Diploma Engineering

Laboratory Manual

(Introduction to Machine Learning)
(4350702)

[Computer Engineering, Semester V]

Enrolment No	
Name	
Branch	
Academic Term	
Institute	

Directorate Of Technical Education
Gandhinagar - Gujarat

DTE's Vision:

- To provide globally competitive technical education;
- Remove geographical imbalances and inconsistencies;
- Develop student friendly resources with a special focus on girls' education and support to weaker sections;
- Develop programs relevant to industry and create a vibrant pool of technical professionals.

DTE's Mission:

Institute's Vision:

Institute's Mission:

Department's Vision:

Department's Mission:

Certificate

This is to certify that Mr./Ms …….. Enrolment No. …….. of ……. Semester of *Diploma in Computer Engineering* of …….. (GTU Code) has satisfactorily completed the term work in course Introduction to Machine Learning (4350702) for the academic year: …… Term: Odd prescribed in the GTU curriculum.

Place:……..

Date: ……..

Signature of Course Faculty **Head of the Department**

Preface

The primary aim of any laboratory/Practical/field work is enhancement of required skills as well as creative ability amongst students to solve real time problems by developing relevant competencies in psychomotor domain. Keeping in view, GTU has designed competency focused outcome-based curriculum -2021 (COGC-2021) for Diploma engineering programmes. In this more time is allotted to practical work than theory. It shows importance of enhancement of skills amongst students and it pays attention to utilize every second of time allotted for practical amongst Students, Instructors and Lecturers to achieve relevant outcomes by performing rather than writing practice in study type. It is essential for effective implementation of competency focused outcome- based green curriculum-2021. Every practical has been keenly designed to serve as a tool to develop & enhance relevant industry needed competency in each and every student. These psychomotor skills are very difficult to develop through traditional chalk and board content delivery method in the classroom. Accordingly, this lab manual has been designed to focus on the industry defined relevant outcomes, rather than old practice of conducting practical to prove concept and theory.

By using this lab manual, students can read procedure one day in advance to actual performance day of practical experiment which generates interest and also, they can have idea of judgement of magnitude prior to performance. This in turn enhances predetermined outcomes amongst students. Each and every Experiment /Practical in this manual begins by competency, industry relevant skills, course outcomes as well as practical outcomes which serve as a key role for doing the practical. The students will also have a clear idea of safety and necessary precautions to be taken while performing experiment.

This manual also provides guidelines to lecturers to facilitate student-centred lab activities for each practical/experiment by arranging and managing necessary resources in order that the students follow the procedures with required safety and necessary precautions to achieve outcomes. It also gives an idea that how students will be assessed by providing Rubrics.

Introduction of machine learning course will help students to build up core competencies in understanding machine learning approaches and students will be able to design and train machine learning modes for various use cases. The lab work of the course is designed to develop crisp understanding of the underpinning theory.

Although we try our level best to design this lab manual, but always there are chances of improvement. We welcome any suggestions for improvement.

Programme Outcomes (POs):

1. **Basic and Discipline specific knowledge:** Apply knowledge of basic mathematics, science and engineering fundamentals and engineering specialization to solve the *engineering* problems.

2. **Problem analysis**: Identify and analyse well-defined *engineering* problems using codified standard methods.

3. **Design/ development of solutions:** Design solutions for *engineering* well-defined technical problems and assist with the design of systems components or processes to meet specified needs.

4. **Engineering Tools, Experimentation and Testing:** Apply modern *engineering* tools and appropriate technique to conduct standard tests and measurements.

5. **Engineering practices for society, sustainability and environment:** Apply appropriate technology in context of society, sustainability, environment and ethical practices.

6. **Project Management:** Use engineering management principles individually, as a team member or a leader to manage projects and effectively communicate about well-defined engineering activities.

