

B.Sc., Data Science

Department of Computer Applications

Syllabus

2024-2025



**SENGAMALA THAYAR EDUCATIONAL TRUST
WOMEN'S COLLEGE (AUTONOMOUS)**

(Affiliated to Bharathidasan University, Tiruchirappalli)
(Accredited by NAAC) | (An ISO 9001:2015 Certified Institution)

**Sundarakkottai, Mannargudi – 614 016,
Thiruvarur (Dt.), Tamil Nadu, India.**

Programme Objectives:

- Graduates will have the analytical abilities necessary to apply the fundamental concepts they have learned in real-world scenarios.
- Graduates with improved knowledge and skills will be able to find employment, start their own businesses, or pursue further education.
- Graduates who have studied contemporary software technologies will be well-suited to contribute as data scientists.
- Graduates who read Allied articles will be able to understand the principles linked to data science.
- Graduates will be instilled with ethical principles and social concerns, to promote a peaceful society

Programme Outcomes:

PO No.	Program Outcomes <i>(Upon completion of the B.Sc. Data Science Degree Program, the Undergraduate will be able to)</i>
PO-1	Disciplinary knowledge: Demonstrate comprehensive knowledge and understanding of for a range of computer applications, computer organization, techniques of Computer Networking, Software Engineering, Web development, Database management and Advance Java that form a part of an undergraduate programme of study in Bachelor of Computer Applications.
PO-2	Critical thinking and Problem Solving: Think critically about the issues and identify, critically analyze and solve problems from the disciplines of concern using appropriate tools and techniques and the knowledge, skills and attitudes acquired and extrapolate the same to real life situations.
PO-3	Scientific reasoning: Analyse, interpret and draw conclusions from quantitative/qualitative data; and critically evaluate ideas, evidence and experiences from an open-minded and reasoned perspective.
PO-4	Digital literacy and Effective Communication: Use ICT in a variety of learning situations and speak, read, write and listen clearly in person and through electronic media in English and in one or more Indian languages, and make meaning of the world by connecting people, ideas, books, media and technology.
PO-5	Individual and Team Work: Effectively accomplish tasks individually as well as work effectively and respectfully as member or leader with diverse teams, facilitate cooperative or coordinated effort on the part of a group, and act together as a group or a team in the interests of a common cause and work efficiently as a member of a team.

PO-6	Multicultural Competence, Social Interaction and Effective Citizenship: Understand the values and beliefs of multiple cultures, global perspectives, engage and interact respectfully with diverse groups and elicit views of others, mediate disagreements and help reach conclusions in group settings, and demonstrate empathetic social concern and equity centred national development.
PO-7	Moral and Ethical Awareness: Embrace moral/ethical values in conducting one's life, formulate a position/argument about an ethical issue from multiple perspectives, and use ethical practices in all work.
PO-8	Environment and Sustainability: Understand the impacts of technology and business practices in societal and environmental contexts, and sustainable development.
PO-9	Human values and Gender Issues: Understand major ideas, values, beliefs, the nature of the individual and the relationship between self and the community and aware of the various issues concerning women and society
PO-10	Self-directed and Lifelong learning: Acquire knowledge and skills, including learning "how to learn", that are necessary for participating in learning activities throughout life and to engage in independent and life-long learning in the broadest context of socio- technological changes.

PSO. No.	PROGRAMME SPECIFIC OUTCOME Students will be able to
PSO 1	Apply the concepts and practical knowledge in analysis, design and development of computing systems and applications to multi-disciplinary problems.
PSO 2	To provide a concrete foundation and enrich their abilities to qualify for Employment, Higher studies and Research in Data science with ethical values

**SENGAMALA THAYAAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**

(Affiliated to Bharathidasan University)

(Accredited by NAAC | An ISO 9001:2015 Certified Institution) SUNDARAKKOTTAI,

MANNARGUDI – 614 016,

TAMILNADU, INDIA.

Department of Computer Applications



B.Sc., Data Science

Course Structure under CBCS

LEARNING OUTCOMES BASED CURRICULUM FRAMEWORK (CBCS - LOCF)

(For the candidates admitted in the academic year 2024-2025)

ELIGIBILITY: A pass in 12th standard with Computer Science or Mathematics or Statistics or Business Mathematics courses.

Sem.	Part	Course	Title	Ins. Hrs	Credit	Exam Hours	Marks		Total
							Int.	Ext.	
I	I	Language Course – I Tamil \$ / Other Languages + #		6	3	3	25	75	100
	II	English Course - I		6	3	3	25	75	100
	III	Core Course – I (CC)	Programming in C and Data Structures	5	5	3	25	75	100
		Core Practical – I (CP)	Programming in C Lab	4	4	3	40	60	100
		First Allied Course – I (AC)		4	4	3	25	75	100
		First Allied Practical (AP)		3	-	-	-	-	-
	IV	Value Education		2	2	3	25	75	100
TOTAL				30	21	-	-	-	600
II	I	Language Course - II Tamil \$ / Other Languages + #		6	3	3	25	75	100
	II	English Course - II		6	3	3	25	75	100
	III	Core Course – II (CC)	Python Programming	5	5	3	25	75	100
		Core Practical – II (CP)	Python Programming – Lab	4	4	3	40	60	100
		First Allied Practical (AP)		3	3	3	40	60	100
		First Allied Course – II (AC)		4	4	3	25	75	100
		Add on Course – I ##	Professional English I	6*	5	3	25	75	100
IV	Environmental Studies		2	2	3	25	75	100	
TOTAL				30	29	-	-	-	800

III	I	Language Course – III Tamil \$ / Other Languages + #		6	3	3	25	75	100
	II	English Course - III		6	3	3	25	75	100
	III	Core Course – III (CC)	Database Systems	5	5	3	25	75	100
		Core Practical - III (CP)	Database Systems Lab	4	4	3	40	60	100
		Second Allied Course – I (AC)		4	4	3	25	75	100
		Second Allied Practical (AP)		3	-	-	-	-	-
		Add on Course – II ##	Professional English II	6*	4	3	25	75	100
IV	Non-Major Elective I @ - Those who choose Tamil in Part I can choose a non-major elective course offered by other departments. Those who do not choose Tamil in Part I must choose either a) Basic Tamil if Tamil language was not studied in school level or b) Special Tamil if Tamil language was studied upto 10 th & 12 th std.		2	2	3	25	75	100	
TOTAL				30	25	-	-	-	700
IV	I	Language Course –IV Tamil \$ / Other Languages + #		6	3	3	25	75	100
	II	English Course – IV		6	3	3	25	75	100
	III	Core Course - IV (CC)	Big Data Analytics with NoSQLys Database	5	5	3	25	75	100
		Core Practical - IV (CP)	Big Data Analytics Lab	4	4	3	40	60	100
		Second Allied Practical (AP)		3	3	3	40	60	100
		Second Allied Course – II (AC)		4	4	3	25	75	100
	IV	Non-Major Elective II Those who choose Tamil in Part I can choose a non-major elective course offered by other departments. Those who do not choose Tamil in Part I must choose either a) Basic Tamil if Tamil language was not studied in school level or b) Special Tamil if Tamil language was studied upto 10 th & 12 th std.		2	2	3	25	75	100
TOTAL				30	24	-	-	-	700

V	III	Core Course - V (CC)	Data Mining, Data Warehousing and Data Visualization	5	5	3	25	75	100
		Core Course – VI (CC)	Network Security	5	5	3	25	75	100
		Core Course – VII (CC)	Cloud Computing	5	5	3	25	75	100
		Core Practical -V (CP)	Data Mining, and Data Visualization Lab	4	4	3	40	60	100
		Major Based Elective – I (Any one)	1. Internet of Things 2. Block Chain Technology	5	4	3	25	75	100
	IV	Skill Based Elective I	Shell Programming	4	2	3	25	75	100
		Soft Skills Development		2	2	3	25	75	100
	TOTAL				30	27	-	-	-
VI	III	Core Course - VIII (CC)	R Programming	6	5	3	25	75	100
		Core Course - IX (CC)	Machine Learning	6	5	3	25	75	100
		Core Practical – VI (CP)	Machine Learning Lab	4	4	3	40	60	100
		Major Based Elective – II (Any one)	1. Computing Technologies for Data Science 2. Augmented Reality and Virtual Reality	5	4	3	25	75	100
		Project		4	4	-	20	80	100
	IV	Skill Based Elective – II	Web Technology	4	2	3	25	75	100
	V	Gender Studies		1	1	3	25	75	100
		Extension Activities **		-	1	-	-	-	-
TOTAL				30	27	-	-	-	700
GRAND TOTAL				180	154	-	-	-	4200

