HINDI MAHAVIDYALAYA

(AUTONOMOUS & NAAC RE-ACCREDITED)

(Affiliated to Osmania University)

Nallakunta, Hyderabad, Telangana

BOARD OF STUDIES

DEPARTMENT OF COMPUTER SCIENCE & APPLICATIONS

B.Sc. (DATA SCIENCE) I/II/III Year

WITH EFFECT FROM

ACADEMIC YEAR 2024 - 25

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)
BOARD OF STUDIES
Department of Computer Science & Applications

Chairperson:
Mr. Avinash Pal Lidlaan
Head of department in Computer Science & Applications
Hindi Mahavidyalaya, Nallakunta, Hyderabad

University Nominee
Professor K. Shyamala
Department of Computer Science
Osmania University-Hyderabad

Members of BOS:

1. Professor P.V. Sudha — Subject Expert
   Department of Computer Science & Engg.
   Osmania University-Hyderabad
2. Mrs.B Ramani — Member of BOS
   Department of Computer Science
   Andhra Mahila Sabha Arts and science college
   Osmania University-Hyderabad
3. Dr. Raghunadh Acharya — Member of BOS
   Professor in Department of Statistics
   Hindi Mahavidyalaya
   Nallakunta, Hyderabad
4. Mr. Aravind Sharma — Industry Expert

PRINCIPAL
HINDI MAHA VIDYALAYA
(AUTONOMOUS)
Arts, Commerce & Science
Nallakunta, Hyderabad-44.

Chairperson

Department of Computer Science
Hindi Mahavidyalaya
(AUTONOMOUS & NAAC REACCREDITED)
Nallakunta, Hyderabad-44.

University Nominee

CHAIRMAN
Board of Studies in CSE (A)
Dept. of Computer Science & Engg.
College Of Engg., O.U. Hyderabad.

Members

1. P.V. Sudha
2.
3.

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

Principal

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)
DEPARTMENT OF COMPUTER SCIENCE & APPLICATIONS
AGENDA OF THE MEETING

1 Welcome address by the chair.

2 Previous Meeting Details.

3 Details of choice-based credit system.

4 Discussion and Distribution of Common Core Syllabus for all the Semester.

5 Marks allotted for internal and end semester exams.

6 Discussion on Pattern and model paper of Semester Exam and internal exam for all the Semester.

7 Discussion on Practical exam model paper for all the Semester.

8 Panel of Examiners

9 Any other matter

10 Vote of thanks

PRINCIPAL
HINDI MAHA VIDYALAYA
(AUTONOMOUS)
Arts, Commerce & Scienc
Nallakunta, Hyderabad-4

Chairperson **University Nominee** **Members** **Principal**

Department of Computer Science
Hindi Mahavidyalaya
(AUTONOMOUS & NAAC REACCREDITED)
Nallakunta, Hyderabad-44

1. P.V. Sudha
PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

2.

3.

CHAIRMAN
Board of Studies in CSE (A)
Dept. of Computer Science & Engg.
College Of Engg., O.U. Hyderabad

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)
DEPARTMENT OF COMPUTER SCIENCE & APPLICATIONS
BOARD OF STUDIES
Academic Year – 2024-2025
Minutes of BOS Meeting

BOS meeting of the Department of Computer science was ~~held on~~ ..... by circulation
The following members were present.

CHAIRMAN
Board of Studies in CSE (A)
Dept. of Computer Science & Engg.
College Of Engg., O.U. Hyderabad.

Prof. K. Shyamala	-	Ex-Officio member-BOS
Prof.P.V.Sudha	-	Subject Expert, O.U
Smt B.Ramani	-	Member of BOS
Mr.Aravind.Sharma	-	Member of BOS
Dr. Raghunadh Acharya	-	Member of BOS
Mr.Avinash Pal Lidlaan	-	Member of BOS

P.V. Sudha
PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

Department of Computer Science
Hindi Mahavidyalaya
(AUTONOMOUS & NAAC REACCREDITED)
Nallakunta, Hyderabad-44.

4.1 Welcome address by the chair

The chair welcomed the University Nominee, Chairperson BOS, O.U. Department of Computer Science and Member of B.O.S

4.2 Previous Meeting details

The CBCS system has been introduced by Osmania University from 2020-21. The theory and practical syllabus of I, II & III years of B.Sc., question paper pattern for theory and practical, internal assessment pattern, practical examination scheme and panel of examiners were discussed and approved by all the BOS Members in previous BOS meeting.

4.3 Details of choice-based credit system.

Members were informed that TSCHE has referred that from the academic year 2020-21 autonomous institutions have to follow CBCS i.e. From the Academic Year 2020-21 Osmania University has instructed all the Degree colleges including Autonomous Degree colleges to follow CBCS under which after passing the exam student will get the Grade in the Result. B.Sc. I YEAR SEM I & II- AND II-YEAR SEM III & IV 4 credits are given for the theory and 1 credit for practical in each semester and III YEAR in V and VI semester 3 credits are given for theory paper and 1 credit is given for practical in each semester.

4.4 Discussion and Distribution of Common Core Syllabus for semester I, II and III.

1. Members were informed by the chair that Department of Computer Science, Hindi Mahavidyalaya is following common core syllabus prescribed by Osmania University B.Sc. III YEARS in all the semesters.
2. The syllabus comprises of 4 units.
3. Syllabus copy for all the semesters is enclosed.
4. Syllabus was approved by the Members of BOS.

4.5 Marks allotted for Internal and end Semester exams.

1. Internal assessment is of 30 marks and 5 marks assignment ,5 marks seminar where students have to answer 20 MCQs in 25 minutes. Each question carries 1 mark. In each Semester two internal assessments of 20 Marks will be conducted and an average of both the internal assessments will be added in the marks of theory exam.
2. Theory Question paper is of 70 marks.
3. Total allotted marks are 100 for each theory paper DSC/DSE (A&B).

The distribution of marks was approved by the Members of BOS.

4.6 Discussion on Pattern and Model Paper of Semester exam and Model Paper of Internal Exam

1. It was informed by the department that in each Semester Two Internal exams will be conducted for 20 marks. The internal assessment will have three sections.

Section – A 20 Multiple choice questions each carries 1 marks (20*1 =20M),
Section – B Assignment – 5 Marks
Section – C Seminar – 5 Marks

Average marks of these two internal exams will be taken.

2. Semester exam will be conducted as per the Almanac which will be provided by the exam branch. Internal exam duration will be 25 Min and Semester exam duration will be of 2½ hrs.

3. Model Question paper for all the Semesters was discussed. Theory paper for each Semester will have 2 sections.

i) Section A contains 8 short Questions. The student must answer six questions.
Each Question carries 3 Marks (6X3=18 Marks)
ii) Section B contains 4 Essay type Questions with internal choice. Each Question carries 13 Marks (4X13=52 Marks)
ii) Section B contains 2 Essay type Questions with internal choice. Each Question carries 15 Marks (2X10=20 Marks)

- Pattern of Model Theory Question Papers for DSC Paper III and Paper IV are enclosed.
- Pattern of Model Theory Question Papers for DSC was approved by Member of BOS

4. It was informed by the department that in each semester one internal exam will be conducted for AECC of 15 marks. The internal assessment will have two sections
   i) Section A 10 mcqs each carries 1 mark(10*1=10m)
   ii)section B Assignment/seminar- 5 marks

5. Model question paper of AECC for Semester I and semester II was discussed.Theory paper for each AECC will have 2 sections
   i)Section A contains 4 short Questions.The student has to answer THREE questions.Each question carries 5marks(3*5=15marks)
   ii) Section B contains 2 Essay type Questions with internal choice.Each question carries 10marks(2*10=20marks)

6. It was informed by the department that in each semester one internal exam will be conducted for SEC of 15 marks. The internal assessment will have two sections
   i) Section A 10 mcqs each carries 1mark(10*1=10m)
   ii)section B Assignment/seminar-5marks

7. Model question paper of SEC for Semester III and semester IV was discussed.Theory paper for each SEC will have 2 sections
   i)Section A contains 4 short Questions.The student has to answer THREE questions.Each question carries 5marks(3*5=15marks)
   ii) Section B contains 2 Essay type Questions with internal choice.Each question carries 10marks(2*10=20marks)

4.7 Discussion on Practical Exam Model paper.

It was decided in BOS meeting that 50 Marks Practical Exam of 2 hrs will be held in each Semester's and 1 credit will be given for Practical in each Semester.

- Pattern of Model Practical Question Papers for each Paper are enclosed.
- Pattern of Model Practical Question Papers was approved by Members of BOS

4.8 Panel of Examiners

The panel of examiners was approved by the members.

- List is enclosed

4.9 Any other matter.

4.10 Vote of Thanks

Meeting concluded with the Vote of Thanks by Mr. Avinash Pal Lidlaan

PRINCIPAL
HINDI MAHA VIDYALAYA
(AUTONOMOUS)
Arts, Commerce & Science
Nallakunta, Hyderabad

Chairperson	University Nominee	Members	Principal
Department of Computer Science Hindi Mahavidyalaya (AUTONOMOUS & NAAC REACCREDITED) Nallakunta, Hyderabad-44	CHAIRMAN Board of Studies in CSE (A) Dept. of Computer Science & Engg. College Of Engg., O.U. Hyderabad.	1. P.V. Sudha PROFESSOR 2. Department of Computer Science & Engineering University College of Engineering (A) 3. Osmania University Hyderabad-500 007.	

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)
DEPARTMENT OF COMPUTER SCIENCE & APPLICATIONS
Academic Year – 2024-2025
B.Sc. (Data Science) with Mathematics Combination
CBCS PATTERN IN SEMESTER SYSTEM-2020-2021

Course Title (B.Sc - Data Science)	Hours/Week		Credits	Marks
	Theory	Practical		
Semester –I				
Problem solving and Python Programming	4	2	4+1	100
Certification Course: Foundations of Data Science	2			50
Semester –II				
Data Base Management Systems (DBMS)	4	2	4+1	100
Fundamentals of Computers (AECC - II)	2		2	50
Certification Course: Data Mining Specialization	2			50
Semester –III				
Communication skills/Professional skills (SEC-I)	2		2	50
Mini Project (SEC - II)	2		2	50
Data Engineering with Python	4	2	4+1	100
Certification Course: Analytics with Excel	2			50
Semester –IV				
Leadership and management skills/universal human values (SEC- III)	2		2	50
Mini Project (SEC - IV)	2		2	50
Machine Learning	4	2	4+1	100
Certification Course: Data Science With R	2			50
Semester –V				
A. Natural Language Processing	4	2	4+1	100
B. No SQL Data Bases	4	2	4+1	100
Data Structures and Algorithms (GE)	4		4	100
Certification Course: Data Science with Machine Learning	2			50
Semester –VI				
A. Big Data	4	2	4+1	100
B. Deep Learning	4	2	4+1	100
Major Project	4		4	100
Certification Course: Data Engineering	2			50

Total

Chairperson
Department of Computer Science
Hindi Mahavidyalaya
(AUTONOMOUS & NAAC REACCREDITED)
Nallakunta, Hyderabad-44.

