineering principles: object oriented design, decomposition, encapsulation, abstraction and testing by using Java Programming. This course has theory and lab component which emphasizes on understanding the implementation and application of object-oriented programming paradigm. All along the course, we discuss how OOP allows software engineers to develop complex software by making high-level abstractions starting from very general classes down to more concrete classes. The Programming methodology emphasizes on software engineering principles, following best practices in software development, which enables the student to build real time applications with industry standards.				
Course Out Comes	On successful completion of the course the students shall be able to: 1) Identify and model the objects and their relationship. 2) Apply the concept of arrays, strings, polymorphism & inheritance for real world scenarios. 3) Implement interface & packages for building applications 4) Apply the error handling and multithreading concepts appropriately. 5) Use collections and generics to create desktop applications. 6) Create GUI and web-based application.				
Course Content:					
Module 1	INTRODUCTION	Assignment	Programming		**No. of Classes:8**
Introduction to Object Oriented Programming, TOKENS: Data types, Variables, Operators, Control Statements, Command Line Arguments. Common Errors, Comments, Pre-conditions and Post-conditions, Decomposition, Importance of Software Engineering, The Right Decomposition. CLASSES, OBJECTS, AND METHODS: Defining a class, access specifiers, instantiating objects, reference variable, accessing class members and methods, constructors, method overloading, static members, static methods, inner class, Wrapper class, Autoboxing and Unboxing.					

Module 2	Arrays, Strings, Inheritance and Package	Assignment	Programming	**No. of Classes:6**
Array, Strings, Inheritance, Interface : extending an interface, Implementing interfaces, Package: Package as Access Protection, Defining a Package, Library Packages, import packages.				
Module 3	Exception Handling &MultiThreading	Assignment	Programming	**No. of Classes:8**
Exception handling: Introduction to Exceptions, Difference between Exceptions & Errors, Types of Exception. Handling Exceptions: Use of try, catch, finally, throw, throws. User Defined Exceptions, Checked and Un-Checked Exceptions. Introduction to threads, life cycle of a thread, creating threads, extending the Thread Class, Implementing the "runnable" interface. Thread Priority, Thread synchronization, Inter communication of Threads				
Module 4	Generics, Collection Framework, JAR File	Assignment	Programming	**No. of Classes:8**
Generics: Introduction, using wildcard, generic method, generic class hierarchies, erasure. Collections: Introduction to Collections, Classification of Collection. Introduction to List, Map and Set Interface.				
Module 5	Graphic Programming & Java Bean, Servlet, JDBC	Assignment	Mini Project	**No. of Classes:10**
Swings: Introduction, Swing GUI Components and Layout Managers, Swing Menus. Java Bean: Introduction, Introspection, Persistence, JavaBeans API. Servlet: Lifecycle, Simple Servlet, Handling HTTP request and responses. JDBC: Driver, connection, statement, ResultSet, Transaction Processing, Metadata.				

List of Laboratory Tasks:

Experiment N0 1: Programming assignment with class, objects and basic control structures. (Application: Build a basic menu driven application). RandomGenerator Program, The RollDice Program.

Level 1: Programming scenarios which use control structures to solve simple case scenarios.

Level 2: Programming assignment which will build menu driven application by identifying the class and its relevant methods.

Experiment No. 2: Programming assignment using Arrays and Strings.

Level 1: Programming scenarios which build single dimensional and multidimensional array, apply the different methods to operate on strings.

Level 2: Programming assignment which will manipulate the data stored in matrices and identify the appropriate usage String methods.

Experiment No. 3: Programming assignment using acm.graphics Package, GCanvas, Methods Common to All GObjects, Interfaces and Methods Defined by Them, The Bouncing Ball Program Example, The Geometry of the GLabel Class, The GArc Class.

Level 1: Programming scenarios which use the concept the Package and usage of Packages

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Level 2: Programming assignment which build application which have Packages for different scenarios.

Experiment No. 4: Programming assignment using Exception Handling

Level 1: Programming assignment on building applications using built in Exceptions.

Level 2: Programming assignment on building application using user defined Exceptions.

Experiment No. 5: Programming assignment using Multithreading. (E.g.: Building an application which performs different arithmetic operations and sharing the resources using threads)

Level 1: Programming scenarios to build a thread, assign priority and use the thread methods to perform operations

Level 2: Programming scenarios for building synchronized applications.

Experiment No. 6: Programming assignment using Collections, Generics. Music store case study: Managing Large Amounts of Data, Principles of Design, Defining the Song Class, Defining the Album Class, Considering the Data Structures Needed, Reusing Data - Shallow Copy vs. Deep Copy. Jar File: Creating a Jar file.

Level 1: Programming scenarios which build applications Using Collections and Generics.

Level 2: Programming scenarios which help in understanding the need and scenarios to use Collections

Experiment No. 7: Programming assignment to build GUI Applications. Building Online Music Store.

Level 1: Programming Scenarios to build GUI for a given scenario using Swings concepts.

Level 2: Understanding and application of Swing and Graphics Concepts to build an Application

Targeted Application & Tools that can be used: Targeted Employment sector is Software application,product development Companies in IT sector and Non IT Sector. The skills include

- Platform independent Application Development
- Secure Application Development
- Data Mining
- Operating Systems.
- Database Management Systems
- Banking software
- Automobiles
- Mobile Applications

Tools: JDK (Java Development Tool kit), Integrated Development Environment (IDE), Apache NetBeans, Eclipse.

Project work/Assignment: Mention the Type of Project /Assignment proposed for this course

After completion of each module a programming based Assignment/Assessment will be conducted. A scenario will be given to the student to be developed as a Java Application.

Text Book

1) **Cay S Horstmann and Cary Gornell, *"CORE JAVA volume I-Fundamentals"*, Pearson.**
2) **Cay S Horstmann and Cary Gornell, *"CORE JAVA volume II-Advanced Features"*, Pearson.**

References

1)Herbert Schildt, *"The Complete Reference Java 2"*, Tata McGraw Hill Education.

REGISTRAR

2)James W. Cooper, *"Java TM Design Patterns – A Tutorial"*, Addison-Wesley Publishers.

3) Jim Keogh, "J2EE Complete reference", Tata McgrawHill.

4) Timothy C. Lethbridge and Robert Laganiere, "Object Oriented Software Engineering : Practical Software Development using UML and Java", Tata McgrawHill.

5) Sarcar, Vaskaran, "Java Design Patterns – A hands on experience with real world examples", Apress.

Catalogue prepared by	Dr. Blessed Prince P
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: **CSE 5002**	**Course Title: Human Computer Interaction** **Type of Course: Open Elective** **Theory Only**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	**Basic knowledge of HTML and web design**				
Anti-requisites	**NIL**				
Course Description	The Course is intended to introduce students about the key concepts of User Interface Design. It will cover the theory and methods that exist in the field. User Interface Design is an interdisciplinary field that integrates theories and methodologies from computer science, cognitive psychology, design, and many other areas. It stresses the importance of good interfaces and the relationship of interface design to effective human interaction with computers. It helps in categorizing the interfaces based on the processes, methods and programming used. It focuses on applications of emerging fields in user Interface Design.				
Course Outcomes	On successful completion of the course the students shall be able to: 1) Identify the factors influencing user interfaces; 2) Apply guidelines, principles, theories and methodologies for designing interfaces; 3) Explain various user interface evaluation methods. 4) Identify the applications of emerging fields in human computer interaction				
Course Content:					

REGISTRAR

Module 1	Importance of User Interface Design	Assignment	Cognitive Frameworks, Ergonomics	**10 Classes**

Topics:

Introduction: Importance of user Interface – definition, importance of good design. Benefits of good design. A brief history of Screen design. Human Perception –Human Thinking, Emotion, Psychology and the design of interactive systems – Cognitive frameworks – Models of interaction, Frameworks and HCI – Ergonomics – Universal usability

Module 2	Interface Design	Assignment	Guidelines, Pillars of design	**10 Classes**

Topics:

Interface design : The software life cycle Usability engineering Iterative design and prototyping, Interaction design – Guidelines – Principles – Theories – The process of design –Prototyping and Construction - Conceptual design – Physical design – The four pillars of design – Development methodologies – Participatory design – Scenarios development – Social impact statement for early design review – Legal issues.

Module 3	Evaluation of interface design	Case Study	Expert Reviews, Usability Testing	**8 Classes**

Topics:

Evaluating interface design Evaluating interface design **–** Evaluation, Goals of evaluation, Expert Reviews, Usability testing and Laboratories, Survey Instruments, Acceptance Tests, Evaluating during Active Use, Controlled Psychologically Oriented Experiments, Choosing an evaluation method.

Module 4	Information Presentation	Assignment	Ubiquitous computing and Augmented Reality	**10 Classes**

Topics:

Information presentation: Information presentation – Data type by task taxonomy, Challenges for Information Visualization –Ubiquitous computing and augmented realities Ubiquitous computing applications research Design Focus: Ambient Wood – augmenting the physical Virtual and augmented reality Design Focus: Shared experience Design Focus: Applications of augmented reality Information and data visualization Design Focus: Getting the size right. Groupware – Goals of collaboration and participation, Design for diversity – Graphical user interfaces.

Targeted Application & Tools that can be used:

Targeted employment sector is Developing Mobile Apps and web Applications vendors like Amazon, Flip kart, Snap Deal, Byjus, eBay etc. Targeted job profiles include HCI Specialist, UX Design etc.

Tools:

- Xampp Server
- Any Text Editor like notepad++

Case Study Analysis/Assignment:

Case Study Analysis :

- Students have to choose any of the Application it can be Mobile App or web Applications and they should relate with User Interface Design concepts in term of Guidelines and Principles of Interface Design etc. to evaluate design with respect to user perspective.

Term Assignments:

REGISTRAR

• **Analysis of Applications with respect to Guidelines and Principles of Interface Design by taking various case studies** • **A Case study on Usability and Social impact Statement of different Applications**
Text Book 1. Ben Shneiderman and Catherine Plaisant, "Designing the User Interface". Addison Wesley. 2. Dix A. et al. "Human-Computer Interaction", Prentice Hall 3. The essential guide to user interface design, Wilbert O Galitz, Wiley Dream Tech.
References 1. Yvonne Rogers, Helen sharp, Jenny Preece, "Interaction Design: Beyond Human Computer Interaction", Wiley. 2. The Essentials of Interaction Design, Fourth Edition by Cooper, Reimann, Cronin, & Noessel (2014). 3. Human – Computer Interaction. Alan Dix, Janet Fincay, Gre Goryd, Abowd, Russell Bealg, Pearson Education
Topics relevant to the "Foundation Skills": Identifying factors which influences User Interface Topics relevant to "Human Values and Professional ethics": Guidelines for User Interface Design and Data collection for Term Assignments and case studies

Catalogue prepared by	Prof. T Ramesh
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE 5003	**Course Title: IOT Applications** **Type of Course: Open Elective** **Theory Only**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	Fundamentals of computer network, wireless sensor network, communication, internet technology, web technology and information security.				
Anti-requisites	**NIL**				
Course Description	This course introduces the core principles of Internet of things. This theory based course emphasizes on understanding the applications areas of IOT. The course will focus on creative thinking of IoT concepts & technologies.				
Course Outcomes	On successful completion of the course the students shall be able to: 1. Understand general concepts of Internet of Things (IoT) (Understand) 2. Recognize various devices, sensors and applications (Knowledge) 3. Apply design concept to IoT solutions (Apply) 4. Evaluate design issues in IoT applications (Evaluate) 5. Create IoT solutions using sensors, actuators and Devices (Create)				

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Course Content:				
Module 1	Introduction to IOT	Assignment	Data Analysis task	**10 Classes**

Topics:
Introduction, Definition & Characteristics of IOT, Physical Design of IoT- Things in IoT, IoT Protocols, Logical design of IoT- IoT functional blocks, IoT Communication Models, IoT Communication APIs, IoT Enabling Technologies- Wireless sensor networks, Cloud computing, Big data Analytics.

Module 2	IOT Protocols	Assignment	Analysis, Data Collection	**10 Classes**

Topics:
Connectivity Protocols: 6LoWPAN, IEEE 802.15.4, Zigbee, Wireless HART, Z-Wave, ISA 100, NFC, RFID.
Communication/Transport Protocols: Bluetooth. Data Protocols: Message Queue Telemetry Transport (MQTT), Constrained Application Protocol (CoAP), Advanced Message Queuing Protocol (AMQP), XMPP – Extensible Messaging and Presence Protocol. RFID: Introduction, Principle of RFID, Components of an RFID system.

Module 3	IOT Application building tools	Assignment	Data analysis task	**10 Classes**

Topics:
Introduction to Arduino Boards, Types of Arduino Boards, Installation of Arduino IDE, Usage of Tinker Cad, Structure of Coding – Embedded C, Hands on session in Blinking of LED, Hands on session in Serial Monitor, Hands on session in Traffic Signal, Hands on session in Arduino UNO Board. Introduction to WiFi Module (ESP8266), Introduction to Node MCU, Pin Configuration of Node MCU, Installation of Node MCU and Blynk Library files, Introduction to Blynk App Open Source Cloud Environment, demonstrate session in connecting Node MCU with Blynk, demonstrate session in Blinking Inbuilt LED in Node MCU, demonstrate session in Blinking of LED. Sensors and its applications.

Module 4	Applications of IOT	Assignment	Analysis, Data Collection	**10 Classes**

Topics:
Overview of IoT applications: Automotive and Transport, Smart factories, Smart buildings, Smart cities, Smart utilities, Security and Surveillance, Smart agriculture, Retail, and Healthcare with suitable examples.
Building IoT Application: Enabling and facilitating the students to take up existing problems and building the solution.

Targeted Application & Tools that can be used:

Targeted employment sector is service provider and control monitor like GE, Siemens, TCS etc. Targeted job profiles include digital domain and Physical system design engineer, IOT engineer etc.

Tools:

- Arduino IDE
- TinkerCad
- NodeMCU
  Tensor Flow and Keras

Project work/Assignment:

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Mini Project:

- A mini project to demonstrate use of IOT tools, techniques and protocols to build and smart real time application.

Term Assignments:

- **Comparative analysis of communication protocols and connectivity protocols**

  Carry out a thorough analysis of the various IOT protocols and its efficiency for given data set.

- **A short survey of applications in IOT**

  Study and analyze few important applications of IOT and then applying the concepts IOT to build a real time application.

Text Book

4. "Internet of Things (A Hands-on-Approach)", by Vijay Madisetti and Arshdeep Bahga, 1st Edition, VPT, 2014
5. "Industry 4.0: The Industrial Internet of Things", by Alasdair Gilchrist (Apress)

References

4. "Industrial Internet of Things: Cyber manufacturing Systems" by Sabina Jeschke, Christian Brecher, Houbing Song, Danda B. Rawat (Springer)

Web Resource References

1. https://www.udemy.com/internet-of-things-iot-for-beginners-getting-started/
2. http://playground.arduino.cc/Projects/Ideas
3. http://runtimeprojects.com

Topics relevant to development of "FOUNDATION SKILLS": Get introduced to AI programming and Interfacing of IOT devices.

Topics relevant to "HUMAN VALUES &PROFESSIONAL ETHICS": Set of standard procedures to build IOT applications.

Catalogue prepared by	Prof. Prakash B Metre, Prof. Mohammed Mujeer Ulla
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE 5004	**Course Title: Programming Essentials in Python** **Type of Course: Open Elective** **Theory Based Course**	**L- P- C**	3	0	3

REGISTRAR

Version No.	1.0				
Course Pre-requisites	Basic knowledge of Computers and Mathematics				
Anti-requisites	Python programming				
Course Description	This course introduces the core concepts of programming using Python. This course has theory component which emphasizes on understanding and programming right from basics to Visualization in Python. It helps the student to explore data by applying these concepts and also for effective problem solving, visualizing and analyzing.				
Course Outcomes	**On successful completion of this course the students shall be able to:** 1. Illustrate the python programming constructs. 2. Explore Data using Python Numpy and Pandas 3. Demonstrate Data Visualization using Matplotlib. 4. Analyze the data using scikit.				
Course Content:					
Module 1	Basics of Python programming	Assignment	Programming	**8 Classes**	
Topics: Data types, Operators and expressions, I/O statements, Control structures- Sequential, selective and Repetitive structures, Functions- user defined and builtin functions.					
Module 2	Data Exploration using Numpy and Pandas	Assignment	Programming	**10 Classes**	
Topics: Installation of **Numpy**, Numpy Basics, Placeholders, Datatypes, Arrays, Basic Statistics, Copying, Slicing & Subsetting, Indexing, Flattening, Reshaping, Resizing, Sorting, Swapping, Dealing with Missing values. **PANDAS** - the PYTHON Data Analysis Library, Motivation, Installation of PANDAS, PANDAS Data Structure, Series, Dataframe, Loading the Data, Descriptive Statistics, Indexing & ReIndexing, Renaming, Iteration, Sorting, Statistical functions, Window functions, Aggregations, Dealing with Missing Data, Groupby Operations, Merging/Joining, Concatenation, Time Series, Working with Categorical Data and Text Data.					
Module 3	I/O Tools and Visualization	Assignment	Mini project	**10 Classes**	
Topics: **I/O API Tools**, CSV and Textual files, Reading data in CSV or text files, Using RegExp to parse txt files, Reading txt files into parts, Writing data in CSV, Reading and Writing HTML files, Reading data from XML, Reading and Writing data from excel file, JSON Data, The format HDF5, Pickle-PYTHON Object Serialization, Serialize a PYTHON object with cpickle The **Matplotlib library**, Installation, A simple interactive chart, Adding elements to the chart, Adding a grid, Adding a legend, Converting the session to an html file, Saving your chart directly as an image, Handling date values, Chart typology, Line charts, Histograms, Bar charts, Horizontal Bar Charts, Multiseries bar charts, Multiseries stacked bar chart, Pie chart.					
Module 4	Sci-kit	Assignment	Mini project	**8 Classes**	

Topics: **The Scikit learn library**, Machine learning, Supervised and Unsupervised learning, Supervised learning with Scikit learn, The iris flower dataset, KNN Classifier, Diabetes dataset, Linear Regression-the least square regression, SVMs, SVC, Non linear SVC, Plotting different SVM classifier using iris dataset.				
Targeted Application & Tools that can be used: **Targeted Application : Web application development, AI, Operating systems** **Tools: Python IDLE, ANACONDA** • **Application Areas:** • Web Development • Game Development • Scientific and Numeric Applications • Artificial Intelligence and Machine Learning • Software Development • Enterprise-level/Business Applications • Education programs and training courses • Operating Systems • Web Scrapping Applications • Image Processing and Graphic Design Applications **Professionally Used Software: Python IDLE, Spyder, Jupyter Notebook, Google Colab**				
Project work/Assignment: - After completion of each module a programming based Assignment/Assessment will be conducted. - A scenario will be given to the students to be developed as a series of Program/ Application. On completion of Module 3 and Module 4, students will be asked to develop a Mini Project using Python.				
Text Book • Charles Dierbach, "Introduction to Computer Science Using Python", Wiley India Edition,2015.				
References 1. E. Balagurusamy, "Introduction to Computing and Problem Solving Using Python", Tata McGraw-Hill, 2016 2. Y. Daniel Liang, "Introduction to Programming Using Python", Pearson, 2017 3. Python Tutor - Visualize Python, Java, C, C++, JavaScript, TypeScript, and Ruby code execution 4. https://practice.geeksforgeeks.org/courses/Python-Foundation				
Topics relevant to development of "Foundation Skills": - Basics of python and data exploration. **Topics relevant to development of "Employability":** - Solve real time problems by analysing and visualising data.				
Catalogue prepared by	Prof. Pallavi M, Prof. Shweta Singh			

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: **CSE 5005**	**Course Title: Artificial Intelligence** **Type of Course: Program Core Theory and Laboratory Integrated**	**L-P-C**	2	2	3
Version No.	1.0				
Course Pre-requisites	1.Strong knowledge of Computer science, programming languages and coding. 2.Python. 3. Probability and Statistics. 4. Strong data analytics skills.				
Anti-requisites	NIL				
Course Description	Artificial intelligence (AI) is a research field that studies how to realize the intelligent human behaviors on a computer. The ultimate goal of AI is to make a computer that can learn, plan, and solve problems autonomously. This course along with Python explores the concepts and algorithms at the foundation of modern artificial intelligence, diving into the ideas that give rise to technologies like game-playing engines, handwriting recognition, etc. This course contains a theory component about the concepts and principles that underlie modern AI algorithms, and a practice component to relate theoretical principles with practical implementation. By course's end, students emerge with experience in libraries for machine learning as well as knowledge of artificial intelligence principles that enable them to design intelligent systems of their own.				
Course Out Comes	On successful completion of the course the students shall be able to: 7) Explain the Concepts and algorithms of Modern Artificial Intelligence 8) Choose appropriate AI Methods in applying scientific method to models of machine learning 9) Apply AI Principles and techniques to real-world problems to develop intelligent systems.				
Course Content:					

Module 1	Python for Artificial Intelligence	Assignment	Programming	**No. of Classes:15**
Topics: Introduction to Python, Advanced Python: Object Oriented, OOPs concept, Class and object, Attributes, Inheritance, Overloading, Overriding, Data hiding, Operations Exception, Exception Handling, Python Libraries, Data migration and visualization: GGPlot, seaborn Pandas and Matplotlib, Database Interaction in Python.				
Module 2	**Introduction to AI**	Assignment	Programming	**No. of Classes:14**

REGISTRAR

Topics: Introduction, A.I. Representation, , Problem Characteristics, Intelligent Agents and Environments, concept of rationality, the nature of environments, structure of agents, problem solving agents, problem formulation

State Space Search – Searching for Solutions – Uniformed Search Strategies –Informed Search Strategies – Heuristic Function, Hill Climbing, Stimulated Annealing, Constraint Satisfaction Problems (CSP) – Backtracking Search and Adversarial Search – Games – Optimal Decisions in Games Greedy Search algorithms – MINMax algorithm -Alpha-Beta Pruning.