List of Allied Courses

First Allied Course

Second Allied Course

Mathematics

Statistics

SUMMARY OF CURRICULUM STRUCTURE OF UG PROGRAMMES

Sl. No.	Part	Types of the Courses	No. of Courses	No. of Credits	Marks
1.	I	Language Courses	4	12	400
2.	II	English Courses	4	12	400
3.	III	Core Courses	8	40	800
4.		Core Practical	7	29	700
5.		Allied Courses I & II	4	16	400
6.		Allied Practical	2	4	200
7.		Major Based Elective Courses	2	8	200
8.		Add on Courses	2	8	200
9.		Project	1	3	100
10.		IV	Non-Major Elective Courses (Practical)	2	4
11.	Skill Based Elective Courses		2	4	200
12.	Soft Skills Development		1	2	100
13.	Value Education		1	2	100
14.	Environmental Studies		1	2	100
15.	V	Gender Studies	1	1	100
16.		Extension Activities	1	1	--
Total			43	148	4200

*For those who studied Tamil upto 10th +2 (regular Stream)

Those who studied Tamil upto 10th +2 but opt for other languages in degree level under Part I should study special Tamil in Part IV

** Extension Activities shall be outside instruction hours.

NME courses are offered to other Under Graduate programmes. For B.Sc Data science programme. NME courses of other department are offered.

NOTE:			
S.NO.	PARTICULARS	CIA	ESE
1.	THEORY	25	75
2.	PRACTICAL	25	75
3. Separate passing minimum is prescribed for Internal and External marks			

**SENGAMALA THAYAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester I-CC-I: Programming in C and Data Structures

Ins.Hrs/Week: 5

Course Credit: 5

Course Code

UNIT –I

(15 HOURS)

Basic of C: History of C and its importance – Structure of a C program – Data Types – Constants and Variables – Operators and Expressions – Order of Precedence, Evaluating of Arithmetic Expressions – Type Conversion- Decision Statements: if, if-else, and nested if statements.

UNIT – II

(15 HOURS)

Loops Structures: For Loop, While, Do-while loop – Arrays: - One Dimensional Array, Two-dimensional Arrays, Character Arrays and Strings – Functions: Function with arrays- Function with decision and looping statements - Recursion.

UNIT - III

(15 HOURS)

Pointers: Introduction – Pointer Expressions – Chain of Pointers – Pointers and Arrays – Array of Pointers – Pointers as function arguments – Functions returning Pointers – Pointers to Functions – Function pointer – Structures - declaration, initialization, Array of Structures – Pointer to structures, Structures and functions – Typed of Enumerated data types, Unions.

UNIT – IV

(15 HOURS)

Strings Processing, Standard string library functions – Files: introduction and files functions – Writing and reading in Text mode – Simple application: Display the contents of a file. Write data to a file. Append data to an existing file – File IO – Reading and writing structures.

Stack: LIFO concept, Stack operations, Array implementation of stack – Queue: FIFO concept, Queue operations, Array implementation of queue – Singly Linked List: concepts, operations – Doubly Linked List: concepts, operations – Trees: General trees, Binary trees.

TEXT BOOK

1. E. Balagurusamy, “Programming in ANSI C”, Tata McGraw Hill, New Delhi, Seventh Edition, 2016.
2. E. Horowitz, S. Sahni and Susan Anderson Freed, “Fundamental Data Structures in C”, 2ed, Orient Black Swan Publisher, 2009.

REFERENCE BOOK

1. Byron S. Gottfried, “Programming with C”, Schaum’s Outline Series, Tata- McGraw Hill Edition, New Delhi, 1991.
2. E. Karthikeyan, “A Textbook on C Fundamentals, Data Structures and Problem Solving”, Prentice-Hall of India Private Limited, New Delhi, 2008.
3. Yashavant Kanetkar, “Let us C”, BPB Publications, Tenth Edition, New Delhi, 2010.
4. Szuhay, Jeff, and Szuhay, Jeff, “Learn C Programming: A Beginner's Guide to Learning C Programming the Easy and Disciplined Way”, Packt Publishing, 2020.
5. Jena, Sisir Kumar, and Jena, Sisir Kumar, “C Programming: Learn to Code”, CRC Press, 2021.

E-RESOURCES

1. <https://www.tutorialspoint.com/cprogramming/index.htm>
2. <https://www.w3schools.in/data-structures/intro>

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Summarize the basic knowledge to develop C programs
2. Manipulate Looping, arrays and functions
3. Apply and write programs for solving real world problems
4. Create open, read, manipulate, write and close files.
5. Understand the basic concepts in data structures.

**SENGAMALA THAYAAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester I-CP-I: PROGRAMMING IN C LAB

Ins.Hrs/Week: 5

Course Credit: 5

Course Code

1. Write a Program
 - (i) To convert temperature from degree Centigrade to Fahrenheit,
 - (ii) Find whether given number is Even or Odd,
 - (iii) Find the greatest of three numbers.
2. Write a Program to display Monday to Sunday using switch statement
3. Write a Program to display first Ten Natural Numbers and their sum.
4. Write a Program to perform Multiplication of Two Matrices.
5. Write a Program
 - (i) To find the maximum number in an Array using pointer.
 - (ii) To reverse a number using pointer.
 - (iii) To add two numbers using pointer.
6. Write a Program to solve Quadratic Equation using functions.
7. Write a Program to find factorial of a number using Recursion.
8. Write a Program to demonstrate Call by Value and Call by Reference.
9. Write a Program to create a file containing Student Details.
10. Write a program to implement a stack using singly linked list, Implement Queue using Linked List.

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Relate the use of language constructs to solve simple programs
2. Develop programs for various concepts in C language
3. Understand and trace the execution of the list of programs
4. Understand the usage of file handling in C programming
5. Solve data problems related to data structures

**SENGAMALA THAYAAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester II-CC-II: PYTHON PROGRAMMING

Ins.Hrs/Week: 5

Course Credit: 5

Course Code

UNIT - I: (18 HOURS)

Object Oriented Programming: Procedural and Object-Oriented Programming – Classes – Working with instances – techniques for designing classes. – Inheritance: introduction to inheritance – Polymorphism

UNIT - II: (18 HOURS)

Python: Installing python- The python Interpreter – Interactive mode – Writing and running programs in script mode- IDLE programming environment – Input, processing and output – Displaying output with print function- print function – strings and string literals-comments-variables.

Reading input from the keyboard- operators- more about output- decision structures and Boolean logic-repetition structures

UNIT III (18HOURS)

Lists, Tuples, Strings, Dictionaries and Set: Sequences – Introduction to Lists – List slicing – ‘in’ operator – list methods and built-in-functions – copying lists – processing lists – Two Dimensional Lists – Tuples. Strings: Basic String Operations – String Slicing Testing, Searching and manipulating strings – Dictionaries and Set: Dictionaries – Set – Serializing Objects.

UNIT - IV: (18 HOURS)

Functions, Modules and File Handling: Functions: introduction to functions – Defining and calling functions – designing a program to use functions – Local variables – passing arguments to functions – Global variable and Global Constants- –Value returning functions: generation – user defined value returning functions – Modules: math module- Storing functions in modules - Files: Introduction to File Input and Output – Using Loops to process files – processing records - Exceptions – Python Standard Library - Regular Expression..

UNIT - V: (18 HOURS)
Data Analysis using Python: Load data into a Data Frame -
Fundamentals of Data Manipulation with Python.

TEXT BOOKS

1. Allen B. Downey, "Think Python: How to Think Like a Computer Scientist", 2nd edition, Updated for Python 3, Shroff/O'Reilly Publishers, 2016
2. Satyanarayana, Radhika Mani, Jagadesh,"Python Programming", Universities Press (India) Pvt. Ltd 2018

REFERENCE BOOKS

1. Wesley Chun "Core python Programming" Pearson Education, 2006,
2. Al Sweigart, "Invent your own computer games with python", 2nd edition, 2008
3. Bill Lubanovic, "Introducing Python", O'Reilly, First Edition-Second Release, 2014.
4. Tony Gaddis, "Starting out with python", 2nd edition, Addison Wesley, Pearson
5. Michael Dawson, "Python programming for the absolute beginner", Premier press, 2003.