7. **Life-long learning:** Ability to analyze individual needs and engage in updating in the context of technological changes *in field of engineering.*

Practical Outcome - Course Outcome matrix

Course Outcomes (COs):

a) *CO1: Describe basic concept of machine learning and its applications*
b) *CO2: Practice Numpy, Pandas, Matplotlib, sklearn library's inbuilt function required to solve machine learning problems*
c) *CO3: Use Pandas library for data preprocessing*
d) *CO4: Apply supervised learning algorithms based on dataset characteristics*
e) *CO5: Apply unsupervised learning algorithms based on dataset characteristics*

S. No.	Practical Outcome/Title of experiment	CO1	CO2	CO3	CO4	CO5
1.	Explore any one machine learning tool. (like Weka, Tensorflow, Scikit-learn, Colab, etc.)	✓				
2.	Write a NumPy program to implement following operation • to convert a list of numeric values into a one-dimensional NumPy array • to create a 3x3 matrix with values ranging from 2 to 10 • to append values at the end of an array • to create another shape from an array without changing its data (3*2 to 2*3)		✓			
3.	Write a NumPy program to implement following operation • to split an array of 14 elements into 3 arrays, each with 2, 4, and 8 elements in the original order • to stack arrays horizontally (column wise)		✓			
4.	Write a NumPy program to implement following operation • to add, subtract, multiply, divide arguments element-wise • to round elements of the array to the nearest integer • to calculate mean across dimension, in a 2D numpy array • to calculate the difference between neighboring elements, element-wise of a given array		✓			
5.	Write a NumPy program to implement following operation • to find the maximum and minimum value of a given flattened array		✓			

	• to compute the mean, standard deviation, and variance of a given array along the second axis					
6.	Write a Pandas program to implement following operation • to convert a NumPy array to a Pandas series • to convert the first column of a DataFrame as a Series • to create the mean and standard deviation of the data of a given Series • to sort a given Series		✓			
7.	Write a Pandas program to implement following operation • to create a dataframe from a dictionary and display it • to sort the DataFrame first by 'name' in ascending order • to delete the one specific column from the DataFrame • to write a DataFrame to CSV file using tab separator		✓			
8.	Write a Pandas program to create a line plot of the opening, closing stock prices of given company between two specific dates.		✓			
9.	Write a Pandas program to create a plot of Open, High, Low, Close, Adjusted Closing prices and Volume of given company between two specific dates.		✓			
10.	Write a Pandas program to implement following operation • to find and drop the missing values from the given dataset • to remove the duplicates from the given dataset			✓		
11.	Write a Pandas program to filter all columns where all entries present, check which rows and columns has a NaN and finally drop rows with any NaNs from the given dataset.			✓		
12.	Write a Python program using Scikit-learn to print the keys, number of rows-columns, feature names and the description of the given data.			✓		
13.	Write a Python program to implement K-Nearest Neighbour supervised machine learning algorithm for given dataset.				✓	

14.	Write a Python program to implement a machine learning algorithm for given dataset. (It is recommended to assign different machine learning algorithms group wise – micro project)					✓

Industry Relevant Skills

The following industry relevant skills are expected to be developed in the students by performance of experiments of this course.

a) *Student will learn to automate variety of task making system more efficient and cost effective*
b) *Student will learn efficient handling of data that will cater to better data analytics*
c) *Student will lean to implement machine learning approaches to varied field of applications from healthcare to e-commerce.*

Guidelines to Course Faculty

1. Couse faculty should demonstrate experiment with all necessary implementation strategies described in curriculum.
2. Couse faculty should explain industrial relevance before starting of each experiment.
3. Course faculty should involve & give opportunity to all students for hands on experience.
4. Course faculty should ensure mentioned skills are developed in the students by asking.
5. Utilise 2 hours of lab hours effectively and ensure completion of write up with quiz also.
6. Encourage peer to peer learning by doing same experiment through fast learners.

Instructions for Students

1. Organize the work in the group and make record of all observations.
2. Students shall develop maintenance skill as expected by industries.
3. Student shall attempt to develop related hand-on skills and build confidence.
4. Student shall develop the habits of evolving more ideas, innovations, skills etc.
5. Student shall refer technical magazines and data books.
6. Student should develop habit to submit the practical on date and time.
7. Student should well prepare while submitting write-up of exercise.