University Nominee

CHAIRMAN
Board of Studies in CSE (A)
Dept. of Computer Science & Engg.
College Of Engg., O.U. Hyderabad.

Members
1. P.V. Subba
2.
3.

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

PRINCIPAL
Principal
HINDI MAHA VIDYALAYA
(AUTONOMOUS)
Arts, Commerce & Science
Nallakunta, Hyderabad

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)
DEPARTMENT OF COMPUTER SCIENCE & APPLICATIONS

Academic Year – 2024-2025

B.Sc. (Artificial Intelligence and Machine Learning)

with Mathematics & Statistics Combination CBCS PATTERN IN SEMESTER SYSTEM-2022-2023

Course Title (B.Sc – AI & ML) Mathematical papers	Hours/Week		Credits	Marks
	Theory	Tutorial		
Semester –I				
Differential & Integral Calculus	5	1	4+1	100
Semester –II				
Differential Equations	5	1	4+1	100
Semester –III				
Communication skills/Professional skills **(SEC-I)**	2		2	50
(A) Theory of Equations (OR) **(B)** Logic & Sets **(SEC II)**	2		2	50
Real Analysis	5	1	4+1	100
Semester –IV				
Leadership and management skills/universal human values **(SEC III)**	2		2	50
(A) Number Theory (OR) **(B)** Vector Calculus **(SEC IV)**	2		2	50
Algebra	5	1	4+1	100
Semester –V				
Linear Algebra	5	1	4+1	100
(A) Basic Mathematics (OR) **(B)** Mathematics for Economics & Finance **(GE)**	4		4	100
Semester –VI				
A. Numerical Analysis	5	1	4+1	100
B. Integral Transforms	5	1	4+1	100
C. Analytical Solid Geometry	5	1	4+1	100
Major Project (Mathematical Modelling)	4		4	100

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)
DEPARTMENT OF COMPUTER SCIENCE & APPLICATIONS

Academic Year – 2024-2025

B.Sc. (Artificial Intelligence and Machine Learning)

with Mathematics & Statistics Combination CBCS PATTERN IN SEMESTER SYSTEM-2022-2023

Course Title (B.Sc – AI & ML) Statistics papers	Hours/Week		Credits	Marks
	Theory	Practical		
Semester –I				
Descriptive Statistics and Probability	4	3	4+1	100
Semester –II				
Probability Distributions	4	3	4+1	100
Semester –III				
Communication skills/Professional skills (SEC-I)	2		2	50
Data Collection, Presentation and Interpretation (SEC II)	2		2	50
Statistical Methods and Theory of Estimation	4	3	4+1	100
Semester –IV				
Leadership and management skills/universal human values (SEC III)	2		2	50
Data Scaling Techniques and Report writing (SEC IV)	2		2	50
Statistical Inference	4	3	4+1	100
Semester –V				
Applied statistics I	3	2	3+1	75T + 25P
A. Statistical Quality Control and Reliability	3	2	3+1	75T + 25P
B. Bio-Statistics - I	3	2	3+1	75T + 25P
C. Actuarial Statistics - I	3	2	3+1	75T + 25P
Basic Statistics-1 **(GE)**	4		4	100
Semester –VI				
Applied statistics II	3	2	3+1	75T + 25P
A. Operations Research	3	2	3+1	75T + 25P
B. Bio-Statistics - II	3	2	3+1	75T + 25P
C. Actuarial Statistics - II	3	2	3+1	75T + 25P
Major Project	4		4	100

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)
DEPARTMENT OF COMPUTER SCIENCE
Academic Year – 2024-2025
B.Sc. (Data Science) with Mathematics Combination
CBCS PATTERN IN SEMESTER SYSTEM-2020-2021

Ist YEAR SEMESTER –I					Semester end exams		Continuous internal evaluation		TOTAL	Practical 2hrs
Code	Course title	Course type	HPW	Credits	Duration in hrs	Marks	Exam duration	Marks		
	Problem solving and Python Programming	DSC-3C	4T+3P	4+1=5	2½	70	25MIN	30	100	50
	Fundamentals of Computers	AECC - I	2T	2	1½	35	15MIN	15	50	-

Ist YEAR SEMESTER –II					Semester end exams		Continuous internal evaluation		TOTAL	Practical 2hrs
Code	Course title	Course type	HPW	Credits	Duration in hrs	Marks	Exam duration	Marks		
	Data Base Management Systems	DSC-3C	4T+3P	4+1=5	2½	70	25MIN	30	100	50
	Python II	AECC - II	2T	2	1½	35	15MIN	15	50	-

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)
DEPARTMENT OF COMPUTER SCIENCE
Academic Year – 2024-2025
B.Sc. (Data Science) with Mathematics Combination
CBCS PATTERN IN SEMESTER SYSTEM-2020-2021

IInd YEAR SEMESTER –III					Semester end exams		Continuous internal evaluation		TOTAL	Practical 2hrs
Code	Course title	Course type	HPW	Credits	Duration in hrs	Marks	Exam duration	Marks		
	Data Engineering with Python	DSC-3C	4T+3P	4+1=5	2½	70	25MIN	30	100	50
	Mini Project (SEC - II)	SEC - II	2T	2	1½	35	15MIN	15	50	-

IInd YEAR SEMESTER –IV					Semester end exams		Continuous internal evaluation		TOTAL	Practical 2hrs
Code	Course title	Course type	HPW	Credits	Duration in hrs	Marks	Exam duration	Marks		
	Machine Learning	DSC-3C	4T+3P	4+1=5	2½	70	25MIN	30	100	50
	Mini Project (SEC - IV)	SEC - IV	2T	2	1½	35	15MIN	15	50	-

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)
DEPARTMENT OF COMPUTER SCIENCE
Academic Year – 2024-2025
B.Sc. (Data Science) with Mathematics Combination
CBCS PATTERN IN SEMESTER SYSTEM-2020-2021

IIIrd YEAR SEMESTER –V					Semester end exams		Continuous internal evaluation		TOTAL	Practical 2hrs
Code	Course title	Course type	HPW	Credits	Duration in hrs	Marks	Exam duration	Marks		
	A. Natural Language Processing B. No SQL Data Bases	DSC-3C	4T+3P	4+1=5	2½	70	25MIN	30	100	50
	Data Structures and Algorithms	GE	4T	4	2½	70	25MIN	30	100	-

IIIrd YEAR SEMESTER –VI					Semester end exams		Continuous internal evaluation		TOTAL	Practical 2hrs
Code	Course title	Course type	HPW	Credits	Duration in hrs	Marks	Exam duration	Marks		
	A. Big Data B. Deep Learning	DSC-3C	4T+3P	4+1=5	2½	70	25MIN	30	100	50
	Major Project	Project	4T	4	-	60	15MIN	15	100	25

CERTIFICATION COURSES

Semester	Theory/Practical	Paper Title	Instruction Hrs/Week	Continuous Evaluation	Total Marks	Credits
				IA and Assign.		
I	Certification Course	Foundations of Data Science	2	I:15 + E:35	50	-
II	Certification Course	Data Mining Specialization	2	I:15 + E:35	50	-
III	Certification Course	Analytics with Excel	2	I:15 + E:35	50	-
IV	Certification Course	Data Science With R	2	I:15 + E:35	50	-
V	Certification Course	Data Science with Machine Learning	2	I:15 + E:35	50	-
VI	Certification Course	Data Engineering	2	I:15 + E:35	50	-

I: Internal exam E: External exam

SEMESTER WISE ADDON COURSES

Semester – I

Foundations of Data Science (40 hrs)

Objective: To acquaint the participants with basics of Data Science

- Introduction to data science concepts
- impact of data
- data applications and workflow
- career and data professional
- course project

Semester – II

Data Mining Specialization (40 hrs)

Objective: To acquaint the participants with data mining concept.

- Data Visualization
- Text Retrieval and Search Engines
- Text Mining and Analytics
- Pattern Discovery in Data Mining
- Cluster Analysis in Data Mining
- Data Mining Project

Semester – III

Analytics with Excel (40 hrs)

Objective: To acquaint the participants, giving an understanding and hands on experince to use excel.

- Introduction to Data Analytics
- Excel Basics for Data Analysis
- Data Visualization and Dashboards with Excel and Cognos
- Assessment for Data Analysis and Visualization Foundations
- course project

Semester – IV

Data Science with R (40 hrs)

Objective: To acquaint the participants with understanding and hands on experince on R Programming.

- Introduction to R Programming for Data Science
- SQL for Data Science with R
- Data Analysis with R
- Data Visualization with R
- Data Science with R - Project

Semester – V

Data Science with Machine Learning (40 hrs)

Objective: To acquaint the participants providing AI concept using data.

- Data Analysis with Python
- Working with different types of Datasets
- DIY Datasets with Web Scraping
- Dashboards with powerBI
- Building ML Models from Scratch
- NLP with NLTK
- Neural Networks with TensorFlow
- Working with Images with OpenCV
- Projects

Semester – VI

Data Engineering (40 hrs)

Objective: To acquaint the participants providing advanced content of R .

- R programming
- Probability, inference, and modeling
- Tidyverse, including data visualization with ggplot2 and data wrangling with dplyr
- Unix/Linux, git and GitHub, and RStudio
- Implement machine learning algorithms.
- Fundamental data science concepts with real-world case studies

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD

(AUTONOMOUS)

B.Sc. (Data Science) – I YEAR SEMESTER- I

Paper-I: Problem Solving and Python Programming

HPW- 4T+2P **Credits – 4** **Marks – 70 + 30**

Objectives

The main objective is to teach Computational thinking using Python.

- To know the basics of Programming
- To convert an algorithm into a Python program
- To construct Python programs with control structures.
- To structure a Python Program as a set of functions
- To use Python data structures-lists, tuples, dictionaries.
- To do input/output with files in Python.
- To construct Python programs as a set of objects.