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Interpret and manipulate the OOPs Concepts
2. Install python and write programs to solve simple problems
3. Explain basic data structures in Python
4. Store and manipulate data using file system
5. Implement Python packages and libraries

**SENGAMALA THAYAAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester II-CP-II: PYTHON PROGRAMMING LAB

Ins.Hrs/Week: 5

Course Credit: 5

Course Code

1. Write Python programs for the following: (using Basics of Python)
 - a. Purposefully raise Indentation Error and correct it.
 - b. Compute distance between two points taking input from the user (use Euclidean distance formula).
 - c. To take numbers as command line arguments and print its sum
2. Write Python programs for implementing the following: (using Control Flow)
 - a. Finding the factorial of a number.
 - b. Print the prime numbers below 100
3. Write Python programs for implementing the following: (using Strings)
 - a. Count the numbers of characters in the string and store them in a dictionary data structure
 - b. Using split and joins methods in the string and trace a birthday with a dictionary data structure.
4. Write Python programs for the following: (using List)
 - a. Finding mean, median, mode for the given set of numbers in a list.
 - b. Function dups to find all duplicates in the list.
5. Write Python programs to do the following (using Methods)
 - a. Create a Python class named Circle constructed by a radius and two methods which will compute the area and the perimeter of a circle.
 - b. Create a Python class named Rectangle constructed by a length and width and a method which will compute the area of a rectangle.
6. Write Python programs to implement
 - a. Constructors to assign the PI_VALUE.
 - b. Polymorphism to print area of square/rectangle depending upon the number of parameters passed.
7. Write Python programs to implement
 - a. Inheritance
 - b. Method overloading and overriding
8. Write a python program (using files)

- a. To open and write "Hello World" into a file.
 - b. To write the content "Hi Python Programming" for the existingfile.
 - c. To import values from a CSV file to create Pandas Data Frame
9. Write a Python program to create an Email slicer.
 10. Write a Python program to generate password.

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

- Write simple programs using control structures, functions and strings
- Develop programs using tuples, lists, sets and dictionary
- Write simple programs using Constructors, Method overloading and inheritance
- Develop programs using files and regular expressions
- Write simple programs using packages and exception handling

**SENGAMALA THAYAAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester III-CC-III: Database Systems

Ins.Hrs/Week: 5

Course Credit: 5

Course Code

UNIT - I:

Databases and Database Users: Introduction - characteristics of data base approach - intended uses of a DBMS- advantages and implication of database approach. Database Systems Concepts and Architecture: Data models- schemas and instances- DBMS architecture and data independence - database languages and interfaces- database system environment- classification of data base management systems

UNIT - II:

Data Modelling :High level conceptual data models for database design- Entity types- entity sets- attributes and keys- Relationships- relationship types- roles and structural constraints - Weak entity types- ER diagrams and design issues.

UNIT - III:

Relational Data Model and Relational Algebra: Brief discussion on CODD rules- relational model concepts- constraints and schemas- Update operation on relations- basic and additional relational algebra operations- queries in relational algebra

UNIT - IV:

Structured Query Language (SQL): Data definition in SQL- Basic and complex queries- SQL Insert- delete- update statements and views in SQL- embedded SQL. Database Design: Design guidelines for relational schemas- functional Dependencies - normalization -1st, 2nd, 3rd, 4th and 5th normal forms- Database design process- factors influencing physical database design guidelines and guidelines for relational systems.

UNIT - V:

Record Storage and Primary File Organizations: Secondary storage devices- buffering of blocks- placing file records on disk- operations on files- heap files and sorted files – hashing techniques. Index Structure of Files: Single – level and multilevel ordered indexes- dynamic multi-level indexes using B-trees and B+ trees.

UNIT VI: Current Contours

(For Continuous Internal Assessment Only):

Contemporary Developments Related to the Course during the Semester Concerned.

TEXT BOOKS

1. Ramez Elmasri and Shamkanth B. Navathe, “Fundamentals of Database Systems”, Pearson Education, 5th Edition, 2008, ISBN: 978- 81- 317 – 1625-0.
2. Abraham Silberschatz, Hendry F. Korth, S Sudharshan,” Database System Concepts”, 6th Edition, McGraw Hill International, 2019.

REFERENCE BOOKS

1. Raghu Ramakrishnan, “Database Management System”, McGraw Hill, 3rd Edition, 1997, ISBN: 0071230572.
2. Gary W. Hansen and James V. Hansen, “Database Management and Design” PHI Pvt. Ltd., 2nd Edition, 1995 ISBN: 81-203-1465-4.
3. Date C J, Kanna A, “Database Systems”, Pearson Education, 8th Edition, 2006, ISBN: 978- 81-775-8556-8.
4. H G Molina, J D Ullman, J Widom, “Database Systems – The Complete Book” ,Pearson; 2nd edition (4 September 2008).

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Understand the basic concepts of database systems
2. Understand the relational data model.
3. Apply the concepts of data base design.
4. Design and create database with normalization
5. Analyse the storage and file organization concepts.

**SENGAMALA THAYAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
**SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.**

**Department of Computer Applications
B.Sc., Data Science**

Semester III-CP-III: DATABASE SYSTEMS LAB

Ins.Hrs/Week: 5

Course Credit: 4

Course Code

1. Perform the following using SQL: 9942369258
 - a. Viewing all databases,
 - b. Creating a Database,
 - c. Viewing all Tables in a Database.
2. DDL: Creating Tables (With and Without Constraints)
 - a. Primary key, Foreign Key,
 - b. Unique, Default
 - c. Null, Not null
 - d. Enable Constraints, Disable Constraints Drop Constraints
3. DML: Inserting/Updating/Deleting Records in a Table
4. TCL: Saving (Commit) and Undoing (rollback)
5. Perform the following:
 - a. Altering a Table,
 - b. Dropping a Table
 - c. Truncating a Table
 - d. Renaming a Table
 - e. Backing up / restoring a Database.
6. For a given set of relation schemes, create tables and perform the following
 - a. Simple Queries,
 - b. Simple Queries with Aggregate functions,
 - c. Queries with Aggregate functions (group by and having clause),
7. For a given set of relation schemes, create tables and perform the following
 - a. Queries involving- Date Functions
 - b. Queries involving -String Functions
 - c. Queries involving - Math Functions
8. Perform the following with creating necessary tables.
 - a. Join Queries- Inner Join
 - b. Outer Join
 - c. Sub queries
9. Perform the following with creating necessary tables
 - a. With IN clause
 - b. With EXISTS clause
10. For a given set of relation tables perform the following
 - a. Creating Views (with and without check option),
 - b. Dropping views,
 - c. Selecting from a view

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Write queries to manipulate data.
2. Demonstrate the aggregate functions and set operations.
3. Apply the join operations.
4. Know about usage of nested sub queries
5. Understand the method to create views

**SENGAMALA THAYAAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
**SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.**

**Department of Computer Applications
B.Sc., Data Science**

Semester IV-CC-IV: Big Data Analytics with NOSQL Databases

Ins.Hrs/Week: 5

Course Credit: 5

Course Code

UNIT - I:

Introduction: Introduction to big data– sources of big data – real time application of big data – challenges in collecting and validating big data.

UNIT - II:

Hadoop: Hadoop – Hadoop ecosystem for processing big data –Hadoop cluster- Hadoop distributed file system – working with files in HDFS – map reduce technique for big processing – Joining data from different sources using map reduce.

