Sinhgad Institutes

Sinhgad Technical Educational Society's

SINHGAD INSTITUTE OF TECHNOLOGY

LONAVALA

Laboratory Manual

Subject: - 314457: DS & BDA Laboratory

Class:-T.E. Information Technology

Subject Incharge:

Dr. P. C. Latane

(Information Technology Department)

Teaching Scheme

Practical: 2Hrs/Week

Examamination Scheme

Termowork: 25 Marks

Practical: 25 Marks

ACADEMIC YEAR 2023-24 SEM II

Vision and Mission of Institute

VISION

उत्तमपुरुषान् उत्तमाभियंतृन् निर्मातुं कटीबध्दा: वयम्।

We are committed to produce not only good engineers but good human beings, also.

MISSION

Holistic development of student and teacher is what we believe in and work for.

- We strive to achieve this by imbibing a unique value system, transparent work culture, excellent academic and physical environment conducive to learning, creativity and technology transfer. Our mandate is to generate, preserve and share knowledge for developing a vibrant Society.

Vision and Mission of Department VISION

The department of IT visualizes teaching-learning methodologies to create groomed, technically competent, skilled intellectual IT professionals to achieve the challenges of dynamic needs of local as well global industry and professional community.

MISSION

- The department offers an inviting, nurturing, and challenging environment for self-learning, which is responsive to the intellectual, social and cultural needs of a diverse learning community.
- The success of the department is reflected in well-being of its alumni, who are known for their leadership, adaptability, and commitment to high professional standards.

Short Term Goals

To continuously upgrade existing resources through qualified and experienced faculty and state of the art laboratories.

- To strengthen the institute-industry relationship for mutual benefits.
- To initiate value addition programs and certification examinations for life skills development of students.
- To establish project-based learning model for real life learning to improve academic performance of students and maintain high placement record.

Long Term Goals

- To provide dynamic curriculum to meet the requirements of industries.
- To strengthen patent-based research and product development efforts for supporting indigenous market.
- To foster research in the field of Information Technology for the benefits of society.
- To create better entrepreneurs in the IT Sector.

Program Educational Objectives: PEOs

i. Possess strong fundamental concepts in mathematics, science, engineering and Technology to address technological challenges.
ii. Possess knowledge and skills in the field of Computer Science and Information Technology for analyzing, designing and implementing complex engineering problems of any domain with innovative approaches.
iii. Possess an attitude and aptitude for research, entrepreneurship and higher studies in the field of Computer Science and Information Technology.
iv. Have commitment to ethical practices, societal contributions through communities and lifelong learning.
v. Possess better communication, presentation, time management and teamwork skills leading to responsible & competent professionals and will be able to address challenges in the field of IT at global level.
vi. Possess strong fundamental concepts in mathematics, science, engineering and Technology to address technological challenges.

vii. Possess knowledge and skills in the field of Computer Science and Information Technology for analyzing, designing and implementing complex engineering problems of any domain with innovative approaches.

viii. Possess an attitude and aptitude for research, entrepreneurship and higher studies in the field of Computer Science and Information Technology.

ix. Have commitment to ethical practices, societal contributions through communities and lifelong learning.

x. Possess better communication, presentation, time management and teamwork skills leading to responsible & competent professionals and will be able to address challenges in the field of IT at global level.

Program Specific Outcomes: PSOs

1. To provide a platform to the scholars for software development with new innovative multidisciplinary approach to serve the industry needs.
2. Apply ethical principles and plan to professional ethics and responsibilities and norms of the engineering practice.
3. To apply mathematical methodologies to unravel the real-world problems using appropriate mathematical analysis, arrangement and efficient computer algorithms.
4. To provide effective and efficient knowledge of recent technologies like web technologies, data science, machine learning, AI, Cyber security and IOT etc.