Outcomes:

On completion of the course, students will be able to:

1. Develop algorithmic solutions to simple computational problems.
2. Develop and execute simple Python programs.
3. Develop simple Python programs for solving problems.
4. Structure a Python program into functions.
5. Represent compound data using Python lists, tuples, and dictionaries.
6. Read and write data from/to files in Python Programs

Unit-I

Introduction to Computing and Problem Solving: Fundamentals of Computing – Computing Devices – Identification of Computational Problems – Pseudo Code and Flowcharts – Instructions – Algorithms – Building Blocks of Algorithms. Introduction to Python Programming: Python Interpreter and Interactive Mode– Variables and Identifiers – Arithmetic Operators – Values and Types – Statements, Reading Input, Print Output, Type Conversions, The type() Function and Is Operator, Dynamic and Strongly Typed Language. Control Flow Statements: The if, the if...else, the if...elif...else Decision Control Statements, Nested if Statement, The while Loop, the for Loop, The continue and break Statements.

Unit-II

Functions: Built-In Functions, Commonly Used Modules, Function Definition and Calling the Function, The return Statement and void Function, Scope and Lifetime of Variables, Default Parameters, Keyword Arguments, *args and **kwargs, Command Line Arguments. Strings: Creating and Storing Strings, Basic String Operations, Accessing Characters in String by Index Number, String Slicing and Joining, String Methods, Formatting Strings.

Unit-III

Lists: list operations, list slices, list methods, list loop, mutability, aliasing, cloning lists, list parameters; Tuples: tuple assignment, tuple as return value; Dictionaries: operations and methods; advanced list processing - list comprehension; Illustrative programs: selection sort, insertion sort, merge sort, histogram. Files and exception: text files, reading and writing files, format operator; command line arguments, errors and exceptions, handling exceptions, modules, packages; Illustrative programs: word count, copy file.

Unit-IV

Object-Oriented Programming: Classes and Objects, Creating Classes in Python, Creating Objects in Python, The Constructor Method, Classes with Multiple Objects, Class Attributes versus Data Attributes, Encapsulation, Inheritance The Polymorphism. Functional Programming: Lambda, Iterators, Generators, List Comprehensions.

References:

1. Introduction to Python Programming. Gowrishankar S., Veena A. CRC Press, Taylor & Francis Group, 2019
2. Allen B. Downey, ``Think Python: How to Think like a Computer Scientist'', 2nd edition, Updated for Python 3, Shroff/O'Reilly Publishers, 2016 (http://greenteapress.com/wp/think- python/)

Suggested Reading:

1. Learning To Program with Python. Richard L. Halterman. Copyright © 2011

2. Python for Everybody, Exploring Data Using Python 3. Dr. Charles R. Severance. 2016

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – I YEAR SEMESTER- I

Practical-I: Problem Solving and Python Programming (Lab)

HPW- 2P **Credits- 1** **Marks - 50**

Objective

The main objective of this laboratory is to put into practice computational thinking. The students will be expected to write, compile, run and debug Python programs to demonstrate the usage of

- Variables, conditionals and control structures
- Functions (both recursive and iterative)
- basic data types as well as compound data structures such as strings, lists, sets, tuples, dictionaries
- Object-oriented programming

Installing Python and Setting up the Environment

Python interpreter can be downloaded for Windows/Linux platform using the link below:
https://www.python.org/downloads/windows/

Exercises

I. Programs to demonstrate the usage of operators and conditional statements

1. Write a program that takes two integers as command line arguments and prints the sum of two integers. **2.** Program to display the information: Your name, Full Address, Mobile Number, College Name, Course Subjects
3. Program to find the largest number among 'n' given numbers.
4. Program that reads the URL of a website as input and displays contents of a webpage.

II. Programs to demonstrate usage of control structures

5. Program to find the sum of all prime numbers between 1 and 1000.
6. Program that reads set of integers and displays first and second largest numbers.
7. Program to print the sum of first 'n' natural numbers.
8. Program to find the product of two matrices.
9. Program to find the roots of a quadratic equation

III. Programs to demonstrate the usage of Functions and Recursion

10. Write both recursive and non-recursive functions for the following:
   a. To find GCD of two integers
   b. To find the factorial of positive integer
   c. To print Fibonacci sequence up to given number 'n'
   d. To convert decimal number to Binary equivalent

11. Program with a function that accepts two arguments: a list and a number 'n'. It should display all the numbers in the list that are greater than the given number 'n'.
12. Program with a function to find how many numbers are divisible by 2, 3,4,5,6 and 7 between 1 to 1000

IV. Programs to demonstrate the usage of String functions

13. Program that accepts a string as an argument and returns the number of vowels and consonants the string contains.
14. Program that accepts two strings S1, S2, and finds whether they are equal is not.
15. Program to count the number of occurrences of characters in a given string.
16. Program to find whether a given string is palindrome or not

V. Programs to demonstrate the usage of lists, sets, dictionaries, tuples and files.

17. Program with a function that takes two lists L1 and L2 containing integer numbers as parameters. The return value is a single list containing the pair wise sums of the numbers in L1 and L2.
18. Program to read the lists of numbers as L1, print the lists in reverse order without using reverse function.
19. Program to find mean, median, mode for the given set of numbers in a list.
20. Program to find all duplicates in the list.
21. Program to o find all the unique elements of a list.
22. Program to find max and min of a given tuple of integers.
23. Program to find union, intersection, difference, symmetric difference of given two sets.
24. Program to display a list of all unique words in a text file
25. Program to read the content of a text file and display it on the screen line wise with a line number followed by a colon
26. Program to analyze the two text files using set operations
27. Write a program to print each line of a file in reverse order.

VI. Programs to demonstrate the usage of Object Oriented Programming

28. Program to implement the inheritance
29. Program to implement the polymorphism

VII. Programs to search and sort the numbers

30. Programs to implement linear search and Binary search
31. Programs to implement Selection sort, Insertion sort

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – I YEAR SEMESTER- II

AECC-II: Fundamentals of Computers

HPW-2T **Credits-2** **Marks - 50**

Objectives:

1. **To deal with the basic concepts of computers.**
2. **To discuss about the computer hardware, its components and basic computer architecture.**
3. **To understand the basic computer logic gates concept.**
4. **To introduce the software development process.**
5. **To introduce the basic concept of system hardware.**

Outcomes:

Students should be able to

1. Identify the components of a computer and their functions.
2. Understand the concept of networking, LAN, Internet, and working of www.
3. Understand the notion of problem-solving using computer by programming
4. Understand the notion of computer software process and its features.

Unit-I

Introduction to Computers: what is a computer, characteristics of Computers, Generations of Computers, Classifications of Computers, Basic Computer organization, Applications of Computers. Input and Output Devices: Input devices, Output devices, Softcopy devices, hard copy devices. Computer Memory and Processors: Introduction, Memory Hierarchy, Processor, Registers, Cache memory, primary memory, secondary storage devices, magnetic tapes, floppy disks, hard disks, optical drives, USB flash drivers, Memory cards, Mass storage devices, Basic processors architecture.

Unit-II

Number System and Computer Codes: Binary number system, working with binary numbers, octal number system, hexadecimal number system, working with fractions, signed number representation in binary form, BCD code, and other codes. Boolean algebra and logic gates: Boolean algebra, Venn diagrams, representation of Boolean functions, logic gates, logic diagrams and Boolean expressions using karnaugh map. Computer Software: Introduction to computer software, classification of computer software, system software, application software, firmware, middleware, acquiring computer software, design and implementation of correct, efficient and maintainable programs.

Text Book:

Reema Thareja, Fundamentals of Computers.

References:

1. **V.Rajaraman, 6th Edition Fundamentals of Computers, Neeharika Adabala.**
2. **Anita Goel, Computer Fundamentals.**

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – I YEAR SEMESTER- II

Paper-II: Data Base Management Systems (DBMS)

HPW- 4T+2P **Credits - 4** **Marks – 70 + 30**

Unit - I

Introduction: Database-System Applications, Purpose of Database Systems, View of Data, Database Languages, Relational Databases, Database Design, Data Storage and Querying, Transaction Management, Database Architecture, Database Users and Administrators.

Introduction to the Relational Model: Structure of Relational Databases, Database Schema, Keys, Schema Diagrams, Relational Query Languages, Relational Operations.

Unit - II

Database Design and the E-R Model: Overview of the Design Process, The Entity Relationship Model, Constraints, Removing Redundant Attributes in Entity Sets, Entity Relationship Diagrams, and Reduction to Relational Schemas, Entity-Relationship Design Issues, Extended E-R Features, Alternative Notations for Modeling Data, Other Aspects of Database Design.

Relational Database Design: Features of Good Relational Designs, Atomic Domains and First Normal Form, Decomposition Using Functional Dependencies, Functional- Dependency Theory, Decomposition Using Multivalued Dependencies, Normal Forms-2 NF, 3 NF, BCNF, The Database Design Methodology for Relational Databases.

Unit - III

Introduction to SQL: Overview of the SQL Query Language, SQL Data Definition, Basic Structure of SQL Queries, Additional Basic Operations, Set Operations, Null Values, Aggregate Functions, Nested Subqueries , Modification of the Database. Intermediate SQL: Join Expressions, Views, Transactions, Integrity Constraints, SQL Data Types and Schemas, Authorization.

Advanced SQL: Accessing SQL from a Programming Language, Functions and Procedures, Triggers, Recursive Queries.

Unit - IV

Transaction Management: Transaction Support–Properties of Transactions, Database Architecture, Concurrency Control–The Need for Concurrency Control, Serializability and Recoverability, Locking Methods, Deadlock, Time Stamping Methods, Multi-version Timestamp Ordering, Optimistic Techniques, Granularity of Data Items, Database Recovery– The Need for Recovery, Transactions and Recovery, Recovery Facilities, Recovery Techniques, Nested Transaction Model. Security: Database Security–Threats, Computer Based Controls–Authorization, Access Controls, Views, Backup and Recovery, Integrity, Encryption, RAID.

Text book:

1. Silberschatz, H. Korth and S. Sudarshan, Database System Concepts, 6th Ed., Tata McGraw Hill, 2011

2. Thomas M. Connolly, Carolyn E. Begg, Database Systems–A Practical Approach to Design, Implementation, and Management (6e)

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – I YEAR SEMESTER- II

Practical- II: Data Base Management Systems (DBMS) (Lab)

HPW- 2P **Credits- 1** **Marks - 50**

Note:

- Programs of all the Concepts from Textbook including exercises must be practice and execute.
- In the external lab examination student has to execute two programs with compilation and deployment steps are necessary.
- External Vice-Voce is compulsory.