Module 3	Reasoning	Assignment	Case Study	**No. of Classes:8**

Topics Reasoning – Introduction to Reasoning – Types of Reasoning, Certainty Factors, Reasoning under Uncertainty – Probabilistic Reasoning – Baye's Theorem – Bayesian Network – Hidden Markov Model – Dempster Shafer Theory.

Module 4	Learning	Assignment	Case Study	**No. of Classes:06**

Topics: Learning: Learning from observations, Forms of Learning, Inductive Learning, Learning decision trees, Theory of learning, Learning Probabilistic Models, Learning in Neural and Belief networks

List of Laboratory Tasks:

Experiment No 1: Write a Python program to implement Basic Elements of Python such as Branching, Recursion, Global Variables, Modules, Files, Inheritance , Encapsulation and Information Hiding.

Level 1: Programming Scenarios which use control structures to solve simple case scenarios.

Level 2: Programming assignment to implement python elements to solve relavant case scenarios.

Experiment No. 2: Write a Python Program to Implement Basic Elements of Python such as Tuples, Functions, Dictionaries, Exceptions and Assertions.

Level 1: Programming Scenarios which use control structures to solve simple case scenarios.

Level 2: Programming assignment to implement python elements to solve relevant case scenarios.

Experiment No. 3: Write a python program to perform data visualization on Canada, Titanic and Customer_data dataset using MatplotLib, Seaborn.

Level 1: Programming Scenarios which use the methods of MatplotLib and seaborn to produce various data visualizations on Canada, titanic datasets.

Level 2: Programming assignment which utilizes the methods of MatplotLib and seaborn to produce various data visualizations on wine datasets.

Experiment No. 4: Write a Program to Implement Breadth First Search and Depth First Search using Python.

Level 1: Programming scenario which implements BFS and DFS on different graph models.

Level 2: Programming assignment which implements BFS and DFS on different graph models.

Experiment No. 5: Write a Program to implement A* and AO*algorithm using Python.

Level 1: Programming scenarios to implement A* and AO* search algorithm on given dataset.

Level 2: Programming assignment to implement A* and AO* search algorithm on given Dataset.

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Experiment No. 6: Write a Program to Implement Tic-Tac-Toe game using Python.

Level 1: Programming Scenario to implement AI gaming theory in Tic-Tac-Toe Game

Experiment No. 7: Write a Program to Implement 8-Puzzle problem via Hill Climbing algorithm in Python

Level 1: Programming Scenarios to implement Hill Climbing Algorithm in 8-Puzzle problem.

Experiment No. 8: Write a Program to Implement stimulated annealing algorithm on Canada dataset.

Level 1: Programming Scenarios to implement stimulated annealing algorithm on Canada dataset.

Experiment No. 9: Write a Program to Implement Water-Jug problem using Python

Level 1: Programming Scenarios to implement AI gaming theory in Water-Jug Problem

Experiment No. 10: Write a Program to Implement Travelling Salesman Problem using Python

Level 1: Programming Scenarios to implement AI gaming theory in Travelling Salesman Problem.

Experiment No. 11: Write a Program to Implement Missionaries-Cannibals Problems using Python

Level 1: Programming Scenarios to implement AI gaming theory in Missionaries-Cannibals Problems.

Experiment No. 12: Write a Program to Implement N-Queens Problem using Python.

Level 1: Programming Scenarios to implement AI gaming theory in N-Queens Problem.

Targeted Application & Tools that can be used:

Targeted Applications:

AI applications do not just have the promise to yield better business results but improve the human experience as a whole. The contribution of the technology giants like Microsoft, Google, Apple and IBM in the healthcare sector holds significant importance for the industry. AI is currently being applied for a wide range of healthcare services, Speech recognition, Virtual Agents, AI optimized hardware, Decision Management, Deep Learning, etc. Few of the top recruiters are Amazon, NVIDIA. Microsoft, IBM, Accenture, Facebook, Intel, Samsung, Lenovo, Adobe etc., among numerous others.

AI Tools:

- Scikit Learn.
- TensorFlow.
- Auto ML.
- Theano.
- PyTorch.
- Caffe.
- Google ML Kit.

Project work/Assignment: Mention the Type of Project /Assignment proposed for this course

Project Works:

1. HANDWRITTEN DIGITS RECOGNITION

Digits written by humans vary a lot in curves and sizes as they are hand-drawn and everyone's writing is not the same. It is a great way to start artificial intelligence by building a handwritten digits recognition system that can identify the digit drawn by humans.

REGISTRAR

2. WEBSITE EVALUATION USING OPINION MINING

A website evaluation system where users can comment on a particular website about the genuineness, time taken to deliver products, etc.It will analyze the comments to rate the website on these factors.

3. CUSTOMER RECOMMENDATAION

E-commerce has benefitted dramatically from AI. The finest example is Amazon and its customer recommendation system. This customer recommendation system has helped the platform in enhancing its income tremendously thanks to better customer experience.

4. Any other project ideas mutually decided by students-instructor

After completion of each module a programming based Assignment/Assessment will be conducted. A scenario will be given to the student to be developed as a Gaming Application.

Case Studies:

1. Rolls-royce and Google Partner to Create Smarter Autonomous Ships based on AI.
2. How the US Retail giant is using AI and Robots to prepare for the 4th Industrial Evolution.
3. The amazing ways google uses AI and Satellite Data to prevent Illegal Fishing.
4. AI in china: The Amazing ways Tencent is Driving it's Adoption
5. Any other Case Studies mutually decided by students-instructor

At the end of the Semester, Teams will be formed according to Student's Count and will be assigned with a Debate Topic

1. AI Risks.
2. Emerging trends in AI.
3. AI in Future.
4. Any other topic mutually decided by students-instructor

Text Book

1. John V Guttag. "Introduction to Computation and Programming Using Python", Second Edition, Prentice Hall of India, 2013.
2. O'Reilly, "Python for Data Analysis", Second Edition, O'Reilly Media Inc, 2017
3. Stuart Russell, Peter Norvig, "Artificial Intelligence: A Modern Approach", Third Edition, Pearson Edu pdfcation / Prentice Hall of India, 2010.

References

1. Wesley J. Chun. "Core Python Programming - Second Edition", Prentice Hall, 2006.
2. Kenneth A. Lambert, "Fundamentals of Python – First Programs", CENGAGE Publication
3. Denis Rothman, Matthew Lamons, Rahul Kumar, Abhishek Nagaraja, Amir Ziai, Ankit Dixit, "Python: Beginner's guide to Artificial Intelligence", Packt publishing, 2018
4. 4.Prateek Joshi, "Artificial Intelligence with Python", Packt Publishing, 2017

Topics relevant to development of "Employability": Problem Solving, Search Algorithms

Topics relevant to " HPROFESSIONAL ETHICS": Developing Ethical Rules for Game Applications.

Catalogue prepared by	Dr. Shankar Rammoorthy, Ms. S.Poornima

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE 5006	**Course Title: KNOWLEDGE ENGINEERING AND EXPERT SYSTEM** **Type of Course: Program Core** **Theory Only**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	**Basic Knowledge on Artificial Intelligence and Discrete Mathematics such as Predicative Logic, Logic rules etc.**				
Anti-requisites	**NIL**				
Course Description	Knowledge engineering is a field within artificial intelligence that develops knowledge-based systems. Such systems are computer programs that contain large amounts of knowledge, rules and reasoning mechanisms to provide solutions to real-world problems. A major form of knowledge-based system is an expert system, one designed to emulate the reasoning processes of an expert practitioner. Topics includes: Introduction to Knowledge Engineering, Knowledge based Systems, Types of Knowledge based systems, Knowledge acquisition, Knowledge representation and reasoning: Logic rules and representations, Semantic Networks, frames, Life cycle Methodologies, Uncertain Reasoning with confidence factor, Basic Structure and Architecture of Expert System. Tools used in Expert System.				
Course Outcomes	On successful completion of the course the students shall be able to: 1] Explain the basic concepts in Knowledge Engineering and types of Knowledge based system. (Knowledge) 2] Discuss the process of acquiring the Knowledge from the human expert (Comprehension). 3] Apply the logical rules, Semantic Networks and Frames for representing the knowledge. (Application) 4] Life Cycle and Methodologies applied to support the development of Knowledge based Systems. (Application) 5] Explain how expert system deal with uncertainty and describes architecture and tools used. (Comprehension)				
Course Content:					
Module 1	Introduction to Knowledge Engineering and Knowledge Base	Assignment	Analysis		**10 Classes**

Topics: Data, Information and Knowledge Skills of a Knowledge Engineering. Introduction to Knowledge-Based Systems, Types of Knowledge based systems.				
Module 2	Knowledge Acquisition	Assignment	Analysis, Data Collection	**5 Classes**
Topics: Knowledge acquisition - knowledge acquired from a human expert - purpose and types of Interviews in obtaining knowledge – Types of interviews using techniques such as repertory grids.				
Module 3	Knowledge Representation and Reasoning	Problem Solving	Data analysis task	**9 Classes**
Topics: Using knowledge - Logic, rules and representation- Developing rule-based systems - Semantic networks- Frames.				
Module 4	Life Cycle and Methodologies	Assignment	Analysis	**9 Classes**
Topics: Need for methodologies- Blackboard architectures- Problem Solving Methods (PSMs)- Knowledge Acquisition Design System (KADS)- The Hybrid Methodology (HyM)- Building a well-structured application using Aion BRE.				
Module 5	Uncertain Reasoning and Expert System	Assignment	Analysis	**10 Classes**
Topics: Uncertainty – Confidence factor- Expert System – Basic Structure, Architecture – Tools used Constructing Expert System.				

.

Targeted Application & Tools that can be used:

After Completion of the course student may get an opportunity to be a Knowledge engineer to design and develop Knowledge base with reference to Acquisition and to represent it.
Expert System can be developed on real time application (To highlight a few)
Medical Knowledge Automation, Chemical and Biological Synthesis, Mineral and Oil explorations, Planning and Scheduling. Space Defense, VLSI Design, Air traffic control, Equipment fault Diagnosis. Circuit Diagnosis and So on.

Tools: Programming tools for building Expert System.

- OPS 5
- EMYCIN
- KAS
- TEIRESIAS

Project work/Assignment:

Case Study Analysis: To Study, analyze and develop expert system on applications.

Term Assignments:

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

- **Comparative analysis on methods in Knowledge representations.**
- **A short survey on techniques used to build Knowledge base.**
- **Recent trends used in developing Expert System.**

Text Book

6. "An introduction to knowledge engineering", Simon Kendal, Malcolm creen, Springer, 2007.(with Recent version copyright)
7. "An Overview of Expert System " William B. Gevarter,Dept. of Commerce,U.S , NBS, Washignton,D.C.

References

5. "An introduction to knowledge engineering", Peter Smith, Thomson computer press, 1996.
6. "A guide to an Expert System ",Donald Waterman, Pearson India.

Topics relevant to development of "FOUNDATION SKILLS": Introduction and types of Knowledge base system, Knowledge Acquisition methods, Basic Structure of Expert system.

Topics relevant to "HUMAN VALUES &PROFESSIONAL ETHICS": Collaboration and Data collection for Term assignments and Case Studies.

Topics relevant to "EMPLOYABILITY SKILLS" Develop skills to be an Knowledge Engineer,

Catalogue prepared by	Dr.T.K.Thivakaran.
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: **CSE 5007**	**Course Title: Machine Learning Algorithms** **Type of Course: Program Core Theory and Laboratory Integrated**	**L-P-C**	2	2	3
Version No.	1.0				
Course Pre-requisites	Basic knowledge of a. Statistics, Calculus, Linear Algebra and Probability. b. Programming Language. c. Data Modeling.				
Anti-requisites	NIL				

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Course Description	This course provides a broad introduction to machine learning and statistical pattern recognition. Topics include: supervised learning (generative/discriminative learning, parametric/non-parametric learning, neural networks, support vector machines); unsupervised learning (clustering, dimensionality reduction, kernel methods); learning theory (bias/variance tradeoffs, practical advice); reinforcement learning and adaptive control.
Course Out Comes	On successful completion of the course the students shall be able to: 10) Identify the characteristics of datasets and compare the trivial data for various applications. 11) Understand and apply scaling up machine learning techniques. 12) To design and implement various machine learning algorithms in a range of real-world applications.
Course Content:	

Module 1	Machine Learning Model Fundamentals	Assignment	Programming	**No. of Classes:10**

Topics: Data-generating process, Understanding the structure and properties of good datasets, Scaling datasets, including scalar and robust scaling, Normalization and Whitening, Selecting training, validation and test sets, including cross-validation, Features of a machine learning model, Learnability, Capacity, including Vapnik-Chervonenkis theory, Bias including underfitting, Variance including overfitting and the Cramer-Rao bound,. Regularization, Definig loss and cost functions.

Module 2	Clustering and Unsupervised Models	Assignment	Programming	**No. of Classes:10**

Topics: K-Nearest Neighbors(KNN), based on k-dimensional(k-d) trees and ball tress, K-means and K-means++, Clustering Fundamentals, Clustering basic, Evaluation of clustering models on the ground truth, Fuzzy C-means, Spectral clustering, DBSCAN.

Module 3	Semi- Supervised Learning Algorithms	Assignment	Programming	**No. of Classes:15**

Topics: Introduction to Semi- Supervised Learning, Semi-supervised scenario, The assumptions needed to efficiently operate in such a scenario, The different approaches to semi-supervised learning, Generartive Gaussian Mixture, contrastive pessimistic likelihood estimation approach, Self-Training, Co-Training,

Advanced Semi-Supervised Classification, Contrastive Pessimistic Likelihood Estimation(CPLE), Semi-supervised Support Vector Machines(S3VM), Transductive Support Vector Machines(TSVM).

Module 4	Graph-Based Semi-Supervised Learning	Assignment	Programming	**No. of Classes:12**

Topics: Graph-Based Semi-Supervised Learning, Label propagation, Example of label propagation, Label spreading, Label propagation based on Markov random walks, Manifold Learning.

List of Laboratory Tasks:

Experiment N0 1: Programming assignment for data cleaning..

Level 1: Programming scenarios which handles missing features, data normalization, data scaling.

Level 2: Programming assignment which helps in feature filtering, selection.

Experiment No. 2: Programming assignment for unsupervised learning

Level 1: Implementation of covariance rule.
Implementationof rubner_tavan_network

Level 2: Implementation of sanger_network.

Experiment No. 3: Programming assignment for advanced unsupervised learning

Level 1: Implementation of kNN, K-means.
Implementation of fuzzy cmeans.

Level 2: Implementation of spectral clustering.

Experiment No. 4: Programming assignment for supervised learning.

Level 1: Programming assignment on label_propogation, spreading

Experiment No. 5: Programming assignment for supervised learning.

Level 1: Implementing SVM

Level 2: Implementing TSVM

Experiment No. 6: Programming assignment for Graph-Based Supervised learning.

Level 1: Estimating Gaussian mixture in ICA

Level 2: Estimating parameter using PCA.

Targeted Application & Tools that can be used:

- Data Mining
- Text Mining
- Web Mining
- Medical Industry

Tools: Anaconda for Python or Google Colab for Python.

Project work/Assignment: Mention the Type of Project /Assignment proposed for this course

After completion of each module a programming based Assignment/Assessment will be conducted.
A dataset will be given to the student to practice the learned algorithms
On completion of Module 4, student will be asked to develop a Project for analyzing the given dataset.

Text Book

3) **Giuseppe Bonaccorso, *"Mastering Machine Learing Algorithms"*,Packt.**
4) **Giuseppe Bonaccorso, *"Machine Learning Algorithms"*,Packt.**

References

1) Imran Ahmed, *"40 Algorithms Every Programmer Should Know"*, Packt

Topics relevant to development of "Employability": Real time data analysis.

Catalogue prepared by	Dr. Aditya Saxsena, Ms. Galiveeti Poornima

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: **CSE 5008**	**Course Title: Programming in Data Science** **Type of Course: Program Core** **Theory and Laboratory Integrated**	**L-P-C**	2	2	3
Version No.	1.0				
Course Pre-requisites	Basic knowledge of any structured programming: Data types, variables, constants, operators, conditional & control structures, Loops, arrays & function.				
Anti-requisites	Python, R Programming Language				
Course Description	This course introduces the core concepts of Data Science followed by programming using Python and R. This course has theory and lab component which emphasizes on understanding and programming right from Basics to Visualization in Python and R. It helps the student to explore data by applying these concepts and also for effective problem solving, visualizing and analyzing.				
Course Out Comes	On successful completion of the course the students shall be able to: 13) Discuss about the process involved in Data Science (Knowledge) 14) Explore Data using Python Numpy and Pandas (Application) 15) Demonstrate Data Visualization using Matplotlib (Application) 16) Explore Data using R and Visualize using R Graphics (Application)				
Course Content:					

Module 1	Introduction to Data Science	Assignment	Case Studies	**No. of Classes: 10**

Topics:

Introduction to Data Science –- The field of Data Science – The various Data Science Disciplines, Connecting the Data Science Disciplines, Benefits of each Disciplines, Data Science Techniques and Tools – Types of Data – Measures and Metrics – Descriptive Statistics – Inferential Statistics.