UNIT - III:

Hive and Pig: Hive and pig – need for high-level tools in big data processing – unstructured and structured data – Not Only SQL (NOSQL) commands – use of Hive as an interface to Hadoop – Use of pig as a programming Tool for big data processing

UNIT - IV:

Introduction to NoSQL: An Overview of NoSQL - Defining NoSQL - What NoSQL is and what it is not - List of NoSQL Databases - Characteristics of NoSQL -RDBMS approach – Challenges

UNIT - V:

NoSQL Storage Types: NoSQL Storage Types - Column-oriented databases - Document store – Key value store -Multi-storage type databases - Advantages and Drawbacks - Transactional application - Computational application - Web-scale application

TEXT BOOKS

1. Gaurav Vaish, "Getting Started with NoSQL", Published by Packt Publishing Ltd., 2013, ISBN 978-1-84969-4-988

REFERENCE BOOKS

1. "Big data analysis for Dummies", Dummies press, 2011.
2. Hadoop Fundamentals, packet Publications, 2012.
3. Bart Baesens, Analytics in a Big Data World: The Essential Guide to DataScience and its Applications, Wiley, 2014
4. Soumendra Mohanty, Madhu Jagadeesh, and Harsha Srivatsa, "Big Data Imperatives: Enterprise Big Data Warehouse, BI Implementations and Analytics", Published by Apress Media, 2013.
5. Ankam, Venkat, Big Data Analytics. India, Packt Publishing, 2016.
6. Prabhu, C.S.R., Big Data Analytics: Systems, Algorithms, Applications, Springer, 2019.

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Apply big data concepts in various domains
2. Understand hadoop ecosystems
3. Apply the concepts of pig and hive
4. Implement NoSQL Databases
5. Explain various NoSQL storage types

**SENGAMALA THAYAAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester IV-CP-IV: BIG DATA ANALYTICS LAB

Ins.Hrs/Week: 5

Course Credit: 4

Course Code

1. Set up a pseudo-distributed, single-node Hadoop cluster backed by the Hadoop Distributed File System
2. Installation of Hadoop Ecosystem - Pig
3. Installation of Hadoop Ecosystem - Hive
4. Write a hive query to load the data into table from any local file on the system.
5. Write a hive query, which returns the contents of the whole table.
6. Demonstrate the usage of Hive functions.
7. Installation of MongoDB
8. Reading CSV file and loading into MongoDB
9. Reading JSON file and Loading into MongoDB
10. Reading MongoDB and writing into MySQL

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Install Hadoop distributed file system
2. Install Hadoop ecosystems
3. Write hive queries
4. Implement MongoDB for unstructured data
5. Use read and write various files in MongoDB

**SENGAMALA THAYAAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester V-CC-V: Data Mining, Data Warehousing and Data Visualization

Ins.Hrs/Week: 5 Course Credit: 5 Course Code

UNIT - I:

Introduction: Introduction to data mining: Datamining process- Applications- Techniques-Examples – Future – Limitations-Software - Data Collection and Pre- processing - Outliers - Mining Outliers - Missing Data - Types of Data - Computing Distance - Data Summarising Using Basic Statistical Measurements - Displaying Data Graphically Multidimensional Data Visualization.

UNIT - II:

Association Rules Mining: Basics - The Task and a Naïve Algorithm - The Apriori Algorithm - Improving the Efficiency of the Apriori Algorithm - Apriori-TID - Direct Hashing and Pruning (DHP) - Dynamic Itemset Counting (DIC) - Mining Frequent Patterns without Candidate Generation (FP–Growth) - Performance Evaluation of Algorithms – Case Study: Mining Customer Value: From Association Rules to Direct Marketing.

UNIT - III:

Classification & Cluster Analysis: Classification: Decision Tree – Building a decision tree-Split algorithm – Rules-Naïve Bayes method – Estimating Predictive Accuracy of Classification Methods - Improving Accuracy of Classification Methods Cluster analysis: Desired Features of Cluster Analysis Types of Cluster Analysis Methods - Partitional Methods - Hierarchical Methods - Density-Based Methods - Quality and Validity of Cluster Analysis Methods.

UNIT - IV:

Data warehousing: Data warehousing – Operational Data Stores - Data Warehouses – Data warehouse Design – Guidelines for Data Warehouse Implementation – OLAP - Multidimensional View and Data Cube - Data Cube Implementations - Data Cube Operations.

UNIT - V:

Data Visualization: A Conceptual Framework for Data Visualization: Data, information, knowledge, and insight - Developments in data visualization - Visualization in decision-making - Visualization fit - Visualization plots - Data Analysis and Visualization: Visualization planning - - Visually representing the results - Perception and presentation methods - best practices for visualization.

TEXT BOOKS

1. G.K. Gupta, "Introduction to Data mining with case studies", 3rd Edition, PHI Learning Private Limited, Delhi, 2014 (ISBN-978-81-203-5002-1). (Unit1: chapter2: 2.1-2.10, Unit2: chapter 3: 3.1-3.10, Unit3: chapter 4:4.2-4.4,4.9- 4.11,Chapter5: 5.2-5.6,5.8, Unit4: chapter8: 8.2-8.5, Chapter9: 9.2, 9.5, 9.6, 9.7)

REFERENCE BOOKS

1. Kirthi Raman, "Mastering Python Data Visualization", Packt Publishing, 2015. (Unit5: Chapter1, Chapter2)
2. S. K. Mourya, Shalu Gupta, "Data Mining and Data Warehousing", Alpha Science International Limited, 2013
3. K.P. Soman & Shyam Diwakar and V. Ajay, "Insight to Data Mining Theory and Practice", Prentice Hall of India, 2006. (ISBN -81-203- 2897-3)
4. Jiawei Han and Micheline Kamber, "Data Mining Concepts and Techniques", Elsevier, Second Edition, 2007 (ISBN: 81-312-0535-5)
5. Bhatia, Parteek, Data Mining and Data Warehousing: Principles and Practical Techniques, Cambridge University Press, 2019.
6. Wilke, Claus, Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures, O'Reilly Media, 2019.

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Apply basic concepts of data mining and preprocessing in various domains
2. Implement Association Rule.
3. Evaluate Classification and clustering
4. Explain data warehousing
5. Implement various data visualization techniques

**SENGAMALA THAYAAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester V-CC-VI: NETWORK SECURITY

Ins.Hrs/Week: 5 Course Credit: 5 Course Code

UNIT - I:

Introduction: Security Trends – The OSI Architecture – Security Attacks – Security Services – Security Mechanisms – A model for Network Security – Classic Encryption Techniques – Symmetric Cipher Model – Substitution Techniques – Transposition techniques – Rotor Machines – Steganography.

UNIT - II:

Block Ciphers and Data Encryption Standards: Block Cipher - Principles - Data Encryption Standard – The strength of DES – Differential and Linear Cryptanalysis – Block Cipher design principles – Advanced encryption Standard – The AES Cipher

UNIT - III:

Public Key Encryption and Hash functions: Principles of Public Key Crypto Systems – The RSA algorithm – Message Authentication – Authentication Requirements – Authentication Functions – Message Authentication codes – Hash Functions – Security of Hash Functions and MAC – Whirlpool - Digital Signatures
– Authentication Protocols – Digital Signature standard

UNIT - IV:

Authentication Applications: Kerberos – X.509 Authentication Service – PKI – Electronic Mail Security - Pretty Good Piracy – S/MIME – IP Security – IP Security Overview – IP Security Architecture – Authentication Header – Key Management – Web Security – Web Security Considerations-SSL and Transport Layer Security

UNIT - V:

System Security: Intruders – Intrusion Detection – Password Management – Viruses – DOS and DDOS Attacks – Firewalls – Firewall Design Principles – Trusted Systems – IT Security Evaluation

TEXT BOOKS

1. William Stallings, “Cryptography and network Security”, Principles and Practices, Prentice Hall (Pearson Education), Fourth Edition, 2006. (**Unit 1:** Chapter1, Chapter2; **Unit 2:** Chapter3, Chapter5; **Unit 3:** Chapter9, Chapter11, Chapter12, Chapter13; **Unit 4:** Chapter14, Chapter15, Chapter16, Chapter17; **Unit 5:** Chapter18, Chapter19, Chapter20.