1. Create a database having two tables with the specified fields, to computerize a library system of a University College. Library Books (Accession number, Title, Author, Department, PurchaseDate, Price), IssuedBooks (Accession number, Borrower)

a) Identify primary and foreign keys. Create the tables and insert at least 5 records in each table.

b) Delete the record of book titled "Database System Concepts".

c) Change the Department of the book titled "Discrete Maths" to "CS".

d) List all books that belong to "CS" department.

e) List all books that belong to "CS" department and are written by author "Navathe".

f) List all computer (Department="CS") that have been issued.

g) List all books which have a price less than 500 or purchased between "01/01/1999" and "01/01/2004".

2. Create a database having three tables to store the details of students of Computer Department in your college. Personal information about Student (College roll number, Name of student, Date of birth, Address, Marks(rounded off to whole number) in percentage at 10 + 2, Phone number) Paper Details (Paper code, Name of the Paper) Student's Academic and Attendance details (College roll number, Paper Code, Attendance, Marks in home examination).

a) Identify primary and foreign keys. Create the tables and insert at least 5 records in each table.

b) Design a query that will return the records (from the second table) along with the name of student from the first table, related to students who have more than 75% attendance and more than 60% marks in paper2.

c) List all students who live in "Warangal" and have marks greater than 60 in paper1.

d) Find the total attendance and total marks obtained by each student.

e) List the name of student who has got the highest marks in paper2.

3. Create the following tables and answer the queries given below: Customer (CustID, email, Name, Phone, ReferrerID) Bicycle (BicycleID, DatePurchased, Color, CustID, ModelNo) BicycleModel(ModelNo, Manufacturer, Style) Service (StartDate, BicycleID, EndDate)

a) Identify primary and foreign keys. Create the tables and insert at least 5 records in each table.

b) List all the customers who have the bicycles manufactured by manufacturer "Honda".

c) List the bicycles purchased by the customers who have been referred by Customer "C1".

d) List the manufacturer of red colored bicycles.

e) List the models of the bicycles given for service.

4. Create the following tables, enter at least 5 records in each table and answer the queries given below. Employee (Person_Name, Street, City) Works (Person_Name, Company_Name, Salary) Company (Company_Name, City) Manages (Person_Name, Manager_Name)

a) Identify primary and foreign keys.

b) Alter table employee, add a column "email" of type varchar (20).

c) Find the name of all managers who work for both Samba Bank and NCB Bank.

d) Find the names, street address and cities of residence and salary of all employees who work for "Samba Bank" and earn more than $10,000.

e) Find the names of all employees who live in the same city as the company for which they work.

f) Find the highest salary, lowest salary and average salary paid by each company.

g) Find the sum of salary and number of employees in each company.

h) Find the name of the company that pays highest salary.

5. Create the following tables, enter at least 5 records in each table and answer the queries given below. Suppliers (SNo, Sname, Status, SCity) Parts (PNo, Pname, Colour, Weight, City) Project (JNo, Jname,Jcity) Shipment (Sno, Pno, Jno, Qunatity)

a) Identify primary and foreign keys.

b) Get supplier numbers for suppliers in Paris with status>20.

c) Get suppliers details for suppliers who supply part P2. Display the supplier list in increasing order of supplier numbers.

d) Get suppliers names for suppliers who do not supply part P2.

e) For each shipment get full shipment details, including total shipment weights.

f) Get all the shipments where the quantity is in the range 300 to 750 inclusive.

g) Get part nos. for parts that either weigh more than 16 pounds or are supplied by suppliers S2, or both.

h) Get the names of cities that store more than five red parts.

i) Get full details of parts supplied by a supplier in Hyderabad.

j) Get part numbers for part supplied by a supplier in Warangal to a project in Chennai.

k) Get the total number of project supplied by a supplier (say, S1).

l) Get the total quantity of a part (say, P1) supplied by a supplier (say, S1).

6. Write a PL/SQL Program to demonstrate Procedure.

7. Write a PL/SQL Program to demonstrate Function.
8. Write a PL/SQL program to Handle Exceptions.
9. Write a PL/SQL Program to perform a set of DML Operations.
10. Create a View using PL/SQL program.
11. Write a PL/SQL Program on Statement Level Trigger.
12. Write a PL/SQL Program on Row Level Trigger.

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)
B.Sc. (Data Science) – II YEAR SEMESTER- III

Paper-III: Data Engineering with Python

HPW- 4T+2P **Credits – 4** **Marks – 70 + 30**

Objective:

The main objective of this course is to teach how to extract raw data, clean the data, perform transformations on data, load data and visualize the data

Outcomes:

At the end of the course the student will be able to:

- Handle different types of files and work with text data
- Use regular expression operations
- Use relational databases via SQL
- Use tabular numeric data
- Use the data structures: data series and frames
- Use PyPlot for visualization

Unit – I

Data Science: Data Analysis Sequence, Data Acquisition Pipeline, Report Structure [Reference 1(Chapter 1-Unit1 to Unit 3)]]

Files and Working with Text Data: Types of Files, Creating and Reading Text Data, File Methods to Read and Write Data, Reading and Writing Binary Files, The Pickle Module, Reading and Writing CSV Files, Python os and os.pathModules. [Reference 2, Chapter 9)]

Working with Text Data: JSON and XML in Python[Reference 2, Section12.2]

Unit – II

Working with Text Data: Processing HTML Files, Processing Texts in Natural Languages [Reference 1(Chapter3 –Unit 13, and Unit16)

Regular Expression Operations: Using Special Characters, Regular Expression Methods, Named Groups in Python Regular Expressions, Regular Expression with glob Module [Reference 2-Chapter 10]

Unit – III

Working with Databases: Setting Up a MySQL Database, Using a MySQL Database: Command Line, Using a MySQL Database, Taming Document Stores: MongoDB [Reference 1 (Chapter4-Unit17toUnit20)]

Working with Tabular Numeric Data(Numpy with Python): NumPy Arrays Creation Using array() Function, Array Attributes, NumPy Arrays Creation with Initial Placeholder Content, Integer Indexing, Array Indexing, Boolean ArrayIndexing, Slicing and Iterating in Arrays, Basic Arithmetic Operations on NumPy Arrays, Mathematical Functions in NumPy, Changing the Shape of an Array, Stacking and Splitting of Arrays, Broadcasting in Arrays. [Reference 2: Section 12.3)]

Unit – IV

Working with Data Series and Frames: Pandas Data Structures, Reshaping Data, Handling Missing Data, Combining Data, Ordering and Describing Data, Transforming Data, Taming Pandas File I/O [Reference 1 (Chapter 6-Unit 31 to Unit 37)]

Plotting: Basic Plotting with PyPlot, Getting to Know Other Plot Types, Mastering Embellishments, Plotting with Pandas [Reference 1 (Chapter8-Unit 41 to Unit 44)]

References:

1. Data Science Essentials in Python: Collect, Organize, Explore, Predict, Value. Dmitry Zinoriev, The Pragmatic Programmers LLC, 2016

2. Introduction to Python Programming. Gowrishankar S., Veena A. CRC Press, Taylor & Francis Group, 2019

Suggested Reading

3. Python for Everybody: Exploring Data Using Python 3. Charles R Severance, 2016

4. Python Data Analytics – Data Analysis and Science using Pandas, matplotlib and the Python Programming Language. Fabio Nelli, Apress, 2015

5. Website Scraping with Python. Using BeautifulSoup and Scrapy. GáborLászlóHajba, Apress, 2018

6. Machine Learning with Python Cookbook:.Practical Solutions from Preprocessing to Deep Learning. Chris Albon, O'Reilly 2018

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – II YEAR SEMESTER- III

Practical- II: Data Engineering with Python (Lab)

HPW- 2P **Credits- 1** **Marks – 50**

Objective:

The main objective of this laboratory is to put into practice the ETL (extract, transform, load) pipeline which will extract raw data, clean the data, perform transformations on data, load data and visualize the data. This requires mentoring by TCS.

Libraries:

In this course students are expected to extract, transform and load input data that can be text files, CSV files, XML files, JSON, HTML files, SQL databases, NoSQL databases etc. For doing this, they should learn the following Python libraries/modules: pandas, numpy, BeautifulSoup, pymysql, pymongo, nltk, matplotlib.

Datasets:

For this laboratory, appropriate publicly available datasets, can be studied and used.

Example:

MNIST (http://yann.lecun.com/exdb/mnist/),
UCI Machine Learning Repository(https://archive.ics.uci.edu/ml/datasets.html),
Kaggle(https://www.kaggle.com/datasets)
Twitter Data

Exercises:

1. Write programs to parse text files, CSV, HTML, XML and JSON documents and extract relevant data. After retrieving data check any anomalies in the data, missing values etc.

2. Write programs for reading and writing binary files

3. Write programs for searching, splitting, and replacing strings based on pattern matching using regular expressions

4. Design a relational database for a small application and populate the database. Using SQL do the CRUD (create, read, update and delete) operations.

5. Create a Python MongoDB client using the Python module pymongo. Using a collection object practice functions for inserting, searching, removing, updating, replacing, and aggregating documents, as well as for creating indexes

6. Write programs to create numpy arrays of different shapes and from different sources, reshape and slice arrays, add array indexes, and apply arithmetic, logic, and aggregation functions to some or all array elements.

7. Write programs to use the panda's data structures: Frames and series as storage containers and for a variety of data-wrangling operations, such as:

- Single-level and hierarchical indexing
- Handling missing data
- Arithmetic and Boolean operations on entire columns and tables
- Database-type operations (such as merging and aggregation)
- Plotting individual columns and whole tables
- Reading data from files and writing data to files

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – II YEAR SEMESTER- III

SEC- II: MINI -PROJECT

HPW-2T Credits-2 Marks – 50

Course Objectives:

1. To understand and select the task based on their core skills.
2. To get the knowledge about analytical skill for solving the selected task.
3. To get confidence for implementing the task and solving the real time problems.
4. Express technical and behavioural ideas and thought in oral settings.
5. Prepare and conduct oral presentations

Outcomes:

On the successful completion of the course, student will be able to:

1. Formulate a real-world problem and develop its requirements develop a design solution for a set of requirements

2. Test and validate the conformance of the developed prototype against the original requirements of the problem

3. Work as a responsible member and possibly a leader of a team in developing software solutions

4. Express technical ideas, strategies and methodologies in written form. Self-learn new tools, algorithms and techniques that contribute to the software solution of the project

5. Generate alternative solutions, compare them and select the optimum one.

Aim of the project work:

1. The aim of the project work is to acquire practical knowledge on the implementation of the programming concepts studied.
2. Each student should carry out individually one project work and it may be a work using the software packages that they have learned or the implementation of concepts from the papers studied or implementation of any innovative idea focusing on application-oriented concepts.
3. The project work should be compulsorily done in the college only under the supervision of the department staff concerned.