Continuous Assessment Sheet

Enrolment No: **Name:**

Term:

Sr no	Practical Outcome/Title of experiment	Page	Date	Marks (25)	Sign
1	Explore any one machine learning tool. (like Weka, Tensorflow, Scikit-learn, Colab, etc.)				
2	Write a NumPy program to implement following operation • to convert a list of numeric values into a one-dimensional NumPy array • to create a 3x3 matrix with values ranging from 2 to 10 • to append values at the end of an array • to create another shape from an array without changing its data (3*2 to 2*3)				
3	Write a NumPy program to implement following operation • to split an array of 14 elements into 3 arrays, each with 2, 4, and 8 elements in the original order • to stack arrays horizontally (column wise)				
4	Write a NumPy program to implement following operation • to add, subtract, multiply, divide arguments element-wise • to round elements of the array to the nearest integer • to calculate mean across dimension, in a 2D numpy array • to calculate the difference between neighboring elements, element-wise of a given array				
5	Write a NumPy program to implement following operation • to find the maximum and minimum value of a given flattened array • to compute the mean, standard deviation, and variance of a given array along the second axis				
6	Write a Pandas program to implement following operation • to convert a NumPy array to a Pandas series				

	• to convert the first column of a DataFrame as a Series • to create the mean and standard deviation of the data of a given Series • to sort a given Series				
7	Write a Pandas program to implement following operation • to create a dataframe from a dictionary and display it • to sort the DataFrame first by 'name' in ascending order • to delete the one specific column from the DataFrame • to write a DataFrame to CSV file using tab separator				
8	Write a Pandas program to create a line plot of the opening, closing stock prices of given company between two specific dates.				
9	Write a Pandas program to create a plot of Open, High, Low, Close, Adjusted Closing prices and Volume of given company between two specific dates.				
10	Write a Pandas program to implement following operation • to find and drop the missing values from the given dataset • to remove the duplicates from the given dataset				
11	Write a Pandas program to filter all columns where all entries present, check which rows and columns has a NaN and finally drop rows with any NaNs from the given dataset.				
12	Write a Python program using Scikit-learn to print the keys, number of rows-columns, feature names and the description of the given data.				
13	Write a Python program to implement K-Nearest Neighbour supervised machine learning algorithm for given dataset.				
14	Write a Python program to implement a machine learning algorithm for given dataset. (It is recommended to assign different machine learning algorithms group wise – micro project)				

Date: ……

Practical No.1: Explore any one machine learning tool. (like Weka, Tensorflow, Scikit-learn, Colab, etc.)

A. Objective

Getting familiarized with machine learning tools related to understand machine learning.

B. Expected Program Outcomes (POs)

PO1, PO4, PO6, PO7

C. Expected Skills to be developed based on competency

- Installing and configuring softwares as per the requirements.
- Programming skills.
- Debugging skills.

D. Expected Course Outcomes(Cos)

CO -1

E. Practical Outcome(PRo)

Understand different tools of machine learning.

F. Expected Affective domain Outcome(ADos)

- Handle tools /components/equipment carefully with safety and necessary precaution.
- In software ethics Environment sustainability and environment consciousness whenever suitable.

G. Prerequisite Theory:

- Refer Unit 1 of course curriculum
- https://www.javatpoint.com/machine-learning-tools
- https://www.simplilearn.com/best-machine-learning-tools-article

H. Resources/Equipment Required

Sr. No.	Instrument/Equipment with Broad Specifications
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM Python versions: 2.7.X, 3.6.X

2	Python IDEs and Code Editors (jupyter, spyder, google colab) Open Source: Anaconda Navigator
3	Internet Connection

I. Safety and necessary Precautions followed

- Read the experiment thoroughly before starting and ensure that you understand all the steps and concepts involved from underpinning theory.
- Keep the workspace clean and organized, free from clutter and unnecessary materials.
- Use the software according to its intended purpose and instructions.
- Ensure that all the necessary equipment and software are in good working condition.
- Never eat or drink in the lab, as it can cause contamination and create safety hazards.
- If any accidents or injuries occur, immediately notify the instructor and seek medical attention if necessary.

J. Source code:

Snapshot of installing Machine learning tool.