Data Science Methodology - From Problem to Approach and From Requirements to Collection, From Understanding to Preparation and From Modeling to Evaluation, From Deployment to Feedback.

Data Preprocessing - Data Quality Assessment, Feature Aggregation, Feature Sampling, Dimensionality Reduction, Feature Encoding.

Module 2	Data Exploration using Numpy and Pandas	Assignment	Programming	**No. of Classes: 8**

Topics:

Introduction to Python World, Motivation, Installation of **NUMPY**, Numpy Basics, Placeholder, Datatypes, Arrays, Basic Statistics, Copying, Slicing & Subsetting, Indexing, Flattening, Reshaping, Resizing, Sorting, Swapping, Dealing with Missing values.

REGISTRAR

PANDAS - the PYTHON Data Analysis Library, Motivation, Installation of PANDAS, PANDAS Data Structure, Series, Dataframe, Loading the Data, Descriptive Statistics, Indexing & ReIndexing, Renaming, Iteration, Sorting, Statistical functions, Window functions, Aggregations, Dealing with Missing Data, Groupby Operations, Merging/Joining, Concatenation, Time Series, Working with Categorical Data and Text Data.

Module 3	I/O Tools and Visualization	Assignment	Mini Project	**No. of Classes: 8**

Topics:

I/O API Tools, CSV and Textual files, Reading data in CSV or text files, Using RegExp to parse txt files, Reading txt files into parts, Writing data in CSV, Reading and Writing HTML files, Reading data from XML, Reading and Writing data from excel file, JSON Data, The format HDF5, Pickle-PYTHON Object Serialization, Serialize a PYTHON object with cpickle

The **Matplotlib library**, Installation, A simple interactive chart, Adding elements to the chart, Adding a grid, Adding a legend, Converting the session to an html file, Saving your chart directly as an image, Handling date values, Chart typology, Line charts, Histograms, Bar charts, Horizontal Bar Charts, Multiseries bar charts, Multiseries stacked bar chart, Pie chart.

Module 4	Introduction to R	Assignment	Programming	**No. of Classes: 10**

Topics:

R Environment, Using R Studio, Vectors, List, Matrices, Arrays, Data Frames, Factors. Functions - Conditional Functions, User Defined Functions. Reading Data from files, Handling Missing Data, Installing Packages,

R Graphics – Plot, Line, Scatter Plot, Pie Charts, Bars, Histogram, Box Plots, 3D Pie Charts, 3D Scatter Plot, Visualization with GG Plot.

R Statistics – Dataset, Max & Min, Mean Median Mode, Subgroup Analyses, Probability Distributions, Pipes in R.

List of Laboratory Tasks:

Experiment No 1: Create a Numpy array and perform the following operations on it

Level 1: Basic Statistics, Copying, Slicing & Subsetting, Indexing, Flattening, Reshaping, Resizing,

Level 2: Sorting, Swapping, Dealing with Missing values

Experiment No. 2: Create a PANDAS Data frame and perform the following operations on it

Level 1: Descriptive Statistics, Indexing & ReIndexing, Renaming, Iteration, Sorting, Dealing with Missing Data

Level 2: Statistical functions, Window functions, Aggregations

Experiment No. 3: Create a PANDAS Data frame and perform the following operations on it

Level 1: Group by Operations, Merging/Joining, Concatenation,

Level 2: Time Series, Categorical Data and Text Data

Experiment No. 4: Demonstrate Reading and Writing using IO API tools

Level 1: CSV and EXCEL files, HTML and XML files,

Level 2: HDF5 CPickle

Experiment No. 5: Using Matplotlib, Visualize the Data

Level 1: Visualize the data using Line Chart, Bar Charts, Pie Chart, Histograms, Bar chart, Horizontal Bar Chart

Level 2: Visualize the data using Multiseries Bar Chart, Multiseries Stacked Bar Chart

Experiment No. 6: Install R Studio and perform basic operations

Level 1: Vectors, List, Matrices, Arrays, Data Frames, Factors,

Level 2: Functions and handling Missing Data

Experiment No. 7: Using R graphics perform the following

Level 1: Plot, Line, Scatter Plot, Pie Charts, Bars, Histogram, Box Plots,

Level 2: 3D Pie Charts, 3D Scatter Plot, GG Plot

Experiment No. 8: Using R Statistics perform the following

Level 1: Max & Min, Mean Median Mode, Subgroup Analyses,

Level 2: Probability Distributions and Pipes

Targeted Application & Tools that can be used:

- Data Exploration
- Data Visualization
- Data Analysis

Tools:

- Google Colab
- Anaconda
- R Studio

Project work/Assignment: Mention the Type of Project /Assignment proposed for this course

- After completion of each module a programming based Assignment/Assessment will be conducted.
- A scenario will be given to the students to be developed as a series of Program/ Application.
- On completion of Module 2 and Module 4, students will be asked to develop a Mini Project using Python and R.

Text Book

5) The essentials of Data Science, Knowledge Discovery Using R, Graham J Williams, CRC Press, 2017
6) PYTHON Data Analysis, APRESS Publications, Fabio Nelli, 2015

References

1) Comparative Approaches to using R and PYTHON for Statistical Data Analysis, Information Series Reference, 2018
2) Practical Data Science CookBook, APRESS Publications, 2018

Topics relevant to development of "Foundation Skills":

- Data Exploration using Python and R Programming.

Topics relevant to development of "Employability Skills":

- Data Analysis and Visualization using Python and R Programming.

Catalogue prepared by	Dr. R. Kesavamoorthy, Ms. Napalakshmi

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: **CSE 5009**	**Course Title: Data Analytics and Visualization** **Type of Course: Program Core** **Theory and Laboratory Integrated Course**	**L-P-C**	2	2	3
Version No.	1.0				
Course Pre-requisites	Mathematical background in calculus, linear algebra, and probability & statistics, Programming in Python / R				
Anti-requisites	NIL				
Course Description	The Course consists of two parts where first Part covers advanced analytics that covers topics necessary to give businesses greater insight into their data than they could ordinarily, and the Second Part covers data visualization concepts. Primary concepts include machine learning, data mining, predictive analytics, location analytics, big data analytics, and location intelligence. Visualization for Time series, Geolocated data, Correlations, connections, Hierarchies, networks, and interactivity.				
Course Out Comes	On successful completion of the course the students shall be able to: 17) Analyze data by performing Exploratory Data Analysis. 18) Apply techniques of Machine Learning to build Generalized Predictive Models. 19) Explain basic concepts of Data Visualization. 20) Apply principles of Data Visualizations to provide insights from data.				
Course Content:					
Module 1	Data Analytics	Assignment	Analysis, Data Collection		**No. of Classes:11**
Topics: Characteristics and types of data, Types of Analytics, Location Analytics, Working with Geospatial Data, Feature Engineering and Selection, Dimensionality Reduction Techniques, Data Preparation.					
Module 2	Advanced Analytics	Case Study	Analysis, Data Collection, Programming		**No. of Classes:13**
Topics: Statistical methods for Data Analytics, Advance topics in Supervised and Unsupervised Machine Learning: Cluster Analysis, Hyper-Parameter Tuning, Measuring Performance of the Models, Model Selection, Data Mining techniques.					
Module 3	Introduction to Data Visualization	Assignment	Analysis, Data Collection		**No. of Classes:9**
Topics:					

Fundamentals of Data Visualization, Human Perception, Basic plotting techniques, Interaction concepts, Visualization techniques for Time Oriented data, Visualization techniques Networks. Introduction to Data Visualization Tools

Module 4	Application - Data Visualization	Case Study	Analysis, Data Collection, Programming	**No. of Classes:14**

Topics:

Designing effective Visualizations, Advanced Visualization Tools, Visualizing Geospatial Data, Document Visualization, Visualization Systems, Evaluating Visualizations, Visualization Benchmarking.

List of Laboratory Tasks:

Experiment No 1: Exploratory Data analysis

Level 1: Demonstration of Tools to implement EDA

Level 2: Use the Dataset to analyze and summarize data, analyze anomalies, analyze Outliers, and Missing Value Treatment

Experiment No. 2: Dimensionality Reduction Techniques

Level 1: Implement DR Technique(s)

Experiment No. 3: Machine Learning Methods

Level 1: Implement Supervised Learning Techniques for the given dataset

Level 2: Implement Un-Supervised Learning Techniques for the given dataset and Cluster Analysis

Experiment No. 4: Measure the Performance of the Models

Level 1: Perform Model Selection

Level 2: Regularize the model

Experiment No. 5: Introduction to Data Visualization Tools

Level 1: Implement Basic plotting techniques

Experiment No. 6: Time Oriented data

Level 1: Visualization techniques for Time Oriented data

Experiment No. 7: Trees, Graphs, Networks

Level 1: Visualization techniques for Trees, Graphs, Networks

Experiment No. 8: Advanced Visualization Tools

Level 1: Design effective Visualizations for the given scenario

Level 2: Implement Visualizing of Geospatial Data and Document Visualization

Experiment No. 9: Analyze Visualization Systems

REGISTRAR

Level 1: Analyze Visualization Systems

Targeted Application & Tools that can be used:

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data. In the world of Big Data, data visualization tools and technologies are essential to analyze massive amounts of information and make data-driven decisions.

Tools:

1) R Programming
2) Python
3) Tableau
4) SAS
5) Excel
6) RapidMiner
7) IBM Cognos Analytics
8) Microsoft Power BI

Project work/Assignment: Mention the Type of Project /Assignment proposed for this course

After completion of each module a Data analysis or programming based Assignment/Assessment will be conducted.

Mini Project:

Perform exploratory data analysis on a given dataset and provide insights on the same.

1. Crunchbase – Find business information about private and public companies. You can look up how many investments they had, who the founding members are, and if they had any mergers or acquisitions.

2. Glassdoor Research – Glassdoor offers data related to employment. You can, for example, figure out how much you can save by retaining employees.

3. Open Corporates – Open Corporates is the largest open database of companies and company data in the world. Used by banks and governments, they pride themselves on having the most accurate data.

4. **FBI Uniform Crime Reporting –** The Uniform Crime Reporting compiles statistical crime reports, publications, and data points from thousands of cities, universities, states, and federal law enforcement agencies.

5. **Uppsala Conflict Data Program** – The Uppsala Conflict Data Program (UCDP) provides data on organized crime and civil war around the world.

6. **National Institute on Drug Abuse** – The National Institute on Drug Abuse (NIDA) monitors the prevalence and trends regarding drug abuse in the United States.

7. **DBpedia** – DBpedia aims to make Wikipedia's information easily searchable via SPARQL queries or by downloading their information directly. For instance, you can search for NBA players born in the 80s, in cities with more than 1M inhabitants.

8. **Google Trends** – Google Trends allows you to look at what's going on in the world. It gives you data about what's becoming popular, and how much people are searching for a particular term.

9. **Instagram API** – Facebook allows you to use Instagram's API to quickly access comments, metadata, and metrics.

10. **Comtrade** – Official trade in goods and services data sets managed by the UN COMTRADE database. There are data visualization tools and an API and other extraction tools available.

11. **Datahub – Stock Market** – From gold prices, NASDAQ listings, to S&P 500 companies, you'll find it all on datahub.io

12. **Global Financial Data** – Global Financial Data gives you exactly what it says on the tin: data about the finances of the world. Ranges from real estate, global macro data, to market data.

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

13. **IMF Data** – The IMF, or International Monetary Fund, is an organization that aims to foster monetary collaboration between countries. You can find data on trade, government finance, and financial development.
14. **The Atlas of Economic Complexity** – The Atlas of Economic Complexity provides data about global trade dynamics over time. Want to know the quantity of textiles China exported to South Korea? Easy.
15. **World Bank** – Not only does the World Bank provide financial data about countries, but it also provides data on education and health.
16. **Financial Times Data** – Here you'll find cold, hard numbers about the different markets in the world. Data include fluctuations in currency, yield rates of bonds, and commodity prices.

Text Book

7) Wilke, Claus O. Fundamentals of data visualization: a primer on making informative and compelling figures. O'Reilly Media, 2019.
8) Ward, Grinsten, Keim. Interactive Data Visualization: Foundations, Techniques, and Applications, A K Peters/CRC Press,2nd Edition, 2015

References

1) Mohammed J. Zaki, and Wagner Meira Jr., "Data Mining and Analysis: Fundamental Concepts and Algorithms", Cambridge University Press, 2016
2) I.H. Witten and E. Frank, Data Mining: Practical Machine learning tools and techniques Morgan Kaufmann publishers; 3rd Edition, 2011

Topics relevant to development of "Foundation Skills": Analysis of Large data and create graphics for easy understanding.

Topics relevant to development of "Employability": Real time decision-making application development using Data visualization tools.

Catalogue prepared by	Dr.S.Thiruselvan
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE 5010	**Course Title: Robotic Process Automation** **Type of Course: Discipline Elective** **Theory only**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	Basic knowledge of computer and internet, Basic programming skill, Logical Thinking.				
Anti-requisites	**NIL**				

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

<table>
<tr><td>Course Description</td><td colspan="4">The purpose of this course is to enable the students to appreciate the need for Robotic Process Automation and course offers comprehensive knowledge and professional-level skills focused on developing and deploying software robots using UiPath Platforms. The course is both conceptual and Practical in nature and needs basic knowledge of Computer Programming. The course assumes no prior knowledge of RPA. It begins by refreshing basic programming skills and introducing basic RPA concepts. The course develops skills to identify task which can be automated and develop it with UiPath Studio. The course also enhances the programming abilities through assignments.</td></tr>
<tr><td>Course Outcomes</td><td colspan="4">On successful completion of the course the students shall be able to:
1. Explain the concept of automation.
2. Describe various programming constructs in RPA.
3. Identify different applications which can be automated.
4. Apply automation to various concepts related to AI and ML algorithms.</td></tr>
<tr><td>Course Content:</td><td colspan="4"></td></tr>
<tr><td>Module 1</td><td>Introduction to Programming Concepts and RPA Basics</td><td>Assignment</td><td>Data Analysis</td><td>10 Classes</td></tr>
<tr><td colspan="5">Topics:
Programming Concepts Basics-1: Software applications, Introduction to Programming, Data and data structure, Algorithms, Sequence and Flow, Software Development Guidelines. Programming Concepts Basics-2: Compiler and execution, Scripting and Macro, Frameworks and Languages, Information Sharing Mechanism, Variables and Arguments, Files and File Types, Access Control.
RPA Basics: Automation and RPA, Programming Constructs in RPA, Robots in RPA, RPA in Business and Technology.</td></tr>
<tr><td>Module 2</td><td>RPA Advanced Concepts</td><td>Assignment</td><td>Build own bots</td><td>10 Classes</td></tr>
<tr><td colspan="5">Topics:
RPA Advanced Concepts: Setting up the Center of Excellence, RPA Project Methodology, The RPA Journey, RPA in the Emerging Ecosystem.
Introduction to UiPath: The Basics of UiPath Studio Installation, The User Interface, the various steps involved in the automation projects, The installation of UiPath extensions.
Variables: Variables, Types of Variables, Variables in UiPath, Arguments, Namespaces.
Control Flow: Control Flow & Universal Statements, Control Flow Statements in UiPath, Practical Exercise</td></tr>
<tr><td>Module 3</td><td>Automation with various Data</td><td>Assignment</td><td>Build own bots</td><td>10 Classes</td></tr>
<tr><td colspan="5">Topics:
Data Manipulation: Basics of Data Manipulation, Data Types, Data Manipulation Operations, Text Manipulation
Recording and Advanced UI Interaction: UiPath Recording, Input/Output Methods, Data Scraping, Advance Scraping Techniques.
Selectors: Selectors, Types of Selectors, Customization, Debugging
Image Text Advanced Citrix: Image and text-based automation, Keyboard based, automation, Information Retrieval, Native Citrix Automation challenges, Best Practices
Excel Data Tables and PDF: Basics of Excel and Data Table, Extracting Data from Data table, Anchors</td></tr>
<tr><td>Module 4</td><td>Advanced Automation and Orchestrator</td><td>Case Study</td><td>Data Collection and Team Project</td><td>10 Classes</td></tr>
<tr><td colspan="5">Topics:</td></tr>
</table>

REGISTRAR

Email Automation: Introduction to Email Automation, Email Automation in UiPath Studio, Practice retrieving and sending emails
Debugging and Exception Handling: Exception Handling, Debugging Tools, Workflow Designs, Catching errors
Project Organization: Project Organization, Process, Library, Robotic Enterprise Framework
Orchestrator: Introduction to Orchestrator, Processes, Robots in Orchestrator, Working with Orchestrator
Future Trends: Artificial Intelligence, Autonomous things, Digital Assistant, Computing

Targeted Application & Tools that can be used:

Targeted employment sector is service provider and control monitor like GE, Siemens, TCS etc. Targeted job profiles include digital domain and Service based indusrty etc.

Tools:

- UiPath Studio/StudioX

Project work/Assignment:

Term Assignments:

Project 1: Sales order entry Robot
Project 2: E-Mail auto responder Robot
Project 3: Disk Monitoring Robot

Text Book

8. "Robotic Process Automation using UiPath StudioX", Adeel Javed, Anum Sundrani, Nadia Malik, Sidney Madison Prescott, Apress, 2021

References

1. "Learning Robotic Process Automation", Alok Mani Tripathi, Packetz, 2018.
2. https:// academy.uipath.com/

Topics relevant to development of "FOUNDATION SKILLS": Get introduced to RPA Studio and RPA developer Tools .

Topics relevant to "HUMAN VALUES &PROFESSIONAL ETHICS": Set of standard procedures to build RPA - BOTs.

Catalogue prepared by	Prof. Raghavendra T S Prof. Kukkala Prudhvi Raj
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE 5011	**Course Title: Data Science with Cloud Computing** **Type of Course: Discipline Elective** **Theory Only**	**L- P- C**	3	0	3

REGISTRAR

Version No.	1.0				
Course Pre-requisites	Having familiarity with the Basic concepts of Mathematics(Linear Algebra, Calculus) and Statistics (Descriptive Statistics, Inferential Statistics),Programming Knowledge(R,Python,Excel)				
Anti-requisites	**NIL**				
Course Description	This course introduces new Transformative, more collaborative way of doing Data Science. It helps in understanding End to End Data Pipeline, Ingesting Data in a server less way and work our way through Data Exploration, Dashboards, Streaming Data all the way to training and making operational a Machine Learning Model.				
Course Outcomes	On successful completion of the course the students shall be able to: 1. Define Data Science and its fundamentals and the process in Data Science. 2. Explain the process of Ingesting Data into the Cloud Platform. 3. Analyze real world problems with Accuracy . 4. Demonstrate the overall organization of Data and Storage.				
Course Content:					
Module 1	**Making Better Decisions Based on Data**	**Assignment**	**Case Study**	**10 Classes**	
Topics: Many Similar Decisions, Role of Data Engineers, The Cloud Makes Data Engineers Possible, The Cloud Turbocharges Data science, Case Studies. Ingesting Data into the Cloud, Airline on Time Performance Data, Scheduling Monthly Downloads.					
Module 2	**Creating Compelling Dashboards**	**Assignment**	**Case Study**	**10 Classes**	
Topics: Explain your Model with Dashboards, Why build a Dashboard First, Loading Data into google Cloud SQL, Creating Google cloud Instance, Interacting with Google cloud Platform, Build our First Model.					
Module 3	**Streaming Data: Publication and Ingest**	**Assignment**	**Case Study**	**10 Classes**	
Topics: Designing the Event Feed, Time Correction, Apache Beam/Cloud Dataflow, Publishing an Event Stream to Cloud Pub/Sub, Real Time Stream processing, Interactive Data Exploration, Exploratory Data Analysis, Loading Flights Data into Big Query, Exploratory Data Analysis in Cloud AI Platform Notebooks, Quality control, Arrival Delay conditioned on Departure Delay, Evaluating the Model.					
Module 4	**Cloud Dataproc**	**Assignment**	**Case Study**	**10 Classes**	
Topics:					

REGISTRAR

Bayes Classifier on Cloud Dataproc, Map Reduce and Hadoop Eco System, Quantization using Spark SQL, Bayes Classification using Pig

Targeted Application & Tools that can be used:

Targeted Industries like Banking, Transport, e-commerce, healthcare and many more are using data science to make optimal Decisions. The usage of data science helps in rising sales. It can explore historic data, make comparisons and analysis of the market and provide recommendations.
Target Jobs Data Scientist, Data Architect, Data Engineer, Statistician.