Viva Voce:

1. Viva-Voce will be conducted at the end of the year by both Internal (Respective Guides) and External Examiners, after duly verifying the Annexure Report available in the College, for a total of 50 marks at the last day of the practical session.
2. Out of 50 marks, 35 marks for project report and 15 marks for Viva Voce.

Project Work Format

PROJECT WORK

TITLE OF THE DISSERTATION

Bonafide Work Done by

STUDENT NAME

REG. NO.

Dissertation submitted in partial fulfillment of the requirements for the award of

<Name of the Degree>

Of Hindi MahaVidyalaya, Hyderabad - 500044.

College Logo

Signature of the Guide

Signature of the HOD

Submitted for the Viva-Voce Examination held on _____

Internal Examiner

External Examiner

Month – Year

CONTENTS

3. System Design and Development

3.1 File Design

3.2 Input Design

3.3 Output Design

3.4 Database Design

3.5 System Development

3.5.1 Description of Modules (Detailed explanation about the project work)

4. Testing and Implementation

5. Conclusion Bibliography Appendices

A. Data Flow Diagram

B. Table Structure

C. Sample Coding

D. Sample Input

E. Sample Output

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – II YEAR SEMESTER- IV

Paper-IV: Machine Learning

HPW- 4T+2P **Credits – 4** **Marks – 70 + 30**

Objectives:
The main objective of this course is to teach the principles and foundations of machine learning algorithms.

Outcomes:
At the end of the course the student will be able to understand
- Basics of Machine Learning and its limitations
- Machine Learning Algorithms: supervised, unsupervised, bio-inspired
- Probabilistic Modeling and Association Rule Mining

Unit-I
Introduction: What does it mean to learn, Some canonical Learning Problems, The Decision Tree Model of Learning, Formalizing the Learning Problem **[Reference 1]**, ID3 Algorithm **[[Reference 2]**
Limits of Learning: Data Generating Distributions, Inductive Bias, Not Everything is learnable, Underfitting and Overfitting, Separation of training and test Data, Models, parameters and Hyperparameters, Real World Applications of Machine Learning **[Reference 1]**

Geometry and Nearest Neighbors: From Data to Feature Vectors, k-Nearest Neighbors, Decision Boundaries, k-means Clustering, High Dimensions **[Reference 1]**

Unit-II
The Perceptron: Bio-inspired Learning, The Perceptron Algorithm, Geometric Interpretation, Interpreting Perceptron Weights, Perceptron Convergence and Linear Separability, Improved Generalization, Limitations of the Perceptron **[Reference 1]**

Practical Issues: Importance of Good Features, Irrelevant and Redundant Features, Feature Pruning and Normalization, Combinatorial Feature Explosion, Evaluating Model Performance, Cross Validation, Hypothesis Testing and Statistical Significance, Debugging Learning Algorithms, Bias Variance tradeoff **[Reference 1]**

Linear Models: The Optimization Framework for Linear Models, Convex Surrogate Loss Functions, Weight Regularization, Optimization and Gradient Descent, Support Vector Machines **[Reference 1]**

Unit-III
Probabilistic Modeling: Classification by Density Estimation, Statistical Estimation, Naïve Bayes Models, Prediction **[Reference 1]**
Neural Networks: Bio-inspired Multi-Layer Networks, The Back-propagation Algorithm, Initialization and Convergence of Neural Networks, Beyond two layers, Breadth vs Depth, Basis Functions **[Reference 1]**

Unit IV

Unsupervised Learning: Clustering Introduction, Similarity and Distance Measures, Agglomerative Algorithms, Divisive Clustering, Minmum Spanning Tree **[Reference 2]**

Association Rules: Introduction, large Itemsets, Apriori Algorithm **[Reference 2]**

References:

1. A Course in Machine Learning (CIML). Hal Daume III, 2017 (freely available online) http://ciml.info/
2. Data Mining: Introductory and Advanced Topics. Margaret H Dunham, Pearson Education, 2003

Suggested Reading:

3. Hands on Machine Learning with SciKit-Learn, Keras and Tensor Flow. AurélienGéron. O'Reily, 2019
4. Machine Learning with Python Cookbook. Chris Albo, O'Reily, 2018
5. Introduction to Machine Learning with Python: A guide. Andreas C Miller, Sarah Guido. O'Reily, 2017

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – II YEAR SEMESTER- IV

Practical- IV: Machine Learning (Lab)

HPW- 2P **Credits- 1** **Marks – 50**

Objective:
The main objective of this laboratory is to put into practice the various machine learning algorithms for data analysis using Python and Weka.

ML Toolkits:

Students are expected to learn

1. Scikit-learn(https://scikit-learn.org/) an open source machine learning Python library that supports supervised and unsupervised learning. It also provides various tools for model fitting, data preprocessing, model selection and evaluation, and many other utilities.

2. Weka (http://www.cs.waikato.ac.nz/ml/weka/)is another widely used ML toolkit.

Datasets:

1. The sklearn.datasets package embeds small toy datasets. It includes utilities to load these datasets. It also includes methods to load and fetch popular reference datasets and features some artificial data generators. Students are expected to study and make use of these datasets.

2. Weka also has provides various data sets.

References:

1. scikit-learn user guide.https://scikit-learn.org/stable//_downloads/scikit-learn-docs.pdf

2. Ian Witten, Eibe Frank, and Mark Hall, Chris Pal. DATA MINING: Practical Machine Learning Tools and Techniques, 4th Edition. Morgan Kaufmann.

Exercises:

8. Write a Python program using Scikit-learn to split the iris dataset into 70% train data and 30% test data. Out of total 150 records, the training set will contain 120 records and the test set contains 30 of those records. Print both datasets

9. Write Python program to use sklearn'sDecisionTreeClassifier to build a decision tree for the sklearn's datasets. Implement functions to find the importance of a split (entropy, information gain, gini measure)

10. Write a Python program to implement your own version of the K-means algorithm. Then apply it to different datasets and evaluate the performance.

11. Design a perceptron classifier to classify handwritten numerical digits (0-9). Implement using scikit or Weka.

12. Write a Python program to classify text as spam or not spam using the Naïve Bayes Classifier

13. Use WEKA and experiment with the following classifiers: Association Rule Mining (Apriori), Agglomerative and Divisive Clustering.

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – II YEAR SEMESTER- IV

SEC- IV: MINI -PROJECT

HPW-2T Credits-2 Marks – 50

Course Objectives:

1. To understand and select the task based on their core skills.
2. To get the knowledge about analytical skill for solving the selected task.
3. To get confidence for implementing the task and solving the real time problems.
4. Express technical and behavioural ideas and thought in oral settings.
5. Prepare and conduct oral presentations

Outcomes:

On the successful completion of the course, student will be able to:

1. Formulate a real-world problem and develop its requirements develop a design solution for a set of requirements

2. Test and validate the conformance of the developed prototype against the original requirements of the problem

3. Work as a responsible member and possibly a leader of a team in developing software solutions

4. Express technical ideas, strategies and methodologies in written form. Self-learn new tools, algorithms and techniques that contribute to the software solution of the project

5. Generate alternative solutions, compare them and select the optimum one.

Aim of the project work:

1. The aim of the project work is to acquire practical knowledge on the implementation of the programming concepts studied.
2. Each student should carry out individually one project work and it may be a work using the software packages that they have learned or the implementation of concepts from the papers studied or implementation of any innovative idea focusing on application-oriented concepts.
3. The project work should be compulsorily done in the college only under the supervision of the department staff concerned.

Viva Voce:

1. Viva-Voce will be conducted at the end of the year by both Internal (Respective Guides) and External Examiners, after duly verifying the Annexure Report available in the College, for a total of 50 marks at the last day of the practical session.
2. Out of 50 marks, 35 marks for project report and 15 marks for Viva Voce.

Project Work Format

PROJECT WORK

TITLE OF THE DISSERTATION

Bonafide Work Done by

STUDENT NAME

REG. NO.

Dissertation submitted in partial fulfillment of the requirements for the award of

<Name of the Degree>

Of Hindi MahaVidyalaya, Hyderabad - 500044.

College Logo

Signature of the Guide

Signature of the HOD

Submitted for the Viva-Voce Examination held on _____

Internal Examiner

External Examiner

Month – Year

CONTENTS

3. System Design and Development

3.1 File Design

3.2 Input Design

3.3 Output Design

3.4 Database Design

3.5 System Development

3.5.1 Description of Modules (Detailed explanation about the project work)

4. Testing and Implementation

5. Conclusion Bibliography Appendices

A. Data Flow Diagram

B. Table Structure

C. Sample Coding

D. Sample Input

E. Sample Output

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD

(AUTONOMOUS)

B.Sc. (Data Science) – III YEAR SEMESTER- V

Paper-V (A): Natural Language Processing

HPW- 4T+2P **Credits – 4** **Marks – 70 + 30**

Objective:

The main objective of this course is to give a practical introduction to NLP. It deals with morphological processing, syntactic parsing, information extraction, probabilistic NLP and classification of text using Python's NLTK Library.