REFERENCE BOOKS

1. Kevin Mitnick (Author), Robert Vamosi Mikko Hypponen, “The Art of Invisibility: The World's Most Famous Hacker Teaches You How to Be Safe in the Age of Big Brother and Big Data Hardcover” , 2017 (ISBN-13: 9780316380522)
2. AtulKahate, “Cryptography and Network Security”, Tata McGraw Hill Publications, New Delhi.
3. Stallings, “Data and Computer Communications”, Pearson Education 2012, 7th Edition.
4. B. A. Forouzan, “Data Communications and Networking”, Tata McGraw Hill 2007, 4th Edition.
5. F. Halsall, “Data Communications, Computer Networks and Open Systems”, Pearson Education 2008

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Describe the fundamentals of network security
2. Identify the various types of ciphers
3. Describe encryption standards
4. Ensure authentication applications
5. Identify and apply system security

**SENGAMALA THAYAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester V-CC-VII: CLOUD COMPUTING

Ins.Hrs/Week: 5

Course Credit: 5

Course Code

UNIT - I:

Cloud Computing Foundation: Introduction to Cloud Computing – Move to Cloud Computing – Types of Cloud – Working of Cloud Computing

UNIT - II:

Cloud Computing Architecture: Cloud Computing Technology – Cloud Architecture – Cloud Modeling and Design - Virtualization: Foundation – Grid, Cloud and Virtualization – Virtualization and Cloud Computing

UNIT - III:

Data Storage and Cloud Computing: Data Storage – Cloud Storage – Cloud Storage from LANs to WANs – Cloud Computing Services: Cloud Services – Cloud Computing at Work

UNIT - IV:

Cloud Computing and Security: Risks in Cloud Computing – Data Security in Cloud – Cloud Security Services – Cloud Computing Tools: Tools and Technologies for Cloud – Cloud Mashups – Apache Hadoop – Cloud Tools

UNIT - V:

Cloud Applications – Moving Applications to the Cloud – Google Cloud Applications – Amazon Cloud Services – Cloud Applications

Text book:

1. A. Srinivasan and J.Suresh, “Cloud Computing – A Practical Approach for Learning and Implementation”, Pearson India Publications, 2014. (**Unit 1:** Chapter1, Chapter2, Chapter3, Chapter4; **Unit 2:** Chapter5, Chapter6, Chapter7, Chapter8, Chapter9, Chapter10; **Unit 3:** Chapter11, Chapter12, Chapter13, Chapter14, Chapter16, Chapter17; **Unit 4:** Chapter18, Chapter19, Chapter20, Chapter24, Chapter25, Chapter26,Chapter27; **Unit 5:** Chapter28, Chapter30, Chapter31, Chapter32)

Reference books:

1. Rajkumar Buyya, James Broberg, Andrzej, “Cloud Computing: Principles and Paradigms”, Wiley India Publications, 2011.
2. Anthony T.Velte , Toby J. Velte Robert Elsenpeter, “Cloud computing a practical approach”, TATA McGraw- Hill , New Delhi – 2010
3. Michael Miller, “Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online” - Que 2008
4. Judith Hurwitz , Robin Bloor , Marcia Kaufman ,Fern Halper, “Cloud computing for dummies”, Wiley Publishing, Inc, 2010
5. Comer, Douglas, The Cloud Computing Book: The Future of Computing Explained, CRC Press, 2021.

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Describe various types of cloud
2. Identify the cloud computing basics and its architecture
3. Implement data storage and security
4. Explore various cloud applications
5. Describe various cloud services.

**SENGAMALA THAYAAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
**SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.**

**Department of Computer Applications
B.Sc., Data Science**

Semester V-CP-V: DATA MINING AND VISUALIZATION LAB

Ins.Hrs/Week: 5

Course Credit: 4

Course Code

1. Introduction to Python libraries for Data Mining : NumPy, SciPy, Pandas, Matplotlib, Scikit-Learn Write a Python program to do the following operations:
(Library: numpy)
 - a. Create multi-dimensional arrays and find its shape and dimension
 - b. Create a matrix full of zeros and ones
 - c. Reshape and flatten data in the array
 - d. Append data vertically and horizontally
 - e. Apply indexing and slicing on array
2. Write a Python program to do the following operations: (Library: Pandas)
 - a. Loading data from CSV file
 - b. Compute the basic statistics of given data - shape, no. of columns, mean
3. Write a python program to load the dataset and understand the input data
(Library : Scipy)
 - a. Load data and describe the given data
 - b. Identify missing data items and outlier data items
4. Write a python program to impute missing values with various techniques on given dataset.
 - a. Remove rows/ attributes
 - b. Replace with mean or mode
5. Write a python program to find rules that describe associations by using Apriori algorithm between different products with appropriate data set. (Libraries: NumPy, SciPy, Matplotlib, Pandas)
 - a. Display top 5 rows of data
 - b. Find the rules with min_confidence : .2, min_support= 0.0045, min_lift=3, min_length=2
6. Write a python program to (use Iris data set)
Iris data set consists of 50 samples from each of three species of Iris: Iris setosa, Iris virginica and Iris versicolor. Four features were measured from each sample: the length and the width of the sepals and petals, in centimetres. (Libraries: import numpy as np)
 - a. Calculate Euclidean Distance.
 - b. Get Nearest Neighbors
 - c. Make Predictions
7. Write a python program (use Iris data set)

- a. To build a decision tree classifier to determine the kind of flower by using given dimensions.
 - b. Training with various split measures(Gini index, Entropy and Information Gain)
 - c. Compare the accuracy
8. Write a Python program to
 - a. Generate histogram & describe cumulative property, bins and range.
 - b. Generate multiple histograms
 9. Write a python program to visualize data using bar chart, line chart, pie chart and Scatter plot
 10. Write a python program to understand the data using Boxplot and describe the summary from boxplot.

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Handle missing data using python package
2. Identify data outliers
3. Implement python pandas package for data set handling, classification
4. Apply various python packages for data visualization
5. Generate histograms

**SENGAMALA THAYAAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester V-MBE-I: INTERNET OF THINGS

Ins.Hrs/Week: 5

Course Credit: 4

Course Code

UNIT -1:

Introduction: History of IoT - Importance of Internet of Things – IoT devices, connectivity and system – Technologies behind Internet of Things - IoT applications – benefits of IoT to organizations - pros and cons of IoT – IoT standards and frameworks – IoT security and privacy issues

UNIT - II:

IoT Markets and Trends: Introduction to Internet of Things - Surveying IoT Markets and Trends - Identifying the Wireless Standards That Enable the IoT-The goal: Interoperability - CHIP to the rescue - Ultra-wideband - 5G and Wi-Fi 6/6E

UNIT - III:

IoT Architecture and Technology Protocols: IoT Architecture and Technology Protocols - Digging into the IoT Ecosystem - Examining the IoT Architecture - IoT Protocols

UNIT - IV:

The Smart Home: Surveying Today's Smart Home - The Mesh Network/One Pod Per Room - Overcoming the Challenges – Interoperability - RF interference - Maintaining a seamless, energy-efficient, and reliable connection - Creating an ease-of-use self-management IoT network - Network security and privacy - Power consumption - Importance of Always-on Connectivity

UNIT - V:

Next-Generation IoT Opportunities and Use Cases: Next-Generation Smart Home - Manufacturing/Industry 4.0 – Transportation – Retail – Smart Energy -Healthcare - Smart Cities – Agriculture

Text book:

1. Vijay Madiseti and Arshdeep Bahga, –Internet of Things: (A Hands-On Approach), Universities Press (INDIA) Private Limited 2014, 1st Edition. (Unit1: Chapter1).
2. Gillis, Alexander, "What is internet of things (IoT)?" IOT Agenda, 2021 (Unit1).
3. Cees Link, "Internet of Things For Dummies", Qorvo 2nd Special Edition, John Wiley & Sons, Inc., 2021 (Unit2: Chapter1; Unit3: Chapter2; Unit4: Chapter3; Unit5: Chapter4).

Reference books:

1. Honbo Zhou, "The Internet of Things in the Cloud: A Middleware Perspective", CRC Press, 2012.
2. Dieter Uckelmann, Mark Harrison, Michahelles, Florian (Eds), "Architectingthe Internet of Things", Springer, 2011.
3. David Easley and Jon Kleinberg, "Networks, Crowds, and Markets: Reasoning About a highly Connected World", Cambridge University Press, 2010.
4. Olivier Hersent, Omar Elloumi and David Boswarthick, "The Internet of Things: Applications to the Smart Grid and Building", Wiley, 2012.
5. Jamil Y. Khan, Mehmet R. Yuce, Internet of Things (IoT): Systems and Applications. Singapore, Jenny Stanford Publishing, 2019.

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Explore the fundamentals of IoT.
2. Identify IoT markets and trends.
3. Describe IoT Architecture and protocols.
4. Understand the challenges in IoT.
5. Identify opportunities in IoT.

