GONDWANA UNIVERSITY, GADCHIROLI

DIRECTION NO. 05 of 2025

"Admission of Students and Conduct of Examinations Leading to the Award of Three Years' Degree Program entitled "B. Sc. Data Science" under Credit System and as per NEP 2020 in the Faculty of Science &Technology, Direction 2025."

Whereas, The Maharashtra Public Universities Act, 2016 (Maha. Act No. VI of 2017) (hereinafter the "Act") governs the Gondwana University, Gadchiroli (hereinafter the "University");

AND

Whereas, the National Education Policy (NEP) 2020 focuses on education and skill development as per the needs of the community and as per Maharashtra State Government Resolution of Higher and Technical Education Dipartment No. NEP-2020/Pr.kr.09/UE-3/SHIKANA, dated 20 April 2023, therefore, the University has introduced Three Years' Degree Programme entitled **"B. Sc. Data Science"** under credit system and as per NEP-2020 in the Faculty of Science and Technology and is being offered by the Model Degree College, a constituent College of the University;

AND

Whereas, as per provisions of section 73(1) and 73(5) of the Act, an ordinance is required to be framed to lay down the conditions under which students can be admitted to courses of study for award of a Certificate, Diploma, Degree and Academic distinctions and conduct of examination, other tests and evaluations and the manner of assessment, but since, making of an ordinance was a time-consuming process and there was an urgency for introduction of Three Years' Degree Programme entitled **"B. Sc. Data Science"** under credit system and as per NEP 2020 in the Faculty of Science &Technology, therefore Vice-Chancellor of the University had issued direction no. 30 of 2023 entitled **"Admission of Students and Conduct of Examinations Leading to the Award of Three Years' Degree Program entitled "B. Sc. Data Science" under Credit System and as per NEP 2020 in the Faculty of Science &Technology, Direction 2023",** which has lapsed by virtue of the proviso to Section 12(8) of the Act. Yet there is no ordinance regarding this matter and there is urgency to take examinations of the said programme.

Now, therefore, I, Dr. Prashant Shridhar Bokare, Vice-Chancellor of the University, in exercise of powers vested in me under section 12(8) of the Act, do hereby issue the following directions:

1. This Direction shall be called **"Admission of Students and Conduct of Examinations Leading to the Award of Three Years' Degree Program entitled "B. Sc. Data Science" under Credit System and as per NEP 2020 in the Faculty of Science &Technology, Direction 2025".**
2. This direction shall come into force from the date of its issuance.
3. Definitions: -In this Direction, unless the context requires otherwise, the words and phrases shall have the meaning given hereunder.
   a) "Program" means the full-time Three Years' Degree Programme entitled **"B. Sc. Data Science"**

b) "Application Form" means a form prescribed by the University for seeking admission to Program under this direction.
c) "Competent Authority" means the Authority appointed by the Vice-Chancellor, for any specific purpose of the Program under this Direction.
d) "Course" means a subject or a paper offered in any semester under this Program.
e) "Credit" refers to the weightage given to a course, in terms of the number of instructional hours assigned to it. In this direction one credit means 15 hours of teaching work or 30 hours of practical work in a semester.
f) "Credit System" means, the system in which weightage of credits is spread over to different semesters during the period of study.
g) 'Grade letter' is an index to indicate the performance of a student in particular course. It is the depiction of actual marks secured by a student by a letter, the Grade letters are as given in **Table 3.**
h) 'Grade point' is the weightage allotted to each grade letter depending on the range of marks awarded in a course.
i) "HSSC" means the Higher Secondary School Certificate (Standard XII) examination conducted by Maharashtra State Board of Secondary and Higher Secondary Education or its equivalent certificate awarded by any other recognized Board.
j) "Qualifying Examination" means an examination on the basis of which a candidate becomes eligible for admission to this Program.

4. In order to conduct the admission process for admitting students to this Program, the Vice Chancellor shall appoint the "Competent Authority".

5. Intake capacity, Eligibility for application, Admission fees, Curriculum, Examination fees for this Program will be as shown in Table 1 below:

Table: 1

Sr. No.	Subtitle	Details
1	Intake Capacity	20
2	Eligibility for application	HSSC Science
3	Admission Fees	Admission fees shall be as prescribed/ revised and notified from time to time by the University.
4	Curriculum	As specified in **Annexure-I**
5	Examination Fees	The Examination fees shall be as prescribed/ revised and notified from time to time by the University.

6. Objectives of the Program: -

(a) To extract knowledge from large volumes of data which is structured or unstructured and is continuation of data mining and predictive analytics.
(b) To instill object oriented programming concepts.
(c) To foster problem-solving skills using data structure and enable data analysis and visualization techniques for effective information communication and computational tasks.

7. EVALUATION OF THE PROGRAM

The internal and university assessment of student performance shall be carrying the weightage as mentioned in the Teaching and Examination Scheme given below:

Table 2: Teaching and Examination Scheme
"B. Sc. Data Science"
Teaching & Examination Scheme
Semester - I

Sr. No.	Subject	Credit	Teaching scheme Hrs/week			Examination Scheme Total Marks The./Pra.									
						Theory					Practical				
			The oy	Pra.	Total	Duration Hrs.	Max. Theory		Total	Min Pass	Dur. Hrs.	Max Marks		Min.Pass Marks	Total Marks
							UA	CA				UA	CA		
1	Data Structure and Algorithom using python	4	4	-	4	3	80	20	100	40	-	-	-	-	100
2	DiscreteMathematics	2	2	-	2	2	40	10	50	20	-	-	-	-	50
3	Introduction to Programming	2	2	-	2	2	40	10	50	20	-	-	-	-	50
4	Web Technology	2	2	-	2	2	40	10	50	20	-	-	-	-	50
5	Business Communication and Information Ethics	2	2	-	2	2	40	10	50	20	-	-	-	-	50
6	Environment Protection Law	2	2	-	2	-	-	50	50	20	-	-	-	-	50
	Precalculus	2	2	-	2	-	-	50	50	20	-	-	-	-	50
7	Web Technology Practical	2	-	4	4	-	-	-	-	-	2	30	20	25	50
8	ICT Practical	2	-	4	4	-	-	-	-	-	2	30	20	25	50
Total		**20**	**16**	**8**	**24**	-	**240**	**160**	**400**	-	-	**60**	**40**	-	**500**

Semester - II

Sr. No.	Subject	Credit	Teaching scheme Hrs/week			Examination Scheme Total Marks The./Pra.									
						Theory					Practical				
			Theoy	Pra.	Total	Duration Hrs.	Max. Theory		Total	Min Pass	Dur. Hrs.	Max Marks		Min.Pass Marks	Total Marks
							UA	CA				UA	CA		
1	Data base Management System	4	4	-	4	3	80	20	100	40	-	-	-	-	100
2	R Programming	2	2	-	2	2	40	10	50	20	-	-	-	-	50
3	Data Analysis	2	2	-	2	2	40	10	50	20	-	-	-	-	50
4	Internet of Things embedded system	2	2	-	2	2	40	10	50	20	-	-	-	-	50
5	Morden Indian Language (Marathi/Hindi)	2	2	-	2	2	40	10	50	20	-	-	-	-	50
6	Project Presentation on Data Science in Environmental Science.	2	2	-	2	2	-	50	50	20	-	-	-	-	50
7	Data Base Management	2	-	4	4	-	-	-	-	-	2	30	20	25	50
8	R Programming	2	-	4	4	-	-	-	-	-	2	30	20	25	50
9	Data Base Management System	2	-	4	4	-	-	-	-	-	2	30	20	25	50
Total		**20**	**14**	**12**	**26**	-	**240**	**110**	**350**	-	-	**90**	**60**	-	**500**

Note:

1. Th=Theory; Pr = Practical; PR=Project; INT=Internship IA = Internal Assessment UA = University Assessment.

Credit Calculations: (1) One credit would mean equivalent of 15 periods of 60 minutes each for Theory.(2) For practical /project /internship/Field work, the Credit Weightage for equivalent hours shall be 50% of that for theory. (3) The strength of Batch of Practical /Workshop / internship / Field visit / Project shall be 20. (4) 10 contact hours equals to 3 credits per semester and 6 credits for two semesters viz one year duration for Project/Field Visit/Industrial Training/Internship

*On Job Training

a. Marks to Letter Grade & Grade Point Conversion

The marks scored by the examinees in their courses/heads of passing of the program shall be converted into Letter Grade and Grade Point as per Table given below:

Table 3: Conversion of marks into letter grade and grade points

LetterGrade	GradePoint
O(out standing)	10
A+(Excellent)	9
A(Verygood)	8
B+(Good)	7
B(Aboveaverage)	6
C(Average)	5
P(Pass)	4
F(Fail)	0
Ab(Absent)	0

*Note: As such, the lowest passing Grade in any passing head shall be 'P'.

a) Calculation of Grade Point Average (GPA)

The Grade Point Average (GPA) shall be calculated for the program and shall be evaluated as mentioned below:

$$\text{GPA} = \frac{\sum_{i=1}^{n}(C_i \times G_i)}{\sum_{i=1}^{n} C_i}$$

Where C_i is the number of credits of the i^{th} course and G_i is the grade point scored by the student in the i^{th} course.

The percentage of marks scored based on obtained GPA can be evaluated using below given formula.

$$Percentage = (GPA - 0.75) * 10$$

8. Division of Passing

The Division of Passing shall be based on GPA secured by an Examinee as shown in the Table 3 below:

Table 4: Interpretation of GPA into Division of Passing

INTERVAL OF GPA	DIVISION OF PASSING
$GPA \geq 8.25$	First with Distinction
$6.75 \leq GPA < 8.25$	First
$6.00 \leq GPA < 6.75$	Second
$5.00 \leq GPA < 6.00$	Pass

9. Declaration of result is based on the Grade Point Average (GPA) earned towards the end of the program as given in Table 4. The names of the successful examinees passing the examination as a whole in the minimum prescribed period and obtaining prescribed number of places securing the grades as per adopted credit-grade system shall be arranged in order of merit as provided in ordinance relating to examinations in general.

10. Provisions with respect to grace marks for passing in a particular course/ head of passing and improvement of Division (Higher Class) and getting Distinction in any course shall be as per relevant Direction/Ordinance of the University.
11. An examinee who does not qualify in examination or remain absent for the examination, shall be eligible to appear in the same re-examination, on payment of re-examination fee and such other fees as may be prescribed from time to time, within 30 days from the date of result.
12. Successful examinees who secure minimum prescribed registered credits (120) for the program durationshall be entitled to receive aDegreeof full time Three Years' Degree Programme entitled **"B. Sc. Data Science"** in the Faculty of Science & Technology signed by the Vice Chancellor of the University on payment of prescribed fees.
13. In the event of any query regarding interpretation/application of any provision of this direction, the Director of Board of Examinations and Evaluation shall refer the matter for the decision of the Dean of the Faculty of Science & Technologyor alternatively to the Board of Deans if found necessary.
14. For any other matter pertaining to this Program and its final award which is beyond the purview of this Direction, it shall be referred to the Vice-Chancellor and that the decision of the Vice-Chancellor shall be final and binding on all the concerned.

Place: Gadchiroli
Date: 27/01 /2025

(Dr. Prashant S. Bokare)
Vice-Chancellor

Gondwana University, Gadchiroli
Academic Session 2024-2025
BSc. Data Science Semester- I

Sr. No	Subject	Credit	Teaching Scheme			Theory					Practical					Total Marks
			Th	Pr.	Total	Duration Hrs	Max Marks UA	CA	Total	Min Pass marks	Duration Hrs	Max Marks UA	CA	Total	Min Pass marks	
1	Data structure and Algorithms using Python	4	4		4	3	80	20	100	40	3			0		100
2	Discrete Mathematics	2	4		4	2	40	10	50	20	2			0		50
3	Introduction to Programming	2	4		4	2	40	10	50	20	2			0		50
4	Web Technology	2	4		4	2	40	10	50	20	2			0		50
5	Business Communication & Information Ethics	2	4		4	2	40	10	50	20	2			0		50
6	Enviroment Protection Law	2	2		2			50	50	20				0		50
7	Precalculus	2	2		2			50	50	20				0		50
8	Lab course in Web Technology	2		4	4							30	20	50	25	50
9	ICT Practicals	2		4	4				0			30	20	50	25	50
	Total	**20**	**24**	**8**	**32**		**240**	**160**	**400**		**11**	**60**	**40**	**100**	**50**	**500**

Gondwana University, Gadchiroli
Academic Session 2024-2025
BSc. Data Science Semester- II

Sr. No	Subject	Credit	Teaching Scheme			Theory					Practical					Total Marks
			Th	Pr.	Total	Duration Hrs	Max Marks UA	CA	Total	Min Pass marks	Duration Hrs	Max Marks UA	CA	Total	Min Pass marks	
1	Database Management Systems	4	4		4	3	80	20	100	40	3			0		100
2	R Programming	2	4		4	2	40	10	50	20	2			0		50
3	Data Analysis	2	4		4	2	40	10	50	20	2			0		50
4	Internet Of things	2	4		4	2	40	10	50	20	2			0		50
5	Modern language- Marathi	2	4		4	2	40	10	50	20	2			0		50
6	Enviromental Science-(Seminar + Project)	2	2		2			50	50	20				0		50
7	Lab course on Data Analysis through R	2		2	2							30	20	50	25	50
8	Lab Course on DBMS	2		4	4							30	20	50	25	50
9	Lab course in R Programming	2		4	4				0			30	20	50	25	50
	Total	**20**	**22**	**10**	**32**		**240**	**110**	**350**		**11**	**90**	**60**			**500**

Note:- 1. Th=Theory ; Pr=Practical ; UA-University Assessment; CA- College Assessment;
Credit Calculations: (1) One credit would mean equivalent of 15 periods of 60 minutes each of theory. (2) For Practical the credit weightage for equivalent hours shall be 50% that of theory.