Outcomes:

At the end of the course the student will be able to
- Write Python programs to manipulate and analyze language data
- Understand key concepts from NLP and linguistics to describe and analyze language
- Understand the data structures and algorithms that are used in NLP
- Classify texts using machine learning and deep learning

Unit-I

Language Processing and Python: Computing with Language: Texts and Words, A Closer Look at Python: Texts as Lists of Words, Computing with Language: Simple Statistics, Back to Python: Making Decisions and Taking Control, Automatic Natural Language Understanding **[Reference 1]**

Accessing Text Corpora and Lexical Resources: Accessing Text Corpora, Conditional Frequency Distributions, Lexical Resources, WordNet **[Reference 1]**

Unit-II

Processing Raw Text: Accessing Text from the Web and from Disk, Strings: Text Processing at the Lowest Level, Text Processing with Unicode, Regular Expressions for Detecting Word Patterns, Useful Applications of Regular Expressions, Normalizing Text, Regular Expressions for Tokenizing Text, Segmentation, Formatting: From Lists to Strings. **[Reference 1]**

Categorizing and Tagging Words: Using a Tagger, Tagged Corpora, Mapping Words to Properties Using Python Dictionaries, Automatic Tagging, N-Gram Tagging, Transformation-Based Tagging, How to Determine the Category of a Word **[Reference 1]**

Unit-III

Learning to Classify Text: Supervised Classification, Evaluation, Naive Bayes Classifiers **[Reference 1]**

Deep Learning for NLP: Introduction to Deep Learning, Convolutional Neural Networks, Recurrent Neural Networks, Classifying Text with Deep Learning **[Reference 2]**

Unit-IV

Extracting Information from Text Information Extraction, Chunking, Developing and Evaluating Chunkers, Recursion in Linguistic Structure, Named Entity Recognition, Relation Extraction. **[Reference 1]**

Analyzing Sentence Structure Some Grammatical Dilemmas, What's the Use of Syntax. Context-Free Grammar, Parsing with Context-Free Grammar, [Reference 1]

References:

1. Natural Language Processing with Python. Steven Bird, Ewan Klein, and Edward Lope, O'Reily, 2009

2. Natural Language Processing Recipes: Unlocking Text Data with Machine Learning and Deep Learning using Python. Akshay Kulkarni, AdarshaShivananda, Apress, 2019

Suggested Reading:

3. Allen James, Natural Language Understanding, Benjamin/Cumming,1995.

4. Charniack, Eugene, Statistical Language Learning, MIT Press, 1993.

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – III YEAR SEMESTER- V

Practical- V (A): Natural Language Processing (Lab)

HPW- 2P | Credits- 1 | Marks – 50

Objective:
The main objective of this laboratory is to write programs that manipulate and analyze language data using Python.

This lab requires mentoring sessions from TCS.

Python Packages
Students are expected to know/ learn the following Python NLP packages
- NLTK (www.nltk.org/ (http://www.nltk.org/))
- Spacy (https://spacy.io/)
- TextBlob (http://textblob.readthedocs.io/en/dev/
- Gensim (https://pypi.python.org/pypi/gensim)
- Pattern (https://pypi.python.org/pypi/Pattern)

Datasets:
1. NLTK includes a small selection of texts from the Project Gutenberg electronic text archive, which contains some 25,000 free electronic books, hosted at http://www.gutenberg.org/.

2. The Brown Corpus contains text from 500 sources, and the sources have been categorized by genre, such as news, editorial, and so on (http://icame.uib.no/brown/bcmlos.html).

3. Wikipedia Articles Or any other dataset of your choice

Reference:
Jacob Perkins. Python 3 Text Processing with NLTK 3 Cookbook. Packt Publishing. 2014

Exercises:
1. Text segmentation: Segment a text into linguistically meaningful units, such as paragraphs, sentences, or words. Write programs to segment text (in different formats) into tokens (words and word-like units) using regular expressions. Compare an automatic tokenization with a gold standard

2. Part-of-speech tagging: Label words (tokens) with parts of speech such as noun, adjective, and verb using a variety of tagging methods , e.g., default tagger, regular expression tagger, unigram tagger, and n-gram taggers.

3. Text classification: Categorize text documents into predefined classes using Naïve Bayes Classifier and the Perceptron model

4. Chunk extraction, or partial parsing: Extract short phrases from a part-of-speech tagged sentence. This is different from full parsing in that we're interested in standalone chunks, or phrases, instead of full parse trees.

5. Parsing: parsing specific kinds of data, focusing primarily on dates, times, and HTML.

Make use of the following preprocessing libraries:
- dateutil which provides datetime parsing and timezone conversion
- lxml and BeautifulSoup which can parse, clean, and convert HTML
- charade and UnicodeDammit which can detect and convert text character encoding

6. Sentiment Analysis: Using Libraries TextBlob and nltk, give the sentiment of a document.

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – III YEAR SEMESTER- V

Paper-V (B): NoSQL Data Bases

HPW- 4T+2P **Credits – 4** **Marks – 70 + 30**

Objective:
The main objective of this course is to cover core concepts of NoSQL databases, along with an example database for each of the key-value, document, column family, and graph databases

Outcomes:
At the end of the course the student will be able to
- Understand the need for NoSQL databases and their characteristics
- Understand the concepts of NoSQL databases
- Implement the concepts of NoSQL databases using four example databases: Redis for key-value databases, MongoDB for document databases, Cassandra for columnfamily databases, and Neo4J for graph databases.

Unit-I
Why NoSQL: The Value of Relational Databases, Impedance Mismatch, Application and Integration Databases, Attack of the Clusters, The Emergence of NoSQL Aggregate Data Models: Aggregates, Column-Family Stores, Summarizing AggregateOriented Databases More Details on Data Models: Relationships, Graph Databases, Schemaless Databases, Materialized Views, Modeling for Data Access

Unit-II
Distribution Models: Single Server, Sharding, Master-Slave Replication, Peer-to-Peer Replication, Combining Sharding and Replication Consistency: Update Consistency, Read Consistency, Relaxing Consistency, Relaxing Durability, Quorums Version Stamps: Business and System Transactions, Version Stamps on Multiple Nodes Map-Reduce: Basic Map-Reduce, Partitioning and Combining, Composing Map-Reduce Calculations.

Unit-III
Key-Value Databases: What Is a Key-Value Store, Key-Value Store Features, Suitable Use Cases, When Not to Use Document Databases: What Is a Document Database, Features, Suitable Use Cases, When Not to Use

Unit-IV
Column-Family Stores: What Is a Column-Family Data Store, Features, Suitable Use Cases, When Not to Use Graph Databases: What Is a Graph Database, Features, Suitable Use Cases, When Not to Use

Reference:

1. Pramod J. Sadalage, Martin Fowler. NoSQL Distilled, Addison Wesley 2013

Suggested Reading

2. Luc Perkins, Eric Redmond, Jim R. Wilson. Seven Databases in Seven Weeks. The Pragmatic Bookshelf, 2018

3. Guy Harrison. Next Generation Databases: NoSQL, NewSQL, and Big Data. Apress, 2015

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – III YEAR SEMESTER- V

Practical- V (B): NoSQL Data Bases (Lab)

HPW- 2P **Credits- 1** **Marks – 50**

Objective: The main objective of this lab is to become familiar with the four NoSQL databases: Redis for key-value databases, MongoDB for document databases, Cassandra for column-family databases, and Neo4J for graph databases

NoSQL Databases:
Redis (http://redis.io)
MongoDB (http://www.mongodb.org)
Cassandra (http://cassandra.apache.org)
Neo4j (http://neo4j.com)

Exercises:

1. Installation of NoSQL Databases: Redis,MongoDB, Cassandra, Neo4j on Windows & Linux

2. Practice CRUD (Create, Read, Update, and Delete) operations on the four databases: Redis,MongoDB, Cassandra, Neo4j

3. Usage of Where Clause equivalent in MongoDB

4. Usage of operations in MongoDB – AND in MongoDB, OR in MongoDB, Limit Records and Sort Records. Usage of operations in MongoDB – Indexing, Advanced Indexing, Aggregation and Map Reduce.

5. Practice with ' macdonalds ' collection data for document oriented database. Import restaurants collection and apply some queries to get specified output.

6. Write a program to count the number of occurrences of a word using MapReduce

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – III YEAR SEMESTER- V

Paper-VI : GE : Data Structures and Algorithms

HPW- 4T+2P **Credits – 4** **Marks – 70 + 30**

- To introduce the time and space complexities of algorithms.
- To discuss the linear and non-linear data structures and their applications.
- To introduce the creation, insertion and deletion operations on binary search trees and balanced binary search trees.
- To introduce various internal sorting techniques and their time complexities

Outcomes:

Students will be

- Able to analyze the time and space complexities of algorithms.
- Able to implement linear, non-linear data structures and balanced binary trees
- Able to analyse and implement various kinds of searching and sorting techniques.
- Able to find a suitable data structure and algorithm to solve a real world problem.

UNIT–I

Performance and Complexity Analysis: Space Complexity, Time Complexity, Asymptotic Notation (Big-Oh), Complexity Analysis Examples. Linear List-Array Representation: Vector Representation, Multiple Lists Single Array. Linear List-Linked Representation: Singly Linked Lists, Circular Lists, Doubly Linked Lists, Applications (Polynomial Arithmetic). Arrays and Matrices: Row and Column Major Representations, Sparse Matrices. Stacks: Array Representation, Linked Representation, Applications (Recursive Calls, Infix to Postfix, Postfix Evaluation). Queues: Array Representation, Linked Representation. Skip Lists and Hashing: Skip Lists Representation, Hash Table Representation, Application- Text Compression.

UNIT- II

Trees: Definitions and Properties, Representation of Binary Trees, Operations, Binary Tree Traversal.

Binary Search Trees: Definitions, Operations on Binary Search Trees. Balanced Search Trees: AVL Trees, and B-Trees.

UNIT –III

Graphs: Definitions and Properties, Representation, Graph Search Methods (Depth First Search and Breadth First Search)

Application of Graphs: Shortest Path Algorithm (Dijkstra), Minimum Spanning Tree (Prim's and Kruskal's Algorithms).

UNIT –IV

Searching: Linear Search and Binary Search Techniques and their complexity analysis.
Sorting and Complexity Analysis: Selection Sort, Bubble Sort, Insertion Sort, Quick Sort, Merge Sort, and Heap Sort.
Algorithm Design Techniques: Greedy algorithm, divide-and-conquer, dynamic programming.

Suggested Reading:

1. Michael T. Goodrich, Roberto Tamassia, David M. Mount, Data Structures and Algorithms Python John Wiley & Sons,2013.

2. Mark Allen Weiss, Data Structures and Problem Solving using C++, Pearson Education International,2003.

3. SartajSahni, Data Structures--Algorithms and Applications in C++, 2nd Edition, Universities Press (India) Pvt. Ltd.,2005.