**SENGAMALA THAYAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester V-MBE-I - BLOCK CHAIN TECHNOLOGY

Ins.Hrs/Week: 5 Course Credit: 4 Course Code

UNIT – I Block chain Fundamentals:

Tracing Block chain's Origin - The shortcomings of current transaction systems - The emergence of Bit coin - The birth of block chain - Revolutionizing the Traditional Business Network. Block chains - The Structure of Block chains - Block chain Applications - The Block chain Life Cycle

UNIT – II Consensus:

The Driving Force of Block chains - Block chains in Use -Block chain applications - Picking a Block chain - Block chains Substance - determining needs - Defining goal - Choosing a Solution - Drawing a block chain decision tree Comparing Popular Block chain Implementations - Making digital currency work - Providing computation control -Scaling to meet enterprise needs

UNIT – III Diving into the Bit coin Block chain:

Creating Bit coin wallet - Generating a Bit coin vanity address - Transferring vanity address - Making an entry into the Bit coin block chain - Reading a block chain entry in Bit coin - Using Smart Contracts with Bit coin - Building your first smart bond -Checking the status of contract - Block chain For Dummies Building a Private Block chain with Docker - Preparing your computer - Building your block chain.

UNIT – IV Beholding the Bit coin Block chain:

Getting a Brief History of the Bit coin Block chain - Debunking Some Common Bit coin Misconceptions - Bit coin: The New Wild West - Fake sites - Mining for Bit coins - Making First Paper Wallet - The Ten Rules to Never Break on theBlock chain

UNIT – V Block chain in Action: Use Cases:

Determining How Block chain Fits in Industry - Identifying Speed Bumps in Business Processes- Determining How Block chain Can Help - Choosing an Appropriate Use Case - Determining the Goal of Your Block chain Network - Identifying Dependencies - Choosing a Block chain Provider and Platform - Developing and Deploying Smart Contracts - Testing and Fine-Tuning Your Application and Network

Text book:

1. Manav Gupta, "BlockchainFor Dummies", IBM Limited Edition, John Wiley & Sons, Inc., 2018. (Unit1: chapter1; Unit5: chapter 6)
2. Bashir, Imran, "Mastering Blockchain: Deeper insights into Decentralization, cryptography, Bit coin, and popular Block chain Frameworks", 2017.

Reference Books:

1. Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller, and Steven Goldfeder, "Bitcoin and cryptocurrency technologies: a comprehensive introduction" Princeton University Press, 2016.
2. Joseph Bonneau et al, "SoK: Research perspectives and challenges for Bitcoin and cryptocurrency", IEEE Symposium on security and Privacy, 2015.
3. Srinivas Mahankali, Blockchain for Non IT Professionals: An Example Driven, Metaphorical Approach, Notion Press, 2020.
4. Laurence, Tiana. Introduction to Blockchain Technology: The many faces of blockchain technology in the 21st century, Van Haren, 2019.

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Explore the fundamentals of block chain
2. Explain Consensus in block chain
3. Describe bit coins
4. Apply abstract models for Block chain Technology.
5. Identify major research challenges and technical gaps.

**SENGAMALA THAYAAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester V-SBE-I: SHELL PROGRAMMING

Ins.Hrs/Week: 2 Course Credit: 2 Course Code

UNIT – I INTRODUCTION:

Types of shells - Creating files – Listing files and directories – Masking file permissions – Directory permissions – access files – Storage of files – Disk related commands - Directory related commands – Miscellaneous commands.

UNIT – II ESSENTIAL COMMANDS:

Introduction – Password – cal is for calendar – The banner command – The touch command – printouts – File compression – I/O redirection and piping

UNIT – III PROCESSES:

Introduction – What`s running right now – Still more processes – Background processes – The nohup command – Killing a processes – Changing processes priorities – Scheduling of processes.

UNIT – IV SHELL PROGRAMMING INTRODUCTION:

When to use shell scripts – The first shell script – Interactive shell scripts – Shell variables – Shell keywords – Another way of assigning values to variables - The if-then-fi statement – The if-then-else-fi statement – Nested if-else –Use of logical operators – Loops –The break statement – The continue statement.

UNIT – V SYSTEM ADMINISTRATION:

Introduction – Adding and removing users – Starting up the system – Shutting the system down – Disk management – Formatting a disk – Making a file system – Mounting a file system – Unmounting a file system – Using a raw disk – Monitoring system usage – Ensuring system security – Providing assistance to users.

TEXT BOOK:

1. Yashavant P. Kanetkar, "Unix Shell Programming", BPB Publications, New Delhi, 2003. (Unit-1: Chapter:1,2 Unit-2: Chapter: 4,5 Unit-3: Chapter:7 Unit-4: Chapter: 9,10,11 Unit-5: Chapter 15)

REFERENCE BOOKS:

1. Wood, Patrick H., and Kochan, Stephen G. UNIX Shell Programming. United Kingdom, Sams, 2003.
2. Parker, Steve. Shell Scripting Tutorial. United States, Createspace Independent Pub, 2014.
3. Tansley, D. S. W., and Tansley, David V. Linux and UNIX Shell Programming, Addison-Wesley, 2000.
4. Newham, Cameron, and Rosenblatt, Bill. Learning the Bash Shell: UNIX Shell Programming, O'Reilly Media, 2005.
5. Wood, Patrick, and Kochan, Stephen G. Shell Programming in UNIX, Linux and OS X: The Fourth Edition of UNIX Shell Programming. United Kingdom, Pearson Education, 2016.

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Write shell scripts
2. Create and explore files
3. Develop shell scripts to solve problems
4. Implement some standard utilities such as ls
5. Add/remove users

**SENGAMALA THAYAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
(Accredited by NAAC | An ISO 9001:2015 Certified Institution)
SUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester V-CC- VIII RPROGRAMMING

Ins.Hrs/Week: 5 Course Credit: 5 Course Code

UNIT - I:

Introduction: History of R- Benefits of Using R - Working with code Editor: RGui and RStudio - Starting your First R Session - Sourcing a script - Navigating the workspace- Vectors: Creating vectors - Discovering the properties of vector- combining vectors repeating vector - Getting values in and out of vectors- working with logical vectors - Math with Vector functions - working with numbers, infinity and missing values

UNIT - II:

Using character vector for text data - Manipulating text - Factoring in Factors - Working with more dimensions: Adding a second dimension - Using the indices - Naming matrix rows and columns - Calculating with matrices - Adding more dimensions: Creating an array - Combining different types of values in Data Frame - Manipulating values in a Data Frame- List: Creating a list - Extracting elements from lists - Changing elements in lists

UNIT - III:

Control Structures: Conditional control structures: if statement - if..else statement - switch statement - Loops: for, while and repeat loops - break and next statement. Functions: The Function Keyword - Arguments - Return Values- Functions as Arguments - Anonymous Functions-Properties of Functions - Argument Order and Named Arguments. Computing basic statistics: mean, median, mode, correlation and covariance.

UNIT - IV:

Getting data into and out of R: Getting data into R: Entering data in the R text editor - Using clipboard to copy and paste - Reading data in CSV files and excel files- Working with other data types - Getting your data out of R - Working with Files and Folders. Packages: Finding packages, installing packages, loading packages, updating package and unloading packages.

UNIT - V:

Introduction to Graphical Analysis: Box-Whiskers plots - Scatter plots - Pairs plots - Line charts - Pie charts - Cleveland dot charts - Bar charts: single category bar charts and multiple category bar charts. Creating Faceted Graphics with Lattices: Creating lattice plot - changing plot option.