Tools:

- Apache Spark
- Jupyter
- Weka

Project work/Assignment:

Mini Project:

Walmart Sales Forecasting in Cloud

- Predict the sales across various departments in each store.
- Predict the effect of markdowns on the sales during the holiday seasons.

Term Assignments:

Consider a Dataset on Bird communities that needs to be analyzed. The data has three columns, a date, a common name, and a count of the number of individuals.

- Count the total number of individuals of each species that were seen in each data file.
- Sort based on the total number of individuals.

Text Book

1. "Data Science on the Google Cloud Platform: Implementing End-to-End Real-Time Data Pipelines: From Ingest to Machine Learning"-Valiappa Lakshmanan,1st Edition, January 2018.
2. "Data Analysis in The Cloud"- Domenico Talia ,1st Edition, September 2015

References

1. Doing Data Science, Straight Talk from the Frontline. O'Reilly. 2014.

Topics relevant to development of "FOUNDATION SKILLS": Data Extraction, Data wrangling

Topics relevant to "HUMAN VALUES &PROFESSIONAL ETHICS": Developing Ethical Rules in Building a Model.

Catalogue prepared by	Mrs Sridevi S Mrs Anitha Prem Kumar

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: **CSE 5012**	**Course Title: Artificial Intelligence in Cloud Computing** **Type of Course: Discipline Elective** **Theory Only**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	• Probability and Statistics • Knowledge on programming language • Familiarity with Fundamentals of AI and Cloud Computing				
Anti-requisites	**NIL**				
Course Description	This Course is designed to acquire the ability to deliver intelligent solutions to problems in a variety of domains and business applications such as natural language processing, text mining, robotics, reasoning and problem-solving in AI. The inclusion of AI in the cloud can lead to a more effective synthesis of data systems for identifying valuable information. This information can then be applied practically in business operations. AI in cloud computing can provide users with seamless data access. AI uses data to get things done, which makes it well-suited to cloud environments as they can hold large amounts of data. Topics Includes: AI Cloud Services, Applications of AI, AI Chatbots , Types of Chatbots, Applications of Chatbot, Cloud platforms –Google cloud, Microsoft Azure, AWS, Developing AI Application using AWS sagemaker				
Course Outcomes	On successful completion of the course the students shall be able to: 1. Gain the knowledge on AI Cloud services (Knowledge Level) 2. Understand the various applications of AI (Comprehensive Level) 3. Explain the factors that lead to the growing popularity of chatbots. (Comprehensive Level) 4. Develop the cloud AI application using AWS SageMaker (Application Level)				

REGISTRAR

Course Content:				
Module 1	AI Cloud Services	Assignment	Cloud API	**10 Classes**
Topics: Introduction to AI cloud, why AI cloud, Technologies that support AI platform for business like IBM Watson, Google Cloud Vision, Microsoft Cognitive Services or Natural Language application programming interfaces allow abstract complex AI capabilities via simple API calls.				
Module 2	AI applications	Use case study	Speech Recognition	**10 Classes**
Topics: Language Models – Information Retrieval- Information Extraction – Natural Language Processing – Machine Translation – Speech Recognition – Robot – Hardware –Perception – Planning – Moving				
Module 3	AI chatbot	Assignment	Applications of chatbots	**8 Classes**
Topics: Explaining what a chatbot is, Describe common applications of chatbots, Identifying factors that drive the growing popularity of chatbots , two main systems in use that bots use to recognize intent and extract entities: Rules-based systems and Natural language understanding Listing examples of tools and services that you can use to create chatbots, Designing a chatbot conversation.				
Module 4	**Cloud-native AI application development**	use case study	Create and deploy AI Application using AWS cloud platform	**10 Classes**
Topics: MLOps: Train, test, and deploy Deep Learning models using containers on a cloud server - Hands-on end-to-end cloud AI applications development and deployment using AWS SageMaker,-Background and cloud platform introduction (AWS, Google Cloud Platform, and Azure Platform),- Hands-on AI application development with APIs provided by the main cloud platforms				
Targeted Application & Tools that can be used: • Google Vertex AI is an integrated suite of machine learning tools and services for building and using ML models with AutoML or custom code. It offers both novices and experts the best workbench for the entire machine learning development lifecycle.				
Project work/Assignment:				

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Mini Project: Build a dynamic mobile chatbot powered with AI

1. Create Watson services with IBM Cloud.
2. Update the details in the back-end application.
3. Deploy the back-end application.
4. Set up IBM Cloud Functions.
5. Set up Watson Assistant.
6. Set up IBM Mobile Foundation Server and CLI.
7. Set up Google Cloud Anchors.
8. Configure the Android mobile app.
9. Build and run the Android mobile app.

Term Assignments:

IBM Watson: IBM Watson is that developers can use this platform to build their AI applications. It's an open AI for any cloud environment, and it's pre-integrated and pre-trained on flexible information architecture. This will expedite the development and deployment of AI application.

Text Book

1. **Micheal Lanham "Practical AI on the Google Cloud Platform", O'Reilly Media, 2020**
   E.Book-Practical AI on the Google Cloud Platform (21h.io)

2. Anand Deshpande, Manish Kumar, Vikram Chaudhari , "Hands-On Artificial Intelligence on Google Cloud Platform: Build intelligent applications powered by TensorFlow, Cloud AutoML, BigQuery, and Dialogflow" , Kindle Edition, 2020

References

1. "Cloud Computing: Principles and Paradigms" by Rajkumar Buyya (Editor), James Broberg (Editor), Andrzej M. Goscinski (Editor), WILEY, First Edition, March 29, 2011
2. Deepak Khemani "Artificial Intelligence", Tata Mc Graw Hill Education 2013.
3. Stuart Russel and Peter Norvig "AI – A Modern Approach", 2nd Edition, Pearson Education 2007.

Topics relevant to development of "Employability": Data Scientist using the Cloud - Data Scientists have to work with a variety of data (structured, semi-structured, unstructured), analytics tools, and programming languages by leveraging the cloud platform.

Topic relevant to HUMAN VALUES & PROFESSIONAL ETHICS": Naming and coding convention for Project Development.

Catalogue prepared by	Mrs Anitha Premkumar & Mrs Sridevi
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE 5013	**Course Title: Soft Computing** **Type of Course: Discipline Elective** **Theory Only**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	**Calculus, Probability, Linear Algebra and Basic Programming Skills**				
Anti-requisites	**NIL**				
Course Description	Soft computing is an emerging approach in computing that mimics the human mind's remarkable ability to reason and learn in an environment of uncertainty and imprecision. Soft computing is based on biologically inspired methodologies such as genetics, evolution, ant behaviors, particle swarming, human nervous systems, etc. Soft computing is the only solution when we don't have any mathematical modeling of problem-solving (i.e., algorithm), needs a solution to a complex problem in real-time, and easily adapts with changing scenarios and is implemented with parallel computing. It has enormous applications in many application areas such as medical diagnosis, computer vision, handwritten character reconditions, pattern recognition, machine intelligence, weather forecasting, network optimization, VLSI design, etc.				
Course Outcomes	On successful completion of the course the students shall be able to: 1. Define the concept and applications of Soft Computing. 2. Discuss Fuzzy logic concepts and its applications. 3. Demonstrate Artificial Neural Networks concepts and its applications. 4. Apply Evolutionary algorithms and hybrid soft computing techniques.				
Course Content:					
Module 1	Introduction Soft Computing	Assignment	Analysis		**9 Classes**
Topics: Introduction to Soft Computing: Concept of computing systems, "Soft" computing versus "Hard" computing, Characteristics of Soft computing, Applications of Soft computing techniques.					
Module 2	Fuzzy Logic	Assignment	Analysis, Data Collection		**12 Classes**
Topics:					

REGISTRAR

Fuzzy Logic: Introduction to Fuzzy logic. Fuzzy sets and membership functions. Operations on Fuzzy sets. Fuzzy relations, rules, propositions, implications and inferences. Defuzzification techniques. Fuzzy logic controller design. Some applications of Fuzzy logic.

Module 3	Neural Networks	Case Study	Analysis, Data Collection	**10 Classes**

Topics:
Neural Network: Biological and Artificial Neuron, Neural Networks, Supervised and Unsupervised Learning. Single Layer Perceptron, Multilayer Perceptron, Backpropagation Learning.

Neural Networks as Associative Memories: Hopfield Networks, Bidirectional Associative Memory. Topologically Organized Neural Networks: Competitive Learning, Kohonen Maps.

Module 4	Evolutionary Computing	Assignment	Analysis, Data Collection	**10 Classes**

Topics:
Evolutionary Computing: Concept of "Genetics" and "Evolution". Genetic Algorithm and Optimization, The Schema Theorem, GA operators: Encoding, Crossover, Selection, Mutation, etc. Introduction to ant colony optimization and particle swarm optimization. Integration of genetic algorithm with neural network and fuzzy logic.

Targeted Application & Tools that can be used:

In recent times, engineers have very well accepted soft computing tools such as Fuzzy Computing, ANN, Neuro-Computing and Evolutionary Computing, etc., for carrying out various numerical simulation studies. In the last two decades, these tools independently and in hybrid forms have been successfully applied to varieties of problems. The main objective is to introduce students to the latest soft computing tools. The training of these tools will be helpful to develop rigorous applications in the engineering domain.

Tools:

- MATLAB
- PYTHON
- C

Project work/Assignment:

Mini Project:

- Training of known/classified datasets representing some objects/pattern using various ANN learning methods including Perceptron, BPN, Adaline, Associative memory networks, Hopfield, kohenen networks.
- Classification of new input feature set/pattern based on training & learning
- Applying GA search to optimize the solutions. Implementation of the GA procedure.

Term Assignments:

- Applications of soft computing techniques in solving day today problems.
- Solving Traveling salesman problem using Genetic Algorithm and comparing different mutation operators with the same

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Text Book

1. Principles of Soft computing, Shivanandam, Deepa S. N Wiley India, 3rd Edition 2019 2. Timothy J. Ross, "Fuzzy Logic with Engineering Applications", Third Edition, Wiley.
References 1. Kumar S., "Neural Networks - A Classroom Approach", Tata McGraw Hill, 2nd Edition 2017. 2. Eiben A. E. and Smith J. E., "Introduction to Evolutionary Computing", Second Edition, Springer, Natural Computing Series, 2nd Edition, 2015. 3. Fakhreddine O. Karray, and Clarence W. De Silva. Soft computing and intelligent systems design: theory, tools, and applications. Pearson Education, 2009.
Topics relevant to development of "Employability": ": Solving real world problems with uncertainty using Nature Inspired Algorithms

Catalogue prepared by	Dr.S.Thiruselvan
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE 5014	**Course Title: Ontology Engineering for the Semantic Web** **Type of Course: Discipline Elective** **Theory Only**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	Basics Programming Skills Basics of XML-Syntax, Schema, Elements, Attributes and Namespaces				
Anti-requisites	**NIL**				
Course Description	This course presents the basics of semantic web and Ontology engineering. This course consist of the detailed description RDF frameworks. This course is designed with theoretical material on ontology design, Description Logics, and developing ontologies using OWL. The course uses the Protege-OWL environment.				
Course Outcomes	On successful completion of the course the students shall be able to: 1] Understand the semantic web basics, architecture and technologies. 2] Describe the semantic relationships among the data elements using Resource Description Framework (RDF) 3] Analyze the conventional web with semantic web. 4] Able to design and implement real-world applications that "discovers" the data and/or other web services via the semantic web				
Course Content:					

Module 1	Introduction	Assignment	Analysis, Data Collection	**9 Classes**

Topics:
Introduction to the Syntactic web and Semantic Web, Evolution of the Web, The visual and syntactic web, Levels of Semantics, Metadata for web information, The semantic web architecture and technologies, Contrasting Semantic with Conventional Technologies, Semantic Modeling -Potential of semantic web solutions and challenges of adoption.

Module 2	Ontological Engineering	Assignment	Analysis, Data Collection	**9 Classes**

Topics:
Ontologies, Taxonomies, Topic Maps, Classifying Ontologies, Terminological aspects: concepts, terms, relations between them, Complex Objects, Subclasses and Sub properties, definitions, Upper Ontologies, Quality, Uses, Types of terminological resources for ontology building, Methods and methodologies for building ontologies, Multilingual Ontologies, Ontology Development process and Life cycle, methods for Ontology Learning, Ontology Evolution, Versioning.

Module 3	Describing the Web Resources	Assignment	Data analysis task	**9 Classes**

Topics:
RDF Overview, The basic elements of RDF, RDF triples, Fundamental rules of RDF Aggregation and distributed information, RDF tools, RDFS, Taxonomy, and Ontology, Need for RDFS, Core elements of RDFS

Module 4	Web Ontology Language and Real world examples	Case Study	Analysis, Data Collection	**11 Classes**

Topics:
Requirements for Ontology Languages, OWL Sub languages, Description of the OWL Language, Layering of OWL, Examples for OWL, OWL in OWL, Namespaces, Classes of Classes, Class Equivalence, Building Classes from Other Classes, Restricting Properties of Classes.

SWOOGLE and FOAF: basics, architecture, usage and examples.

Targeted Application & Tools that can be used:

Enterprise applications. A more concrete example is SAPPHIRE (Health care) or Situational Awareness and Preparedness for Public Health Incidences and Reasoning Engines which is a semantics-based health information system capable of tracking and evaluating situations and occurrences that may affect public health.

Geographic information systems bring together data from different sources and benefit therefore from ontological metadata which helps to connect the semantics of the data.

Domain-specific ontologies are extremely important in biomedical research, which requires named entity disambiguation of various biomedical terms and abbreviations that have the same string of characters but represent different biomedical concepts.

Tools:

- **Protégé**
- **Neon Toolkit**
- **SWOOP**
- **Vitro**

Project work/Assignment:

Mini Project:

- Ontology-Based Model For The “Ward-round” Process in Healthcare
  To design an ontology based model that can fix information flow problems in the ward-round process of hospital unit. This can used to provide relevant information to the domain users according to their

needs and demands. The domain users profiling and describes their roles, information demands with competencies: skills, qualifications and experiences. The ontology based model will be implemented in OWL language that can be used in an application to support ward-round activities for achieving effective patient's treatment process.

Team Assignments:

- **Write a Short survey about Ontology Languages**

  Study and write few points about classifications of Ontology Languages. The classification can be done by structure and syntax.

- **Develop a simple ontology model using protégé tool.**

Text Book/

9. Grigoris Antoniou, Frank Van, "Semantic Web Primer", MIT Press, 2008
10. Karin K. Breitman, Marco Antonio Casanova and Walter Truszowski, "Semantic Web Concepts: Technologies and Applications", Springer, 2007

References Books

1. LiyangYu , "Introduction to the Semantic Web and Semantic web services" Chapman & Hall/CRC, Taylor & Francis group, 2007
2. Peter Mika, "Social networks and the Semantic Web", Springer, 1st edition 2007
3. Robert M. Colomb, "Ontology and the Semantic Web", Volume 156 ,Frontier in Artificial Intelligence and Applications, IOS Press, 2007
4. Michael C. Daconta, Leo J. Obrst, and Kevin T. Smith, "The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management", Fourth Edition, Wiley Publishing, 2003.

Topics relevant to "ONTOLOGY ENGINEERING and " SEMANTIC WEB"
Employability Skills

Catalogue prepared by	Dr. S. Pravinth raja
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE 5015	**Course Title: Data Security and Access Control** **Type of Course: Discipline Elective** **Theory Only**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	**Basics of Cryptography, Networking**				
Anti-requisites	**NIL**				

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Course Description	This course describes fundamental issues and problems in data security and provides technical solutions or facets to the problem of achieving data security. The course also deals with the security of statistical databases, discusses authorization systems, and covers the fundamental ideas of cryptography.			
Course Outcomes	On successful completion of the course the students shall be able to: 1] Describe the basic concepts of a Data Security 2] Apply appropriate techniques for security Algorithms 3] Explain the Access Controls mechanisms 4) Simulate data security algorithms for achieving access control			
Course Content:				
Module 1	Fundamentals of Data Security	Assignment	Algorithms	**8 Classes**
Topics: Introduction to Data Security, Confidentiality, Integrity, Availability, Visibility, Security as Code, Automation, Monitoring, Models and Methodology				
Module 2	Data Security Techniques	Assignment/ Case Study	Presentation	**10 Classes**
Topics: Introduction, data masking, data erasure, and backup storage, Anti-malware protection, operating system updates, Security in Key specified model, Security in Characteristic specified model				
Module 3	Authorization Mechanisms in Data Security	Assignment/ Case Study	Coding	**12 Classes**
Topics: Introduction, concept of Un-decidability, Authorization Systems with Tractable Safety Problem, Authorization Systems with Tractable Safety Problem, Grammatical Authorization Systems, Regular Authorization Systems				
Module 4	An Overview of Data Security Tools, Data Security Policies	Assignment/ Case Study	Simulation of DS tools	**8 Classes**
Topics: Introduction to tools available for Data Security, Demonstration of Security features in Linux platform, Simulate using more than two computers, demonstrate data leakage during transmission, GDPR (General Data Protection Regulation), Comparative study with India regulation, Data Privacy Act, Role Based Access Control				
Targeted Application & Tools that can be used: Anomaly Deduction , Inclusion Prevention Systems, Firewall, Email Security **Tools:** SAGE Mathematical Library package, VPN				
Project work/Assignment:				

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Term Assignments:

1. Implement Cryptographic algorithms using SAGE
2. Comparative Study on Various Data Security Tools
3. Case Study on GDPR - General Data Protection Regulation
4. Identify Data Leakage in LINUX environment using Authorization Mechanisms

Text Book

1. Data Privacy and Security, David Solomon, Springer,

2. Principles Of Data Security, Ernst L. Leiss, Plenum Press. New York And London

References

1. Intelligence and Security Informatics for International Security, Chen, Hsinchun, Springer Publication 2006
2. Certified Information Security Professional (CSIP) web portal
3. https://www.sage.com/en-us/products/sage-toolops/
4. https://gdpr.eu/

Topics relevant to development of "Employability": ": Email Security, Web Security

Catalogue prepared by	Mr. Murthy D H R
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE5016	**Course Title: Essentials for Machine Learning (ML)** **Type of Course: Theory Course**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	**NIL**				
Anti-requisites	**NIL**				