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – III YEAR SEMESTER- VI

Paper-VII (A): Big Data

HPW- 4T+2P **Credits – 4** **Marks – 70 + 30**

UNIT – I

Getting an overview of Big Data: Introduction to Big Data, Structuring Big Data, Types of Data, Elements of Big Data, Big Data Analytics, Advantages of Big Data Analytics. Introducing Technologies for Handling Big Data: Distributed and Parallel Computing for Big Data, Cloud Computing and Big Data, Features of Cloud Computing, Cloud Deployment Models, Cloud Services for Big Data, Cloud Providers in Big Data Market.

UNIT – II

Understanding Hadoop Ecosystem: Introducing Hadoop, HDFS and MapReduce, Hadoop functions, Hadoop Ecosystem. Hadoop Distributed File System- HDFS Architecture, Concept of Blocks in HDFS Architecture, Namenodes and Datanodes, Features of HDFS. MapReduce.

Introducing HBase- HBase Architecture, Regions, Storing Big Data with HBase, Combining HBase and HDFS, Features of HBase, Hive, Pig and Pig Latin, Sqoop, ZooKeeper, Flume, Oozie.

UNIT- III

Understanding MapReduce Fundamentals and HBase: The MapReduceFramework ,Exploring the features of MapReduce, Working of MapReduce, Techniques to optimize MapReduce Jobs, Hardware/Network Topology, Synchronization, File system, Uses of MapReduce, Role of HBase in Big Data Processing- Characteristics of HBase.

Understanding Big Data Technology Foundations: Exploring the Big Data Stack, Data Sources Layer, Ingestion Layer, Storage Layer, Physical Infrastructure Layer, Platform Management Layer, Security Layer, Monitoring Layer, Visualization Layer.

UNIT – IV

Storing Data in Databases and Data Warehouses: RDBMS and Big Data, Issues with Relational Model, Non – Relational Database, Issues with Non Relational Database, Polyglot Persistence, Integrating Big Data with Traditional Data Warehouse, Big Data Analysis and Data Warehouse.

NoSQL Data Management: Introduction to NoSQL, Characteristics of NoSQL, History of NoSQL, Types of NoSQL Data Models- Key Value Data Model, Column Oriented Data Model, Document Data Model, Graph Databases, Schema-Less Databases, Materialized Views, CAP Theorem.

Reference

1. BIG DATA, Black Book TM, DreamTech Press, 2016 Edition.

Suggested Reading:

2. Seema Acharya, SubhasniChellappan , "BIG DATA and ANALYTICS", Wiley publications, 2016
3. Nathan Marz and James Warren, "BIG DATA- Principles and Best Practices of Scalable Real-Time Systems", 2010

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – III YEAR SEMESTER- VI

Practical- VII (A): Big Data (Lab)

HPW- 2P **Credits- 1** **Marks – 50**

Objectives:

- Installation and understanding of working of HADOOP
- Understanding of MapReduce program paradigm.
- Writing programs in Python using MapReduce
- Understanding working of Pig, Hive
- Understanding of working of Apache Spark Cluster

1. Setting up and Installing Hadoop in its two operating modes:
- Pseudo distributed,
- Fully distributed.

2. Implementation of the following file management tasks in Hadoop:
- Adding files and directories
- Retrieving files
- Deleting files

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

4. Map Reduce Program for Stop word elimination:
- Map Reduce program to eliminate stop words from a large text file.

5. Map Reduce program that mines weather data. Weather sensors collecting data every hour at many locations across the globe gather large volume of log data, which is a good candidate for analysis with MapReduce, since it is semi structured and recordoriented.

Data available at: https://github.com/tomwhite/hadoopbook/tree/master/input/ncdc/all.
- Find average, max and min temperature for each year in NCDC data set?
- Filter the readings of a set based on value of the measurement, Output the line of input files associated with a temperature value greater than 30.0 and store it in a separate file.

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

7. Write a Pig Latin script for finding TF-IDF value for book dataset (A corpus of eBooks available at: Project Gutenberg)

8. Install and Run Hive then use Hive to create, alter, and drop databases, tables, views, functions, and indexes.

9. Install, Deploy & configure Apache Spark Cluster. Run apache spark applications using Scala.

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

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – III YEAR SEMESTER- VI

Paper-VIII (B): Deep Learning

HPW- 4T+2P **Credits – 4** **Marks – 70 + 30**

Objective:

The main objective of this course is to give a practical introduction to Deep Learning using Keras. It covers the concepts of deep learning and their implementation.

Outcomes:
At the end of the course the student will be able to

1. Understand the basics of deep learning

2. Understand the usage of tensors in deep learning

3. Use Python deep-learning framework Keras, with Tensor-Flow as a backend engine.

Unit-I

Introduction: History, Hardware, Data, Algorithms Neural Networks, Data representations for neural networks, Scalars (0D tensors), Vectors
(1D tensors), Matrices (2D tensors), 3D tensors and higher-dimensional tensors, Key attributes,. Manipulating tensors in Numpy,The notion of data batches, Real-world examples of data tensors, Vector data, Timeseries data or sequence data, Image data, Video data

Unit-II
Tensor operations: Element-wise operations, Broadcasting, Tensor dot,. Tensor reshaping, Geometric interpretation of tensor operations, A geometric interpretation of deep learning,

Unit-III
Gradient-based optimization, Derivative of a tensor operation, Stochastic gradient descent,. Chaining derivatives: the Backpropagation algorithm
Neural networks: Anatomy, Layers, Models, Loss functions and optimizers

Unit-IV
Introduction to Keras, Keras, TensorFlow, Theano, and CNTK Recurrent neural networks: A recurrent layer in Keras, Understanding the LSTM and GRU layers

Reference:

1. François Chollet. Deep Learning with Python. Manning Publications, 2018

Suggested Reading:

2. AurélienGéron. Hands on Machine Learning with SciKit-Learn, Keras and Tensor Flow. O'Reily, 2019

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

B.Sc. (Data Science) – III YEAR SEMESTER- VI

Practical- VIII (B): Deep Learning (Lab)

HPW- 2P **Credits- 1** **Marks – 50**

Objectives:

The main objective of this lab is to develop deep learning models using Keras

Deep Learning Tools

Students are expected to learn Keras deep-learning framework (https://keras.io), which is open source and free to download. They should have access to a UNIX machine; though it's possible to use Windows, too. It is also recommended that they work on a recent NVIDIA GPU

Exercises:

Note: The exercises should following Keras workflow consisting of four steps
1. Define your training data: input tensors and target tensors
2. Define a network of layers (or model) that maps your inputs to your targets
3. Configure the learning process by choosing a loss function, an optimizer, and some metrics to monitor
4. Iterate on your training data by calling the fit() method of your model

Exercise 1:

Dataset:
IMDB dataset, a set of 50,000 highly polarized reviews from the Internet Movie Database. They're split into 25,000 reviews for training and 25,000 reviews for testing, each set consisting of 50% negative and 50% positive reviews. the IMDB dataset comes packaged with Keras

Binary Classification Task:
Build a network to classify movie reviews as positive or negative, based on the text content of the reviews.

Exercise 2:

Dataset:
Reuters dataset, a set of short newswires and their topics, published by Reuters in 1986. It's a simple, widely used toy dataset for text classification. There are 46 different topics; some topics are more represented than others, but each topic has at least 10 examples in the training set. Reuters dataset comes packaged as part of Keras.

Single-label Multi class Classification Task:
Build a network to classify Reuters newswires into 46 mutually exclusive topics. Each data point should be classified into only one category (in this case, topic),. The problem is more specifically an instance of single-label, multiclass classification.

Exercise 3:

Dataset:
The Boston Housing Price dataset has an interesting difference from the two previous examples. It has relatively few data points: only 506, split between 404 training samples and 102 test samples. And each feature in the input data (for example, the crime rate) has a different scale. For instance, some values are proportions, which take values between 0 and 1; others take values between 1 and 12, others between 0 and 100, and so on.

Regression Task:
The two previous examples were classification problems, where the goal was to predict a single discrete label of an input data point. Another common type of machine-learning problem is regression, which consists of predicting a continuous value instead of a discrete label. You'll attempt to predict the median price of homes in a given Boston suburb in the mid-1970s, given data points about the suburb at the time, such as the crime rate, the local property tax rate, and so on.

3. More exercises can be defined on similar lines.

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD

(AUTONOMOUS)

B.Sc Data Science- III Year's

Semester – I to VI-Paper – I to VI

Theory Model Question Paper

Max. Marks: 70

Time: 2 ½ hrs

SECTION A

I Write short notes on any Six of the following: **6 X 3 = 18 Marks**

1. A question from Unit I
2. A question from Unit I
3. A question from Unit II
4. A question from Unit II
5. A question from Unit III
6. A question from Unit III
7. A question from Unit IV
8. A question from Unit IV

SECTION B

II Answer all the Questions. Each question carries 13 marks **4 X 13 = 52 Marks**

9 (a) A question from Unit I

(OR)

(b) A question from Unit I

10 (a) A question from Unit II

(OR)

(b) A question from Unit II .

11 (a) A question from Unit III.

(OR)

(b) A question from Unit III.

12 (a)A question from Unit IV

(OR)

(b) A question from Unit IV.

Chairperson **University Nominee** **Members**

1.

2.

3.

Department of Computer Science
Hindi Mahavidyalaya
(AUTONOMOUS & NAAC REACCREDITED)
Nallakunta, Hyderabad-44

CHAIRMAN
Board of Studies in CSE (A)
Dept. of Computer Science & Engg.
College Of Engg., O.U. Hyderabad.

Principal
PRINCIPAL
HINDI MAHA VIDYALAYA
(AUTONOMOUS)
Arts, Commerce & Science
Nallakunta, Hyderabad-44.

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD

(AUTONOMOUS)

B.Sc. Data Science- III Year's

Semester – I to VI-Paper – I to VI

Practical Model Question Paper

Max. Marks: 50

Time:2 hrs

I. Answer any two questions: (30 Marks)
   1.program1.
   2.program 2
   3.program 3
   4.program 4

   Program execution (10 Marks)

II. Record (10 Marks)

III. Viva

Chairperson University Nominee Members

Department of Computer Science
Hindi Mahavidyalaya
(AUTONOMOUS & NAAC REACCREDITED)
Nallakunta, Hyderabad-44.

CHAIRMAN
Board of Studies in CSE (A)
Dept. of Computer Science & Engg.
College Of Engg., O.U. Hyderabad.

1. P.V. Sudhir

PROFESSOR
Department of Computer Science & Engineering
2. University College of Engineering (A)
Osmania University
Hyderabad-500 007.