K. Practical related Quiz.

1. Differentiate between machine learning and human learning
2. List application of machine learning

L. References / Suggestions (lab manual designer should give)

https://www.simplilearn.com/best-machine-learning-tools-article
https://www.edureka.co/blog/top-machine-learning-tools/
https://www.mygreatlearning.com/blog/most-popular-machine-learning-tools you-should-know/
https://techvidvan.com/tutorials/machine-learning-tools/

M. Assessment-Rubrics

Criteria	Total Marks	Exceptional (5 - Marks)	Satisfactory (4 to 3 - Marks)	Developing (2 - Marks)	Limited (1 -Mark)
Engagement	5	Performed practical him/her self	Performed practical with others help	Watched other students performing practical but not tried him/her self	Present in practical session but not attentively participated in performance
Accuracy	5	Accurately done	1-2 errors/mistakes found	3-5 errors/mistakes identified	More than 5 errors/mistakes committed
Documentation	5	No errors, Program is well Executed and Documented Properly.	Complete write-up and output tables but presentation is poor	Some of the commands missing with missing outputs	Poor write-up and diagram or missing content
Understanding & Explanation	5	Fully understood the performance & can explain perfectly	Understood the performance but cannot explain	Partially understood the performance &can give little explanation	Partially understood and cannot give explanation
Time	5	Completed thework within 1 week	Work is submitted later than 1 week but by the end of 2ndweek	Work done after 2nd week but before the end of3rd week	Work submittedafter 3 week time
Total Marks:			**Signature with Date:**		

Date: ……

Practical No.2: Write a NumPy program to implement following operation

- to convert a list of numeric values into a one-dimensional NumPy array
- to create a 3x3 matrix with values ranging from 2 to 10
- to append values at the end of an array
- to create another shape from an array without changing its data (3*2 to 2*3)

A. Objective

Getting familiarized with python machine learning libraries.

B. Expected Program Outcomes (POs)

PO1, PO2, PO3, PO4, PO6, PO7

C. Expected Skills to be developed based on competency:

- Programming skills.
- Debugging skills.

D. Expected Course Outcomes(Cos)

CO -2

E. Practical Outcome(PRo)

Store and represent data using python libraries

F. Expected Affective domain Outcome(ADos)

- Handle tools /components/equipment carefully with safety and necessary precaution.
- In software ethics Environment sustainability and environment consciousness whenever suitable.

G. Prerequisite Theory

- Refer Unit 2 of course curriculum
- https://www.w3resource.com/python-exercises/numpy/python-numpy-exercise-3.php
- https://www.w3schools.com/python/numpy/default.asp
- https://www.javatpoint.com/numpy-tutorial

H. Resources/Equipment Required

Sr. No.	Instrument/Equipment with Broad Specifications
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM Python versions: 2.7.X, 3.6.X
2	Python IDEs and Code Editors (jupyter, spyder, google colab) Open Source: Anaconda Navigator

I. Safety and necessary Precautions followed

- Read the experiment thoroughly before starting and ensure that you understand all the steps and concepts involved from underpinning theory.
- Keep the workspace clean and organized, free from clutter and unnecessary materials.
- Use the software according to its intended purpose and instructions.
- Ensure that all the necessary equipment and software are in good working condition.
- Never eat or drink in the lab, as it can cause contamination and create safety hazards.
- If any accidents or injuries occur, immediately notify the instructor and seek medical attention if necessary.

J. Source code and Output:

- to convert a list of numeric values into a one-dimensional NumPy array

Output:

- to create a 3x3 matrix with values ranging from 2 to 10

Output:

- to append values at the end of an array

Output:

- to create another shape from an array without changing its data (3*2 to 2*3)

Output:

K. Practical related Quiz.

1. How are NumPy arrays better than Python's lists?
2. What are ndarrays in NumPy?

L. References / Suggestions (lab manual designer should give)

https://www.w3resource.com/python-exercises/numpy/python-numpy-exercise-2.php
https://www.edureka.co/blog/python-numpy-tutorial/
https://www.geeksforgeeks.org/python-numpy-practice-exercises-questions-and-solutions/
https://www.datasciencelearner.com/learn-python-library-numpy-practical-functions/