Gondwana University, Gadchiroli
Academic Session 2024-2025
BSc. Data Science Semester- III

Sr. No	Subject	Credit	Teaching Scheme			Theory					Practical					Total Marks
			Th	Pr.	Total	Duration Hrs	Max Marks UA	Max Marks CA	Total	Min Pass marks	Duration Hrs	Max Marks UA	Max Marks CA	Total	Min Pass marks	
1	Web Application Development	4	4		4	3	80	20	100	40	3			0		**100**
2	Probability and Statistics	2	4		4	2	40	10	50	20	2			0		**50**
3	Fundamentals in Data Science	2	4		4	2	40	10	50	20	2			0		**50**
4	Java Programming	2	4		4	2	40	10	50	20	2			0		**50**
5	Operating System and Information Security	2	4		4	2	40	10	50	20	2			0		**50**
6	Seminar- 1	2	2		2			50	50	20				0		**50**
7	Optimization Techniques	2	2		2			50	50	20				0		**50**
8	Lab course -Web Application Development	2		4	4							30	20	50	25	**50**
9	Lab course - Java Programming	2		4	4				0			30	20	50	25	**50**
	Total	**20**	**24**	**8**	**32**		**240**	**160**	**400**		**11**	**60**	**40**	**100**	**50**	**500**

Gondwana University, Gadchiroli
Academic Session 2024-2025
BSc. Data Science Semester- IV

Sr. No	Subject	Credit	Teaching Scheme			Theory					Practical					Total Marks
			Th	Pr.	Total	Duration Hrs	Max Marks UA	Max Marks CA	Total	Min Pass marks	Duration Hrs	Max Marks UA	Max Marks CA	Total	Min Pass marks	
1	Machine Learning	4	4		4	3	80	20	100	40	3			0		**100**
2	Data Warehouse and Data Mining	2	4		4	2	40	10	50	20	2			0		**50**
3	Internet Programming with Big data	2	4		4	2	40	10	50	20	2			0		**50**
4	Deep Learning	2	4		4	2	40	10	50	20	2			0		**50**
5	Data Analytics and Visualization	2	4		4	2	40	10	50	20	2			0		**50**
6	Mini Project based on data science	2	2		2			50	50	20				0		**50**
7	Lab course -Data Warehouse and Data Mining	2		2	2							30	20	50	25	**50**
8	Lab course -Machine Learning	2		4	4							30	20	50	25	**50**
9	Lab course- Deep Learning	2		4	4				0			30	20	50		**50**
	Total	**20**	**22**	**10**	**32**		**240**	**110**	**350**		**11**	**90**	**60**			**500**

Note:- 1. Th=Theory ; Pr=Practical ; UA-University Assessment; CA- College Assessment;
Credit Calculations: (1) One credit would mean equivalent of 15 periods of 60 minutes each of theory. (2) For Practical the credit weightage for equivalent hours shall be 50% that of theory.

Gondwana University, Gadchiroli

NEP 2020 U.G. PROGRAMME SESSION 2024-25

Faculty of Science and Technology

B.Sc. (Data Science)
Semester – I & Semester -II
Syllabus

Note: For Details about assessment, practical examination, Paper Pattern, kindly refers Appendix

B.Sc. - I (Data Science)
Semester – I
Syllabus

B.Sc. I (Data Science)
For Semester I
Paper I- Data Structure and Algorithms using Python

Total Marks - 100		Max. Time- 3 Hrs
End Sem Exam- 80 Marks Min. Passing Marks- 40	Internal Assessment- 20 Marks Min. Passing- 50 %	Credits- 4 Lectures- 60

COURSE OBJECTIVES:

This course will enable students to

1. Implement Object Oriented Programming concepts in Python.
2. Understand Lists, Dictionaries and Regular expressions in Python.
3. Understanding how searching and sorting is performed in Python.
4. Understanding how linear and non-linear data structures work.
5. To learn the fundamentals of writing Python scripts.

UNIT – I OOPs Concepts-

class, object, constructors, types of variables, types of methods. Inheritance: single, multiple, multi-level, hierarchical, hybrid, Polymorphism: with functions and objects, with class methods, with inheritance,Abstraction: abstract classes.

UNIT – II Data Structures –

Definition,Linear Data Structures,Non-Linear Data Structures Python Specific Data Structures: List,Tuples, Set, Dictionaries, Comprehensions and its Types,Strings,slicing.

UNIT -III Arrays -

Overview, Types of Arrays, Operations on Arrays, Arrays vs List. Searching -Linear Search and Binary Search. Sorting - Bubble Sort, Selection Sort, Insertion Sort, Merge Sort, Quick Sort.

UNIT -IV Linked Lists –

Implementation of Singly Linked Lists, Doubly Linked Lists, Circular Linked Lists. Stacks - Overview of Stack, Implementation of Stack (List & Linked list), Applications of Stack Queues:Overview of Queue, Implementation of Queue(List & Linked list), Applications of Queues, Priority Queues.

UNIT -V Graphs -Introduction, Directed vs Undirected Graphs, Weighted vs Unweighted Graphs, Representations, Breadth First Search, Depth First Search. Trees - Overview of Trees, Tree Terminology, Binary Trees: Introduction, Implementation, Applications. Tree Traversals, Binary Search Trees: Introduction, Implementation, AVL Trees: Introduction, Rotations, Implementation.

TEXTBOOKS:

1. Data structures and algorithms in python by Michael T. Goodrich
2. Data Structures and Algorithmic Thinking with Python by Narasimha Karumanchi

B.Sc. I (Data Science)
For Semester I
Paper II - Discrete Mathematics

Total Marks - 50		Max. Time- 2 Hrs
End Sem Exam- 40 Marks Min. Passing Marks- 20	Internal Assessment- 10 Marks Min. Passing- 50 %	Credits- 2 Lectures- 60

Course Objectives:

1) To provide an overview of the theory of discrete objects, starting with relations and partially ordered sets.
2) Study about recurrence relations, generating functions and operations on them.
3) Give an understanding of graphs and trees, which are widely used in software.
4) Provide basic knowledge about models of automata theory and the corresponding formal languages.

UNIT – I:

Recurrence Relations

(a) Functions: Definition of function. Domain, co domain and the range of a function. Direct and inverse images. Injective, surjective and bijective functions. Composite and inverse functions. (b) Relations: Definition and examples. Properties of relations , Partial Ordering sets, Linear Ordering Hasse Diagrams, Maximum and Minimum elements, Lattices

UNIT – II:

Recurrence Relations: Definition of recurrence relations, Formulating recurrence relations, solving recurrence relations- Backtracking method, Linear homogeneous recurrence relations with constant coefficients. Solving linear homogeneous recurrence relations with constant coefficients of degree two when characteristic equation has distinct roots and only one root, Particular solutions of non linear homogeneous recurrence relation, Solution of recurrence relation by the method of generation functions, Applications- Formulate and solve recurrence relation for Fibonacci numbers, Tower of Hanoi, Intersection of lines in a plane, Sorting Algorithms

UNIT – III:

Counting Principles, Languages and Finite State Machine (a) Permutations and Combinations: Partition and Distribution of objects, Permutation with distinct and indistinct objects, Binomial numbers, Combination with identities : Pascal Identity, Vandermonde's Identity, Pascal triangle,

Binomial theorem, Combination with indistinct objects. (b) Counting Principles: Sum and Product Rules, Two-way counting, Tree diagram for solving counting problems, Pigeonhole Principle (without II proof); Simple examples, Inclusion Exclusion Principle (Sieve formula) (Without proof). (c) Languages, Grammars and Machines: Languages, regular Expression and Regular languages, Finite state Automata, grammars, Finite state machines, Gödel numbers, Turing machines.

Unit IV:-

Graphs and Trees

(a) Graphs : Definition and elementary results, Adjacency matrix, path matrix, Representing relations using diagraphs, Warshall's algorithm - shortest path, Linked representation of a graph, Operations on graph with algorithms - searching in a graph; Insertion in a graph, Deleting from a graph, Traversing a graph- Breadth-First search and Depth-First search.

(b) Trees: Definition and elementary results. Ordered rooted tree, Binary trees, Complete and extended binary trees, representing binary trees in memory, traversing binary trees, binary search tree, Algorithms for searching and inserting in binary search trees, Algorithms for deleting in a binary search tree

Books and References:

1. Discrete Mathematics and Its Applications, Seventh Edition by Kenneth H. Rosen, McGraw Hill Education (India) Private Limited. (2011)
2. Norman L. Biggs, Discrete Mathematics, Revised Edition, Clarendon Press, Oxford 1989.
3. Data Structure Seymor Lipschutz, Schaum's out lines, McGraw-Hill Inc.
4. Elements of Discrete Mathematics: C.L. Liu , Tata McGraw- Hill Edition .
5. Concrete Mathematics (Foundation for Computer Science): Graham, Knuth, Patashnik Second Edition, Pearson Education.
6. Discrete Mathematics : Semyour Lipschutz, Marc Lipson, Schaum's out lines, McGraw - Hill Inc.
7. Foundations in Discrete Mathematics: K.D. Joshi, New Age Publication, New Delhi.

B.Sc. I (Data Science)
For Semester I
Paper III - Introduction to Programming

Total Marks - 50		Max. Time- 2 Hrs
End Sem Exam- 40 Marks Min. Passing Marks- 20	Internal Assessment- 10 Marks Min. Passing- 50 %	Credits- 2 Lectures- 60

Course Objectives:

- Learn Programming fundamentals using Python
- Understand the concepts and usage data types,variables and other basic elements
- Learn about using operators and control statements in Python
- Learn about using arrays and strings in Python.
- Learn about using IPython architecture for Python.
- Introduce data Science Tools and plot data using appropriate Python visualization libraries

Unit I -

Introduction to Python Language: Overview, Features of Python, Execution of a Python Program, Innards of Python, Frozen Binaries, Python Interpreter, Comparison of Python with C and Java, InstallingPython,Writing & Executing, IDLE

Data Types, Variables And Other Basic Elements: Comments, Docstrings,Data types-Numeric, Compound, Boolean, Dictionary, Sets, Mapping, Basic Elements of Python, Variables

Input and Output Operations: Input Function, Output Statements,CommandLineArguments

Control Statements: Control Statements- Loop Statement, The elseSuite, break Statement, continue Statement, pass Statement, assertStatement,returnStatement

Unit II -

Functions: Defining & Calling a Function, Returning Results, Returning Multiple Values, Built in Functions, Parameters and Arguments, Recursive Functions, Anonymous or Lambda Functions

Operators: Arithmetic operators, Assignment operators, Unary minus operator, Relational operators, Logical operators, Bitwise operators, Membership operators, Identity operators, Precedence of Operators, Associativity of Operators

Arrays: Creating Arrays, Indexing and Slicing, Basic Array Operations, Arrays Processing, Mathematical Operations on Array, Aliasing Arrays, Slicing and Indexing in NumPy Arrays, Basic Slicing. Advanced Indexing. Dimensions of Arrays, Attributes of an Array

Strings: Creating Strings, Functions of Strings, Working with Strings, Length of a String, Indexing & Slicing, Repeating & Concatenation of Strings, Checking Membership, Comparing Strings, Removing Spaces, Finding Substrings, Counting Substrings, Strings are Immutable, Splitting and Joining Strings, Changing Case, Checking Starting and Ending of a String, Sorting & Searching in the Strings, Formatting the Strings, Working with Characters

