



D Y PATIL
DEEMED TO BE
UNIVERSITY
— **RAMRAO ADIK** —
INSTITUTE OF TECHNOLOGY
NAVI MUMBAI

Department of Computer Science & Engineering

Programme: M. Tech in CSE (Artificial Intelligence)

Lab Manual

First Year Semester-I Course Name: Lab Practice-1

Odd Semester

Institutional Vision and Mission

Our Vision

To foster and permeate higher and quality education with value added engineering, technology programs, providing all facilities in terms of technology and platforms for all round development with societal awareness and nurture the youth with international competencies and exemplary level of employability even under highly competitive environment so that they are innovative adaptable and capable of handling problems faced by our country and world at large.

Our Mission

The Institution is committed to mobilize the resources and equip itself with men and materials of excellence thereby ensuring that the Institution becomes pivotal center of service to Industry, academia, and society with the latest technology. RAIT engages different platforms such as technology enhancing Student Technical Societies, Cultural platforms, Sports excellence centers, Entrepreneurial Development Center and Societal Interaction Cell. To develop the college to become an autonomous Institution & deemed university at the earliest with facilities for advanced research and development programs on par with international standards. To invite international and reputed national Institutions and Universities to collaborate with our institution on the issues of common interest of teaching and learning sophistication.

Our Quality Policy

ज्ञानधीनं जगत् सर्वम् ।
Knowledge is supreme.

Our Quality Policy

It is our earnest endeavour to produce high quality engineering professionals who are innovative and inspiring, thought and action leaders, competent to solve problems faced by society, nation and world at large by striving towards very high standards in learning, teaching and training methodologies.

Our Motto: If it is not of quality, it is NOT RAIT!

Departmental Vision and Mission

Vision

To excel in emerging fields of Computer Science and Engineering by imparting knowledge, practical skills, and core human values, transforming students into valuable contributors capable of driving innovation through advanced computing in real-world situations

Mission

M1: To promote academic excellence by providing a rigorous curriculum, encouraging critical thinking, and supporting ongoing learning in emerging fields, thereby contributing to the advancement of computing.

M2: To create a learning environment that prioritizes the practical application of knowledge, ethical conduct, and effective communication, preparing graduates to face the challenges of the constantly evolving tech industry.

M3: To broaden the scope of knowledge by supporting interdisciplinary research, fostering collaborations with industry and academic institutions, and promoting publication of research findings.

Departmental Program Educational Objectives (PEOs)

Program Educational Objectives (PEOs)

PEO 1:

Graduates will apply interdisciplinary approaches to develop and implement AI based solutions that require expertise in multiple domains

PEO 2:

Graduates will engage in lifelong learning to adapt to evolving AI and data science technologies while showcasing strong professional skills and ethical standards in leading industries and organizations.

PEO 3:

Graduates will be equipped to perform innovative research in the field of Data Sciences and Artificial Intelligence, significantly contributing to the advancement of knowledge and solving emerging challenges in the field.

Program Outcomes (POs)

PO1: Engineering knowledge: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

PO2: Problem analysis: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

PO3: Design/development of solutions: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

PO4: Conduct investigations of complex problems: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

PO5: Modern tool usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

PO6: The engineer and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

PO7: Environment and sustainability: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

PO8: Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

PO9: Individual and team work: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

PO10: Communication: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

PO11: Project management and finance: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

PO12: Life-long learning: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

Program Specific Outcomes (PSOs)

PSO 1:

To possess skill sets and expertise required to design, develop and deploy intelligent systems and data science solutions by continuously adapting and integrating new technologies.

PSO 2:

To effectively collaborate with professionals from diverse disciplines, integrating AI and Data Sciences methodologies into various domains to provide impactful solutions

PSO 3:

To design and actively participate in providing innovative solutions for data driven complex real-world problems with the help of statistical and computational techniques.