M. Assessment-Rubrics

Criteria	Total Marks	Exceptional (5 - Marks)	Satisfactory (4 to 3 - Marks)	Developing (2 - Marks)	Limited (1 -Mark)
Engagement	5	Performed practical him/her self	Performed practical with others help	Watched other students performing practical but not tried him/her self	Present in practical session but not attentively participated in performance
Accuracy	5	Accurately done	1-2 errors/mistakes found	3-5 errors/mistakes identified	More than 5 errors/mistakes committed
Documentation	5	No errors, Program is well Executed and Documented Properly.	Complete write-up and output tables but presentation is poor	Some of the commands missing with missing outputs	Poor write-up and diagram or missing content
Understanding & Explanation	5	Fully understood the performance & can explain perfectly	Understood the performance but cannot explain	Partially understood the performance &can give little explanation	Partially understood and cannot give explanation
Time	5	Completed thework within 1 week	Work is submitted later than 1 week but by the end of 2ndweek	Work done after 2nd week but before the end of3rd week	Work submittedafter 3 week time
Total Marks:			**Signature with Date:**		

Date: ……

Practical No.3: Write a NumPy program to implement following operation

- to split an array of 14 elements into 3 arrays, each with 2, 4, and 8 elements in the original order
- to stack arrays horizontally (column wise)

A. Objective

Getting familiarized with python machine learning libraries.

B. Expected Program Outcomes (POs)

PO1, PO2, PO3, PO4, PO6, PO7

C. Expected Skills to be developed based on competency:

- Programming skills.
- Debugging skills.

D. Expected Course Outcomes(Cos)

CO -2

E. Practical Outcome(PRo)

Split the data using python libraries

F. Expected Affective domain Outcome(ADos)

- Handle tools /components/equipment carefully with safety and necessary precaution.
- In software ethics Environment sustainability and environment consciousness whenever suitable.

G. Prerequisite Theory

- Refer Unit 2 of course curriculum
- https://www.w3resource.com/python-exercises/numpy/python-numpy-exercise-61.php
- https://www.w3schools.com/python/numpy/default.asp
- https://www.javatpoint.com/numpy-tutorial

H. Resources/Equipment Required

Sr. No.	Instrument/Equipment with Broad Specifications
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM Python versions: 2.7.X, 3.6.X
2	Python IDEs and Code Editors (jupyter, spyder, google colab) Open Source: Anaconda Navigator

I. Safety and necessary Precautions followed

- Read the experiment thoroughly before starting and ensure that you understand all the steps and concepts involved from underpinning theory.
- Keep the workspace clean and organized, free from clutter and unnecessary materials.
- Use the software according to its intended purpose and instructions.
- Ensure that all the necessary equipment and software are in good working condition.
- Never eat or drink in the lab, as it can cause contamination and create safety hazards.
- If any accidents or injuries occur, immediately notify the instructor and seek medical attention if necessary.

J. Source code and Output:

- to split an array of 14 elements into 3 arrays, each with 2, 4, and 8 elements in the original order

Output:

- to stack arrays horizontally (column wise)

Output:

K. Practical related Quiz.

1. Explain stack() function in Numpy.
2. Describe array_split() function in Numpy.

L. References / Suggestions (lab manual designer should give)

https://www.w3resource.com/python-exercises/numpy/python-numpy-exercise-61.php
https://www.edureka.co/blog/python-numpy-tutorial/
https://www.geeksforgeeks.org/python-numpy-practice-exercises-questions-and-solutions/
https://www.datasciencelearner.com/learn-python-library-numpy-practical-functions/

M. Assessment-Rubrics

Criteria	Total Marks	Exceptional (5 - Marks)	Satisfactory (4 to 3 - Marks)	Developing (2 - Marks)	Limited (1 -Mark)
Engagement	5	Performed practical him/her self	Performed practical with others help	Watched other students performing practical but not tried him/her self	Present in practical session but not attentively participated in performance
Accuracy	5	Accurately done	1-2 errors/mistakes found	3-5 errors/mistakes identified	More than 5 errors/mistakes committed
Documentation	5	No errors, Program is well Executed and Documented Properly.	Complete write-up and output tables but presentation is poor	Some of the commands missing with missing outputs	Poor write-up and diagram or missing content
Understanding & Explanation	5	Fully understood the performance & can explain perfectly	Understood the performance but cannot explain	Partially understood the performance &can give little explanation	Partially understood and cannot give explanation
Time	5	Completed thework within 1 week	Work is submitted later than 1 week but by the end of 2ndweek	Work done after 2nd week but before the end of3rd week	Work submittedafter 3 week time
Total Marks:			**Signature with Date:**		

Date: ……

Practical No.4: Write a NumPy program to implement following operation

- to add, subtract, multiply, divide arguments element-wise
- to round elements of the array to the nearest integer
- to calculate mean across dimension, in a 2D numpy array
- to calculate the difference between neighboring elements, element-wise of a given array

A. Objective

Getting familiarized with python machine learning libraries.