Unit III-

Lists and Tuples: Lists, List Functions and Methods, List Operations, Tuples
Dictionaries: Creating a Dictionary, Operators in Dictionary, Dictionary Methods,Using for Loop with Dictionaries, Operations on Dictionaries, Ordered Dictionaries
Regular Expressions: What is a Regular Expression? Sequence Characters in Regular Expressions, Quantifiers in Regular Expressions,Special Characters in Regular Expressions, Using Regular Expression on Files, Retrieving Information from an HTMLFile
Date and Time in Python: Date and Time, Date and Time Now, Combining Date and Time, Formatting Dates and Times, Finding Durations using "timedelta", Comparing Two Dates, Sorting Dates, Stopping Execution Temporarily , Knowing the Time taken by a Program, Working with Calendar Module

Unit IV-

IPython: Beyond Normal Python, Help and Documentation in IPython, Keyboard Shortcuts in the IPython Shell, IPython Magic Commands, Input and Output History, IPython and Shell Command s, Errors and Debugging, Profiling and Timing Code
Introduction to NumPy: Understanding Data Types in Python, The Basics of NumPy Arrays, Computation on NumPy Arrays: Universal Functions,
Aggregations: Min, Max, and Everything In Between,
Computation on Arrays: Broadcasting, Comparisons, Masks, and Boolean Logic, Fancy Indexing, Sorting Arrays,
Structured Data: NumPy's Structured Arrays

Books and References:

1. Python Programming by Krishna Karoo HSRA PUBLICATIONS; 1st edition
2. Programming through Python M. T. Savaliya, R.KMaurya,G.M Magar Star edu Solutions
3. Python Data Science Handbook Jake Vander Plas O'Reilly Media 1st edition

B.Sc. I (Data Science)
For Semester I
Paper IV - Web Technology

Total Marks - 50		Max. Time- 2 Hrs
End Sem Exam- 40 Marks Min. Passing Marks- 20	Internal Assessment- 10 Marks Min. Passing- 50 %	Credits- 2 Lectures- 60

Program Learning Outcomes:

1. Learn the basic concepts of web design and development.
2. Understand the structure and components of a web page.
3. Gain knowledge about the internet, web servers, and web browsers.
4. Develop the ability to create well-structured HTML documents.

UNIT – I: Introduction to HTML

Basic of HTML and Tag Introduction to HTML - Introduction, Features of HTML, Advantages & Disadvantages of HTML, HTML Editors, Step to Create and View HTML Document, Basic Structure of HTML Program Tags & Attributes-Nesting of Tags, Classification of HTML Tags, Block Formatting Tags. List - Introduction to Lists, Unordered List, Ordered List, Definition List, Nested List, Difference Between Ordered and Unordered List.
Working with HTML Linking - Introduction, Type of Hyperlink Creation, Working with Links, Pathname and Types, Types of Linking or Anchors. Graphics in Web Page - Image Tag, Align Images, Embedding Inline Images and External Images.

UNIT – II: HTML Tables and Forms

Tables - Basic table tags and their related attribute, Creating Forms ,Tables, HTML5 Semantics, Advanced HTML Form designs, Form Controls, Text controls, password fields, radio buttons, and check boxes. Reset and submit buttons, form control selection, option processing and text area

UNIT – III: Introduction to CSS

Introduction to CSS: Concept of CSS, Creating of Style sheet, CSS Properties,Three ways to use CSS, CSS Styling (Background, Text Format, Controlling Fonts), Working with block elements

and objects, Working with Lists and Tables with styling , CSS Id and Class, Box Model (Introduction, Border properties, Padding Properties, Margin properties), CSS frameworks, Bootstrap Introduction, uses and application

UNIT – IV: Scripting language JavaScript

JavaScript: Introduction, Client-SideJavaScript, Server Side JavaScript, JavaScript Objects, JavaScript Security Core JavaScript (Properties and Methods of Each) :
Array, Boolean, Date, Function, Math, Number, Object, String, regExp Document and its associated objects: document, document object methods, Link, Area, Anchor, Image, Layer Events and Event Handlers: General Information about Events, Defining Event Handlers, event, onAbort, onBlur, onChange,onClick,onDblClick,onDragDrop,onError,onFocus,onKeyDown, onKeyPress, onKeyUp, onLoad, onMouseDown, onMouseMove, onMouseOut , onMouseOver, onMouseUp, onMove, onReset, onResize, onSelect, onSubmit, onUnload

Text Books:

1) "Scripting languages and information retrieval " by Dr. Krishna Karoo 1 st edition HSRA PUBLICATIONS - 7892793054
2) The World of Scripting Languages, David Barron,Wiley Publications

References:

1. The Ruby Programming Language, David Flanagan and Yukihiro Matsumoto, O'Reilly Publications.
2. Beginning JavaScript with Dom Scripting and AJAX, RussFerguson, Christian Heilmann,Apress.

B.Sc. I (Data Science)
For Semester I
Paper V- Business Communication and Information Ethics

Total Marks - 50		**Max. Time- 2 Hrs**
End Sem Exam- 40 Marks **Min. Passing Marks-20 Marks**	**Internal Assessment- 10 Marks** **Min. Passing- 50%**	**Credits- 2** **Lectures- 60**

Course Objectives:

- Verbal communication techniques enhance communication and explain the barriers to communication
- To discuss various business activities which are essential at the workplace.
- To explain business communication covering the structure and layout of a letter, planning of a letter and use of language.
- To explain the use of agenda and minutes for effective functioning of any organization.
- To direct the learners attention to the significance of effective writing and the importance and structure of reports.

Unit- I Interpretation of Communication

Basics of communication, Types of Communication, Channels of Communication
Non-verbal communication, Barriers to communication.

Unit-II Business communication at workplace

Letter components and layout
Planningaletter, Process of letter writing,
Email Communication, Memos and memoreports, Employment communication. Notice, Agenda and minutes of meeting. Brochures.

Unit- III Report Writing

Effective writing, Types of business reports, Structure of reports, Gathering Information. Organization of material writing abstracts and summaries, Writing definitions, Visual aids, User Instruction Manual.

Unit-IV Presentation Skills

Importance of Presentation Skills, Planning and Structuring Presentations, Designing Visual Aids

Elements of Presentation skills.

Books and References:

1. Koneru, Arana. *Title Professional Communication* . Tata Mcgraw Hill , 2008.
2. Floridi, Luciano. *The Ethics of Information*. Oxford University Press, 2013.
3. Krizan Patricia Merrier, JoyceLogan, Karen Williams. *Business Communication*. Thomson, 2008.

B.Sc. I (Data Science)
For Semester I
Subject- Environment Protection Law

Total Marks - 50		**Max. Time-**
End Sem Exam- **Min. Passing Marks-**	**Internal Assessment- 50 Marks** **Min. Passing- 20 Marks**	**Credits- 2** **Lectures- 30**

Course Objectives:

- To learn and sensitize learners their environment
- To Know About natural resources, ecology and ecosystem
- To learn insights of biodiversity , pollution and its impact
- Explore about Social Issues and the Environment
- To learn about Environment Management and sustainable development

Unit I- Introduction to Environmental Studies

Importance of Environmental Education, Environmental Literacy, Environmental Engineering, Environmentalism, Components of Environment and their Interactions, Man and the Biosphere, Impacts of Development On Environment Natural Resources Forest Resources , Dams, Water Resources, Food Resources Energy Resources, Land Resources

Unit II- Ecology Ecosystem, Ecosystem

-Anthropo system Comparison ,Biome and Ecosystem, Energy Flow through the Ecosystem, Ecological Succession,Food Chains and Webs, Ecological Pyramids, Biological Magnification or Biomagnification, Human versus Natural Food Chains, Biogeochemical Cycles,
The Water Cycle (Hydrologic Cycle), Carbon Cycle, Oxygen Cycle, Nitrogen Cycle, Forest Ecosystems, Grassland Ecosystems, Aquatic Ecosystems Biodiversity Genetic Diversity, Species Diversity, Ecosystem Diversity, Value of Biodiversity, Value of Genes, Biopiracy, Biogeographical Classification of India, India a Mega Diverse Nation, Endemic Species of India, Threats to Biodiversity ,Hotspots of Biodiversity, Endangered Species, Conservation of Biodiversity, Genetic Engineering and Biodiversity

Unit III- Environmental Pollution

Air Pollution, Water Pollution, Soil Pollution, Marine Pollution, Noise Pollution,Thermal Pollution, Solid Waste Management, Hazardous Waste Management, Pollution Prevention,

Disaster Management Human Population and the Environment Population Growth, Human Rights,Value Education, HIV/AIDS,Environment and Human Health, Family Welfare Programmes, Women and Child Welfare, Role of Information Technology in Environment and Human Health

Unit-IV- Social Issue and Environment

From Unsustainable to Sustainable Development, Urban Problems Related to Energy, Water Conservation, Watershed Management, Resettlement and Rehabilitation, Environmental Ethics, Acid Rain, Ozone Layer Depletion, Greenhouse Effect, Global Warming and Climate Change, Pollution Control Boards and Control Pollution Acts in India, Nuclear Hazards and Accidents, Environmental Impact Assessment, Risk Management, Precautionary Principle, Polluter-Pays Principle,The Beneficiary-Pays Principle, Role of Non-Government Organizations, ISO 14000 Series of Environmental Management Standards, Economy And Environment

Unit V- Environmental Management

Environmental Impact Assessment, Methodology, Waste As Resource, Environmental Laws, Requirements of a Contract, Environmental Legislations, Powers and Functions of Pollution Control Boards, Case Studies, Environmental Management Plan, Environmental Audit, Policies for Quality Improvement, Problems, Policy SustainableDevelopment Ethics, Laws of Nature, Progress, Environmental Stress, Sustainability, Self purification and Regeneration, ActionPlan, Computerization and Information Technology

Books and References:

1. Environmental Studies - Benny Joseph McGrawHill third edition
2. Principles of Environmental Science and Engineering- P.Venugopala Rao- PHI Learning- 6th edition
3. Introduction to Environmental Engineering - GilbertMMaster- Pearson third edition
4. Environmental Ethics: A Very Short Introduction- Robin Attfield- Oxford- 1st edition
5. Fundamental Concept in Environmental Studies- D.D.Mishra- S.Chand- Revised Edition- 2010

B.Sc. I (Data Science)
For Semester I
Subject - Precalculus

Total Marks - 50		Max. Time-
End Sem Exam- Min. Passing Marks-	Internal Assessment- 50 Marks Min. Passing- 50 %	Credits- 2 Lectures- 30

CourseObjectives:

- To master the number fundamentals, equations and different types of mathematical functions .
- To review and explain trigonometry gain expertise in trigonometric identities.
- To understand analytical trigonometry and inverse functions.
- To Give The Detailed knowledge about complex numbers, vectors and matrices.
- To Understand The Conics,sequences and series

UNIT – I:

Fundamentals:Real Numbers, Exponents and Radicals, Algebraic Expressions, Rational Expressions, Equations, Modeling with Equations, Inequalities , Coordinate Geometry, GraphingCalculators;SolvingEquations and Inequalities Graphically, Lines, Making Models UsingVariation.

Functions: What is function? Graphs of Functions, Getting Information from the Graph of a Function, Average Rate of Change of a Function, Transformations of Functions, Combining Functions, One-to One Functions and Their Inverses. Polynomial and Rational Functions: Quadratic Functions and Models, Polynomial Functions and Their Graphs, Dividing Polynomials,

Real Zeros of Polynomials, Complex Numbers, Complex Zeros and the Fundamental Theorem of Algebra, Rational Functions.

UNIT – II:

Exponential and Logarithmic Functions: Exponential Functions, The Natural Exponential Function, Logarithmic Functions, Laws of Logarithms, Exponential and Logarithmic Equations, Modeling with Exponential And Logarithmic Functions.

Trigonometric Functions: Unit Circle Approach: The Unit Circle,Trigonometric Functions of Real Numbers, Trigonometric Graphs, Inverse Trigonometric Functions and Their Graphs, Modelling Harmonic Motion

Trigonometric Functions: Right Triangle Approach: Angle Measure,Trigonometry of Right Triangles, Trigonometric Functions of Angles, Inverse Trigonometric Functions and Right Triangles, The Law of Sines, The Law of Cosines.