Index

Sr. No.	Contents	Page No.
1.	List of Experiments	07
2.	Course Objective, Course Outcome & Experiment Plan	08
3.	CO-PO Mapping & CO-PSO Mapping	09
4.	Study and Evaluation Scheme	10
5.	Experiment No. 1	11
6.	Experiment No. 2	16
7.	Experiment No. 3	19
8.	Experiment No. 4	23
9.	Experiment No. 5	28
10.	Experiment No. 6	32
11.	Experiment No. 7	35
12.	Experiment No. 8	40

List of Experiments

Sr. No.	Experiments Name
1	To implement Singular Value Decomposition (SVD)
2	Study of probability density functions (pdfs) and various statistics
3	Study of logistic regression
4	Performance analysis of multivariate linear regression
5	Exploratory Data Analysis using Seaborn
6	Study of ROC curve and AUC computation
7	Implementation of Principal Component Analysis
8	Case Study of algorithmic trading model for intraday trading on stock markets

Course Objective, Course Outcome & Experiment Plan

Course Objective:

1	To enable students to understand different techniques related to data science..
2	To make students become acquainted with numerical methods and matrix computation strategies.
3	To gain basic knowledge about the key algorithms and theory that forms the foundation of data science.

Course Outcomes:

CO1	Understand the fundamentals data analysis.
CO2	Analyze predictive modeling as applied to data science.
CO3	Analyze fitting model to data.
CO4	Understand model performance parameters and its use.
CO5	Apply and explore data visualization methods.
CO6	Analyze and create data analysis applications.

Experiment Plan

Sr. No.	Week No.	Experiments Name	Course Outcome	Weightage
1	W1	Study of Correlation and Scatter Plots	CO1	10
2	W2	Study of linear regression analysis	CO2	05
3	W3	Study of logistic regression	CO2	05
4	W4, W5	Performance analysis of multivariate linear regression	CO3	10
5	W6	Exploratory Data Analysis using Seaborn	CO4	10
6	W7, W8	Study of Data Visualization using RapidMiner	CO5	10
7	W9, W10	Study of ROC curve and AUC computation	CO6	04
8	W11 W12	Study of SVM classifier	CO6	04
9	W13	Case Study of algorithmic trading model for intraday trading on stock markets	CO6	02

Mapping of COs with POs:

Subject Weight	Course Outcomes (Weightage-100%)		Program Outcomes											
			PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12
PRACTICAL 100%	CO1	Understand the fundamentals data analysis.	1	1	1	1	2				1		2	1
	CO2	Analyze predictive modeling as applied to data science.		2	1		1	1			1	1	2	1
	CO3	Analyze fitting model to data.		2			2				1	1	2	2
	CO4	Understand model performance parameters and its use.		1	1		2				2		2	2
	CO5	Apply and explore data visualization methods.	1	1	1	1	2				1		1	2
	CO6	Analyze and create data analysis applications.		1	1	1	2				1		2	2

Mapping of COs with PSOs:

Course Outcomes		Program Specific Outcomes		
		PSO1	PSO2	PSO3
CO1	Understand the fundamentals data analysis.	3	1	1
CO2	Analyze predictive modeling as applied to data science.	2	1	2
CO3	Analyze fitting model to data.	2	1	2
CO4	Understand model performance parameters and its use.	3	1	2
CO5	Apply and explore data visualization methods.	2	2	1
CO6	Analyze and create data analysis applications.	3	2	2

Study and Evaluation Scheme

Experiment No.: 1

To implement Singular Value
Decomposition (SVD)

Experiment No. 1

1. **Aim:** To implement Singular Value Decomposition (SVD)
2. **Objectives:** a) To understand Singular Value Decomposition (SVD)
3. **Course Outcomes:** a) Apply fundamental mathematical concepts to analyze and solve problems in AI.
4. **Hardware / Software Required:** Python and related libraries
5. **Theory:**

Write theory about SVD
Explain few applications of SVD in detail.
Give analytical solution of finding SVD of given matrix
6. **Procedure and Algorithm:**
Algorithm:
Explain Step-by-step process or sequence of operations to be followed for SVD decomposition
And printing SVD matrices
Code/Design: Include Detailed code with output
7. **Results/Output Analysis:**
Students need to attach their programming codes/design/simulation results as part of the output. Discuss results
8. **Conclusions:** In this lab we implemented Singular Value Decomposition (SVD) using libraries like Numpy and scikit-learn. SVD is very useful for dimensionality reduction.

Experiment No. 2

1. **Aim:** Study of probability density functions (pdfs) and various statistics
2. **Objectives:** a) To understand probability density functions (pdfs) and various statistics
3. **Course Outcomes:** a) Apply probability theory and statistics to model uncertainty and analyze data in AI
4. **Hardware / Software Required:** Python and related libraries
5. **Theory:**

In theory include following:

A detailed explanation of various pdfs including Gaussian, Uniform, Exponential.