B. Expected Program Outcomes (POs)

PO1, PO2, PO3, PO4, PO6, PO7

C. Expected Skills to be developed based on competency:

- Programming skills.
- Debugging skills.

D. Expected Course Outcomes(Cos)

CO -2

E. Practical Outcome(PRo)

Used maths and Statistics functions on Data using Python libraries

F. Expected Affective domain Outcome(ADos)

- Handle tools /components/equipment carefully with safety and necessary precaution.
- In software ethics Environment sustainability and environment consciousness whenever suitable.

G. Prerequisite Theory

- Refer Unit 2 of course curriculum
- https://www.w3resource.com/python-exercises/numpy/python-numpy-math-exercise-1.php
- https://www.w3schools.com/python/numpy/default.asp
- https://www.javatpoint.com/numpy-tutorial

H. Resources/Equipment Required

Sr. No.	Instrument/Equipment with Broad Specifications
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM Python versions: 2.7.X, 3.6.X
2	Python IDEs and Code Editors (jupyter, spyder, google colab) Open Source: Anaconda Navigator

I. Safety and necessary Precautions followed

- Read the experiment thoroughly before starting and ensure that you understand all the steps and concepts involved from underpinning theory.
- Keep the workspace clean and organized, free from clutter and unnecessary materials.
- Use the software according to its intended purpose and instructions.
- Ensure that all the necessary equipment and software are in good working condition.
- Never eat or drink in the lab, as it can cause contamination and create safety hazards.
- If any accidents or injuries occur, immediately notify the instructor and seek medical attention if necessary.

J. Source code and Output:

- to add, subtract, multiply, divide arguments element-wise

Output:

- to round elements of the array to the nearest integer

Output:

- to calculate mean across dimension, in a 2D numpy array

Output:

- to calculate the difference between neighboring elements, element-wise of a given array

Output:

K. Practical related Quiz.

1. List mathematical functions on Numpy arrays.
2. List statistical functions on Numpy arrays.

L. References / Suggestions (lab manual designer should give)

https://www.w3resource.com/python-exercises/numpy/python-numpy-exercise-61.php
https://www.edureka.co/blog/python-numpy-tutorial/
https://www.geeksforgeeks.org/python-numpy-practice-exercises-questions-and-solutions/
https://www.datasciencelearner.com/learn-python-library-numpy-practical-functions/

M. Assessment-Rubrics

Criteria	Total Marks	Exceptional (5 - Marks)	Satisfactory (4 to 3 - Marks)	Developing (2 - Marks)	Limited (1 -Mark)
Engagement	5	Performed practical him/her self	Performed practical with others help	Watched other students performing practical but not tried him/her self	Present in practical session but not attentively participated in performance
Accuracy	5	Accurately done	1-2 errors/mistakes found	3-5 errors/mistakes identified	More than 5 errors/mistakes committed
Documentation	5	No errors, Program is well Executed and Documented Properly.	Complete write-up and output tables but presentation is poor	Some of the commands missing with missing outputs	Poor write-up and diagram or missing content
Understanding & Explanation	5	Fully understood the performance & can explain perfectly	Understood the performance but cannot explain	Partially understood the performance &can give little explanation	Partially understood and cannot give explanation
Time	5	Completed thework within 1 week	Work is submitted later than 1 week but by the end of 2ndweek	Work done after 2nd week but before the end of3rd week	Work submittedafter 3 week time
Total Marks:			**Signature with Date:**		

Date: ……

Practical No.5: Write a NumPy program to implement following operation

- to find the maximum and minimum value of a given flattened array
- to compute the mean, standard deviation, and variance of a given array along the second axis

A. Objective

Getting familiarized with python machine learning libraries.

B. Expected Program Outcomes (POs)

PO1, PO2, PO3, PO4, PO6, PO7

C. Expected Skills to be developed based on competency:

- Programming skills.
- Debugging skills.