Unit III:-

Analytic Trigonometry: Trigonometric Identities, Addition and Subtraction Formulas, Double-Angle, Half-Angle, and Product Sum Formulas,Basic Trigonometric Equations,More Trigonometric Equations

SinusoidalFunctions:A special class of functions,Sketching Sinusoidal graph, Functions not in standard sinusoidal form, sinusoidal behavior. Inverse Circular Functions: Solving three equations, inverse Circular Functions,applications, solving trigonometric equations

Unit IV:-

Polar Coordinates and Parametric Equations: Polar Coordinates,Graphs of Polar Equations, Polar Form of Complex Numbers; DeMoivre's Theorem, Plane Curves And Parametric Equations Vector sin Two and Three Dimensions: Vector In Two Dimensions,The Dot Product, **Three-Dimensional Coordinate Geometry,** Vectors Three Dimensions, The Cross Product, Equations of Lines and Planes Systems of Equations and Inequalities: Systems of Linear Equations in Two Variables, Systems of Linear Equations in Several Variables, Matrices and Systems of Linear Equations, The Algebra of Matrices, Inverses of Matrices and Matrix Equations, Determinants and Cramer's Rule, Partial Fractions, System Of Nonlinear Equations, System Of Inequalities

Books and References:

1. Precalculus– Mathematics for Calculus- James Stewart,Lothar Red lin,Saleem Watson Cengage Learning-2013
2. Precalculus- David H. Collingwood, K. David Prince, Matthew M. Conroy Free Software Foundation-2011
3. Precalculus Demystified- Rhonda Huettenmueller Tata McGraw wHill -2005
4. Contemporary Precalculus: A Graphing Approach- ThomasW. Hungerford,Douglas J. Shaw Thomson Higher Education-2009

B.Sc. I (Data Science)
For Semester I
Lab Course - Web Technology

Total Marks - 50		Max. Time- 2 Hrs
End Sem Exam- 30 Marks Min. Passing Marks- 20	Internal Assessment- 20 Marks Min. Passing- 50 %	Credits- 2 Practicals - 16

Practicals

Lab: Practical Questions (Use any Text Editor, like Notepad++, Sublime text,Visual Studio code etc.)

1. Write a HTML Code for the Demonstration of Logical and physical (Formatting) style tags.
2. Write a HTML Code for the Demonstration of Level of Headings and Block Alignment.
3. Write a HTML Code for the Demonstration of the Font Face, Color and Size. And address tag.
4. Write a HTML Code for the Demonstration of HR Tag and Alignment.
5. Write a HTML Code for the Demonstration of Ordered and unordered list.
6. Write a HTML Code for the Demonstration of internal linking.
7. Write a HTML Code for the Demonstration of table use.
8. Write a HTML Code for the Demonstration of browsing by category.
9. Write a HTML Code for the Demonstration of designing a simple form.
10. Write a HTML Code for the Demonstration of the Master page to link another page.
11. Write a HTML Code for the Demonstration of Link to Web Page.
12. Write a HTML Code for the Demonstration of Compose Mail.
13. Write a HTML Code for the Demonstration to show or load inline images, say waterfall.JPG.
14. Write a HTML Code for the Demonstration of Image Hyperlink.
15. Write a HTML Code for the Demonstration of cell padding attributes.
16. Write a HTML Code for the Demonstration of the use of element selector, id selector and class selector with CSS.
17. Write a HTML Code for the Demonstration of Navigation (with Dropdown) with CSS.
18. Write a HTML Code to Create a CSS Grid.
19. Write a HTML Code to style the element which is not empty with CSS.
20. Write a HTML Code to Create a CSS based Zebra Striped table.

B.Sc. I (Data Science)
For Semester I
Lab Course - ICT Practicals

Total Marks - 50		Max. Time- 2 Hrs
End Sem Exam- 30 Marks Min. Passing Marks- 20	Internal Assessment- 20 Marks Min. Passing- 50 %	Credits- 2 Practicals - 16

Program Learning Outcomes:

- To help the learners become competent and confident users of ICT who can make efficient, effective and creative use of basic application software in their everyday activities.
- To encourage the learners to become critical and reflective users of ICT who can evaluate the capabilities and limitations of the technology and of social, technical, political, ethical, organizational and economic principles associated with its use.
- To prepare the learners for the society of tomorrow by making them adaptable users of ICT who have the necessary openness and flexibility of mind to be able to adjust to future changes in the technology.
- To encourage the learners to develop the appropriate social skills that are essential for cooperative and collaborative learning based around ICT.
- To empower ICT disadvantaged learners by ensuring sufficient access for those learners who have little out-of-school opportunities to use the technology.

List of Practicals:

1. Word Processor: Using word processor for letters,documentation per formatting and advanced features.
2. Presentation Software: Using presentation software for preparing elegant presentations with voice and videos and giving different effects to make it more interesting and catchy.
3. Using Google docs and forms
4. Using Online Collaboration And Video Conferencing Tools
5. Using Screen Recording And Audio Tools
6. Using video editing tools
7. Using Google Maps, Google Street view, Bing Maps
8. Using Social Media (Facebook,Instagram,Twitter,Linkedin,youtube,snapchat, reddit,quora,dig, Pinterest, flipboard, Wordpress, Tumblr, Medium)for business and learning

B.Sc. - I (Data Science)
Semester – II
Syllabus

B.Sc. I (Data Science)
For Semester II
Paper I - Database Management Systems

Total Marks - 100		Max. Time- 2 Hrs
End Sem Exam- 80 Marks Min. Passing Marks- 40	Internal Assessment- 20 Marks Min. Passing- 50 %	Credits- 4 Lectures- 60

Course Objectives:

This course concentrates the concept of the DBMS with respect to principles, design and implementation of DBMS. It aims to specify the functional and data requirements for a typical database application and to understand creation, manipulation and querying of data in databases.

- To understand Organizing, structuring and storing data.
- Understand Database as Relationalmodel.
- To understand SQL to retrieve data and the concept of redundancy.
- To specify the functional and data requirements for a typical database application.
- To understand creation,manipulation and querying data in databases.

Unit-I-

Introduction & DBMS Architecture: Why Databases? Data versus Information, Introducing the Database, Role and Advantages of the DBMS,Types of Databases, Why Database Design Is Important, Evolution of File System Data Processing, Problems with File System DataProcessing, DatabaseSystems Data Models – Data Modeling and Data Models, The Importance of Data Models, Data Model Basic Building Blocks, Business Rules, The Evolution of Data Models, Degrees of Data Abstraction Entity Relationship

Model: Entities, attributes, Relationships, Connectivity and Cardinality, Existence Dependence, Relationship Strength, Weak Entities, Relationship Participation, Relationship Degree, Recursive Relationships, Associative(Composite) Entities, Developing an ER Diagram, Database Design Challenges: Conflicting Goals

Unit-II

Advanced Data Modeling: The Extended Entity Relationship Model,Entity Clustering, Entity Integrity: Selecting Primary Keys,

Design Cases: Learning Flexible Database Design Normalization of Database Tables:Database Tables and Normalization, The Need for Normalization, The Normalization Process, Improving

the Design, Surrogate Key Considerations, Higher Level Normal Forms, Normalization and Database Design, Denormalization, Data-Modeling Checklist

Unit-III

Structured Query Language (SQL): Introduction to SQL, Basic SELECT Queries, SELECT Statement Options, FROM Clause Options, ORDERBY Clause Options, WHERE Clause Options, Aggregate Processing, Subqueries, SQLFunctions, Relational Set Operators, Crafting SELECT Queries

AdvancedSQL: Data Definition Commands, Creating Table Structures, Altering Table Structures, Data Manipulation Commands,Virtual Tables: Creating a View, Sequences, ProceduralSQL, EmbeddedSQL Database

Design: The Information System, The Systems Development LifeCycle, The Database Life Cycle, Conceptual Design, DBMS Software Selection, Logical Design, Physical Design, Database Design Strategies, Centralized versus Decentralized Design

Unit-IV

Transaction Management and Concurrency Control: What Is a Transaction? Concurrency Control with Locking Methods, Concurrency Control with Time Stamping Methods, Concurrency Control with Optimistic Methods, ANSI Levels of Transaction Isolation, Database Recovery Management

Database Performance Tuning and Query Optimization: Database Performance-Tuning Concepts, Query Processing, Indexes and Query Optimization, Optimizer Choices, SQL Performance Tuning, Query Formulation, DBMS Performance Tuning, Query Optimization Examples

Database Administration and Security: Data as a Corporate Asset, The Need for a Database and Its Role in an Organization, Introduction of a Database: Special Considerations, The Evolution of Database Administration, The Database Environment's Human Component, Security, Database Administration Tools, Developing a Data Administration Strategy, The DBA's Role in the Cloud, The DBA atWork:Using Oracle for Database Administration

Books and References:

1. Fundamentals of Database Systems- Elmasri Ramez and Navathe Shamkant B, Pearson Education 6th Edition
2. Database System Concepts Silberschatz,Korth, Sudarshan, McGraw Hill, 5 Edition,
3. Database Management Systems, Ramakrishnam, Gehrke, McGraw- Hill,
4. Murach's MySQL Joel Murach,-Murach,

B.Sc. I (Data Science)
For Semester II
Paper II - R Programming

Total Marks - 50		Max. Time- 2 Hrs
End Sem Exam- 40 Marks Min. Passing Marks- 20	Internal Assessment- 10 Marks Min. Passing- 50 %	Credits- 2 Lectures- 48

Course Objectives:

1. Recognize various disciplines that contribute to a successful data science effort.
2. Understand the processes of data science - identifying the problem to be solved, data collection, preparation, modeling, evaluation and visualization.
3. Be aware of the challenges that arise in data sciences.
4. Develop and appreciate various techniques for data modeling and mining.
5. Be cognizant of ethical issues in many data science tasks.

Unit-I-

Basics of R Programming:

Evolution of R, Features of R, Local Environment support, R Command prompt, R Script File, Comment, R Data types, R Variables, R Operators-function. Understanding data: Introduction – Types of Data: Numeric – Categorical – Graphical – High Dimensional Data – Classification of digital Data: Structured, Semi-Structured and Unstructured - Example Applications. Sources of Data: Time Series – Transactional Data – Biological Data – Spatial Data – Social Network Data – Data Evolution.

Unit-II

R Fundamentals :

Introduction to R- Features of R - Environment - R Studio. R-Decision Making:- R-If statement, R-If....else statement, R- The if....else if...else statement-Switch Statement, RLoop:- Repeat loop, While loop, for loop, Loop ,Control statement:- Break, Next. Basics of R-Assignment - Modes -Operators - special numbers - Logical values – Basic Functions - R help functions - R Data Structures - Control Structures. Vectors: Definition- Declaration - Generating - Indexing - Naming - Adding & Removing elements - Operations on Vectors - Recycling - Special Operators - Vectorized if- then else-Vector Equality – Functions for vectors - Missing values - NULL values - Filtering & Subsetting.

Unit-III

Data Structures in R :

Matrices - Creating Matrices - Adding or Removing rows/columns - Reshaping - Operations - Special functions on Matrices. Lists - Creating List – General List Operations - Special Functions - Recursive Lists. Data Frames - Creating Data Frames - Naming - Accessing - Adding - Removing - Applying Special functions to Data Frames - Merging Data Frames Factors and Tables.

Unit-IV

Input / Output – Reading and Writing datasets in various formats - Functions - Creating User defined functions - Functions on Function Object - Scope of Variables - Accessing Global, Environment - Closures - Recursion. Exploratory Data Analysis - Data Preprocessing - 12 Descriptive Statistics - Central Tendency - Variability - Mean - Median - Range - Variance - Summary - Handling Missing values and Outliers - Normalization Data Visualization in R : Types of visualizations - packages for visualizations - Basic Visualizations, Advanced Visualizations and Creating 3D plots.