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

<table>
<tr><td>Course Description</td><td colspan="4">Machine learning has been emerged as a promising paradigm in the field of Computer science having applications in a wide variety of fields such as communication networks, bioinformatics, image processing, antenna design etc. Now a days people from multiple discipline are interested in Machine learning due to its applicability in predicting behaviors of highly complex systems, which is otherwise difficult based on traditional optimization techniques in a time bound fashion. The goal of this course is to provide the mathematical prerequisite for starting any Machine learning course to the students coming from various engineering disciplines. This course does not require any prerequisite. The goal of the course is:
1. To introduce basic probability and statistics concepts.
2. To introduce basic Linear Algebra concepts.
3. To enable the students to understand Machine Learning/Deep learning concepts in future.</td></tr>
<tr><td>Course Outcomes</td><td colspan="4">On successful completion of this course the students shall be able to:
1. Understand the basic concepts of Probability and Statistics.
2. Understand the basic concepts of Linear Algebra.
3. Peruse courses on Machine learning/Deep learning in future.</td></tr>
<tr><td>Course Content:</td><td colspan="4"></td></tr>
<tr><td>Module 1</td><td>Probability</td><td>Assignment</td><td>Sample space and Events, Interpretation and axioms of Probability, Conditional Probability, Multiplication and total Probability rules, Independence, Bayes’ theorem</td><td>08 Classes</td></tr>
<tr><td colspan="5">Topics:
Sample space and Events, Interpretation and axioms of Probability, Conditional Probability, Multiplication and total Probability rules, Independence, Bayes’ theorem.</td></tr>
<tr><td>Module 2</td><td>Random variables</td><td>Assignment</td><td>Probability distribution, Probability mass function, Probability density function, Cumulative distribution function, Mean and variance of a random variable, Binomial, Poisson and Normal random variables, relation between them.</td><td>08 Classes</td></tr>
<tr><td colspan="5">Topics:

Probability distribution, Probability mass function, Probability density function, Cumulative distribution function, Mean and variance of a random variable, Binomial, Poisson and Normal random variables, relation between them.</td></tr>
</table>

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Module 3	Introduction to Statistics	Assignment	Pie Chart, Bar chart, Box and whisker plot, Mean, Median, Mode, AM, GM, HM, Quartiles, Deciles, Percentiles, Moments, Skewness, Kurtosis, Measures of Central tendency, Software demonstration.	**08 Classes**
Topics: Pie Chart, Bar chart, Box and whisker plot, Mean, Median, Mode, AM, GM, HM, Quartiles, Deciles, Percentiles, Moments, Skewness, Kurtosis, Measures of Central tendency, Software demonstration.				
Module-4	Estimation of Parameters	Assignment	Point estimation, Sampling distribution, Central Limit Theorem, Unbiased estimators, Method of point estimation, Method of moments, method of maximum likelihood, confidence interval estimates of population parameter, student's t distribution, Testing of hypothesis, Chi square distribution, Degrees of freedom	**06 Classes**
Topics: Point estimation, Sampling distribution, Central Limit Theorem, Unbiased estimators, Method of point estimation, Method of moments, method of maximum likelihood, confidence interval estimates of population parameter, student's t distribution, Testing of hypothesis, Chi square distribution, Degrees of freedom				
Module-5	Linear Algebra	Assignment	Scalar, Vector, Matrices and Tensors, Norms, Span, Eigen Value, Eigen Vector, The trace operator, Determinant, Proximity measure, Example: Principal Component Analysis.	**06 Classes**
Topics: Scalar, Vector, Matrices and Tensors, Norms, Span, Eigen Value, Eigen Vector, The trace operator, Determinant, Example: Principal Component Analysis				
Project work/Assignment: **1. Assignment 1 on Probability Theory (Module 1 and Module 2).** **2. Assignment 2 on Statistics (Module 3 and Module 4).** **3. Assignment 3 on Linear Algebra (Module 5).**				
REFERENCE MATERIALS: Text Book(s): 1. Douglas C. Montgomery and George C. Runger, "Applied Statistics and Probability for Engineers", Sixth Edition, Wiley, 2016				

2. Dimitri P. Bertsekas and John N. Tsitsiklis, “Introduction to probability”, MIT press, FALL 2000.
3. Murry R Spiegel and Larry J Stephens, “STATISTICS”, Fourth Edition, Schaum’s outlines, 2008.
4. Narsingh Deo, “System simulation with digital computer”, PHI.
5. G. Strang, “Introduction to Linear Algebra”, Fifth Edition, 2016, Wellesley-Cambridge Press, ISBN: 978-09802327-7-6.

Topics relevant to development of **“Foundation Skills”:** Fundamentals of Probability, **“Skill Development” –** Stochastic simulation techniques, **“Employability”**-

Catalogue prepared by	Dr. Shankar K. Ghosh, Dr. Alamelu Mangai Jothidurai.
Recommended by the Board of Studies on	BOS NO: 13th BOS, held on 08/12/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 17, Dated 11/12/2021

Course Code: CSE 5017	**Course Title: Machine Vision** **Type of Course: Discipline Elective** **Theory Only**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	**Basic Mathematics (**Linear algebra, vector calculus, and probability)**, MAT LAB / Open CV**				
Anti-requisites	**NIL**				
Course Description	This course provides an introduction to computer vision including fundamentals of image formation, camera imaging geometry, feature detection and matching, stereo, motion estimation and tracking, image classification and scene understanding. We'll explore methods for depth recovery from stereo images, camera calibration, automated alignment, tracking, boundary detection, and recognition. We'll use both classical machine learning and deep learning to approach these problems. The focus of the course is to develop the intuitions and mathematics of the methods in lecture, and then to learn about the difference between theory and practice in the projects.				
Course Outcomes	On successful completion of the course the students shall be able to: 1. Describe Image formation and Camera Models [Knowledge] 2. Classify techniques for Local feature extraction and tracking [Comprehension]				

REGISTRAR

	3. Apply the different category of calibration methods and dimension reconstruction approach for computer vision[Application]			
Course Content:				
Module 1	Basic Concept of Image Processing	Mini Project	Mapping Facial Features	**12 Classes**
Introduction to Image Processing-Basic mathematical concepts: Image enhancement: Grey level transforms, Spatial filtering. Extraction of special features: edge and corner detection. Morphological processing, Image transforms, Discrete Fourier Transform, Fast Fourier Transform. Frequency domain enhancement.				
Module 2	Image Segmentation	Mini Project	Hand gesture recognition	**14 Classes**
Image Segmentation Algorithms: contextual, non-contextual segmentation, texture segmentation. Feature Detectors and Descriptors, Feature Matching-Object Recognition, The Use of Motion in Segmentation Optical Flow & Tracking Algorithms, Face detection (Viola Jones), Face Recognition.				
Module 3	Image Dimensions	Mini Project	Surveillance	**14 Classes**
2D and 3D feature-based alignment, Pose estimation, Geometric intrinsic calibration, -Camera Models and Calibration: Camera Projection Models – orthographic, affine, perspective, projective models. Projective Geometry, transformation of 2D and 3D, Internal Parameters, Lens Distortion Models, Calibration Methods – linear, direct, indirect and multi plane methods. Visual servo. Stereo correspondence-Epipolar geometry, Fundamental matrix, Introduction to SLAM (Simultaneous Localization and Mapping).				
Targeted Application & Tools that can be used: Computer Vision applications are **used for traffic sign detection, surveillance and recognition**. Vision techniques are applied to segment traffic signs from different traffic scenes (using image segmentation) and algorithms to recognize and classify traffic signs. **Tools:** • MAT Lab/Open CV				
Project work/Assignment:				
Project Work: 1. **Detect the faces of humans by mapping facial features from a video or an image. There are several steps involved in these projects, such as mapping features.** 2. **Hand gesture recognition is one of the critical topics for human-computer interaction. In this project, there are several tasks which are needed to be performed. This includes the hand region, which is to be extracted from the background, followed by segmenting the palms and fingers to detect finger movements.** 3. **Count the number of people passing through a specific scene. The applications of this project include civilian surveillance, pedestrian tracking, pedestrian counting, etc.** 4. **Design, implement and test on several regions on a set of images based on the segmentation algorithms.**				
Text Book 1. R. C. Gonzalez, R. E. Woods, 'Digital Image Processing', Pearson,2017				

REGISTRAR

2. Introduction to Computer Vision and its Application, Richard Szelinski,2021
References 1 . Emanuele Trucco and Alessandro Verri, "*Introductory Techniques for 3-D Computer Vision*", Prentice Hall, 1998. 2 Olivier Faugeras, "*Three Dimensional Computer Vision*", MIT Press, 1993. 3 Richard Szeliski, "*Computer Vision: Algorithms and Applications*", Springer, 2011. 4 Milan Sonka, Vaclav Hlavac and Roger Boyle, "*Image Processing, Analysis and Machine Vision*", Third Edition, CL Engineering, 2013. 5.Marco Treiber, "An Introduction to Object Recognition Selected Algorithms for a Wide Variety of Applications", Springer, 2010. 6.Forsyth and Ponce, "Computer Vision – A Modern Approach", Second Edition, Prentice Hall, 2011.
Topics relevant to development of "FOUNDATION SKILLS", " IMAGE SEGEMENTATION and DIMENSIONS of Image Processing- We compare IMAGE PROCESSING/ COMPUTER VISION jobs with Information Technology service oriented jobs then obviously there is relatively limited scope. But things are changing very fast as time is changing. Scope of image processing/computer vision jobs is increasing day to day.

Catalogue prepared by	Dr. R Vignesh and Dr. Arul Murugan Ramu
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: **CSE 6001**	**Course Title: Deep Learning** **Type of Course: Program Core Theory and Laboratory Integrated**	**L-P-C**	2	2	3
Version No.	1.0				
Course Pre-requisites	• Data Mining and Machine Learning fundamentals • Basic working knowledge of Statistics and Probability				

REGISTRAR

	• Familiarity with programming languages and hands on coding
Anti-requisites	NIL
Course Description	The course introduces the core intuitions behind Deep Learning, an advanced branch of Machine Learning involved in the development and application of Artificial Neural Networks that function by simulating the working principle of human brain. Deep learning algorithms extract layered high-level representations of data in a way that maximizes performance on a given task. The course includes theory and lab components which emphasizes on understanding the implementation and application of deep neural networks in various prominent problem domains like speech recognition, sentiment analysis, recommendations, and computer vision etc. The course facilitates the students to interpret and appreciate the successful application of deep neural nets in various prediction and classification tasks of ML.
Course Out Comes	On successful completion of the course the students shall be able to: 21) Apply basic concepts of Deep Learning to develop feed forward models 22) Apply Supervised and Unsupervised Deep Learning techniques to build effective models for prediction or classification tasks 23) Identify the deep learning algorithms which are more appropriate for various types of learning tasks in various domains of Machine Learning and Machine vision. 24) Analyze performance of implemented Deep Neural models
Course Content:	

Module 1	Introduction to Deep Learning	Assignment	Programming	**No. of Classes:10**
Topics: Machine Learning in a nutshell, Fundamentals of deep learning and neural networks, Deep Neural Network, Feedforward Neural Network, , Perceptron, MLP Structures, Activation Functions, Loss Functions, Gradient Descent, Back-propagation, Training Neural Networks Building your Deep Neural Network: Step by Step, Deep Neural Network for Classification.				
Module 2	Improving Deep Neural Networks	Assignment	Programming	**No. of Classes:09**
Topics: Hyperparameter tuning, Initialization, Overfitting and Underfitting, Regularization and Optimization, Dropout, Batch Normalization				
Module 3	Deep Supervised Learning Models	Assignment	Programming	**No. of Classes:10**
Topics:				

REGISTRAR

Convolutional neural network, Prediction of image using Convolutional Neural Networks, Deep learning in Sequential Data, RNN & LSTM, GRU, Sentiment Analysis				
Module 4	Deep Unsupervised Learning	Assignment	Programming	**No. of Classes:10**

Topics:

Basics of Deep unsupervised learning, Auto encoders, Restricted Boltzmann Machine, Recommender systems

List of Laboratory Tasks:

Experiment No. 1: Programming assignment to implement a single layer feed forward neural network from scratch (Application: A basic neural network).

Level 1: Programming scenario to implement a basic single layer feed-forward neural network perceptron.

Level 2: Programming scenario to implement a basic single layer feed-forward neural network with a single hidden layer having ReLU activation function and sigmoid in the output layer.

Experiment No. 2: Programming assignment to build an Artificial Neural Network by implementing the Backpropagation algorithm and test the same using appropriate data sets.

Level 1: Programming scenario to use the Backpropagation algorithm to build an ANN and run it on a dataset for few epochs.

Level 2: Programming scenario to use the Backpropagation algorithm to build an ANN and run it on a dataset for few epochs and interpret the accuracy, loss and other evaluation parameters.

Experiment No. 3: Programming assignment to build a multiple layer neural network with specific model parameters and hyperparameters on a given real life dataset.

Level 1: Programming assignment to implement a MLP with

- possibility to use 2-4 layers
- ReLU for the hidden layer
- Sigmoid in the output layer
- optimization via gradient descent (GD)

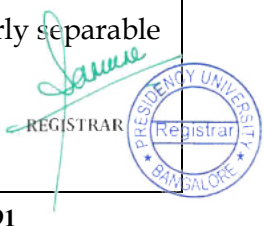
Level 2: Programming assignment to implement the neural network and add some more hyperparameters in the perceptron model

- softmax output layer
- optimization via stochastic gradient descent (SGD)
- Gradient checking code (!!!)

Generate the confusion matrix

Experiment No. 4: Programming assignment to implement classification of linearly separable Data with a Deep neural network (Application: Binary classification).

Level 1: Programming scenarios to build a binary classifier with a deep ANN.



Level 2: Programming scenarios to build a binary classifier with a deep ANN

- Weight initialization with random noise (!!!) (use normal distribution with changing std. deviation for now)
- implement dropout, *l2* regularization
- implement a different optimization scheme (RPROP, RMSPROP, ADAGRAD)
- employ batch normalization

Experiment No. 5: Programming assignment to implement a basic Convolution Neural Network.

Level 1: Programming scenarios which use the concept of convolution and pooling to implement a CNN.

Level 2: Programming scenarios which use the concept of convolution and pooling to implement a CNN and also specify some parameters like number of filters, length of feature detector, stride etc.

Experiment No. 6: Programming assignment to perform image segmentation and object detection using CNNs.

Level 1: Programming assignment to instantiate a CNN (that uses FullyConnectedLayers) and train the neural network using the training data from MNIST data set.

Level 2: Programming assignment to instantiate a CNN (that uses FullyConnectedLayers) and train the neural network using the training data from MNIST data set. Choose appropriate hyper parameters for the training of the neural network. Plot the cost versus training iterations using different mini-batch sizes: 16; 64; 256; 1024. Record the test accuracy in percentage and total training time you spent in seconds. Implement Adam Optimizer. To obtain full marks, the network should be able to achieve a test accuracy of 90% or more across many different random seeds.

Experiment No. 7: Programming assignment to employ CNN in image classification from given dataset.

Level 1: Programming scenario to instantiate a CNN (with at least one convolutional layer) and train the neural network using the training data from CIFAR10 data. Choose appropriate hyperparameters for the training of the neural network. The network should be able to achieve a test accuracy of at least 50% within 10 training epochs.

Level 2: Programming scenario to build a CNN (with more than one convolutional layer) and train the neural network using the training data from CIFAR10 data. Choose appropriate hyperparameters for the training of the neural network. The network should be able to achieve a test accuracy of at least 50% within 10 training epochs. Continue to train further and examine training and testing performance. Report hyperparameters (learning rate, number of hidden layers, number of nodes in each hidden layer, batch size and number of epochs) of the Deep Neural Network. Also, explain the observations.

Experiment No. 8: Programming assignment to perform Sentence (text) Classification using Convolutional Neural Networks.

Level 1: Programming Scenarios to utilize CNN to categorize text data in given datasets like SST movie reviews.

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Level 2: Programming Scenarios to utilize CNN to categorize text data in given datasets like SST and MR movie reviews.

Experiment No. 9: Programming assignment to apply Recurrent Neural Networks for sentiment analysis of text data.

Level 1: Programming scenario to build a model to perform sentiment analysis of IMDB movie reviews using. Reviews are categorized into two polarities: positive and negative.

Level 2: Programming scenario to build a model to perform sentiment analysis of IMDB movie reviews. Reviews are categorized into three polarities: positive, negative and neutral.

Experiment No. 10: Programming assignment to create a generative model for text, character-by-character using Recurrent neural networks.

Level 1: Programming scenario to implement a multi-layer Recurrent Neural Network like LSTM for training/sampling from character-level language models, which takes one text file as input and trains an RNN that learns to predict the next character in a sequence. The RNN can then be used to generate text character by character that will look like the original training data.

Level 2: Programming scenario to implement a multi-layer Recurrent Neural Network utilizing both LSTM and GRU in turns for training/sampling from character-level language models, which takes one text file as input and trains an RNN that learns to predict the next character in a sequence. The RNN can then be used to generate text character by character that will look like the original training data. Train the model and use it to generate new text.

Experiment No. 11: Programming assignment to implement RNN models for multivariate time series forecasting.

Level 1: Programming scenario to implement a many-to-one Recurrent Neural Network for Stock Price forecasting, i.e. trained with a certain number of day's data, the model should predict the stock price of the next day.

Level 2: Programming scenario to implement a many-to-one Recurrent Neural Network for Stock Price forecasting, i.e. trained with a certain number of day's data, the model should predict the stock price of the next day. Students are free to use RNN, GRU, or LSTM (or compare between) and any number of layers and architecture. In the testing, plot the ground truth and your predicted values for 100 days.

Experiment No. 12: Programming assignment to implement Autoencoders and deep Boltzmann's machines.

Level 1: Programming scenario to implement a basic recommender system using deep Boltzmann's machines.

Level 2: Programming scenario to build a recommender system with Collaborative filtering algorithm using deep Boltzmann's machines,

REGISTRAR

Targeted Application & Tools that can be used:

Targeted employment sector is not restricted to any single domain. Today, ML and Dl have been employed for data analysis and improved business intelligence in every sector. Targeted job profiles include Data Analyst, Data Scientist, Data Engineer, Neuroinformatician, Bioinformatician, Image Recognition, Research Analyst, Full Stack Developer for Deep Learning, Natural Language Process Engineer, Business Analyst etc. Few of the top recruiters are Amazon, NVIDIA. Microsoft, IBM, Accenture, Facebook, Intel, Samsung, Lenovo, Adobe etc., among numerous others.

Tools: Neural Designer, AutoML, AutoDL, Keras, TensorFlow, Torch, Google Colaboratory, Spider, Jupiter Notebook

Project work/Assignment:

Throughout the progression in each module, students will have to submit scenario based programming Assignments/Experiments as listed in "List of Lab Tasks". On completion of each module, students will be asked to develop a Mini Project, similar to the following:

- **Music genre classification system**
  This is one of the interesting deep learning project ideas. This is an excellent project to nurture and improve one's deep learning skills. The aim is to create a deep learning model that uses neural networks to classify the genre of music automatically. For this project, students will use an FMA (Free Music Archive) dataset. FMA is an interactive library comprising high-quality and legal audio downloads. It is an open-source and easily accessible dataset.
  However, it is noteworthy that before one can use the model to classify audio files by genre, he/she will have to extract the relevant information from the audio samples (like spectrograms, MFCC, etc.)

- **Image Caption generator**
  This is one of the trending deep learning project ideas. This is a Python-based deep learning project that leverages Convolutional Neural Networks and LTSM (a type of Recurrent Neural Network) to build a deep learning model that can generate captions for an image. An Image caption generator combines both computer vision and natural language processing techniques to analyze and identify the context of an image and describe them accordingly in natural human languages (for example, English, Spanish, Danish, etc.). This project will strengthen one's knowledge of CNN and LSTM, and one will learn how to implement them in real-world applications as this.