D. Expected Course Outcomes(Cos)

CO -2

E. Practical Outcome(PRo)

Used maths and Statistics functions on Data using Python libraries

F. Expected Affective domain Outcome(ADos)

- Handle tools /components/equipment carefully with safety and necessary precaution.
- In software ethics Environment sustainability and environment consciousness whenever suitable.

G. Prerequisite Theory

- Refer Unit 2 of course curriculum
- https://www.w3resource.com/python-exercises/numpy/python-numpy-stat-exercise-1.php
- https://www.w3schools.com/python/numpy/default.asp
- https://www.javatpoint.com/numpy-tutorial

H. Resources/Equipment Required

Sr. No.	Instrument/Equipment with Broad Specifications
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM Python versions: 2.7.X, 3.6.X
2	Python IDEs and Code Editors (jupyter, spyder, google colab) Open Source: Anaconda Navigator

I. Safety and necessary Precautions followed

- Read the experiment thoroughly before starting and ensure that you understand all the steps and concepts involved from underpinning theory.
- Keep the workspace clean and organized, free from clutter and unnecessary materials.
- Use the software according to its intended purpose and instructions.
- Ensure that all the necessary equipment and software are in good working condition.
- Never eat or drink in the lab, as it can cause contamination and create safety hazards.
- If any accidents or injuries occur, immediately notify the instructor and seek medical attention if necessary.

J. Source code and Output:

- to find the maximum and minimum value of a given flattened array

Output:

- to compute the mean, standard deviation, and variance of a given array along the second axis

Output:

K. Practical related Quiz.

1. The value most recurrent in the sample set is known as _____.
   A. Mean
   B. Median
   C. Mode
   D. Standard Deviation

2. What is true about Statistics?
   A. Statistics is used to process complex problems in the real world
   B. Statistics is used to process simple problems in the virtual world
   C. Statistics is used to process simple problems in the real world
   D. None of the aboveList statistical functions on Numpy arrays.

3. What is the main role of Statistical functions, principles, and algorithms?
   A. to analyze raw data
   B. build a Statistical Model
   C. predict the result
   D. All of the above

4. A variable may also be called a _____.
   A. Data Set
   B. Data Item
   C. Data Value
   D. Data variable

5. Which language is commonly used with Statistics?
   A. C
   B. C++
   C. Ruby
   D. Python

6. Which of the following Numpy operation are correct?
   A. Mathematical and logical operations on arrays.
   B. Fourier transforms and routines for shape manipulation.
   C. Operations related to linear algebra.
   D. All of the above

L. References / Suggestions (lab manual designer should give)

https://www.w3resource.com/python-exercises/numpy/python-numpy-stat-exercise-7.php
https://www.edureka.co/blog/python-numpy-tutorial/
https://www.geeksforgeeks.org/python-numpy-practice-exercises-questions-and-solutions/
https://www.datasciencelearner.com/learn-python-library-numpy-practical-functions/

M. Assessment-Rubrics

Criteria	Total Marks	Exceptional (5 - Marks)	Satisfactory (4 to 3 - Marks)	Developing (2 - Marks)	Limited (1 -Mark)
Engagement	5	Performed practical him/her self	Performed practical with others help	Watched other students performing practical but not tried him/her self	Present in practical session but not attentively participated in performance
Accuracy	5	Accurately done	1-2 errors/mistakes found	3-5 errors/mistakes identified	More than 5 errors/mistakes committed
Documentation	5	No errors, Program is well Executed and Documented Properly.	Complete write-up and output tables but presentation is poor	Some of the commands missing with missing outputs	Poor write-up and diagram or missing content
Understanding & Explanation	5	Fully understood the performance & can explain perfectly	Understood the performance but cannot explain	Partially understood the performance &can give little explanation	Partially understood and cannot give explanation
Time	5	Completed thework within 1 week	Work is submitted later than 1 week but by the end of 2ndweek	Work done after 2nd week but before the end of3rd week	Work submittedafter 3 week time
Total Marks:			**Signature with Date:**		

Date: ……

Practical No.6: Write a Pandas program to implement following operation

- to convert a NumPy array to a Pandas series
- to convert the first column of a DataFrame as a Series
- to create the mean and standard deviation of the data of a given Series
- to sort a given Series

A. Objective

Getting familiarized with python machine learning libraries.