Books and References:

1. Fundamentals of Database Systems- Elmasri Ramez and Navathe Shamkant B, Pearson Education 6th Edition
2. Database System Concepts Silberschatz,Korth, Sudarshan, McGraw Hill, 5 Edition,
3. Database Management Systems, Ramakrishnam, Gehrke, McGraw- Hill,
4. Murach's MySQL Joel Murach,-Murach,

B.Sc. I (Data Science)
For Semester II
Paper III- Data Analysis

Total Marks - 50		Max. Time- 2 Hrs
End Sem Exam- 40 Marks Min. Passing Marks- 20	Internal Assessment- 10 Marks Min. Passing- 50 %	Credits- 2 Lectures- 48

Course Outcomes

On Completion of this course, student will be able to -

- Introduction to data science technologies and techniques
- Introduction to data acquisition techniques
- Data analysis using R programming
- How to wrangle data using combining and merging data sets

Unit-I - Foundation of Data Science

1.1 Data Science as a discipline, data to data science , Data growth issue and data challenges

1.2 Foundation of data science, Tools for data science , Application of data science

1.3 Data Science Process, Messy data , Cleaning data

1.4 Anomalies and Artifacts in Datasets , Owner of data

1.5. Five C-s of data Science , Future trends in data Science

Unit-II - Data Acquisition

2.1 Introduction to data acquisition

2.1.1 Data Preprocessing and Techniques

2.1.2 Data Integration , Transformation and Data interpretation

2.2 Data Reduction

2.2.1 Dimensionality Reduction

2.2.2 Numerosity Reduction

2.2.3 Data Compression

2.3 Data Mining

2.3.1 Data Mining Applications

2.3.2 Combining Data, Ordering and Describing Data, Transforming Data,

Unit III- Data Wrangling

3.1 Introduction to data wrangling - steps and tools

3.2 Combining and merging datasets

3.3 Reshaping data sets using melt(), stack(), Unstack(), and pivote() functions

3.4 Data Transformation using data frame creation, missing value , encoding, inserting and splitting values
3.5 Regular expressions - findall(), search(), split() and sub() function

Unit-IV - Data Manipulation using R Programming
4.1 Why to use R for data Analysis?
4.2 Something more about R data types
4.3 Reading and writing data in txt,csv and other file formats
4.5 Simple Data Visualization
4.6 Common Statistical Processes
4.7 Simple Linear regression
4.8 Bootstrapping

Reference Books
1.Data Analysis Became the master in data analysis- by Richard Dorsey
2.Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking
3. Statistical Data Analysis: A Practical Guide- Book by Jiri Militky and Milon Meloun

B.Sc. I (Data Science)
For Semester II
Paper IV - Internet Of Things

Total Marks - 50		Max. Time- 2 Hrs
End Sem Exam- 40 Marks Min. Passing Marks- 20	Internal Assessment- 10 Marks Min. Passing- 50 %	Credits- 2 Lectures- 48

Course Outcomes

On Completion of this course, student will be able to -

1. Explain the definition and usage of the term —Internet of Things‖ in different contexts
2. Understand the key components that make up an IoT system
3. Differentiate between the levels of the IoT stack and be familiar with the key technologies and protocols employed at each layer of the stack
4. Apply the knowledge and skills acquired during the course to build and test a complete, working IoT system involving prototyping, programming and data analysis
5. Discover where the IoT concept fits within the broader ICT industry and possible future trends

Unit-I- Introduction To IOT

Introduction – Definition and Characteristics of IoT, Physical Design of IoT; Things in IOT, Logical Design of IoT; IoT Functional Blocks, IoT Communication APIs, IoT Enabling Technologies; WSN, Cloud Computing, Big Data Analysis, Communication Protocols, Embedded Systems

Unit-II- IoT Hardware

IoT Hardware, Devices and Platforms – Basics of Arduino Hardware, The Arduino IDE, Basic Arduino Programming, Basics of Raspberry pi; Introduction to Raspberry pi, Programming

Unit-III- IoT Protocols

IoT Protocols – IoT Data link Protocols, Network Layer Routing Protocols, Network Layer Encapsulation Protocols, Session Layer Protocols, IoT Security Protocols, Service Discovery Protocols, Infrastructure Protocols.

Unit-IV- IoT Programming

IoT Programming – Arduino Programming: Serial Communications – Getting Input from Sensors, Visual, Physical and Audio Outputs, Remotely Controlling External Devices, Wireless Communication, Programming with Raspberry pi: Basics of python Programming, Python Packages of IoT, IoT Programming with CADC IoT devices.

Unit-V- Domain Specific IoT

Domain Specific IoT – Home automation, smart cities, Smart Environment, IoT in Energy, Logistics, Agriculture, industry and Health & Life style sensors, Case Studies: A Case Study of Internet of Things Using Wireless Sensor Networks and Smart Phones, Security Analysis of Internet-of-Things: A Case Study of August Smart Lock, Open IoT Platform.

Reference Books-

1. Vijay Madisetti and ArshdeepBahga, —Internet of Things (A Hands-on-Approach), 1st Edition, VPT, 2014.
2. Margolis, Michael. —Arduino CooK book: Recipes to obegin, Expand and Enhance Your Projects‖. O'Reilly Media Inc.2011.
3. Monk, Simon. Raspberry Pi Cookbook: Software and hardware problems and Solutions. O'Reilly Media,Inc. 2016.

B.Sc. I (Data Science)
For Semester II
Paper V- Modern Indian Language- Marathi

Total Marks - 50		Max. Time- 2 Hrs
End Sem Exam- 40 Marks Min. Passing Marks-20 Marks	Internal Assessment- 10 Marks Min. Passing- 50%	Credits- 2 Lectures- 48

5- Marathi / Hindi

(1) मराठी आवश्यक

गद्य विभाग:—
१. अखेरचे कीर्तन— गाडगेबाबा
२. लोकशाहीचे भवितव्य— डॉ. बाबासाहेब आंबेडकर
३. विज्ञानयुगात भारत. जयंत नारळीकर
४. भटक्या— कचरू जनार्दन गिऱ्हे
५. माणूस— उत्तम कांबळे

पद्य विभाग:—
१. पसायदान— ज्ञानेश्वर
२. विद्यार्थ्यांप्रत— केशवसुत
३. भंगू दे काठीण्य माझे— बा. सी. मर्ढेकर
४. स्वप्नांची समाप्ती— कुसुमाग्रज
५. लेखनीच्या तलवारी— उषाकिरण आत्राम

व्यावहारिक मराठी व व्याकरण:—

१. भाषिक कौशल्य आणि व्यक्तीमत्व विकास
२. भाषा, लिपी आणि वर्णविचार

पाठ्यपुस्तक:— 'अक्षरधारा' (संपादित), राघव प्रकाशन, नागपूर.

B.Sc. I (Data Science)
For Semester II
Environmental Science for Data Science- Project and Seminar

Total Marks - 50		Max. Time-
End Sem Exam- Marks Min. Passing Marks- 25	Internal Assessment- 50 Marks Min. Passing- 50 %	Credits- 2 Practicals - 16

Note- Students must choose data science related topics from the Environmen Protection Law curriculum and should prepare a seminar and Project presentation based on topic.

Gondwana University, Gadchiroli
Choice Based Credit System (CBCS)
U.G. PROGRAMME SESSION 2024-25
Faculty of Science and Technology

B.Sc. (Data Science) Semester – III & Semester - IV Syllabus

Note: For Details about assessment, practical examination, Paper Pattern, kindly refers Appendix

B.Sc. - II (Data Science)
Semester – III
Syllabus

B.Sc.–II
SEMESTER-III
(Data Science)
Paper I- Web Application Development

Total Marks - 100		**Max. Time- 3 Hrs**
End Sem Exam- 80 Marks **Min. Passing Marks- 40**	**Internal Assessment- 20 Marks** **Min. Passing- 50 %**	**Credits- 4** **Lectures- 60**

Learning Objectives:

The aim of this course is to apply basic design principles to present ideas, information, products, and services on websites applications.

Learning Outcomes:

- Student will be able to understand use of web technology in web application development
- Student will be able to understand server side scripting languages
- Student will be able to understand javascript library applications
- Students will learn to handle the backend data handling with node.js

Unit-1- Javascript JSON
JSON-Introduction, Syntax, data , JSON objects and arrays , JSON data values, Comparison of JSON and XML, JSON parse(), stringify(), JSON server, JSON PHP, JSON HTML, JSONP
JQuery- Javascript Vs. JQuery , JQuery Selectors

Unit-2- Server Side Scripting using PHP
PHP Introduction, Installation, Syntax, Comments, Variables in PHP, echo and print statements, Data types in PHP- Strings Operations on strings , Numbers, PHP casting, Math, Operators , if, if-else , nested if , switch , loops, while,do while, for, foreach,break, continue, functions, Arrays,operations on arrays, php global, php servers,
PHP form handling and validation, php file handling, file operations, php cookie, php sessions , php filters, PHP exceptions, PHP Ajax

Unit-3- Javascript library Applications- ReactJs

Introduction to React, React History and features, Concept of DOM, React Dom, Comparison DOM and Virtual DOM, Setting up React Environment, React versions ,
React ES6 Features- Classes, Arrow functions, React variables, Array methods, rest and spread operators , Ternary Operators ES6 Modules,
React JSX, React Render, React Components- Function based, Class based, React props, React constructor, React life cycles, React Events, React Lists, React Forms, React Router,
React Hooks- useState, useEffect, useContext, useRef, useReducer,useReducer,useMemo

Unit-4- Node.js server environment
Introduction to Node.js- What and Why Node js? Applications, Setting up environment ,
Module in Nodejs- Http module,file System module , URL module, Events module, send email, NPM
Nodejs MySQL- Create database, create tables, selection from tables, selection filters, CRUD operation on table,
Nodejs-MongoDB- Create database and collection, select and find in MongoDB, CRUD in MongoDB

Reference Books:

1. Advanced Web Development with React - Mehul Mohan - bpb publications-
2. Web Development with Nodejs and express- Ethan brown - O'Really publication
3. Mastering Nodejs- Sandro Pasquali Pact publication
4. PHP 8 basics - Gunnard Engbirth Satej kumar Sahu, Apress publication

B.Sc.–II
SEMESTER-III
(Data Science)
Paper II- Probability and Statistics

Total Marks - 50		Max. Time- 2 Hrs
End Sem Exam- 40 Marks Min. Passing Marks- 20	Internal Assessment- 20 Marks Min. Passing- 50 %	Credits- 2 Lectures- 60

Learning Objectives:

To make students able to
1. Learn the basics and advanced concepts of Probability theory
2. Learn various sampling techniques
3. Find and understand the applications of Probabilities in data science.

Learning Outcomes:
After completion of this course successfully the students will be able to:
1. Solve problem related to probability
2. Understand and use various Probabilities theory to solve real problem of data science
3. Understand and use various Probability distributions for different machine learning related tasks.
4. Understand the sampling techniques and use them for preparing effective datasets.

Unit I- Introduction to Probability

Basic Probability - Random Experiments - Sample Spaces Events - The Concept of Probability -The Axioms of Probability - Some Important Theorems on Probability - Assignment of Probabilities -Conditional Probability -Theorems on Conditional Probability – Independent Events -Bayes' Theorem or Rule Combinatorial Analysis - Fundamental Principle of Counting - Tree Diagrams –Permutations. Introduction to Statistics – Primary and Secondary data – Nominal, Ordinal, Ratio, and Interval scale (with examples) - Graphical Representation of data – Bar-charts, Pie-diagrams, Histograms, Frequency polygon, Ogives.
Central Limit Theorem and Confidence Interval: Introduction, Sampling Variability and CLT, CLT (for the mean) examples, Confidence Interval (for a mean), Accuracy vs. Precision, Required Sample Size for ME, CI (for the mean) examples.

Unit II- Random Variables and Probability Distributions

- Random Variables - Discrete Probability Distributions -Distribution Functions for Random Variables - Distribution Functions for Discrete Random Variables - Continuous Random Variables – Graphical Interpretations Joint Distributions Independent Random Variables - Change of Variables - Probability Distributions of Functions of Random Variables – Convolutions – Conditional Distributions Applications to Geometric Probability.

Measures of central tendency :– properties – merits and demerits – weighted means– graphical location of median,quartiles, deciles, percentiles, and mode – relation between arithmetic mean, geometric mean and harmonic mean.

Measures of dispersion : – characteristics – Coefficient of dispersion – Coefficient of variation – Moments –Relation between moments about mean in terms of moments about point – Pearson's coefficients.

Inference and Significance: Introduction to Inference, Hypothesis Testing (for a mean), HT (for the mean) examples, Inference for Other Estimators, Decision Errors, Significance vs. Confidence Level, Statistical vs. Practical Significance.

Unit III-

Mathematical Expectation - Definition of Mathematical Expectation - Functions of Random Variables - Theorems on Expectation - Variance & Standard Deviation - Theorems on Variance - Standardized Random Variables - Special Probability Distributions - Binomial Distribution - Normal Distribution - Poisson distribution.

Skewness and Kurtosis – Pearson's coefficient of skewness – Bowley's coefficient of skewness – coefficient of skewness based upon moments

Curve fitting – Principle of least squares – Fitting of straight line, parabola, exponential and power curve.

Inference for Comparing Means: Introduction, t-distribution, Inference for a mean, Inference for comparing two independent means, Inference for comparing two paired means, Power, Comparing more than two means, ANOVA, Conditions for ANOVA, Multiple comparisons, Bootstrapping.