- **Visual tracking system**
  A visual tracking system is designed to track and locate moving object(s) in a given time frame via a camera. It is a handy tool that has numerous applications such as security and surveillance, medical imaging, augmented reality, traffic control, video editing and communication, and human-computer interaction. This system uses a deep learning algorithm to analyze sequential video frames, after which it tracks the movement of target objects between the frames. The two core components of this visual tracking system are Target representation and localization

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

- **Traffic Signal Classification**
  The traffic sign classification project is useful for all autonomous vehicles. Machines are able to identify traffic signs from the image. Students can use the GTSRB dataset that contains 43 different traffic sign classes. This is a good project to understand image classification.

- **Driver Drowsiness Detection**
  The driver drowsiness detection is a project which can detect whether a person is sleeping or not while driving. We can implement a model for drivers and it can also prevent accidents from happening.

- **Autocolouring old Black and white images**
  The idea of this project is to make a model that is capable of colorizing old black and white images to colorful images. Digital artists take a few hours to color the image but now with Deep Learning, it is possible to color an image within seconds.

Text Book

9) Ian Goodfellow, Yoshua Bengio, Aaron Courville, "Deep Learning", MIT Press, 2017

References

1. Duda, R.O., Hart, P.E., and Stork, D.G. Pattern Classification. Wiley-Inderscience, 2nd Edition. 2013
2. Theodoridis, S. and Koutroumbas, K. Pattern Recognition. Edition 4, Academic Press, 2015
3. Russell, S. and Norvig, N. Artificial Intelligence: A Modern Approach. Prentice Hall Series in Artificial Intelligence, 2013
4. Bishop, C. M. Neural Networks for Pattern Recognition, Oxford University Press, 2008.

Topics relevant to development of "Employability": Real time Data Analysis using Deep learning.

Topics relevant to "PROFESSIONAL ETHICS": Naming and coding convention for Data Science Project Development using ML/DL.

Catalogue prepared by	Prof.Tapas Guha, Prof.Nappa Lakshmi
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE 6002	**Course Title: Natural Language Processing**	**L- P- C**	2	2	3

REGISTRAR

	Type of Course: Program Core **Theory and Laboratory Based Course**				
Version No.	1.0				
Course Pre-requisites	**[1] Basic Algebra [2] Probability [3] Deep learning** **Familiarity with basics of probability and statistics, and linear algebra would be essential. Prior exposure to Deep learning techniques would be desirable.**				
Anti-requisites	**NIL**				
Course Description	**This course introduces a basics of Natural Language Processing methods with specific emphasis on modern applications. The course will cover pre-processing techniques of textual data like stemming, lemmatization, tokenization etc. Different word Vectorization Techniques like Bag of Words, TF-iDF etc. followed by basics of Probability for building language models. Basics of Neural Network, LSTM Recurrent Neural Network, Applications of NLP like Information Extraction, Emotion Extraction from text, sentiment analysis etc.**				
Course Outcomes	**On successful completion of this course the students shall be able to:** **1] Understanding the fundamentals of NLP techniques.** **2] Apply Language modelling techniques for predictions.** **3] Apply Deep learning Techniques to build NLP Model** **4] Outline the application of NLP Techniques.**				
Course Content:					
Module 1	**pre-processing techniques**	Assignment	Apply all **the pre-processing techniques to the corpus of your choice.**	**14 Hours**	
Topics: **Introduction to Natural Language Processing, terminologies, empirical rules, why NLP is hard , why NLP is useful, Natural Language generation , natural Language Understanding, Corpus Cleaning techniques – word tokenization, sentence tokenization, word frequency distribution, stemming, lemmatization, word sense disambiguation, dictionary, Part of Speech Tagging, Textual Pre Processing techniques – Stop words removal, regular expression, lower case, text standardization.**					
Module 2	Lan**guage Model**	Assignment	Build n-**gram language model for future word predictions.**	**11 Hours**	
Topics: **Word Embeddings** techniques- bag of words, Tf-iDF, Word2Vec and optimization. **The role of language models. Hidden Markov Models Simple N-gram models. Estimating parameters and smoothing. Negative Sampling Evaluating language models. (Forward and Viterbi algorithms and EM training), Baum Welch Algorithm, Maximum Entropy Models,**					

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Module 3	Deep Learning **techniques for NLP models**	Assignment	Build model for spam detection using mail subject as Corpus	**11 Hours**
Topics: **Introduction to Neural Network, Perceptron, back Propa**gation, Recurrent Neural network, LSTM, Attention Models, Transformer , BERT (Bidirectional **E**ncoder **R**epresentation from **T**ransformer), Reformer				
Module 4	Application of NLP	Assignment	Paper Review of State of the Art NLP Tec**h**nique	**11 Hours**
Topics: Application of NLP- **Lexical semantics and word-sense disambiguation. Named entity recognition and relation extraction. IE using sequence labeling. Machine Translation, Emotion Extraction. tExt Summarization.**				

Targeted Application & Tools that can be used:

1. **Application Area Sentiment Analysis , Text Classification , Chatbots & Virtual Assistants , Text Extraction , Machine Translation , Text Summarization , Market Intelligence , Auto-Correct , Intent Classification , Urgency Detection , Speech Recognition**

Professionally Used Software: Anaconda Navigator, Python Packages, NLP toolkit

List of Laboratory Task

1. **Experiment No. 1: Apply all preprocessing technique to corpus of choice and plot word frequency.**
2. **Experiment No. 2: Word Embedding using Bag of words**
3. **Experiment No. 3: Word Embedding using TF-iDF**
4. **Experiment No. 4: Word Embedding using Word2Vec Continuous Bag of words**
5. **Experiment No. 5: Word Embedding using Word2Vec Skip gram Model**
6. **Experiment No. 6: Build language Model using n- gram.**
7. **Experiment No. 7: Build NLP model using LSTM**
8. **Experiment No. 8: Build NLP model using BERT**
9. **Experiment No. 9: Build NLP model using Reformer to show optimization.**

Project work/Assignment:

Project Assignment: NIL

Assignment 1: Paper Review of the state of the art NLP Technique

Text Books

1. Daniel Jurafsky, James H. MartinSpeech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech, Pearson Publication, 2014.

REGISTRAR

2. Steven Bird, Ewan Klein and Edward Loper, Natural Language Processing with Pythonll, First Edition, OReilly Media, 2009.

References

1. **Breck Baldwin, Language Processing with Java and LingPipe Cookbook, Atlantic Publisher, 2015.**
2. **Richard M Reese, Natural Language Processing with Javall, OReilly Media, 2015.**
3. **Nitin Indurkhya and Fred J. Damerau, Handbook of Natural Language Processing, Second Edition, Chapman and Hall/CRC Press, 2010.**
4. **Tanveer Siddiqui, U.S. Tiwary, Natural Language Processing and Information Retrieval, Oxford University Press, 2008.**

Catalogue prepared by	Ms. Mitali Halder
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE 6003	**Course Title:** Big Data Analytics Tools and Techniques **Type of Course: Program Core** **Theory and Lab Integrated Course**	**L- P- C**	2	2	3
Version No.	1.0				
Course Pre-requisites	Relational Database management System: Data, Database, Relational Data, Relational Model, Entity Relational model, SQL Clauses, SQL Queries				
Anti-requisites	**NIL**				
Course Description	This course is designed to provide the fundamental knowledge to equip students being able to handle real world big data problems including the three key resources of Big Data: people, organizations and sensor. With the advancement of IT storage, processing, computation and sensing technologies. It helps the students to interpret the operational concepts of computer technology as well as performance enhancement.				
Course Outcomes	On successful completion of the course the students shall be able to: • Understand managing big data using Hadoop analytical tools and technologies • Understand map-reduce analytics using Hadoop and related tools				

REGISTRAR

	• Preparing for data summarization, query, and analysis. • Applying data modeling techniques to large data sets • Building a complete business data analytic solution			
Course Content:				
Module 1	Introduction to Hadoop and HDFS	Assignment	Data Collection and Analysis	**8 Classes**

Topics:
Meet Hadoop: Data!, Data Storage and Analysis, Querying All Your Data, Beyond Batch, Comparison with Other Systems: Relational Database Management Systems, Grid Computing, Volunteer Computing Hadoop Fundamentals Map Reduce A Weather Dataset: Data Format, Analyzing the Data with Unix Tools, Analyzing the Data with Hadoop: Map and Reduce, Java Map Reduce, Scaling Out: Data Flow, Combiner Functions, Running a Distributed Map Reduce Job, Hadoop Streaming

The Hadoop Distributed File system: The Design of HDFS, HDFS Concepts: Blocks, Name nodes and Data nodes, HDFS Federation, HDFS High-Availability, The Command-Line Interface, Basic File system Operations, Hadoop File systems Interfaces, The Java Interface, Reading Data from a Hadoop URL, Reading Data Using the File System API, Writing Data, Directories, Querying the File system, Deleting Data, Data Flow: Anatomy of a File Read, Anatomy of a File Write.

Module 2	YARN and Hadoop I/O	Assignment	Data Collection and Analysis	**8 Classes**

Topics:
YARN Anatomy of a YARN Application Run: Resource Requests, Application Lifespan, Building YARN Applications, YARN Compared to Map Reduce, Scheduling in YARN: The FIFO Scheduler, The Capacity Scheduler, The Fair Scheduler, Delay Scheduling, Dominant Resource Fairness

Hadoop I/O: Data Integrity, Data Integrity in HDFS, Local File System, Checksum File System, Compression, Codecs, Compression and Input Splits, Using Compression in Map Reduce, Serialization, The Writable Interface, Writable Classes, Implementing a Custom Writable, Serialization Frameworks, File-Based Data Structures: Sequence File

Module 3	Map Reduce Applications	Case Study	Data analysis	**8 Classes**

Topics:
Developing a Map Reduce Application: The Configuration API, Combining Resources, Variable Expansion, Setting Up the Development Environment, Managing Configuration, Generic Options Parser, Tool, and Tool Runner, Writing a Unit Test with MR Unit: Mapper, Reducer, Running Locally on Test Data, Running a Job in a Local Job Runner, Testing the Driver, Running on a Cluster, Packaging a Job, Launching a Job, The Map Reduce Web UI, Retrieving the Results, Debugging a Job, Hadoop Logs, Tuning a Job, Profiling Tasks, Map Reduce Workflows: Decomposing a Problem into Map Reduce Jobs, Job Control, Apache Oozie

How Map Reduce Works: Anatomy of a Map Reduce Job Run, Job Submission, Job Initialization, Task Assignment, Task Execution, Progress and Status Updates, Job Completion, Failures: Task Failure, Application Master Failure, Node Manager Failure, Resource Manager Failure, Shuffle and Sort: The Map Side, The Reduce Side, Configuration Tuning, Task Execution: The Task Execution Environment, Speculative Execution, Output Committers

Module 4	Map Reduce Types and Formats, Flume	Case Study	Data analysis	**10 Classes**

Topics:

REGISTRAR

Map Reduce Types, Input Formats: Input Splits and Records Text Input, Binary Input, Multiple Inputs, Database Input (and Output) Output Formats: Text Output, Binary Output, Multiple Outputs, Lazy Output, Database Output

Flume Installing Flume, An Example, Transactions and Reliability, Batching, The HDFS Sink, Partitioning and Interceptors, File Formats, Fan Out, Delivery Guarantees, Replicating and Multiplexing Selectors, Distribution: Agent Tiers, Delivery Guarantees, Sink Groups, Integrating Flume with Applications, Component Catalog

Module 5	**Hive, Pig, Spark Analytical Tools**	Case Study	Data analysis	**10 Classes**

Topics:

Hive Apache Hive with Hive Installation, Hive Data Types, Hive Table partitioning, Hive DDL commands, Hive DML commands, Hive sort by vs order by, Hive Joining tables

Pig Installing and Running Pig, Execution Types, Running Pig Programs, Grunt, Pig Latin Editors, An Example: Generating Examples, Comparison with Databases, Pig Latin: Structure, Statements, Expressions, Types, Schemas, Functions, Data Processing Operators: Loading and Storing Data, Filtering Data, Grouping and Joining Data, Sorting Data, Combining and Splitting Data.

Spark An Example: Spark Applications, Jobs, Stages and Tasks, A Java Example, A Python Example, 20082020 / 31 Resilient Distributed Datasets: Creation, Transformations and Actions, Persistence, Serialization, Shared Variables, Broadcast Variables, Accumulators, Anatomy of a Spark Job Run, Job Submission, DAG Construction, Task Scheduling, Task Execution, Executors and Cluster Managers: Spark on YARN

List of Laboratory Tasks:

1. (i) Perform setting up and Installing Hadoop in its two operating modes:
   - Pseudo distributed,
   - Fully distributed.

   (ii) Use web based tools to monitor your Hadoop setup.

   Level 1: Programming assignment to install the Hadoop environment tools.

2. (i) Implement the following file management tasks in Hadoop:
   - Adding files and directories
   - Retrieving files
   - Deleting files

   (ii) Benchmark and stress test an Apache Hadoop cluster

   Level 1: Programming assignment to maintain the Hadoop Distributed File System.

3. Run a basic Word Count Map Reduce program to understand Map Reduce Paradigm.
   - Find the number of occurrence of each word appearing in the input file(s)
   - Performing a Map Reduce Job for word search count (look for specific keywords in a file)

   Level 1: Programming scenario to use map reduce programming to perform the analysis.

   Level 2: Programming assignment to analyze the data for any given data file.

4. Stop word elimination problem:

   Input:
   - A large textual file containing one sentence per line
   - A small file containing a set of stop words (One stop word per line)

   Output:

   A textual file containing the same sentences of the large input file without the words appearing in the small file.

5. 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 Map Reduce, since it is semi structured and record-oriented. Data available at: https://github.com/tomwhite/hadoopbook/tree/master/input/ncdc/all.

- Find average, max and min temperature for each year in NCDC data set?
- 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.

Level 2: Programming assignment to analyze the social media data for business analytics.

6. For Purchases.txt Dataset , 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?

  i. Toys
  ii. Consumer Electronics

- Find the monetary value for the highest individual sale for each separate store
  What are the values for the following stores?
  i. Reno
  ii. Toledo
  iii. Chandler
- Find the total sales value across all the stores, and the total number of sales.

Level 2: Programming assignment to analyze and find the maximum sales, minimum sales and average sales in each store.

7. Install and Run Pig then write Pig Latin scripts to sort, group, join, project, and filter your data.

Level 1: Programming scenarios to perform the grouping, filtering and Joining.

8. Write a Pig Latin scripts for finding TF-IDF value for book dataset (A corpus of eBooks available at: Project Gutenberg)
Level 2: Programming Assignment to analyze the data from the given text file using Pig latin script.

9. Install and Run Hive then use Hive to create, alter, and drop databases, tables, views, functions, and indexes.
Level 1: Programming scenario to analyze the data from the given text file to perform SQL operations.

10. Install, Deploy & configure Apache Spark Cluster. Run apache spark applications using Scala.
Level 1: Programming scenario to analyze a dataset using spark.

11. Data analytics using Apache Spark on Amazon food dataset, find all the pairs of items frequently reviewed together.

Write a single Spark application that:

- Transposes the original Amazon food dataset, obtaining a Pair RDD of the type:
- Counts the frequencies of all the pairs of products reviewed together;
- 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.

Level 2: Programming assignment to analyze the data using spark.

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Targeted Application & Tools that can be used:

- **Business Analytical Applications**
- **Social media Data Analysis**
- **Predictive Analytics**
- **Government Sector for analyzing the data**
- **Improve the business through analytics**

Tools: Hadoop Framework tools like map reduce, Hive, Hbase, Spark, Pig, Flume.

Project work/Assignment:

After completion of each module a programming based Assignment/Assessment will be conducted.
A scenario will be given to the student to be developed as a data analysis application.
On completion of Module 5, student will be asked to develop a project for Data Analysis.

Text Book

1. Hadoop: The Definitive Guide Tom White O'Reilley Third Edition, 2012

References

7. SPARK: The Definitive Guide MateiZaharia and Bill Chambers Oreilly 2018
8. Apache Flume: Distributed Log Collection for Hadoop . D'Souza and Steve Hoffman Oreilly 2014

Topics relevant to development of "Employability": Real time application development using Hadoop Ecosystem tools.
Topics relevant to "HUMAN VALUES &PROFESSIONAL ETHICS"": Naming and coding convention for Project Development.

Catalogue prepared by	Dr. Manujakshi B C, Dr.Senthilkumar S
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE 6004	**Course Title: Time Series Analysis and Forecasting** **Type of Course: Discipline Elective Theory Only**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	**R,** Calculus, Linear Algebra, Probability and Statistics				
Anti-requisites	**NIL**				
Course Description	The course will provide a basic introduction to time series analysis. This theory based course covers topics in time series analysis and some statistical techniques on forecasting. Time series regression, exploratory data analysis, AR models, Seasonal Models, GARCH Models and Box-Jenkins approach are the major topics covering in this course. R and RStudio will be required for this class.				
Course Outcomes	On successful completion of the course the students shall be able to 1. Select appropriate model, to fit parameter values and make concise decisions based on forecasts obtained 2. Demonstrate an understanding of the principles behind modern forecasting techniques. 3. Apply concepts to real time series data using packages.				

REGISTRAR

Course Content:				
Module 1	Introduction	Assignment	Data Analysis task	**9 Classes**

Topics:
Examples of Time Series, Objectives of Time Series Analysis, Characteristics of Time Series, Approaches used for time series forecasting, ETS (Error, Trend, Seasonality) models to make forecasts, Decomposition method, Irregularity concept in decomposition method, Case study on decomposition method, Model forecast theory, Model forecast hands-on, stochastic process.

Module 2	Time Series Regression and Exploratory Data Analysis	Assignment	Data analysis	**10 Classes**

Topics:
Classical Regression in the Time Series Context, Exploratory Data Analysis, Stationary Models and the Autocorrelation Function, Detrending and De-seasonalizing Smoothing, Fundamental Statistical Concepts, Introduction to Time Series Analysis with R

Module 3	AR models	Assignment	Data analysis	**10 Classes**

Topics:
Models for Stationary Time Series, Models for Non-Stationary Time Series, Identification, Forecasting, ARIMA (Autoregressive, Integrated, Moving Average) models, ARMA models

Module 4	Additional models, Spectral Analysis and packages	Case Study	Data analysis	**10 Classes**

Topics:
Seasonal Models, Time Series Regression Models, GARCH Models, Box-Jenkins approach, Introduction to Spectral Analysis, Estimating the Spectrum
Preparing model using ITSM, Time series using astsa, ARIMA models is to use sarima from astsa

Targeted Application & Tools that can be used:
Targeted Applications: Time series analysis on economics, finance, natural sciences, health care and more
Tools:

- R package astsa (Applied Statistical Time Series Analysis)
- The package ITSM2000 (https://extras.springer.com/)

Project work/Assignment:

Mini Project:
Choose any suitable real time dataset and build time series forecast models.
Example: In the Air Passengers dataset set, go back 12 months in time and build the ARIMA forecast for the next 12 month. Investigate following questions

Is the series stationary? If not what sort of differencing is required?
What is the order of your best model?
What is the AIC of your model?
What is the order of the best model predicted by auto_arima() method?

Term Assignments:
Understand and implement ARMA and ARIMA models in Python/R for time series forecasting.