TEXT BOOK

1. Andrie Devries And Jorismeys , "R Programming For Dummies", Wiley Publications, Isbn:978-81-265-5201-6. (Unit 1 : chapter-2,3,4) (Unit 4: chapter 12,14) (Unit 2 chapter 5,6,7)
2. Dr.Mark Gardener, "Beginning R- The Statistical Programming Language", Wiley Publications, Isbn: 978-81-265-4120-1.
3. Paul Teetor, "R Cook Book",O'relly Publications, First Edition, 2011, Isbn: 978-0-596-80915-7 (Unit 5: Chapter 10)

REFERENCE BOOK

1. Joseph Adler, "R In Nutshell A Desktop Quick Reference",Isbn:978-0-596-80170-0
2. Tilman M .Davis "The Book of R", No Starch Press, 2016 (Unit 3: chapter 10,11,3)
3. Roger D. Peng, "R Programming For Data Science", Leanopub, 2015
4. Matloff, Norman, And Matloff, Norman S, The Art Of R Programming: A Tour Of Statistical Software Design, No Starch Press, 2011.
5. Lovelace, Robin, And Gillespie, Colin. Efficient R Programming: A Practical Guide to Smarter Programming, O'reilly Media, 2016.
6. Alex Nordeen, Learn R Programming In 24 Hours: Complete Guide ForBeginners, Guru99, 2020.
7. Jeeva Jose, Beginner's Guide For Data Analysis Using R Programming., Khanna Publishing House,2019

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Understand the basics of R programming
2. Work with vectors, matrices and data frames
3. Acquire the knowledge of various control structures
4. Parse data files using built-in functions
5. Apply the various statistical functions and produce high quality graphics

**SENGAMALA THAYAAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
Accredited by NAAC | An ISO 9001:2015 Certified Institution
MUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester VI-CC- IX- MACHINE LEARNING

Ins.Hrs/Week: 5 Course Credit: 5 Course Code

UNIT - I:

Introduction: Introduction, easy for human hard for machines, a simple predicting machine, classifying is not very different from predicting, training a simple classifier, one classifier is not enough, Types of machine learning, Applications of Machine Learning, Perspectives and issues in machine learning.

UNIT - II:

Probabilistic and Stochastic Models: Bayesian Learning – Bayes theorem, Concept learning, Maximum likelihood, Bayes optimal classifier, Gibbs algorithm, Naive Bayes classifier, Expectation maximization and Gaussian Mixture Models, Hidden Markov models

UNIT - III:

Supervised learning: Introduction, Regression, Linear regression, Classification: Decision trees, k-Nearest Neighbours, Support Vector Machine, Logistic regression, Random Forest. Artificial Neural Network: Introduction, Perceptrons, multi-layer networks and back propagation.

UNIT - IV:

Unsupervised learning: Introduction, Supervised vs Unsupervised Cluster Analysis, K-means clustering, Hierarchical clustering. Dimension reduction: Principal Component Analysis, Linear Discriminant Analysis

UNIT - V:

Modelling and evaluation: Building the model, Training a model, evaluating a model, improving a model. Performance metrics - accuracy, precision, recall, sensitivity, specificity, AUC, RoC, Bias Variance decomposition.

TEXT BOOK

1. SaikatDutt, Subramanian Chandramouli, Amit Kumar Das, –Machine Learning, Pearson Education, 2018(Unit 1: Chapter1:4,1. 5,1. 7,1. 9; Unit 2: Chapter6; Unit 3: Chapter7,8,10; Unit 4 Chapter-9 (9.1,9.4), Unit 5:Chapter3;)
2. Ethem Alpaydin,"Introduction to Machine Learning", MIT Press, Prentice Hall of India, Third Edition 2014 (Unit2: Chapter7, Chapter 15)
3. Tariq Rashid, "Make your own neural network", Create Space IndependentPublishing Platform, US 2016, ISBN:978-1-5308-2660-5 (Unit1, Part1)

REFERENCE BOOK

1. Shai Shalev-Shwartz, Shai Ben-David, Understanding Machine Learning:From Theory to Algorithms, Cambridge University Press, 2015
2. T. Hastie, R. Tibshirani and J. Friedman, "Elements of Statistical Learning",Springer, 2013
3. Charu C. Aggarwal, "DATA CLUSTERING Algorithms and Applications", CRCPress, 2014.
4. C. Bishop, "Pattern Recognition and Machine Learning", Springer,2006
5. Sebastian Raschka and Vahid Mirjalili, "Python Machine Learning", PacktPublishing, Third Edition, 2019

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Explain machine learning
2. Apply machine learning concepts in various domains
3. Implement supervised, unsupervised learning techniques
4. Differentiate supervised and unsupervised learning techniques
5. Create and evaluate models

**SENGAMALA THAYAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
Accredited by NAAC | An ISO 9001:2015 Certified Institution
UNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester VI- CP-VI-MACHINE LEARNING LAB

Ins.Hrs/Week: 5 Course Credit: 4 Course Code

1. Write R script for the implementation of
 - a. vector data objects operations
 - b. matrix, array and factors
 - c. data frames in R
2. Write R script to find basic descriptive statistics using summary, str, quartilefunction on mtcars & cars datasets.
3. Write R script to find Correlation coefficient rank correlation, partial and multiple correlation
4. Write R script to find subset of dataset by using subset (), aggregate () functions on iris dataset.
5. Perform the following using R : Reading and writing different types of datasets
 - a. Reading different types of data sets (.txt, .csv) from Web, disk and writing in file in specific disk location.
 - b. Reading Excel data sheet in R.
 - c. Reading XML dataset in R.
6. Data Manipulation with dplyr package
7. Predicting house prices using simple linear regression (use the Boston dataset, which contains data about the housing and price data in the Boston area)
8. Predicting house prices using multiple features – multiple linear regression (use the Boston dataset, which contains data about the housing and price data in the Boston area)
9. Perform Decision Tree Learning Methods in R
10. Perform Naïve Bayes Learning Methods in R

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Perform various statistical operations using R
2. Write R script to predict unknown values
3. Read various data formats in R
4. Find correlation using R scripts
5. Perform decision tree method

**SENGAMALA THAYAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
Credited by NAAC | An ISO 9001:2015 Certified Institution)
MUNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester VI- MBE II-Computing Technologies for Data Science

Ins.Hrs/Week: 5 Course Credit: 4 Course Code

UNIT – I Mobile Computing:

Introduction to Mobile Computing — Applications of Mobile Computing-
Spread spectrum -MAC Protocols — SDMA- TDMA- FDMA- CDMA -
GSM — Services & Architecture — Protocols — Connection Establishment
— Frequency Allocation — Routing – GPRS.

UNIT – II Edge Computing Definition and Use Cases:

Introduction to Edge Computing Scenario's and Use cases - Edge computing
purpose and definition, Edge computing use cases, Edge computing hardware
architectures, Edge platforms, Edge vs Fog Computing, Communication
Models - Edge, Fog and M2M.

UNIT – III Quantum Computing:

Introduction to quantum computing – quantum computing need - – Quantum
Measurements Density Matrices - Positive-Operator Valued Measure -
Fragility of quantum information: Decoherence - Quantum Superposition and
Entanglement - Quantum Gates and Circuits - Deutsch's algorithm

UNIT – IV Robotic Process Automation:

RPA Basics: History of Automation, RPA vs Automation, Processes &
Flowcharts, Programming Constructs in RPA, processes Types of Bots,
Workloads which can be automated - Excel Data Tables: Data Tables in RPA,
Excel and Data Table basics, Data Manipulation in excel, Email Automation:
Email Automation, Incoming Email automation, Sending Email automation

UNIT – V Distributed Ledger Technology:

Introduction – distributed ledgers in a nutshell - distributed identity –
decentralized network – distributed ledger – network consensus – future of DLT

TEXT BOOKS:

1. Prasant Kumar Pattnaik, Rajib Mall, "Fundamentals of Mobile Computing", PHI Learning Pvt. Ltd, New Delhi – 2012.(Unit1)
2. IoT and Edge Computing for Architects - Second Edition, by Perry Lea, Publisher: Packt Publishing, 2020, ISBN: 9781839214806.(Unit2)
3. M. A. Nielsen and I. L. Chuang. Quantum computation and quantum information. Cambridge, 2010 (Unit3)
4. Alok Mani Tripathi, Learning Robotic Process Automation, Publisher: Packt Publishing Release, 2018, ISBN: 9781788470940.(Unit4)
5. Alexander Lipton and Adrien Treccani, "Blockchain and Distributed Ledgers Mathematics, Technology, and Economics, World Scientific, 2021.(Unit5)

REFERENCE BOOKS

1. Dharma Prakash Agarwal, Qing and An Zeng, "Introduction to Wireless and Mobile systems", Thomson Asia Pvt Ltd, 2005.
2. Uwe Hansmann, Lothar Merk, Martin S. Nicklons and Thomas Stober, "Principles of Mobile Computing", Springer, 2003.
3. Fog and Edge Computing: Principles and Paradigms by Rajkumar Buyya, Satish Narayana Srirama, Wiley publication, 2019, ISBN: 9781119524984.
4. Swan, Melanie. Blockchain: Blueprint for a New Economy. United States, O'Reilly Media, 2015.
5. Jaime Lloret Mauri, M. Bala Krishna, Advances in Mobile Computing and Communications: Perspectives and Emerging Trends in 5G Networks Taylor & Francis, 2016.