UNIT-IV-

Sampling Theory - Population and Sample - Statistical Inference- Sampling With and Without Replacement Random Samples - Random Numbers - Population Parameters - Sample Statistics - Sampling Distributions - Sample Mean - Sampling Distribution of Means - Sampling Distribution of Proportions - Sampling Distribution of Differences and Sums – Sample Variance - Sampling Distribution of Variances - Computation of Mean, Variance, and Moments for Grouped Data - The Least-Squares Parabola - Multiple Regression Standard Error of Estimate

The Linear Correlation Coefficient Generalized CorrelationCoefficient Rank Correlation **Inference for Proportions:** Introduction, Sampling Variability and CLT for Proportions, Confidence Interval for a Proportion, Hypothesis Test for a Proportion, Estimating the Difference Between Two Proportions, Hypothesis Test for Comparing Two Proportions, Small Sample Proportions, Examples, Comparing Two Small Sample Proportions, Chi-Square GOF Test, The Chi-Square Independence Test.

Correlation and Regression: Simple correlation – Karl Pearson's coefficient. of correlation – Rank correlation –Simple Regression – lines of regression – properties of regression coefficient –Multiple and Partial correlation coefficient in three variables. Hypothesis Testing : Estimation and Hypothesis testing, t-test, chi-square test, ANOVA

Reference Books

1. Murray R. Spiegel, John J. Schiller & R. Alu Srinivasan, "Probability and Statistics", Schaum outlines, McGraw Hill, 3rd edition, 2009.
2. S. P. Gupta, Statistical Methods, S. Chand and Sons.
3. S. C Gupta and V. K. Kapoor, "Fundamentals of Mathematical Statistics", 11th edition, S.Chand and Sons.
4. Agarwal.B.L (1996): Basic Statistics, 3/e, New Age International (P) Ltd.,.
5. Sanjay Arora & Bansilal (2002): New Mathematical statistics, Meerat Publications, New Delhi
6. Hooda.R.P.(2003): Statistics for Business and Economics, 3/e, Mac Millan.

B.Sc.–II
SEMESTER-III
(Data Science)
Paper III- Fundamentals in Data Science

Total Marks - 50		**Max. Time- 2 Hrs**
End Sem Exam- 40 Marks **Min. Passing Marks- 20**	**Internal Assessment- 20 Marks** **Min. Passing- 50 %**	**Credits- 2** **Lectures- 60**

Learning Objectives:

To make students able to
- Learn the fundamental concepts of data science
- Know the various domain and vertices of data science
- Learn the usage and application of data science

Learning Outcomes:
After completion of this course successfully the students will be able to:
1. Understand the key difference between various areas of data science.
2. Understand the fundamental concepts of tools and techniques available in data science.
3. Understand the fundamental algorithms available in Artificial Intelligence.
4. Understand the key algorithms available in data mining and machine learning.

Unit I- Introduction to Data Science
Foundation of Data science, Evolution of Data Science – Data Science Roles – Stages in a Data Science Project – Applications of Data Science in various fields – Data Security Issues. Area and Scope of Data Science, Steps of Data Science Process: Data collection, Pre- processing, Data Collection Strategies – Data Pre-Processing Overview – Data Cleaning – Data Integration and Transformation – Data Reduction – Data Discretization. Training and testing. Use cases in various domains such Image, Natural Language, Audio and Video.

Unit II- Introduction to Artificial Intelligence:
Introduction Artificial Intelligence, The Foundations of AI, AI Technique, Production system characteristics, Production systems: 8-puzzle problem. Searching: Uninformed search strategies – Breadth first search, depth first search.

Unit III-Searching Algorithms and Learning:

Local Search Algorithms: Generate and Test, Hill climbing, simulated annealing search, Constraint satisfaction problems, Greedy best first search, A* search, AO* search. Self-Learning: Propositional logic - syntax & semantics Game Playing: Overview, Minimax algorithm, Alpha-Beta pruning, Additional Refinements.

Model Development Simple and Multiple Regression – Model Evaluation using Visualization – Residual Plot – Distribution Plot – Polynomial Regression and Pipelines – Measures for In-sample Evaluation – Prediction and Decision Making.

UNIT-IV- Introduction to Data Mining and Machine Learning:

Introduction to Data Mining and Machine Learning, Supervised, Unsupervised and Reinforcement learning. Prediction vs Classification v/s Clustering. Association Rule Mining, classification and regression techniques,clustering, Scalability and data management issues in data mining algorithms, measures of interestingness. Model Evaluation Generalization Error – Out-of-Sample Evaluation Metrics – Cross Validation – Overfitting – Underfitting and Model Selection – Prediction by using Ridge Regression – Testing Multiple Parameters by using Grid Search.

Reference Books

1. Jojo Moolayil, "Smarter Decisions : The Intersection of IoT and Data Science", PACKT, 2016. 2. Cathy O'Neil and Rachel Schutt , "Doing Data Science", O'Reilly, 2015.
3. David Dietrich, Barry Heller, Beibei Yang, "Data Science and Big data Analytics", EMC 2013
4. Raj, Pethuru, "Handbook of Research on Cloud Infrastructures for Big Data Analytics", IGI Global

B.Sc.–II
SEMESTER-III
(Data Science)
Paper IV- Java Programming

Total Marks - 50		**Max. Time- 2 Hrs**
End Sem Exam- 40 Marks **Min. Passing Marks- 20**	**Internal Assessment- 20 Marks** **Min. Passing- 50 %**	**Credits- 2** **Lectures- 60**

Learning Objectives:

1. To understand the fundamental concepts of Java Programming.
2. To study the object oriented programming features of Java Programming.
3. To understand the concepts of exception and file handling in Java.
4. Understanding concepts of Applets and toolkits.

Learning Outcomes:

On successful completion of this course, the students are able:
1. To understand basic programming concepts in Java..
2. To describe OOPs features of Java programming language.
3. To handle exceptions and multithreading in Java.
4. Learning Advance Java concepts.

Unit I- Introduction to Java

History of Java, Features of Java, JDK Environment, Java Virtual Machine, Garbage Collection **Programming Concepts of Basic Java:** Identifiers and Keywords, Data Types in Java, Java coding Conventions, Expressions in Java, Control structures, decision making statements, Arrays and its methods

Unit II- Objects and Classes

Object Fundamentals, Pass by value, 'this' reference, Data Hiding and Encapsulation, Overloading, Overriding Constructors, Finalization, Subclasses (Inheritance), Relationship between super class object and subclass object, implicit subclass object to super class object Conversion, Dynamic method dispatch.

Language Features: Scope rules, Static data, Static methods, Static blocks, Modifiers of Class, Method, Data Members and Variable, Abstract Classes, Interfaces, Packages, Importing Packages and Classes, User define packages.

Unit III- Exception Handling & Multithreading

Types of Exceptions try, catch, finally, throws keywords, creating your own exception, exceptions and Inheritance Multithreading: Multithreading Concept, Thread Life Cycle, Creating multithreading Application, Thread Priorities, Thread synchronization.

UNIT-IV- Abstract Window Toolkit & Applets

Abstract Window Toolkit: Components and Graphics, Containers, Frames and Panels, Layout Managers-Border Layout, Flow Layout, Grid Layout, Card Layout, AWT all Components, Event Delegation Model, Event Source and Handlers, Event Categories, Listeners, Applets-Applet Life Cycle, Applet Context, Inter applet communication.

Recommended Books:

1) E. Balaguruswamy, "Programming with Java - A Primer", The Sun Microsystems Press, New Delhi, ISBN 81-265-0931-7

2) Peter Van der Linden, "Just Java", The Sun Microsystems Press, New Delhi, ISBN, 0130897930

3) Advance Java by Dr. Krishna Karoo and Vikas Chitte HSRA Publication

References:

1) Deitel and Deitel, "Java How to Program", Prentice Hall Upper Saddle River, New Jersey 07458 (US). ISBN 0-13-034151-7

2) Jerry R Jackson Alan L, "Java by Example 1.2", McClellan Publication

B.Sc.–II
SEMESTER-III
(Data Science)
Paper V- Operating System and Information Security

Total Marks - 50		**Max. Time- 2 Hrs**
End Sem Exam- 40 Marks **Min. Passing Marks- 20**	**Internal Assessment- 20 Marks** **Min. Passing- 50 %**	**Credits- 2** **Lectures- 60**

Learning Objectives:

1. To understand the fundamental concepts and techniques of Operating Systems.
2. To study the concepts in process management and concurrency control mechanisms.
3. To understand the concepts in memory management and deadlocks solutions.
4. To study on file management and storage structures

Learning Outcomes:

On successful completion of this course, the students are able:
1. To understand basic concepts of operating systems.
2. To describe process management, scheduling and concurrency control mechanisms.
3. To analyze memory management and deadlocks.
4. To compare various file systems with operating systems examples.

Unit I- Introduction:

1. Basic OS functions, resource abstraction, types of operating systems–programming systems, batch systems, time sharing systems; operating systems for personal computers & workstations, process control & real time systems.
2. **Operating System Organization** Processor and user modes, kernels, system calls and system courses.
3.**Process Management:** System view of the process and resources, process abstraction, process hierarchy, threads, threading issues, thread libraries; Process Scheduling, non-pre-emptive and preemptive scheduling algorithms; concurrent and processes, critical section, semaphores, methods for interprocess communication; deadlocks

Unit II- OS Management :

1.**Memory Management:** Physical and virtual address space; memory allocation strategies – fixed and variable partitions, paging, segmentation, virtual memory.

2.**File and I/O Management**: Directory structure, file operations, files allocation methods, device management.

3. **Protection and Security**: Policy mechanism, Authentication, Internal access Authorization.

Unit III-The Security Problem in Computing:

The meaning of computer Security, Computer Criminals, Methods of Defense, Elementary Cryptography: Substitution Ciphers, Transpositions, Making "Good" Encryption algorithms, The Data Encryption Standard, The AES Encryption Algorithms, Public Key Encryptions, Uses of Encryption. Program Security: Secure Programs, Nonmalicious Program Errors, viruses and other malicious code, Targeted Malicious code, controls Against Program Threats, Protection in General- Purpose operating system protected objects and methods of protection memory and addmens protection, File protection Mechanisms, User Authentication Designing Trusted O.S: Security policies, models of security, trusted O.S design, Assurance in trusted O.S. Implementation examples

UNIT-IV- Database Security:

Security requirements, Reliability and integrity, Sensitive data, Inference, multilevel database, proposals for multilevel security. Security in Network: Threats in Network, Network Security Controls, Firewalls, Intrusion Detection Systems,Secure EMail.

Administering Security: Security Planning, Risk Analysis, Organizational Security policies, Physical Security. Legal Privacy and Ethical Issues in Computer Security: Protecting Programs and data, Information and the law, Rights of Employees and Employers, Software failures, Computer Crime, Praia, Ethical issues in Computer Security, case studies of Ethics.

Recommended Books:

1. A Silberschatz, P.B. Galvin, G. Gagne, Operating Systems Concepts, 8th Edition, John Wiley Publications 2008.

2. A.S. Tanenbaum, Modern Operating Systems, 3rd Edition, Pearson Education 2007.

3. G. Nutt, Operating Systems: A Modern Perspective, 2nd Edition Pearson Education 1997.

4. W. Stallings, Operating Systems, Internals & Design Principles, 5th Edition, Prentice Hall of India. 2008.

5. M. Milenkovic, Operating Systems- Concepts and design, Tata McGraw Hill 1992.

B.Sc. II
For Semester III
Subject - Seminar -I

Total Marks - 50		Max. Time-
End Sem Exam- Min. Passing Marks-	Internal Assessment- 50 Marks Min. Passing- 50 %	Credits- 2 Lectures - 60

Note- Students must choose data science related topics from the Fundamentals in Data Science curriculum and should prepare a seminar presentation based on topic.

B.Sc.–II
SEMESTER-III
(Data Science)
Subject- Optimization Techniques

Total Marks - 50		**Max. Time- 2 Hrs**
End Sem Exam- **Min. Passing Marks-**	**Internal Assessment- 50 Marks** **Min. Passing- 50 %**	**Credits- 2** **Lectures- 60**

Learning Objectives:
The objective of this course is to teach various optimization techniques and highlight its usage in data science related applications.