REGISTRAR

Text Book

11. Montgomery DC, Jennings CL, Kulahci M. Introduction to time series analysis and forecasting. John Wiley & Sons; 2015 Apr 21.
12. Brockwell & Davis (2016) Introduction to Time Series and Forecasting, 3rd edition, Springer.
13. Shumway & Stoffer (2011) Time Series Analysis and its applications, with examples in R , 3rd edition, Springer.

References

9. Box GE, Jenkins GM, Reinsel GC, Ljung GM (2015) Time series analysis: forecasting and control. John Wiley & Sons
10. Cryer & Chan (2008) Time Series Analysis with Applications in R, Springer
11. Prado & West (2010) Time Series: Modeling, Computation, and Inference Chapman & Hall

Topics relevant to development of "FOUNDATION SKILLS":

Topics relevant to "HUMAN VALUES &PROFESSIONAL ETHICS": **Analysis of time domain datasets like daily climate variables, health records, stock exchange data etc**

Catalogue prepared by	Prof. Jobin Thomas, Dr. Jacob A.
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE 6005	**Course Title: Intelligent Information Retrieval** **Type of Course: Discipline Elective** **Theory Only**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	**Core programming and algorithm skills**				
Anti-requisites	**NIL**				
Course Description	This Course studies the theory, design, implementation and evaluation of information retrieval systems. The focus is on the core concepts of Text- based information systems, statistical characteristics of text, representation of information needs and documents. Several important retrieval models, algorithms, and Recommender System. Also examined is how an effective information search and retrieval is interrelated with the organization and description of information to be retrieved. Throughout the course, current literature from the viewpoints of both research and practical retrieval technologies on the World Wide Web will be examined.				
Course Outcomes	On successful completion of the course the students shall be able to: 1] Define basic concepts of information Retrieval and Recommender System				

REGISTRAR

	2] Evaluate the effectiveness and efficiency of different information retrieval methods 3] Explain the standard methods for Web indexing and retrieval 4] Develop Methods for implementing recommender system			
Course Content:				
Module 1	INTRODUCTION	Assignment	Term Paper	**8 Classes**
Topics: Information Retrieval – Early Developments – The IR Problem – The Users Task – Information versus Data Retrieval – The IR System – The Software Architecture of the IR System – The Retrieval and Ranking Processes – The Web – The e-Publishing Era – How the web changed Search – Practical Issues on the Web – How People Search – Search Interfaces Today – Visualization in Search Interfaces.				
Module 2	MODELING AND RETRIEVAL EVALUATION	Assignment	Term Paper	**12 Classes**
Topics: Basic IR Models – Boolean Model – TF-IDF (Term Frequency/Inverse Document Frequency) Weighting – Vector Model – Probabilistic Model – Latent Semantic Indexing Model – Neural Network Model – Retrieval Evaluation – Retrieval Metrics – Precision and Recall – Reference Collection – User-based Evaluation – Relevance Feedback and Query Expansion – Explicit Relevance Feedback.				
Module 3	WEB RETRIEVAL AND WEB CRAWLING	Assignment	Term Paper	**10 Classes**
Topics: The Web – Search Engine Architectures – Cluster based Architecture – Distributed Architectures – Search Engine Ranking – Link based Ranking – Simple Ranking Functions – Learning to Rank – Evaluations — Search Engine Ranking – Search Engine User Interaction – Browsing – Applications of a Web Crawler – Taxonomy – Architecture and Implementation – Scheduling Algorithms – Evaluation.				
Module 4	RECOMMENDER SYSTEM	Assignment	Term Paper	**10 Classes**
Topics: Recommender Systems Functions – Data and Knowledge Sources – Recommendation Techniques – Basics of Content-based Recommender Systems – High Level Architecture – Advantages and Drawbacks of Content-based Filtering – Collaborative Filtering – Matrix factorization models – Neighbourhood models.				
Targeted Application & Tools that can be used: • Information Retrieval Applications • Machine Learning Applications **Tools:**				

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

- Bow Toolkit
- GATE
- Lemur
- MG
- Smart (System for the Mechanical Analysis and Retrieval of Text) Information Retrieval System is an information retrieval system **developed at Cornell University** in the 1960s.

Project work/Assignment:

Term Paper: Students are expected to submit a Term Paper on application IIR

- **MIDTERM - Phase 1: survey paper on Current trends and tools used for IIR (Intelligent Information Retrieval) and propose a model**
- **ENDTREM - Phase 2: Implementation of the proposed model and Paper should be published in the reputed conference / journal**

Text Book

14. Ricardo Baeza-Yates and Berthier Ribeiro-Neto, —Modern Information Retrieval: The Concepts and Technology behind Search, Second Edition, ACM Press Books, 2011. Link: https://people.ischool.berkeley.edu/~hearst/irbook/
15. Ricci, F, Rokach, L. Shapira, B.Kantor, —Recommender Systems Handbook, First Edition, 2011.
16. C. Manning, P. Raghavan, and H. Schütze, —Introduction to Information Retrieval, Cambridge University Press, 2008. Link: https://nlp.stanford.edu/IR-book/

References

12. Mikhail Klassen, Matthew A. Russell, Mining the Social Web,O'Reilly Media, Inc., 3rd Edition (2019)
13. Stefan Buettcher, Charles L. A. Clarke and Gordon V. Cormack, —Information Retrieval: Implementing and Evaluating Search Engines, The MIT Press, 2010.
14. Ceri, S., Bozzon, A., Brambilla, M., Della Valle, E., Fraternali, P. and Quarteroni, S., 2013. Web information retrieval. Springer Science & Business Media.

Topics relevant to development of "EMPLOYMENT SKILLS": Software Development Engineer(Flipkart), Architect, Information Retrieval Officer, Research Scientist – IBM Research, Machine Learning Application Developer and Lead Engineer / Module Lead – Java / Python

Topics relevant to "HUMAN VALUES &PROFESSIONAL ETHICS": NA

Catalogue prepared by	Prof. Zafar Ali Khan N, Dr. R Mahalakshmi
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Course Code: CSE 6006	**Course Title: AI in Internet of Things** **Type of Course: Discipline Elective** **Theory Only**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	Basic knowledge of computer and internet, Basic Python programming, Probability and Linear Algebra				
Anti-requisites	**NIL**				
Course Description	This course introduces the core principles of Internet of things and Artificial Intelligence from basic to intermediate level. This theory based course emphasizes on understanding the application of AI in IOT. The course will focus on creative thinking of AI & IoT concepts & technologies.				
Course Outcomes	On successful completion of the course the students shall be able to: 1] Understand building blocks of Internet of Things and characteristics. 2] Describe IoT Protocols 3] Compare and contrast from a range of AI techniques when implementing smart systems. 4] Identify and Apply techniques in areas of AIIoT.				
Course Content:					
Module 1	Introduction to AI	Assignment	Data Analysis task		**10 Classes**
Topics: Introduction to Artificial Intelligence, Basics of Python , Example programs in Python, Introduction to Machine Learning, Types of Machine Learning Algorithms, Introduction to Linear Algorithm, Genetic Algorithms, Adversarial Search, Constraint Satisfaction, Propositional Logic & Satisfiability, Uncertainty in AI, Bayesian Networks.					
Module 2	Introduction to IOT	Assignment	Data Collection		**10 Classes**
Topics: Introduction to Internet of Things, What is Microcontrollers?, Difference between Microprocessor and Microcontroller, Introduction to Arduino Boards, Types of Arduino Boards, Installation of Arduino IDE, Usage of Tinker Cad, Types of Sensors, Working Principle, Pin Configuration of Sensors, Demonstrating of Ultrasonic Sensor, IR Sensor, Gas Sensor and PIR Sensor. Structure of Coding – Embedded C, Demonstrating - Blinking of LED and in Serial Monitor, intensity of LED using a potentiometer, Traffic Signal.					
Module 3	AI algorithms for sensors	Assignment	Data Collection		**10 Classes**
Topics: Algorithms of Artificial Intelligence in Sensors- Classification algorithms Data clustering, Evolutionary algorithms in sensing, Data pattern recognition, Maintenance and production scheduling, Artificial intelligence in predictive and proactive scheduling, Energy efficient scheduling, Stochastic models in artificial intelligence, Queuing theory-based approach, Genetic					

programming, Project scheduling, Artificial intelligence in assembly line balancing, Disassembly line balancing.

Module 4	IOT Protocols and Applications of AI in IOT	Case Study	Data Collection	**10 Classes**

Topics:
Connectivity Protocols: 6LoWPAN, IEEE 802.15.4, Zigbee, Wireless, NFC, RFID. **Communication/Transport Protocols:** Bluetooth. **Data Protocols:** Message Queue Telemetry Transport (MQTT), Constrained Application Protocol (CoAP), Advanced Message Queuing Protocol (AMQP), XMPP – Extensible Messaging and Presence Protocol.**Applications of AI in IOT- Case Study: Smart Retail, Drone Traffic Monitoring , Office Buildings**

Targeted Application & Tools that can be used:

Targeted employment sector is service provider and control monitor like GE, Siemens, TCS etc. Targeted job profiles include digital domain and Physical system design engineer, IOT engineer etc.

Tools:

- Arduino IDE
- TinkerCad
- NodeMCU
- Tensor Flow and Keras

Project work/Assignment:

Term Assignments:

- **Comparative analysis of Machine learning algorithm**

  Carry out a thorough analysis of the various ML algorithms and its efficiency for the given data set.

- **A short survey of the applications of AI in IOT**

  Study and analyze few applications of IOT and then applying the concepts of AI to improve the efficiency.

Text Book

17. "Artificial Intelligence: A Modern Approach", Stuart Russell & Peter Norvig, Prentice-Hall, Third Edition (2009).
18. "Internet of Things: A hands on approach", Arshdeep Bagha & Vijay Madisetti, Universities Press 2015.
19. https://www.tinkercad.com/

References

3. "The internet of Things: Connecting Objects to web", Hakima Chaouchi, Wiley 2017.
4. "Prolog: Programming for Artificial Intelligence", I. Bratko, Fourth edition, AddisonWesley Educational Publishers Inc, 2011.
5. E. RICH, K. KNIGHT, S. B. NAIR (2017), Artificial Intelligence, McGraw Hill Education, 3rd Edition.

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Topics relevant to development of "FOUNDATION SKILLS": Get introduced to AI programming and Interfacing of IOT devices. Topics relevant to "HUMAN VALUES &PROFESSIONAL ETHICS": Set of standard procedures to build AI-IOT applications.	
Catalogue prepared by	Prof. Mohammed Mujeer Ulla Prof. Preethi
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE 6007	**Course Title: IOT Data Analytics** **Type of Course: Discipline Elective Theory Only**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	**Computer networks, Data mining, Data analytics and Basics of IoT.**				
Anti-requisites	**NIL**				
Course Description	This course helps in understanding the context of analytics in IoT data. Strategies to collect IoT data in order to enable analytics. Skills learnt will enable to understand the tradeoffs in streaming and batch processing. Data science techniques such as machine learning, deep learning, and forecasting are applied on IoT data. This course also teaches how to implement machine learning methods and ARIMA forecasting on IoT data. Deep learning will be described along with a way to get started experimenting with it on AWS.				
Course Outcomes	On successful completion of the course the students shall be able to: 1] Discuss the challenges of IoT Analytics. 2] Apply strategies and techniques to collect IoT data. 3] Apply data science techniques on IoT data				
Course Content:					

Module 1	IoT analytics, challenges, devices and networking protocols	Assignment	Data Analysis task	**12 Classes**

Topics:
Defining IoT Analytics and Challenges Defining IoT Analytics. IoT analytics challenges. Business value concerns
IoT Devices and Networking Protocols IoT devices Networking basics IoT networking connectivity protocols
IoT networking data messaging protocols Message Queue Telemetry Transport (MQTT) Hyper-Text Transport Protocol (HTTP) Constrained Application Protocol (CoAP) Data Distribution Service (DDS)

REGISTRAR
PRESIDENCY UNIVERSITY Registrar BANGALORE

Module 2	Data – Strategies, Techniques and Exploring IoT Data	Assignment	Analysis, Data Collection	**12 Classes**
Topics: Collecting All That Data - Strategies and Techniques Designing data processing for analytics Applying big data technology to storage Apache Spark for data processing Exploring IoT Data Exploring and visualizing data				
Module 3	Data Science for IoT Analytics	Case Study	Data analysis task	**13 Classes**
Topics: machine learning Generalization Feature engineering with IoT data Validation methods Understanding the bias–variance trade off Comparing different models to find the best fit Random Forest models Gradient Boosting Machines Anomaly detection, Forecasting				
Targeted Application & Tools that can be used: Employment opportunities are available in Companies like Hexaware, Episteme, Randstad. Siemens, Accenture etc. as IoT Data Engineer Tools R Python Microsoft Azure Stream Analytics. AWS IoT Analytics. SAP Analytics Cloud. Oracle Stream Analytics and Oracle Edge Analytics.				
Project work/Assignment:				
Mini Project: Develop a IoT application for real time data analysis of manufacturing sector. The automated IoT Analytics should aid in using real time data to watch out for certain patterns and send alerts to the concerned departments. It should enable smart manufacturing. **Term Assignments:** Discuss the various IoT data analysis strategies and list the real time applications that they could be used with.				
Text Book "Analytics for the Internet of things (IoT)", Andrew Minteer, Packt, 2017				
References Internet of Things and Big Data Analytics for Smart Generation, Valentina E Balas, Springer				
Topics relevant to development of "Employability": Processing geo-spatial IoT Data				
Catalogue prepared by	Prof. Shashidhar. V, Prof. S. Deepak Raj			
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021			

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE 6009	**Course Title: Probabilistic graph Models** **Type of Course: Discipline Elective** **Theory Only**	**L- P- C**	3	0	3

Version No.	1.0
Course Pre-requisites	**Probability, Probability distributions and graph theory** Basic concepts of probability, conditional probability, Standard probability distributions and basic concepts of undirected and directed graphs along with traversal.
Anti-requisites	**NIL**
Course Description	Probabilistic graphical models are used to model stochasticity (uncertainty) in the world and are extremely popular in AI and machine learning. The course will cover two classes of graphical models: Bayesian belief networks (also called directed graphical models) and Markov Random Fields (undirected models). After introducing the two frameworks the course will focus on recent advances in inferences and learning with graphical models, including topics such as loopy belief propagation, variational approximations, conditional Markov random fields and others.
Course Outcomes	On successful completion of the course the students shall be able to: 1.Apply key concepts of Statistics to solve problems. 2: Analyze the properties of distributions encoded by graphs 3: Illustrate Inference in graphic models 4: Illustrate Learning in graphic models
Course Content:	

Module 1	Fundamentals of Probability and Graph Theory	Assignment	Understanding all standard probability distributions	**9 Classes**
Topics: Fundamentals of Statistics and Probability, Conditional Probability, Conditional Independence, Joint Distributions, Marginalization, Baye's Theorem, Probability Distributions, Fundamentals of Graph Theory - Paths, Cliques, Sub-graphs, Cycles and Loops.				
Module 2	Graphical Models	Assignment	Construction of Markov chain model for real time problems	**9 Classes**
Topics:				

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Directed Models: Bayesian Network; Undirected Models: Markov Random Fields; Parameterization of MRFs, Independencies, D-Separation

Module 3	Inference in Graphical Models	Assignment	Study about some problems based on Monte Carlo method	**9 Classes**

Topics:
Inference in Graph Models, Variable Elimination; Belief Propagation, Sampling Methods: Markov Chain Monte Carlo, Metropolis Hastings Algorithm, Hidden Markov Model, Viterbi Algorithm.

Module 4	Learning in Graphical Models	Assignment	Applications of Naïve Bayes Classifier	**10 Classes**

Topics:
Learning in Graph Models, Maximum Likelihood Estimation, Naïve Bayes Classifier, Conditional Random Fields, Structural SVM

Targeted Application & Tools that can be used:

Targeted employment sector is to acquire knowledge to analyze the given problem to frame Probabilistic graphical models which are a powerful framework for representing complex domains using probability distributions, with numerous applications in machine learning, computer vision, natural language processing and computational biology.

Tools:

- Python
- HUGIN Tool for Learning Bayesian Networks
- MATLAB Toolbox for Bayesian net

Assignment:

Term Assignments:

- **Analysis and Application of Bayesian Network to real time problems**

  Understanding the given problem, analyze accordingly to apply Bayesian network and convert the problem in a Bayesian Network. The answering the required queries.

- **A short survey of the Monte Carlo Method**

  Study and analyze few realistic problems to apply Monte Carlo Technique to answer the solution of the problem.

- **A short survey of the Markov Chain & Hidden Markov Method**

  Study and analyze few realistic problems to convert into Markov chain & Hidden Markov to answer the required problem.

Text Book

20. Kevin P. Murphy, “Machine Learning: A Probabilistic Perspective”, MIT Press, 2012.
21. Koller and Friedman Probabilistic Graphical Models: Principles and Techniques, MIT Press 2009.

REGISTRAR

References 15. S. Lauritzen. Graphical Models. Oxford University Press, 1996. 16. David J.C. Mackay. Information theory, inference, and learning algorithms. Cambridge, UK: Cambridge University Press 2003.

Catalogue prepared by	Dr. M.V Chakradhara Rao
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: **CSE 6009**	**Course Title: ARTIFICIAL NEURAL NETWORK** **Type of Course: Discipline Elective** **Theory only**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	[1] Basic Algebra [2] Probability [3] Statistics Familiarity with basics of probability and statistics, and linear algebra would be essential. Prior exposure to machine learning techniques would be desirable.				
Anti-requisites	**NIL**				
Course Description	The objective of this course is to provide students with a basic understanding of the fundamentals and applications of artificial neural networks. The course will cover techniques in Single layer perceptron classifier and feed forwards network for single layer and multilayer. Along with basic concepts of Associative network and Self organizing map.				
Course Outcomes	On successful completion of this course the students shall be able to: 1] Understand the mathematical foundations of neural network models. 2] Solve real world problems using neural network systems. 3] Explain feed forward network for Single layer and multiple layers. 4] Describe the Knowledge of Associative memories and Self organizing maps.				
Course Content:					
Module 1	Fundamental Concepts of ANN	Assignment	Numerical to observe performance of different learning rule.		09 Classes
Topics:					

Structure of biological neurons relevant to ANNs. Models of ANNs; Feedforward & feedback networks; learning rules; Hebbian learning rule, perception learning rule, delta learning rule, Widrow-Hoff learning rule, correction learning rule, Winner –lake all learning rule

Module 2	Single layer Perception Classifier	Assignment	Build classifier using discrete perceptron algorithm.	12 classes

Topics:
Classification model, Features & Decision regions; training & classification using discrete perceptron, algorithm, single layer continuous perceptron networks for linearly separable classifications.

Module 3	Feed forward Networks	Assignment	STEP BY STEP SOLVE BACK PROPAATION	12 classes

Topics:
MULTI LAYER FEED FORWARD NETWORK: Linearly non-separable pattern classification, Delta learning rule for multi-perceptron layer, generalized delta learning rule, error back-propagation training, learning factors, Examples.
Single LAYER FEED FORWARDS NETWORK: Basic Concepts, Hopfield networks, Training & Examples.

Module 4	ASSOCIATIVE MEMORIES AND SOM	Assignment	Paper Review of State of the Art OPT	10 Classes

Topics:

Linear Association, Basic Concepts of recurrent Auto associative memory: retrieval algorithm, storage algorithm; By directional associative memory, Architecture, Association encoding & decoding, Stability. UN supervised learning of clusters, winner-take-all learning, recall mode, Initialization of weights, separability limitations

Targeted Application & Tools that can be used:

Application Area:
Resource Allocation, Finance and Economics (Risk Analysis and Consumption Assessment), Fraud Detection, Image Segmentation, Dimensionality Reduction, Gene Expression Analysis, Recommender System, Image reconstruction, Large Scale Surveillance.