Give details of various statistics including mean, variance for above pdfs

Explain skewness and Kurtosis computation for pdfs

6. **Algorithm / Design / Procedure / Flowchart / Analysis:**

Algorithm: Explain Step-by-step process or sequence of operations to be followed for
Download Iris dataset

Read various variables, replace missing values

Plot histogram for all variables

Calculate mean, variance skewness and kurtosis for all variables

Find approximate pdf of the variable as per histogram (use approximate curve fitting in histogram and compare shape with pdfs).

Code/Design: Include detailed code

Procedure: write step-by-step guide on how to carry out the experiment.

7. **Results/Output Analysis:**

Include results in the form of figures

Include analysis of output

Attach programming codes and simulation results separately as part of the output.

8. **Conclusions:** Histogram is simplest way of finding approximate pdf of the data variables.
Various statistics are very useful to conclude about shape and distribution of pdfs.

Experiment No. 3

1. **Aim:** Study of logistic regression
2. **Objectives:** a) To understand logistic regression
b) To perform logistic regression
3. **Course Outcomes:** a) Apply probability theory and statistics to model uncertainty and analyze data in AI
4. **Hardware / Software Required:** Python and related libraries

5. **Theory:**

A detailed explanation of the theoretical details of

Logistic regression, Log odds

Logistic function

Fitting logistic curve

Maximum likelihood and Gradient decent methods

Multivariable logistic regression

6. **Algorithm / Design / Procedure / Flowchart / Analysis:**

Algorithm:

Generate a binary classification dataset with `n_samples=1000`, `n_features=20` or use
AirQualityUCI.csv data set

Split the dataset into training and testing sets

Train a Logistic Regression model

Predict the probabilities and classes for the test set

Calculate the accuracy, precision, recall, F1 score

Print values of performance parameters

Code/Design: Include Detailed code

7. **Results/Output Analysis:**

Explain results obtained for logistic regression algorithm.

8. **Conclusions:** Summarize the main findings of the logistic regression algorithm. Explain utility of logistic regression
9. **Viva Questions:** Explain Logistic function
Explain how linear regression is connected to logistic regression
Logistic regression is supervised or unsupervised algorithm, Explain, etc.

Experiment No. 4

1. **Aim:** Performance analysis of multivariate linear regression
2. **Objectives:** a) To understand model performance parameters
b) Analyze fitting model to data and performance of the model
3. **Course Outcomes:** a) Understand the principles of calculus and optimization for machine learning algorithms
4. **Hardware / Software Required:** Python and related libraries
5. **Theory:**

In theory include following:

A detailed explanation of the theoretical concept of multivariable linear regression

Give mathematical derivations for parameter estimation

Parameter estimation steps

Explain various performance parameters for regression analysis

6. **Algorithm / Design / Procedure / Flowchart / Analysis:**

Algorithm/Procedure: Explain Step-by-step process or sequence of operations to be followed

Import and load necessary packages

Load the AirQualityUCI dataset

Drop date and time parameters also drop na values from the dataset

Split the dataset into train and test data

Fit the linear regression model to data

Find various performance parameters and analyse the result

Code/Design: Include Detailed code

7. **Results/Output Analysis:**

Include analysis of output

Attach programming codes and simulation results separately as part of the output.

8. **Conclusions:** A multivariate linear regression is carried out on given data. It fits a straight line to observed data. Various performance parameters are computed and analysis is carried out.

- 9. Viva Questions:** A list of potential questions that might be asked in a viva examination related to the experiment.

What is multi variable linear regression

What are various performance parameters for regression analysis

How one can find best fit line

Experiment No. 5

1. **Aim:** Exploratory Data Analysis using Seaborn
2. **Objectives:** a) To explore seaborn capabilities
b) To explore various data visualization methods
3. **Course Outcomes:** a) Students will gain proficiency in using common AI programming languages.
4. **Hardware / Software Required:** Python and Seaborn library
5. **Theory:**

Give detailed explanation of following:
Describe various theoretical aspects in exploratory data analysis
Describe capabilities of Seaborn library
6. **Algorithm/Flowchart:**
Algorithm: Explain Step-by-step process or sequence of operations to be followed:
Load iris.csv data set
Display Data
Display data information
Calculate various statistics like min, max, mean, SD, IQR, Kurtosis, etc
Using Seaborn plot pairwise relationships in a dataset (use seaborn pairplot)
Note observations about correlations among various variables
Using kdeplot obtain density estimate plots for variables
Summarize about density distributions of various variables

Code/Design: Include Detailed code
7. **Results/Output Analysis:**

Attach programming codes/design/simulation results as part of the output.
8. **Conclusions:** Exploratory data analysis is carried out for iris dataset. Seaborn library is very much useful for carrying out exploratory data analysis.