Learning Outcomes:
After completion of this course successfully the students will be able to:
1. Understand the optimization problems.
2. Solve the optimization problems.
3. Understand the use of genetic algorithms for solving optimization problems.
4. Find the usage the optimization algorithms for data science tasks

Unit I- Basics of optimization :
Basics of optimization —how to formulate the problem, Maxima, minima, convex function, global solution Linear programming, simplex algorithm, Integer programming, Constraint programming, Knapsack Problem

Unit II- Randomized optimization :
Randomized optimization techniques—hill climbing, simulated annealing

Unit III- Introduction Genetic algorithms :
Foundation of Evolutionary theory, Evolutionary Strategies, Evolutionary programming,Evolutionary Algorithms, Evolutionary Algorithm Case Study, Genetic Algorithm, Genetic Representations, Initial Population, Fitness Function, Selection and Reproduction,

UNIT-IV- Genetic Operators :
Genetic Operators(Selection, Crossover, Mutation), Artificial Immune Systems, Other Algorithms Harmony Search, Honey-Bee Optimization, Memetic Algorithms, Co-evolution, MultiObjectiveOptimization, Artificial Life, Constraint Handling

Recommended Books:

1.Optimization Techniques Hardcover,New Age Science Ltd; 1st edition (30 April 2009) by Chander Mohan, Kusum Deep.

2. Optimization Techniques: An Introduction,L. R. Foulds,Springer-Verlag.

3. Optimization Techniques,Chander Mohan and Kusum Deep,New Age Science.

4. Genetic Algorithms in Search, Optimization & Machine Learning, David E. Goldberg,Pearson Education India; 1st edition(1 December 2008)

5. Genetic Algorithms: Concepts and Designs (Advanced Textbooks in Control and Signal Processing), KimFung Man, Kit-Sang Tang, Sam Kwong, Springer.

B.Sc.–II

SEMESTER-III

(Data Science)

Lab course - Web Application Development

Total Marks - 50		**Max. Time- 2 Hrs**
End Sem Exam- 30 Marks **Min. Passing Marks- 20**	**Internal Assessment- 20 Marks** **Min. Passing- 50 %**	**Credits- 2** **Practicals-16**

1. "Hello World" Component: Start with the basics. Create a simple React component that renders "Hello, World!"
2. Counter App: Build a counter application with buttons to increment and decrement the count.
3. Real-time Input: Create a form that captures user input and displays it in real-time as the user types.
4. List Component: Construct a component to display a list of items.
5. Toggle Switch: Implement a basic toggle switch component that changes its state when clicked.
6. API Data Fetch: Develop a component that fetches data from an API and displays it on the page.
7. Calculator App: Build a simple calculator application with basic arithmetic operations.
8. Timer: Create a timer that counts down from a specified time.
9. To-Do List: Develop a to-do list application with features to add and remove tasks.
10. Dynamic Background: Create a component that changes its background color when clicked.
11. Routing Setup: Set up a basic routing system using React Router.
12. Random Quote Generator: Build a component that displays a random quote each time it's rendered.
13. File Uploader: Develop a file uploader component that allows users to upload images.
14. Authentication Form: Implement a basic authentication form with login and registration.
15. Weather App: Build a weather app that fetches and displays weather information based on the user's location.

B.Sc.–II

SEMESTER-III

(Data Science)

Lab course - Java Programming-

Total Marks - 50		**Max. Time- 2 Hrs**
End Sem Exam- 30 Marks **Min. Passing Marks- 20**	**Internal Assessment- 20 Marks** **Min. Passing- 50 %**	**Credits- 2** **Practicals - 16**

Practical List for Java Programming

1. Write a program to accept two numbers and display the result using the command line argument.

2. Write a program for sorting a list of a number using Array.

3. Write a java program to print the following output.

```
A
A B
A B C
A B C D
A B C D E
```

4. Write a java program to no. of evens and no. odd numbers in an array of size 10. Also, calculate the sum of evens and the sum of odds.

5. Write a java program to find the sum of prime numbers ranging from 1 to 100.

6. Write a program to calculate multiplication and division using the static method.

7. Write a program of Constructor Overloading to calculate the area of the Room. i. Default constructor ii. Constructor with one argument. iii. Constructor with three arguments.

8. Write a program to demonstrate Single Inheritance.

9. Write a program to calculate the area of a rectangle and circle using Interface.

10. Write a java program to demonstrate the try...catch mechanism.

11. Write a java program to show the use of throw, throws, and finally keywords.

12. Write a program that throws IO Exceptions. (Accept student Name and age from keyboard and display.

13. Write a program to demonstrate user-defined exception (use division of two no's & throw user define exception if the result is smaller than 0.01)

14. Write a java program to demonstrate Threads using Thread class and also with Runnable interface.

15. Write a program to demonstrate the Linked List class.

B.Sc. - II (Data Science)
Semester – IV
Syllabus

B.Sc.–II
SEMESTER-IV
(Data Science)
Paper I- Machine Learning

Total Marks - 100		Max. Time- 3 Hrs
End Sem Exam- 80 Marks Min. Passing Marks- 40	Internal Assessment- 20 Marks Min. Passing- 50 %	Credits- 4 Lectures- 60

Learning Objectives:

The objective of this course is to introduce machine learning fundamentals to students. This course provides introductory concepts of various machine learning techniques to students which will help to build foundation for further understanding. This course also aims to provide details of various steps involved in the machine learning pipeline such as data collection, pre-processing, feature engineering etc. This course also introduces popular tools used in the area of machine learning.

Learning Outcomes:

After completion of this course successfully the students will be able to:

1. Understand the various processes involve in machine learning.
2. Perform data cleaning and preprocessing
3. Decide and classify the problem as classification, prediction or clustering
4. Train and test machine learning algorithms

Unit-I - Introduction to Machine Learning

Machine Learning Introduction, Deep Learning and Neural Networks , perceptrons, Relation in Artificial Intelligence and Machine Learning, ML languages, Applications of ML,

Supervised and Unsupervised Learning. Getting and Cleaning Data : Obtaining data from the web, from APIs, from databases and from colleagues in various formats. basics of data cleaning and making data —tidy. Data pre-processing : Descriptive Data Summarization, Data Cleaning, Data Integration andTransformation, Data Reduction, Data Discretization and Concept Hierarchy Generation.

Unit-II - ML Algorithms

Perceptons,Testing Pattern recognition, ML Terminologies, Linear regressions

Association Rule : Mining Frequent Patterns, Associations, and Correlations: Basic Concepts and a Road Map, Association Rules, theApriori Algorithm Classification and Prediction

Unit-III - ML Classifications

Classification : Classification, Issues Regarding Classification, Classification by Decision Tree Induction, Bayesian Classification,Rule-Based Classification, Metrics for Evaluating Classifier Performance, Holdout Method and Random Sub sampling

Unit-IV - ML Predictions and Clustering

Prediction : Prediction, Issues Regarding Prediction, Accuracy and Error Measures, Evaluating the Accuracy of a Classifier or Predictor.

Clustering : Cluster Analysis, Agglomerative versus Divisive Hierarchical Clustering,Distance Measures in Algorithmic, Evaluation of Clustering.

Tools and Frameworks : Scikit-learn, Weka and RStudio

Recommended Books:

1.Shalev-Shwartz, Shai, and Shai Ben-David. Understanding machine learning: From theory to algorithms. Cambridgeuniversity press, 2014.

2. Duda, Richard O., Peter E. Hart, and David G. Stork. Pattern classification. John Wiley & Sons, 2012.

3. Witten, Ian H., et al. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2016.

B.Sc.–II
SEMESTER-IV
(Data Science)
Paper II- Data Warehouse and Data Mining

Total Marks - 50		Max. Time- 2 Hrs
End Sem Exam- 40 Marks Min. Passing Marks- 20	Internal Assessment- 20 Marks Min. Passing- 50 %	Credits- 2 Lectures- 60

Learning Objectives:

1. To understand data warehouse concepts, architecture, business analysis and tools
2. To understand data pre-processing and data visualization techniques
3. To study algorithms for finding hidden and interesting patterns in data
4. To understand and apply various classification and clustering techniques using tools.

Learning Outcomes:

After completion of this course successfully the students will be able to:
1. Design a Data warehouse system and perform business analysis with OLAP tools.
2. Apply suitable pre-processing and visualization techniques for data analysis.
3. Apply frequent pattern and association rule mining techniques for data analysis.
4. Apply appropriate classification and clustering techniques for data analysis.

Unit-I - Data Warehouse Fundamentals :

Introduction to Data Warehouse, OLTP Systems; Differences between OLTP Systems and Data Warehouse: Characteristics of Data Warehouse; Functionality of Data Warehouse: Advantages and Applications of Data Warehouse; Advantages, Applications: Top - Down and Bottom-Up Development Methodology: Tools for Data warehouse development: Data Warehouse Types: **Planning and Requirements :** Introduction: Planning Data Warehouse and Key Issues: Planning and Project Management in constructing Datawarehouse: Data Warehouse Project; Data Warehouse Development Life Cycle, Kimball Lifecycle Diagram, Requirements Gathering Approaches: Team organization, Roles, and Responsibilities:
Data Warehouse Architecture : Introductions, Components of Data Warehouse Architecture: Technical Architectures; Data warehouse architectures 1: Data warehouse architecture 2: Data warehouse architecture 3: Tool selection: Federated Data Warehouse Architecture:

Unit-II - Dimensional Modeling :

Introduction: E-R Modeling: Dimensional Modeling: E-R Modeling VS Dimensional Modeling: Data Warehouse Schemas; Star Schema, Inside Dimensional Table, Inside Fact Table, Fact Less Fact Table, Granularity, Star Schema Keys: Snowflake Schema: Fact Constellation Schema: **Extract, Transform and Load :** Introduction: ETL Overview or Introduction to ETL: ETL requirements and steps: Data Extraction; Extraction Methods, Logical Extraction Methods, Physical Extraction Methods: Data Transformation; Basic Tasks in Transformation, Major Data Transformation Types: Data loading; Data Loading Techniques: ETL Tools:

Data Warehouse & OLAP : Introduction: What is OLAP?; Characteristics of OLAP, Steps in the OLAP Creation Process, Advantageous of OLAP: What is Multidimensional Data: OLAP Architectures; MOLAP,

Unit-III - Data Mining

Meta data Management in Data Warehouse : Introductions to Metadata: Categorizing Meta data: Metadata management in practice; Meta data requirements gathering, Meta data classification, Meta data collection strategies: Meta Data Management in Oracle and SAS: Tools for Meta data management:

Introduction to Data Mining : Introduction: Scope of Data Mining: What is Data Mining; How does DataMining Works, Predictive Modeling: Data Mining and Data Warehousing: Architecture for Data Mining: Profitable Applications: Data Mining Tools:

Business Intelligence : Introduction, Business Intelligence, Business Intelligence tools, Business Intelligence Infrastructure, Business Intelligence Applications, BI versus Data Warehouse, BI versus Data Mining, Future of BI.

Unit-IV - Data Pre-processing and Data mining techniques

Data Pre-processing : Introduction, Data Preprocessing Overview, Data Cleaning, Data Integration and Transformation, Data Reduction, Discretization and Concept Hierarchy Gene ration.

Data Mining Techniques- An Overview : Introduction, Data Mining, Data Mining Versus Database Management System, Data Mining Techniques- Association rules, Classification, Regression, Clustering, Neural networks.

Clustering : Introduction, Clustering, Cluster Analysis, Clustering Methods- K means, Hierarchical clustering, Agglomerative clustering, Divisive clustering, clustering and segmentation software, evaluating clusters.

Web Mining : Introduction, Terminologies, Categories of Web Mining – Web Content Mining, WebStructure Mining, Web Usage Mining, Applications of Web Mining, and Agent based and Database approaches, Web mining Software. Applications of Data mining

Recommended Books:

1. Alex Berson and Stephen J.Smith, —Data Warehousing, Data Mining & OLAP‖, Tata McGraw – HillEdition, 35th Reprint 2016.
2. K.P. Soman, Shyam Diwakar and V. Ajay, —Insight into Data Mining Theory and Practice, EasternEconomy Edition, Prentice Hall of India, 2006.
3. Ian H.Witten and Eibe Frank, —Data Mining: Practical Machine Learning Tools and Techniques, Elsevier,Second Edition.

B.Sc.–II
SEMESTER-IV
(Data Science)
Paper III- Internet on Things with Big data

Total Marks - 50		**Max. Time- 2 Hrs**
End Sem Exam- 40 Marks **Min. Passing Marks- 20**	**Internal Assessment- 20 Marks** **Min. Passing- 50 %**	**Credits- 2** **Lectures- 48**

Learning Objectives:
The objective of this course is to know and appreciate the software needs of an IoT project and understand how data is managed in an IoT network. This course also aims to explain how to apply software solutions for different systems and Big Data to IoT concept designs. This course focus on Python to write scripts to manage large data files collected from sensor data and interact with the real world via actuators and other output devices.