Tools:
Anaconda Navigator
Python Packages

Project work/Assignment:

Assignment:
Create the Multi-layer Perceptron Neural Network / Back propagation network from scratch in Python using the weight and bias updating for solving the classification Problem. Also check the linear Separability for AND and XOR problem using the created Multi-layer Perceptron Neural Network algorithm/ Back propagation network.

Text Books

1. Machine Learning by Tom Mitchell, McGraw-Hill Press
2. Pattern Recognition and Machine Learning by Christopher M. Bishop, Springer, 2006

References

1. Neural Networks A Classroom Approach– Satish Kumar, McGraw Hill Education (India) Pvt. Ltd, Second Edition.
2. Iintroduction to Artificial Neural Systems-J.M. Zurada, Jaico Publications 1994.
3. Artificial Neural Networks-B. Yegnanarayana, PHI, New Delhi 1998.

REGISTRAR

Topics relevant to development of "Employability Skills": Concept of feed forward network, Hopfield network, self-organizing map.	
Catalogue prepared by	Ms. Mitali Halder
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/2021

Course Code: CSE6011	**Course Title: Application of Probability theory in Computer Science** **Type of Course: Theory Course**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	**Basic probability concepts**				
Anti-requisites	**NIL**				
Course Description	**For both engineers and researchers in the field of Computer science, it is common to develop models of real-life situations and develop solutions based on those models. In this course, our objective is to give an idea regarding application of probability theory in modelling and analyzing different kinds of computer systems. We particularly focus on time complexity analysis of different algorithms, reliability analysis of networks, physical layer security as well as resource allocation in 5G and beyond. Target audience for this course is Masters' and PhD students.** **The student should have basic Probability concepts as pre-requisite.** **With a good knowledge regarding different techniques of applying Probability theory in modelling/analysing computer systems, the students will be able to develop efficient solutions for complex and challenging real-life problems.**				
Course Outcomes	**On successful completion of this course the students shall be able to:** **1. Develop mathematical models for various computer systems.** **2. Apply an appropriate probability concept to analyse the system.** **3. Apply appropriate Reinforcement learning techniques to solve complex real-life problems.** **4. Apply statistical Inference concepts to estimate parameters which is unknown to the model.**				
Course Content:					

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Module 1	Review on basic Concepts	Assignment	Basic Probability Concepts; Basic idea of hypothesis testing; Applications in VLSI & series-parallel systems, Analysing algorithms & network performances, application in cognitive radio network.	**12 Classes**
Topics: Basic probability concepts, Conditional probability, Expectation, random variables, well known distributions, order statistics, basic idea of hypothesis testing. Applications in reliability analysis of VLSI chips, serial-parallel systems, analyzing randomized quicksort, performance analysis of telephone network and binary communication channels, application in cognitive radio network.				
Module 2	Stochastic processes	Assignment	Understanding Markov process; Application in analysing wireless channel behaviour, cache memory, MAC protocols and data structures.	**12 Classes**
Topics: Markov chain, Birth-death process, application in modelling the behavior of wireless channels, memory interference problem, performance analysis of cache memory, performance analysis of medium access protocols, analyzing the time complexity in the implementation of two stacks using a single array.				
Module 3	Reinforcement learning	Assignment	Understanding different Reinforcement learning techniques, Hidden Markov Model, applications in modelling resource allocation in 5G, physical layer security	**12 Classes**
Topics: Markov decision process, Bellman Equation, value and policy iteration, off-policy and on-policy learning techniques (e.g., SARSA, Q-learning), Multi-arm Bandit problem (MAB), modelling resource allocation in 5G as MAB, Hidden Markov model (HMM), application of HMM in physical layer security.				
Project work/Assignment:				
4. Performance analysis of LRU stack model **5. Modelling multiprocessor systems and analysing the reliability** **6. Modelling handovers in wireless networks and performance analysis of handover algorithms.** **7. A short survey on Monte Carlo simulation techniques.**				
REFERENCE MATERIALS: Text Book(s): 6. Kishore S. Trivedi, "Probability and Statistics with Reliability, Queuing, and Computer Science Applications", PHI.				

7. Dimitri P. Bertsekas and John N. Tsitsiklis, “Introduction to probability”, MIT press, FALL 2000.

8. Richard S. Sutton and Andrew G. Barto, “Reinforcement Learning: An Introduction”, MIT press.

9. Narsingh Deo, “System simulation with digital computer”, PHI

10. Selected research papers.

Topics relevant to development of **“Foundation Skills”:** Fundamentals of Probability, **“Skill Development” –** Stochastic simulation techniques, **“Employability”**-

Catalogue prepared by	Dr. Shankar K. Ghosh
Recommended by the Board of Studies on	BOS NO: 13th BOS, held on 08/12/2021
Date of Approval by the Academic Council	Academic Council Meeting No. 17, Dated 11/12/2021

Course Code: CSE 6012	**Course Title:** Recommender Systems with Machine Learning and AI **Type of Course:** Discipline Elective **Theory Based Course**	**L- P- C**	4	0	4
Version No.	1.0				
Course Pre-requisites	Basic knowledge of Python and Machine Learning				
Anti-requisites	NIL				
Course Description	This course helps understanding from the early days of collaborative filtering to bleeding-edge applications of deep neural networks and modern machine learning techniques for recommending the best items to every individual user.				
Course Outcomes	**On successful completion of this course the students shall be able to:** 1. **Define** recommender systems 2. **Use content-based filtering** using item attributes 3. Build model-based methods including **matrix factorization, SVD.** 4. Apply **deep learning, AI, artificial and recursive neural networks,** for session based recommendations. 5. **Analyse recommendation algorithms using various case studies.**				
Course Content:					

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Module 1	Introduction to Recommendation System	Assignment	Seminar	**12 Classes**
Topics: Introduction to Recommendation systems, Goals of Recommendation systems, Architecture of Recommendation systems, Basic models of Recommendation systems, Implicit Ratings, Explicit Ratings, Collaborative Filtering, Content based Recommendation, Knowledge-Based Recommender Systems, Hybrid Recommendation systems, Demographic Recommendation system, Applications of Recommendation systems.				
Module 2	Content Based Recommender Systems	Assignment	Mini Project	**12 Classes**
Topics: Introduction, Architecture of content based recommendation, Basic components of Content-based systems, Learning User profiles and Filtering- KNN, case-based recommendation, Bayes Classifiers, Rule-based Classifiers, Advantages and Disadvantages of Content based recommendations.				
Module 3	Model Based Collaborative Filtering	Assignment	Mini project	**12 Classes**
Topics: Introduction to collaborative filtering, Decision and Regression Trees, Rule based collaborative Filtering-Item-wise vs User-wise models, Naive Bayes Collaborative filtering, Basic matrix Factorization principle, Singular value Decomposition, Advantages and Drawbacks of Collaborative filtering.				
Module 4	Hybrid Recommendation Systems	Assignment	Mini project	**12 Classes**
Topics: Introduction to Hybrid Recommendation systems, Ensemble methods, Weighted Hybrids, Switching Hybrids, Cascade Hybrids, Meta-Level Hybrids, Mixed hybrids, Advantages and disadvantages of Hybrid Recommendation systems.				
Module 5	Application and Evaluation of RS	Assignment	Seminar	**12 Classes**
Topics: Case study on YouTube Recommendation, case study on Netflix Recommendation system, Offline Evaluation, Online Evaluation, Goals of Evaluation design- Accuracy, Coverage, Confidence and Trust, Diversity, Robustness and Stability, Scalability, Training and testing of Ratings, RMSE, MAE, Evaluating Ranking via Correlation, Utility, Receiver Operating Characteristics, Limitations of Evaluation Measures.				
Targeted Application & Tools that can be used: **Targeted Application : Web application development, AI, Operating systems** **Tools: Python IDLE, ANACONDA** **Application Areas:**				

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

- E-Commerce Application
- E-Learning Applications
- E-Business Services
- Artificial Intelligence and Machine Learning
- Enterprise-level/Business Applications

Professionally Used Software: Python, Spyder, Jupyter Notebook, Tensorflow (TFRS), Amazon Personalize.

Project work/Assignment:

- After completion of each module a programming based Assignment/Assessment will be conducted.
- A scenario will be given to the students to be developed as a series of Program/ Application.

On completion of Module 3 and Module 4, students will be asked to develop a Mini Project using Python.

Textbooks

- Frank Kane - Building Recommender Systems with Machine Learning and AI, First Edition,2018
- Charu C.Aggarwal – Recommender Systems, Springer Publishing Company,2016.

References

1. Katarzyna Tarnowska,Lynn Daniel – Recommender System for improving customer Loyalty,Springer,1st edition,2020.
2. EthemAlpaydin, —Introduction to Machine Learning 3e (Adaptive Computation and Machine Learning Series)‖, Third Edition, MIT Press, 2014.

Catalogue prepared by	Dr. Blessed Prince P
Recommended by the Board of Studies on	
Date of Approval by the Academic Council	

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Course Code: CSE5017	**Course Title:** NoSQL Databases **Type of Course:1] School Core** **2] Laboratory integrated**	**L- P- C**	**2**	**2**	**3**
Version No.	1.0				
Course Pre-requisites	**Basic Knowledge about DBMS**				
Anti-requisites	**NIL**				
Course Description	This course aims to give students an insight into few traditional database concepts and move towards the advanced database concepts. This course will explore the origins of NoSQL databases and the characteristics that distinguish them from traditional relational database management systems. Core concepts of NoSQL databases will be presented, followed by an exploration of how different database technologies implement these core concepts. This course also explores the four main NoSQL data models (key-value, column family, document, and graph), highlighting the business needs that drive the development and use of each database. The associated laboratory provides an opportunity to implement the concepts and enhance critical thinking and analytical skills.				
Course Out Comes	On successful completion of this course the students shall be able to: 1] Understand the aspects of Traditional Databases and the importance of NoSQL databases. 2] Learn the architectures and common features of the main types of NoSQL databases (key-value stores, document databases, column-family stores, graph databases). 3] Explain the detailed architecture, define objects, load data, query data and performance tune Document-oriented NoSQL databases. 4] Demonstrate an understanding of the detailed architecture, define objects, load data, query data and performance tune Column-oriented and Graph Based NoSQL databases.				
Course Content:					
Module 1	Advanced Database Concepts (Comprehension)	Assignment	Knowledge Ability		**9 Hours**
Topics: **An Overview of Traditional DB Concepts:** Transaction Processing - Concurrency and Consistency Control, Parallel, Distributed, Temporal database. Relational Databases and their Limitations, Structured vs. Unstructured data.					

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

The Emerging Trends and Needs of NoSQL DB's: Introduction, Overview, and History of NoSQL Databases, What is NoSQL? Why NoSQL? RDBMS vs. NoSQL, Brief history of NoSQL, CAP theorem (Brewer's Theorem), NoSQL pros/cons, NoSQL Categories, Production deployment. Cloud databases – Streaming Databases - Graph Databases-New SQL.				
Module 2	NoSQL Architectures (Comprehension)	Assignment	Understanding Ability	**6 Hours**
Topics: **Transaction in NoSQL**- BASE for reliable database transactions, Achieving horizontal scalability with Data base sharding, Brewers CAP theorem, NoSQL Data model: Document Data Model, Columnar Data Model, Key-Value Data Model, Graph Data Model.				
Module 3	Document Based Model (Application)	Assignment	DB Querying Ability	**8 Hours**
Topics: **Document Based Model** - Document, Collection, Naming, CRUD operation, querying, indexing, Replication, Sharding, Consistency. Implementation: Distributed consistency, Eventual Consistency, Capped Collection. Case studies-MongoDB/Casandra				
Module 4	Columnar and Graph Based Models (Application)	Assignment	DB Querying Ability	**8 Hours**
Topics: **Columnar Based Models**-Columnar Vs Row Oriented Models, Column-store Architectures: C-Store and Vector-Wise, Column-store internals and, Inserts/updates/deletes, Indexing, Adaptive Indexing and Database Cracking. **Graph Based Models**- Graph Database. Features, Consistency, Transactions, Availability, Query Features, Scaling, Suitable Use Cases.				
List of Laboratory Tasks: **Labsheet -1 [4 Practical Sessions]** **Experiment No 1:** **Level 1:** Write NoSQL queries to create and drop database. **Level 2 -** Write a NoSQL queries to display insert, update and delete a document. **Labsheet -2 [2 Practical Sessions]** **Experiment No. 1:** **Level 1 -** Write NoSQL queries to Implement Aggregation operations. **Level 2 -** Write NoSQL queries to Implement Indexing operations. **Labsheet – 3 [4 Practical Sessions]** **Experiment No. 1:** **Level 1 -** Write a NoSQL query to Create Document / Column Based Database. **Level 2 -** Write a NoSQL query to Create Key Value / Graph based Database. **Experiment No. 2:** **Level 1 -** Write a NoSQL query to create Backup, Replica and Restore of existing database. **Labsheet – 4 [3 Practical Sessions]** **Experiment No. 1:** **Level 1 -** Connect Java with any NoSQL DB and perform insert, retrieve, update and delete operations.				

Experiment No. 2:

Level 1 - Create simple objects and array objects using JSON.

Level 2 - Create objects like string, number, Boolean and null using JSON.

Labsheet – 5 [3 Practical Sessions]

Experiment No. 1:

Level 1 - Implement any Machine Learning algorithm for Big Data.

Targeted Application & Tools that can be used: Mango DB / Casandra

Project work/Assignment: Mention the Type of Project /Assignment proposed for this course

Build a real time database application using suitable frontend tool. Indicative areas include; health care, education, industry, Library, Transport and supply chain, etc.

Text Books

1] A. Silberschatz, H. Korth, S. Sudarshan 2021: "Database system concepts", 7th Edition, McGraw Hill Publications.

2] Sadalage, P., Fowler 2019: "NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence", 1st Edition, Wiley Publications.

References

1] Bradshaw, Chodorow 2019: "MongoDB: The Definitive Guide: Powerful and Scalable Data Storage", 3rd Edition, O'Reilly Publications.

2] Pivert 2018: " NoSQL Data Models: Trends and Challenges", 1st Edition. Wiley Publications.

3] Amit Phaltankar, Juned Ahsan, Michael Harrison, Liviu Nedov 2020: "MongoDB Fundamentals A hands-on guide to using MongoDB and Atlas in the real world" ,1st Edition, Packt publications.

Catalogue prepared by	Dr. K. Sankar Dr. M. Naga Raju
Recommended by the Board of Studies on	TBD
Date of Approval by the Academic Council	TBD

Course Code: CSE 6010	**Course Title: Social Network Analysis**	**L- P- C**	3	0	3
Version No.	1.0				
Course Pre-requisites	CSE5016-Essentials for Machine Learning, CSE6011-Application of Probability theory in Computer Science				
Anti-requisites	**NIL**				
Course Description	The rapid growth of social media has given the mass consumers a powerful tool to create knowledge and propagate opinions. At the same time, social media has created an unprecedented opportunity for companies to engage real-time interactions with consumers. In addition, the size and richness of social media				

	data has provided companies an unusually deep reservoir of consumer insights to transform the business and marketing operations. The social media analytics course will enable students to grasp the analytics tools to leverage social media data. The course will introduce tools such as engagement analytics, sentiment analysis, topic modeling, social network analysis, identification of influencers and evaluation of social media strategy.			
Course Outcomes	**On successful completion of this course the students shall be able to:** 1) Interpret the social network landscape and appreciate the importance of analytics in business. 2) Apply appropriate native analytics and measurement tools to analyze data in different social platforms 3) Use Natural Language Processing for efficient mining of web data 4) Demonstrate meaningful insights with actionable and strategic recommendations.			
Course Content:				
Module 1	**Network Science**	Quiz/Assignment		**9 classes**
Topics: Central Measures, Community Analysis, CPM, Homophily and Triadic Closure, Affiliation Networks, Schelling model of Segregation, Current Social Media landscape, working environment, Getting analyzing and visualizing the data, Getting started with the toolset, Need for SMA, Applications of SMA in different areas. Connecting, Capturing and cleaning of Social Data. APIs in nutshell, Introduction to authenticate techniques, Parsing API outputs, Basic cleaning techniques. Exploring GitHub's API, Analyzing GitHub Interest Graphs, Computing Graph Centrality Measures.				
Module 2	**Analyzing Social graphs and Sentiment**	Quiz	Project Development	**10 classes**
Topics: Exploring Facebook's Social Graph API, Open Graph Protocol, Analyzing Social Graph Connections, Mining your posts, Facebook Pages. Exploring Twitter's API, Analyzing Twitter using sentiment analysis, Frequency Analysis, Examining Patterns in Retweets.				
Module 3	**Mining web pages**	Project Development	Assignment	**11 classes**
Topics: Scraping, Parsing and Crawling the Web: BFS in Web Crawling, Discovering Semantics by Decoding Syntax: NLP Illustrated Step-by-Step, Sentence Detection in Human Language Data, Document Summarization, Entity-Centric Analysis: A Paradigm Shift, Summarizing Human Language Data, Quality of Analytics for Processing Human Language Data Campaigns and Consumer Reaction Analytics on YouTube: Structured and Unstructured, Scope and Process, Getting the data, Data pull, Data processing and Data analysis.				
Module 4	**Recommender Systems and SEO**	Quiz	Group Discussion	**8 classes**
Topics: Content-Based Recommendation and Collaborative Filtering, Search Engines, Google PageRank, IBM HITS				

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE

Targeted Application & Tools that can be used:

The applications of Social Media Analytics have been seen in industrial sector, sports and games, local governments services, tourism and hospitality services, politics, social issues, disaster management, community development issues, commerce and business applications, fashion industry, agricultural activities, online media, medical and health related services as well as supplier chain services.

Tools: Google Colab or Jupyter Notebook(Anaconda).

Project work/Assignment:

On completion of all Modules, students will be given a Mini Project to build a deep learning model for a given application.

Sample mini projects include:

Twitter Summaries

Twitter is famous for its character-limited posts. We can use this social media platform for an innovative summary-writing project. Consolidate the takeaways from a topic or reading discussed. Students should be able to understand the text, coherently organize the points and capture the central idea with 280 words, which is the character limit on Twitter.

Hashtag activism

Information and communication technologies provide a tremendous tool for spreading awareness and highlighting issues that may not be adequately represented in the mainstream media. Hashtag activism, in particular is concerned with driving social media traffic to oft-neglected topics. We can devise a project-based activity to teach our students about social justice, human rights, equality etc.

Text Book(s):

1) **Mathew A. Russell, "*Mining the Social Web*", O'Reilly, 3rd Edition, 2019.**

Reference(s):

1) **Marco Bonzanini, "*Mastering Social Media Mining with Python*", PacktPub, 2016.**

Topics related to development of "Employability Skills": Problem solving, Creative Thinking, Team work, Prototype Development.

Topics related to development of "Entrepreneurship": Effective Communication, Strategic Thinking, Creative Thinking.

Catalogue prepared by	Prof. G Poornima, Dr. Tapas Guha
Recommended by the Board of Studies on	BOS NO: 12th BOS, held on 04/08/21
Date of Approval by the Academic Council	Academic Council Meeting No. 16, Dated 23/10/21

REGISTRAR

PRESIDENCY UNIVERSITY Registrar BANGALORE