PRINCIPAL
HINDI MAHA VIDYALAYA
(AUTONOMOUS)
Arts, Commerce & Science
Nallakunta, Hyderabad-44.

3.

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD

(AUTONOMOUS)

B.Sc. Data Science- III Year's

Semester – I to VI-Paper – I to VI

Practical Model Question Paper

Time:2 hrs

Max. Marks: 50

I. Answer any two questions: (30 Marks)
   1.program1.
   2.program 2
   3.program 3
   4.program 4

Program execution (10 Marks)

II. Record (10 Marks)

III. Viva

Chairperson

Department of Computer Science
Hindi Mahavidyalaya
(AUTONOMOUS & NAAC REACCREDITED)
Nallakunta, Hyderabad-44.

University Nominee

CHAIRMAN
Board of Studies in CSE (A)
Dept. of Computer Science & Engg.
College Of Engg., O.U. Hyderabad.

Members

1. P.V. Sridhar

PROFESSOR
Department of Computer Science & Engineering
2. University College of Engineering (A)
Osmania University
Hyderabad-500 007.

3.

Principal

PRINCIPAL
HINDI MAHA VIDYALA
(AUTONOMOUS)
Arts, Commerce & Scienc
Nallakunta, Hyderabad-44.

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD

(AUTONOMOUS)

B.Sc. Data Science- III Year's

Semester – I to VI-Paper – I to VI

Internal Theory Exam

Max. Marks: 30

Time: 1 1/2 hrs

Two internal exams (one at the middle of the semester and the other at the end) of one- hour duration are to be conducted carrying 20 marks each.

Average of the scores of two exams should be considered.

Following is the examination pattern.

- 20 MCQs (multiple choice questions) of 1 mark each, 20*1=20M
- 5 Marks will be allotted for seminar 5M
- 5 Marks will be allotted for Assignment 5M

Total Internal Assessment Marks **30 Marks**

Chairperson University Nominee Members Principal

Department of Computer Science
Hindi Mahavidyalaya
(AUTONOMOUS & NAAC REACCREDITED)
Nallakunta, Hyderabad-44

CHAIRMAN
Board of Studies in CSE (A)
Dept. of Computer Science & Engg.
College Of Engg., O.U. Hyderabad.

1. P.V. Subba Reddy
2. PROFESSOR
Department of Computer Science & Engineering
3. University College of Engineering (A)
Osmania University
Hyderabad-500 007.

PRINCIPAL
HINDI MAHA VIDYALAYA
(AUTONOMOUS)
Arts, Commerce & Science
Nallakunta, Hyderabad-44

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD

(AUTONOMOUS)

B.Sc. Data Science- I Year

Semester – I & II -Paper – AECC I & II

Theory Model Question Paper

Max. Marks: 40

Time:2 hrs

(4 X 4M = 16 Marks)

Section – A

Answer any four of the following six questions. Each carries four marks.

Q1. From Unit 1 Q2. From Unit 1
03. From Unit 1 Q4. From Unit 2 Q5. From Unit 2
06. From Unit 2

(2 X 12M = 24 Marks)

Section – B

Answer all the following two questions. Each carries fifteen marks.

Q09. (a) or (b) from Unit 1 Q10. (a) or (b) from Unit 2

INTERNAL EXAM (THEORY)

Time: 1/ 2 Hr.

Maximum marks: 10

- One internal exam at the end of the semester, of half an hour duration is to be conducted carrying 10 marks.
- Following is the examination pattern.
- 10 MCQs (multiple choice questions) of one mark each, No assignment is required.

Chairperson **University Nominee** **Members** **Principal**

Department of Computer Science
Hindi Mahavidyalaya
(AUTONOMOUS & NAAC REACCREDITED)
Nallakunta, Hyderabad-44.

PRINCIPAL
HINDI MAHA VIDYALAYA
(AUTONOMOUS)
Arts, Commerce & Science
Nallakunta, Hyderabad-44

CHAIRMAN
Board of Studies in CSE (A)
Dept. of Computer Science & Engg.
College Of Engg., O.U. Hyderabad.

1. P.V. Subba Reddy
PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
2. Osmania University
Hyderabad-500 007.

3.

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD

(AUTONOMOUS)

B.Sc. Data Science- II Year

Semester – III & IV -Paper – SEC II & IV

Theory Model Question Paper

Time:2 hrs

Max. Marks: 40

Section – A

(4 X 4M = 16 Marks)

Answer any four of the following six questions. Each carries four marks.

Q1. From Unit 1 Q2. From Unit 1

03. From Unit 1 Q4. From Unit 2 Q5. From Unit 2

06. From Unit 2

Section – B

(2 X 12M = 24 Marks)

Answer all the following two questions. Each carries fifteen marks.

Q09. (a) or (b) from Unit 1 Q10. (a) or (b) from Unit 2

INTERNAL EXAM (THEORY)

Time: 1/ 2 Hr.

Maximum marks: 10

- One internal exam at the end of the semester, of half an hour duration is to be conducted carrying 10 marks.
- Following is the examination pattern.
- 10 MCQs (multiple choice questions) of one mark each, No assignment is required.

Chairperson	University Nominee	Members	Principal
Department of Computer Science Hindi Mahavidyalaya (AUTONOMOUS & NAAC REACCREDITED) Nallakunta, Hyderabad-44.	CHAIRMAN Board of Studies in CSE (A) Dept. of Computer Science & Engg. College Of Engg., O.U. Hyderabad.	1. P.V. Sudha PROFESSOR Department of Computer Science & Engineering University College of Engineering (A) Osmania University Hyderabad-500 007. 2. 3.	PRINCIPAL HINDI MAHA VIDYALAYA (AUTONOMOUS) Arts, Commerce & Sci Nallakunta, Hyderaba

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD

(AUTONOMOUS)

B.Sc. Data Science- III Year

Semester – V -Paper – GE

Theory Model Question Paper

Max. Marks: 70

Time: 2 ½ hrs

SECTION A

6 X 3 = 18 Marks

I Write short notes on any Six of the following:

1. A question from Unit I
2. A question from Unit I
3. A question from Unit II
4. A question from Unit II
5. A question from Unit III
6. A question from Unit III
7. A question from Unit I
8. A question from Unit III

SECTION B

4 X 13 = 52 Marks

II Answer all the Questions. Each question carries 13 marks

9 (a) A question from Unit I

(OR)

(b) A question from Unit I

10 (a) A question from Unit II

(OR)

(b) A question from Unit II .

11 (a) A question from Unit III.

(OR)

(b) A question from Unit III.

12 (a)A question from Unit II

(OR)

(b) A question from Unit III

Chairperson **University Nominee** **Members** **Principal**

Department of Computer Science
Hindi Mahavidyalaya
(AUTONOMOUS & NAAC REACCREDITED)
Nallakunta, Hyderabad-44.

CHAIRMAN
Board of Studies in CSE (A)
Dept. of Computer Science & Engg.
College Of Engg., O.U. Hyderabad.

1. P.V. Sudha
PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
3. Osmania University
Hyderabad-500 007.

PRINCIPAL
HINDI MAHA VIDYALAYA
(AUTONOMOUS)
Arts, Commerce & Science
Nallakunta, Hyderabad-44

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD

(AUTONOMOUS)

B.Sc. Data Science- III Year

Semester – V -Paper – GE

INTERNAL EXAM (THEORY)

Maximum marks: 30

Time: 1/ 2 Hr.

Two internal exams (one at the middle of the semester and the other at the end) of one- hour duration are to be conducted carrying 20 marks each.

Average of the scores of two exams should be considered.

Following is the examination pattern.

- 20 MCQs (multiple choice questions) of 1 mark each, 20*1=20M
- 5 Marks will be allotted for seminar 5M
- 5 Marks will be allotted for Assignment 5M

.....

Total Internal Assessment Marks **30 Marks**

Chairperson **University Nominee** **Members**

Department of Computer Science
Hindi Mahavidyalaya
(AUTONOMOUS & NAAC REACCREDITED)
Nallakunta, Hyderabad-44.

CHAIRMAN
Board of Studies in CSE (A)
Dept. of Computer Science & Engg.
College Of Engg., O.U. Hyderabad.

1. P. V. Sudha

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

3.

Principal

PRINCIPAL
HINDI MAHA VIDYALAYA
(AUTONOMOUS)
Arts, Commerce & Science
Nallakunta, Hyderabad-44

HINDI MAHAVIDYALAYA, NALLAKUNTA, HYDERABAD
(AUTONOMOUS)

Department of Computer Science & Applications
Panel of Examiners

SNO	NAME	PHONE
1	Smt B. Ramani Asst. prof . Dept. of Computer Science AMS ,O.U Campus,Hyderabad	9441214888
2	Smt G. Aparna Asst.prof .Dept of Computer Science AMS, O.U Campus,Hyderabad	9440137700
3	SmtN.Veena Asst.prof .Dept of Computer Science Nizam college ,Hyderabad	9849743764
4	Sri. Avinash Pal Lidlaan Asst.prof. Dept of Computer Science Indira Priyadarshini women govt degree college, Hyderabad	8790077978
5	SmtVijithamalini Asst.prof .Dept of Computer Science Reddy College, Narayanguda,Hyderabad	9000323206
6	SmtBhaskar Rao Asst.prof .Dept of Computer Science AV college, Domalguda,Hyderabad	9885639321
7	Sri N Bhaskar Asst.prof .Dept of Computer Science Bhavan College, Hyderabad	9347983943
8	Ms Salma Asst.prof .Dept of Computer Science RBVRR College for women, Hyderabad	8712960031
9	Ms.Vijitha Asst.prof .Dept of Computer Science Keshav memorial ,Narayanguda, Hyderabad	9640508855
10	Ms. Kavitha Asst.prof .Dept of Computer Science St Francis College for Women,Hyderabad	9393003871

PRINCIPAL
HINDI MAHA VIDYALAYA
(AUTONOMOUS)
Arts, Commerce & Science
Nallakunta, Hyderabad-44

Chairperson

Department of Computer Science
Hindi Mahavidyalaya
(AUTONOMOUS & NAAC REACCREDITED)
Nallakunta, Hyderabad-44.

University Nominee

Members

1. P.V

PROFESSOR
Department of Computer Science & Engineering
University College of Engineering (A)
Osmania University
Hyderabad-500 007.

2.

CHAIRMAN
Board of Studies in CSE (A)
Dept. of Computer Science & Engg.
College Of Engg., O.U. Hyderabad.

3.