Program Outcomes: POs

POs are statements that describe what students are expected to know and be able to do upon graduating from the program. These relate to the skills, knowledge, analytical ability attitude and behavior that students acquire through the program.

a) An ability to apply knowledge of mathematics, computing, science, engineering and technology.

b) An ability to define a problem and provide a systematic solution with the help of conducting experiments, analyzing the problem and interpreting the data.

c) An ability to design, implement, and evaluate a software or a software/hardware system, component, or process to meet desired needs within realistic constraints.

d) An ability to identify, formulate, and provide systematic solutions to complex engineering/Technology problems.

e) An ability to use the techniques, skills, and modern engineering technology tools, standard processes necessary for practice as a IT professional.

f) An ability to apply mathematical foundations, algorithmic principles, and computer science theory in the modeling and design of computer-based systems with necessary constraints and assumptions.

g) An ability to analyze and provide solution for the local and global impact of information technology on individuals, organizations and society.

h) An ability to understand professional, ethical, legal, security and social issues and responsibilities.

i) An ability to function effectively as an individual or as a team member to accomplish a desired goal(s).

j) An ability to engage in life-long learning and continuing professional development to cope up with fast changes in the technologies/tools with the help of electives, professional organizations and extra-curricular activities.

k) An ability to communicate effectively in engineering community at large by means of effective presentations, report writing, paper publications, demonstrations.

l) An ability to understand engineering, management, financial aspects, performance, optimization and time complexity necessary for professional practice.

Sinhgad Institutes

Sinhgad Technical Educational Society's

SINHGAD INSTITUTE OF TECHNOLOGY

LONAVALA

CERTIFICATE

This is to certify that

Mr. / Miss _____ of Class TE – IT Roll No. _____ has completed all the practical work in the subject Data Science and Big Data Analytics Laboratory satisfactorily in the Department of Information Technology as prescribed by Savitribai Phule University of Pune, in the academic year 2023– 2024.

Subject In-charge	**Head of the Department**	**Principal**
Dr. P. C. Latane	Dr. R. V. Babar	Dr. M. S. Gaikwad

LIST OF ASSIGNMENT

Subject: Data Science and Big Data Analytics Laboratory

Ass. No.	Title	Exp. Date	Performed date	Marks
Group - A				
1	Single node/Multiple node Hadoop Installation.			
2	Design a distributed application using MapReduce (Using Java) which processes a log file of a system. List out the users who have logged for maximum period on the system. Use simple log file from the Internet and process it using a pseudo distribution mode on Hadoop platform.			
3	Write an application using HiveQL for flight information system which will include a. Creating, Dropping, and altering Database tables. b. Creating an external Hive table. c. Load table with data, insert new values and field in the table, Join tables with Hive d. Create index on Flight Information Table e. Find the average departure delay per day in 2008.			
Group - B				
1	Perform the following operations using Python on the Facebook metrics data sets a. Create data subsets b. Merge Data c. Sort Data d. Transposing Data e. Shape and reshape Data			
2	Perform the following operations using Python on the Air quality and Heart Diseases data sets			

	a. Data cleaning b. Data integration c. Data transformation d. Error correcting e. Data model building			
3	Integrate Python and Hadoop and perform the following operations on forest fire dataset a. Data analysis using the Map Reduce in PyHadoop b. Data mining in Hive			
4	Visualize the data using Python libraries matplotlib, seaborn by plotting the graphs for assignment no. 2 and 3 (Group B)			
5	Perform the following data visualization operations using Tableau on Adult and Iris datasets. a. 1D (Linear) Data visualization b. 2D (Planar) Data Visualization c. 3D (Volumetric) Data Visualization d. Temporal Data Visualization e. Multidimensional Data Visualization f. Tree/ Hierarchical Data visualization g. Network Data visualization			
Group - C				
1	Create a review scrapper for any ecommerce website to fetch real time comments, reviews, ratings, comment tags, customer name using Python.			
2	Develop a mini project in a group using different predictive models techniques to solve any real life problem. (Refer link dataset- https://www.kaggle.com/tanmoyie/us-graduate-schools- admission-parameters)			

Subject In-charge

Dr. P. C. Latane