COURSE OUTCOME:

Upon successful completion of this course the students would be able to:

- Explain the basics of mobile telecommunication systems
- Explore edge computing
- Describe quantum computing
- Apply automation concepts
- Illustrate DLT

**SENGAMALA THAYAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
Credited by NAAC | An ISO 9001:2015 Certified Institution
UNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester VI- MBE II

AUGMENTED REALITY AND VIRTUAL REALITY

Ins.Hrs/Week: 5

Course Credit: 4

Course Code

UNIT – I Introduction:

Definition virtual and augmented reality - Introducing virtual reality and augmented reality – Other types of virtual and augmented reality –history – the father of virtual reality – Early virtual reality breaks through – Augmented reality hits the mainstream – Evaluating the technology hype cycle.

UNIT – II State of virtual reality & augmented reality:

Available forms factors – Focusing on features – Considering controllers – Recognizing the current issues with VR – Assessing adoption rates- Recognizing the current issues

UNIT – III Consuming content in virtual reality & augmented reality:

Exploring consumers-Grade virtual reality – identifying near-Future hardware – Comparing current and future options.

UNIT – IV Choosing virtual reality & augmented reality:

Choosing virtual reality, augmented reality – Planning virtual, augmented reality project - Exploring design principles in

virtual reality – Definition your social experience - future of virtual and augmented reality - exploring virtual reality & augmented reality use cases.

UNIT – V Industries, mobile app by virtual and augmented reality:

Travel – Museums – Aerospace – Retail – Military – Education
– Entertainment Real estate – Advertising and marketing - Google translate – Amazon AR view –AR city – Ingress and pokemon go
– Sketch AR – Find your car & car finder AR

TEXT BOOKS:

1. Paul Mealy, “Virtual & Augmented Reality For Dummies” Published by: JohnWiley & Sons, Inc., John Wiley & Sons, Inc., 2018. (Unit 1: chapter 1) (Unit2:chapter 2,3) (Unit 3: chapter 4,5) (Unit 4:chapter 6,7,8,10,11,12,13) (Unit 5: chapter 15,16)

REFERENCE BOOKS

1. Aguilera, P, “Digital info on the real world”. MIT Technology Review, 2009
2. Arnall, T “The web in the world fabric rblg”, 2008
3. Aron, J. “AR goggles make crime scene investigation a desk job. NewScientist.”, 2012
4. “Virtual Reality and Augmented Reality”, 17th EuroVR InternationalConference Proceedings, 2020.
5. Bruno Arnaldi, Guillaume Moreau, Pascal Guitton, Virtual Reality andAugmented Reality: Myths and Realities, Wiley, 2018.

COURSE OUTCOMES:

Upon successful completion of this course the students would be able to:

1. Explain virtual and augmented reality
2. Identify issues and controllers of VR & AR
3. Understand VR & AR use cases
4. Apply AR & VR in various industries
5. Plan and choose VR & AR projects

**SENGAMALA THAYAAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
Accredited by NAAC | An ISO 9001:2015 Certified Institution)

UNDARAKKOTTAI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester VI- PROJECT

Ins.Hrs/Week: 5

Course Credit: 3

Course Code

The candidate shall be required to take up a Project Work by group or individual and submit it at the end of the final year. The Head of the Department shall assign the Guide who, in turn, will suggest the Project Work to the students in the beginning of the final year. A copy of the Project Report will be submitted to the University through the Head of the Department on or before the date fixed by the University.

The Project will be evaluated by an internal and an external examiner nominated by the University. The candidate concerned will have to defend his/her Project through a Viva-voce.

Assessment/Evaluation/Viva Voce:

- I. PROJECT REPORT EVALUATION (Both Internal & External)
 - II. Plan of the Project - 20 marks
 - III. Execution of the Plan/collection of - 45 marks
 - IV. Data / Organisation of Materials / Hypothesis, Testing etc. and presentation of the report.
 - V. Individual initiative - 15 marks
 2. Viva-Voce / Internal & External - 20 marks
- TOTAL - 100 Marks

PASSING MINIMUM:

Project	Vivo-Voce 20 Marks 40% out of 20 Marks (i.e. 8 Marks)	Dissertation 80 Marks 40% out of 80 marks (i.e. 32 marks)
---------	--	--

A candidate who gets less than 40% in the Project must resubmit the Project Report. Such candidates need to defend the resubmitted Project at the Viva-voce within a month. A maximum of 2 chances will be given to the candidate.*****

**SENGAMALA THAYAR EDUCATIONAL TRUST WOMEN'S COLLEGE
(AUTONOMOUS)**



(Affiliated to Bharathidasan University)
Accredited by NAAC | An ISO 9001:2015 Certified Institution)
MANNARGUDI, MANNARGUDI – 614 016,
TAMILNADU, INDIA.

**Department of Computer Applications
B.Sc., Data Science**

Semester VI- SBE II-WEB TECHNOLOGY

Ins.Hrs/Week: 2

Course Credit: 2

Course Code

UNIT - I:

HTML: BasicHTML, TheDocument body, Text, Hyperlinks, Adding more formatting, Lists, Tables, Using colors and images, Images, Multimedia objects, Frames, Forms.

UNIT - II:

Towards interactivity Cascading Style Sheets: Introduction, Using styles: Simple examples, Defining your own styles, Properties and values in styles.Client Side Scripting: JavaScript: JavaScript—The basics, Variables, String manipulation, Mathematical functions, Statements, Operators, Arrays, Functions- Dataand objects in java script, Regularexpressions, Exception Handling, Built in objects,Events.

UNIT - III:

Dynamic HTML with JavaScript: Data validation, Opening a new window, Messages and Confirmations, The status bar, writing to a different frame, Rolloverbuttons, Moving images, multiple pages in a single download, A text-only menu system, Floating logos.

UNIT - IV:

Server Side Scripting: PHP: evolution of PHP– structure and syntax of PHP and integrating the same with HTML – comments – variables – data types – operators – control structures–passing information between pages–Strings–Arrays and Functions.

UNIT - V:

Integration of Apache, MySQL, PHP to design dynamic webpages: MySQL functions in PHP – Connecting and disconnecting from MySQL – Using tables – form design– editing the database.

TEXT BOOK

1. Chris Bates, **“Web Programming Building Internet Applications”**, Third Edition, Wiley, 2007, ISBN-10:0470017759. (Unit1: Chapters: 2,3,4; Unit 2: chapters: 6,7; Unit 3: 8, Unit 4: Chapter 12,13)
2. Timothy Boronczyk, Michael, Elizabeth Naramore, Jason Gerner, Yann Le Scouar-nec, Jeremy Stolz,, Michael K. Glass **“Beginning PHP6, Apache, MySQL Web Development”**, Wiley Publishing, 2009 Edition. ISBN-13: 978- 8126521227. (Unit 5: Chapter 3,4, 5)

REFERENCE BOOK

1. Robin Nixon, **“Learning PHP, MySQL & JavaScript With jQuery, CSS & HTML5”** O’Reilly Media, Fourth edition, December2014, ISBN:978-1-491- 91866-1.
2. David R. Brooks, **“An Introduction to HTML and JavaScript for Scientists and Engineers”**, Springer- Verlag London Limited 2007.
3. Michael K Glass, Yann Le Scouarnec, Elizabeth Naramore, Gary Mailer, Jeremy Stolz, Jason Gerner, **“Beginning PHP, Apache, MySQL Web Development”**, Wiley dream tech press, 2004 edition. ISBN: 9780764557446.
4. Rochkind, Marc. **Expert PHP and MySQL: Application Design and Development**. Netherlands, Apress, 2013.
5. Lockhart, Josh. **Modern PHP: New Features and Good Practices**, O’Reilly Media, 2015.

COURSE OUTCOME:

Upon successful completion of this course the students would be able to:

1. Design a static web page using HTML
2. Apply HTML and style sheet

3. Develop server side scripts using PHP
4. Communicate with MySQL database from PHP
5. Validate the HTML form data using JavaScript