Learning Outcomes:

After completion of this course successfully the students will be able to:
1. Find the applications of IoT in real world and use techniques to build software.
2. Understand the IoT network and sensor data
3. Collect and analyze large data collect through various sensor
4. Use Python for building IoT and big data based applications.

Unit-I - Introduction to Big Data from the IoT :

Introduction to Big Data from the IoT : Develop an understanding of the data generated by IoT, and how it is collected; Recognise the problems involved with gathering data and some approaches for addressing these problems; Gain an overview of data storage

Unit-II - Data at the Edge :
Data at the Edge : Understand the process of data acquisition; Be able to analyse where to process data using Edge, Fog or Cloud; Understand how, when, and where to bundle and store IoT data

Unit-III - Data in the Cloud :

Data in the Cloud : Understand the storage, analysis and cleaning of data; Understand why data is stored and processed in the Cloud; Appreciate the costs and benefits of live data versus stored data.

Unit-IV - Obtaining, Visualising and Analyzing Data :

Obtaining, Visualising and Analyzing Data : Understand some methods for cleaning, summarizing and visualizing a large dataset; Construct and use a simple predictive model for predicting the location of a device using signal strength and orientation.

Learn how to use Python, R and RStudio to performance analysis of a large dataset;

Case studies and projects

Recommended Books:

1. Internet of Things – A Hands-on Approach, Arshdeep Bahga and Vijay Madisetti, Universities Press, 2015, ISBN: 9788173719547

2. Getting Started with Raspberry Pi, Matt Richardson & Shawn Wallace, O'Reilly (SPD), 2014, ISBN: 9789350239759

B.Sc.–II
SEMESTER-IV
(Data Science)
Paper IV- Deep Learning

Total Marks - 50		**Max. Time- 2 Hrs**
End Sem Exam- 40 Marks **Min. Passing Marks- 20**	**Internal Assessment- 20 Marks** **Min. Passing- 50 %**	**Credits- 2** **Lectures- 60**

Learning Objectives:
The objective of this course is to provide advanced knowledge of machine learning techniques. This course mainly focused on Regression and Neural network based Machine learning algorithms. This aims to make students aware of various recent developments in the field of Neural networks such as deep learning.

Learning Outcomes:
After completion of this course successfully the students will be able to:
1. Perform regression analysis
2. Use to use Neural Network based model for classification and other task
3. Used to train and test deep learning based models for various tasks.
4. Use Python for building Deep learning based applications

Unit-I - Linear and Logistic Regression
Linear Regression : Prediction using Linear Regression, Gradient Descent, Linear Regression with one variable, Linear Regression with multiple variables, Polynomial Regression, Feature Scaling/Selection.
Logistic Regression : Classification using Logistic Regression, Logistic Regression vs. Linear Regression, Logistic Regression with one variable and with multiple variables.

Unit-II - Regularization
Regularization : Regularization and its utility: The problem of Overfitting, Application of Regularization in Linear and Logistic Regression, Regularization and Bias/Variance.

Unit-III - Neural Networks
Neural Networks : Introduction, Model Representation, Gradient Descent vs. Perceptron Training, Stochastic Gradient Descent, Multilayer Perceptrons, Multiclass Representation, Backpropagation Algorithm.

Unit-IV - Deep Learning

Deep Learning : History, Scope and specification, why deep learning now, building block of neural network, neural networks, Deep learning hardware. Feedforward neural networks, xor model, cost function estimation (maximum likelihood), units, activation functions, layers, , normalization, hyper-parameter tuning, Convolution neural networks, architecture, recurrent neural networks, architecture, types and overview, GAN (Generative Adversarial Networks). **Deep learning applications :** Computer vision, sentiment analysis, music generation, text generation, neural style transfer, image captioning

Recommended Books:

1. Ethem Alpaydin, "Introduction to Machine Learning" 2nd Edition, The MIT Press, 2009.
2. Tom M. Mitchell, "Machine Learning", First Edition by Tata McGraw-Hill Education, 2013.
3. Christopher M. Bishop, "Pattern Recognition and Machine Learning" by Springer, 2007.
4. Mevin P. Murphy, "Machine Learning: A Probabilistic Perspective" by The MIT Press, 2012.

B.Sc.–II
SEMESTER-IV
(Data Science)
Paper V- Data Analytics and Visualization

Total Marks - 50		**Max. Time- 2 Hrs**
End Sem Exam- 40 Marks **Min. Passing Marks- 20**	**Internal Assessment- 20 Marks** **Min. Passing- 50 %**	**Credits- 2** **Lectures- 60**

Learning Objectives:

1. Make students understand commonly used terms and techniques related to data analytics that are in use today.
2. To discuss and train students on how data and data analytics can be used by managers to make better decisions.
3. Have the student gain perspective and practice by applying data analysis techniques in several settings.
4. Make students learn to use Python and its libraries to perform various Data Analytic tasks

Learning Outcomes:

After completion of this course successfully the students will be able to:
1. Understand the need and importance of decision making in business, its inherent difficulties and pitfalls,
2. Learn the importance of proper data analysis in decision making.
3. Understand how the data environment in various domain is changing
4. Apply common quantitative and visual techniques to enhance decision making.
5. Use Python to analyze data and provide useful information for decision making.
6. Tableau course syllabus will help you to become master in Business Intelligence (BI) tool, Data Visualization, reporting and SQL with real-life industry Projects in Health care, Retail and Banking domains.Latest Tableau course content to pass Tableau Desktop, Analyst and Server certification exams.

Unit-I - Data Definitions and Analysis Techniques:
Data Definitions and Analysis Techniques: Elements, Variables, and Data categorization, Levels of Measurement, Data management and indexing,Introduction to statistical learning.
Basic analysis techniques: Statistical hypothesis generation and testing, Chi-Square test , t-Test , Analysis of variance , Correlation analysis ,Maximum likelihood test.

Data analysis techniques: Regression analysis , Classification techniques , Clustering , Association rules analysis

Unit-II - Tableau Basics :

Learn Tableau Basic Reports Parameters, Grouping Example 1, Grouping Example 2, Edit Groups, Set, Combined Sets, Creating a First Report, Data Labels, Create Folders, Sorting Data, Add Totals, Sub Totals and Grand Totals to Report.

Learn Tableau Charts Area Chart, Bar Chart, Box Plot, Bubble Chart, Bump Chart, Bullet Graph, Circle Views, Dual Combination Chart, Dual Lines Chart, Funnel Chart, Traditional Funnel Charts, Gantt Chart, Grouped Bar or Side by Side Bars Chart, Heatmap, Highlight Table, Histogram, Cumulative Histogram, Line Chart, Lollipop Chart, Pareto Chart, Pie Chart, Scatter Plot, Stacked Bar Chart, Text Label, Tree Map, Word Cloud, Waterfall Chart, Geographic map, Filled map, Crosstab, Combines axis, Motion chart, Reference lines

Unit-III - Tableau Advanced Reports:

SQL Convert to Custom SQL, Learn Tableau Advanced Reports: Dual Axis Reports, Blended Axis, Individual Axis, Add Reference Lines, Reference Bands, Reference Distributions, Basic Maps, Symbol Map, Use Google Maps, Mapbox Maps as a Background Map, WMS Server Map as a Background Map

Learn Tableau Calculations & Filters Calculated Fields, Basic Approach to Calculate Rank, Advanced Approach to Calculate Rank, Calculating Running Total, Filters Introduction, Quick Filters, Filters on Dimensions, Conditional Filters, Top and Bottom Filters, Filters on Measures, Context Filters, Slicing Filters, Data Source Filters, Extract Filters.

Unit-IV - Tableau dashboards and UI

Learn Tableau Dashboards

Create a Dashboard, Format Dashboard Layout; Create a Device Preview of a Dashboard, Create Filters on Dashboard, Dashboard Objects, and Create a Story.

Tableau data server: Physical architecture overview, User access, Component functions & processes, Tableau server on-premises, Tableau reader, Tableau online v tableau server. Tableau **Server UI:** Tableau server user interface, Users, Site roles, Groups, Schedules, Tasks, Tableau server menu, Content display options.

Recommended Books:

1. All of statistics: a concise course in statistical inference. Larry Wasserman. Springer, 2004.
2. C. Bishop, Pattern Recognition and Machine Learning, Springer 2007
3. Hastie, Trevor, et al. The elements of statistical learning. Vol. 2. No. 1. New York: springer, 2009.

B.Sc. II
For Semester IV
Subject - Mini Project based on data Science

Total Marks - 50		Max. Time-
End Sem Exam- Min. Passing Marks-	Internal Assessment- 50 Marks Min. Passing- 50 %	Credits- 2 Lectures - 60

Note- Students must choose data science related topics from the Fundamentals in Data Science and Python Programming curriculum and should prepare a mini project presentation based on the topic.

B.Sc.–II

SEMESTER-IV

(Data Science)

Lab course- Data warehouse and Data Mining

Total Marks - 50		**Max. Time-**
End Sem Exam- 30 Marks **Min. Passing Marks- 20**	**Internal Assessment- 20 Marks** **Min. Passing- 50 %**	**Credits- 2** **Practicals-16**

List Of Practicals:

I. Build Data Warehouse and Explore WEKA 2
II. Perform data preprocessing tasks and Demonstrate performing association rule mining on data sets 3
III. Demonstrate performing classification on data sets 4
IV. Demonstrate performing clustering on data sets 5
V. Demonstrate performing Regression on data sets 6

Task 1: Credit Risk Assessment. Sample Programs using German Credit Data 7

Task 2: Sample Programs using Hospital Management System 8 Beyond the Syllabus -Simple Project on Data Preprocessing

A. Build Data Warehouse/Data Mart (using open source tools like Pentaho Data Integration Tool, Pentaho Business Analytics; or other data warehouse tools like Microsoft-SSIS, Informatica,Business Objects,etc.,)

A.(i) Identify source tables and populate sample data. The data warehouse contains 4 tables:

1. Date dimension: contains every single date from 2006 to 2016.

2. Customer dimension: contains 100 customers. To be simple we'll make it type 1 so we don't create a new row for each change.

3. Van dimension: contains 20 vans. To be simple we'll make it type 1 so we don't create a new row for each change.

4. Hire fact table: contains 1000 hire transactions since 1st Jan 2011. It is a daily snapshot fact table so that every day we insert 1000 rows into this fact table. So over time we can track the changes of total bill, van charges, satnav income, etc. Create the source tables and populate them So now we are going to create the 3 tables in HireBase database: Customer, Van, and Hire. Then we populate them.

B.Sc.–II
SEMESTER-IV
(Data Science)
Lab course- Machine Learning

Total Marks - 50		**Max. Time-**
End Sem Exam- 30 Marks **Min. Passing Marks- 20**	**Internal Assessment- 20 Marks** **Min. Passing- 50 %**	**Credits- 2** **Practicals - 16**

List Of Practicals:

1. Write a program to implement the naïve Bayesian classifier for a sample training data set stored as a .CSV file. Compute the accuracy of the classifier, considering few test data sets.

2. Assuming a set of documents that need to be classified, use the naïve Bayesian algorithm.

3. Classifier model to perform this task. Built-in Java classes/API can be used to write the program. Calculate the accuracy, precision, and recall for your data set.

4. Write a program to implement k-Nearest Neighbour algorithm to classify the iris. print both correct and wrong predictions. Java/Python ML library classes can be used for this problem.

5. Write a program to implement a Logistic Regression algorithm to classify the housing price data set. Print both correct and wrong predictions. Java/Python ML library classes can be used for this problem.

6. Write a program to implement and compare SVM, KNN and Logistic regression algorithms to classify the iPhone purchase records data set. Print both correct and wrong predictions. Java/ Python ML library classes can be used for this problem.

B.Sc.–II
SEMESTER-IV
(Data Science)
Lab course- Deep Learning

Total Marks - 50		**Max. Time- 2 Hrs**
End Sem Exam- 30 Marks **Min. Passing Marks- 20**	**Internal Assessment- 20 Marks** **Min. Passing- 50 %**	**Credits- 2** **Practicals - 16**

LIST OF EXERCISES:

1. Setting up the Spyder IDE Environment and Executing a Python Program
2. Installing Keras, Tensorflow and Pytorch libraries and making use of them
3. Artificial Neural Networks
4. Convolutional Neural Networks
5. Image Transformations
6. Image Gradients and Edge Detection
7. Image Contours
8. Image Segmentation
9. Harris Corner Detection
10. Face Detection using Haar Cascades
11. Chatbot Creation