Experiment No. 6

1. **Aim:** Study of ROC curve and AUC computation
2. **Objectives:** a) To understand model visualization methods
b) To plot ROC and compute AUC and their applications
3. **Course Outcomes:** a) Develop mathematical reasoning skills for analyzing AI algorithms and models.
4. **Hardware / Software Required:** Python/MATLAB and related libraries

5. **Theory:**

A detailed explanation of the theoretical concept of:

Model performance parameters

Explain details of ROC

Explain AUC and method for computing AUC

State applications of ROC and AUC

6. **Algorithm / Procedure**

Generate a binary classification dataset with n_samples=1000, n_features=20

Split the dataset into training and testing sets

Train a Logistic Regression model

Predict the probabilities and classes for the test set

Calculate the accuracy, precision, recall, F1 score, AUC - ROC score

Plot the ROC curve and compute AUC.

Code/Design: Include detailed code

7. **Results/Output Analysis:**

Include output and write about analysis of the results or output

Students should attach their programming codes and simulation results separately as part of the output.

8. **Conclusions:**

ROC curves are a very general way to represent and compare the performance of different models. The AUC can be calculated from the area under the receiver operator characteristic

(ROC) curve. ROC is a graphical plot that illustrates a binary classification system's diagnostic ability.

- 9. Viva Questions:** A list of potential questions that might be asked in a viva examination related to the experiment.

State various performance parameters for ML models

Explain ROC

State significance of AUC

What is lift curve

Experiment No. 7

1. **Aim:** Implementation of Principal Component Analysis
2. **Objectives:** a) To understand working of PCA
b) To explore use of PCA for various applications
3. **Course Outcomes:** a) Simulate various statistical algorithms that are widely used in AI
4. **Hardware / Software Required:** Python and related libraries
5. **Theory:**

A detailed explanation of the theoretical concept of PCA, computational aspects with proper illustration with graphs.

Include relevant equations, principles, and background information necessary to understand the PCA.

6. Algorithm / Procedure:

Algorithm:

1. Load the Iris Dataset
2. Standardize the data
1. Implement PCA
2. Project the 4 dimensional input data to 2 dimensions
5. Visualize 2D projection
7. Compute the explained variance

Code/Design: Include detailed code

Analysis: Explain the results and carry out analysis.

7. Results/Output Analysis:

Present and analyze the results and output

Students should attach their programming codes/design/simulation results separately as part of the output

8. **Conclusions:** Much of the variance is contained in first component and then in the remaining second component. Some of the variance, usually negligible is lost. But PCA reduces dimensions effectively and helps in data reduction greatly.

Experiment No. 8

1. **Aim:** Case Study of algorithmic trading model for intraday trading on stock markets
2. **Objectives:** a) To understand functioning of stock market
b) To explore technical analysis methods for stocks
3. **Course Outcomes:** a) To analyse and create AI applications
b) Analyze predictive modelling as applied to AI
4. **Hardware / Software Required:** Python/R and related libraries

5. **Theory:**

Students are provided with the reading material about algorithmic trading model for intraday trading on stock markets based on technical analysis methods. They are required to go through the material and study various methods used for the study.

Summarize various methods used and conclude about utility of the said algorithms for trading. You can add your perspectives/opinions in this lab work.

6. **Algorithm / Design / Procedure / Flowchart / Analysis:**

Summarize various methods used in the provided study material (see the link provided at the end)

7. **Results/Output Analysis:**

Present your observations about utility of technical analysis in stock market trading after analyzing the technical analysis

8. **Conclusions:** Summarize the main findings of the experiment

9. **Reference:**

Reading material URL

<https://core.ac.uk/download/pdf/79105923.pdf>