B. Expected Program Outcomes (POs)

PO1, PO2, PO3, PO4, PO6, PO7

C. Expected Skills to be developed based on competency:

- Programming skills.
- Debugging skills.

D. Expected Course Outcomes(Cos)

CO -2

E. Practical Outcome(PRo)

Used Pandas libraries for data analysis tasks in Python

F. Expected Affective domain Outcome(ADos)

- Handle tools /components/equipment carefully with safety and necessary precaution.
- In software ethics Environment sustainability and environment consciousness whenever suitable.

G. Prerequisite Theory

- Refer Unit 2 of course curriculum
- https://www.w3resource.com/python-exercises/pandas/python-pandas-data-series-exercise-6.php
- https://www.w3schools.com/python/pandas/pandas_intro.asp
- https://www.geeksforgeeks.org/introduction-to-pandas-in-python/

H. Resources/Equipment Required

Sr. No.	Instrument/Equipment with Broad Specifications
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM Python versions: 2.7.X, 3.6.X
2	Python IDEs and Code Editors (jupyter, spyder, google colab) Open Source: Anaconda Navigator

I. Safety and necessary Precautions followed

- Read the experiment thoroughly before starting and ensure that you understand all the steps and concepts involved from underpinning theory.
- Keep the workspace clean and organized, free from clutter and unnecessary materials.
- Use the software according to its intended purpose and instructions.
- Ensure that all the necessary equipment and software are in good working condition.
- Never eat or drink in the lab, as it can cause contamination and create safety hazards.
- If any accidents or injuries occur, immediately notify the instructor and seek medical attention if necessary.

J. Source code and Output:

- to convert a NumPy array to a Pandas series

Output:

- to convert the first column of a DataFrame as a Series

Output:

- to create the mean and standard deviation of the data of a given Series

Output:

- to sort a given Series

Output:

K. Practical related Quiz.

1. Describe Python Pandas.
2. How to import pandas and check the version?
3. How to create a series from a list, numpy array and dict?

L. References / Suggestions (lab manual designer should give)

https://www.w3resource.com/python-exercises/pandas/python-pandas-data-series-exercise-8.php
https://www.javatpoint.com/python-pandas
https://pandas.pydata.org/docs/getting_started/tutorials.html
https://www.kaggle.com/learn/pandas

M. Assessment-Rubrics

Criteria	Total Marks	Exceptional (5 - Marks)	Satisfactory (4 to 3 - Marks)	Developing (2 - Marks)	Limited (1 -Mark)
Engagement	5	Performed practical him/her self	Performed practical with others help	Watched other students performing practical but not tried him/her self	Present in practical session but not attentively participated in performance
Accuracy	5	Accurately done	1-2 errors/mistakes found	3-5 errors/mistakes identified	More than 5 errors/mistakes committed
Documentation	5	No errors, Program is well Executed and Documented Properly.	Complete write-up and output tables but presentation is poor	Some of the commands missing with missing outputs	Poor write-up and diagram or missing content
Understanding & Explanation	5	Fully understood the performance & can explain perfectly	Understood the performance but cannot explain	Partially understood the performance &can give little explanation	Partially understood and cannot give explanation
Time	5	Completed thework within 1 week	Work is submitted later than 1 week but by the end of 2ndweek	Work done after 2nd week but before the end of3rd week	Work submittedafter 3 week time
Total Marks:			**Signature with Date:**		

Date: ……

Practical No.7: Write a Pandas program to implement following operation

- to create a dataframe from a dictionary and display it
- to sort the DataFrame first by 'name' in ascending order
- to delete the one specific column from the DataFrame
- to write a DataFrame to CSV file using tab separator

A. Objective

Getting familiarized with python machine learning libraries.

B. Expected Program Outcomes (POs)

PO1, PO2, PO3, PO4, PO6, PO7

C. Expected Skills to be developed based on competency:

- Programming skills.
- Debugging skills.

D. Expected Course Outcomes(Cos)

CO -2

E. Practical Outcome(PRo)

Used Pandas libraries for data analysis tasks in Python

F. Expected Affective domain Outcome(ADos)

- Handle tools /components/equipment carefully with safety and necessary precaution.
- In software ethics Environment sustainability and environment consciousness whenever suitable.

G. Prerequisite Theory

- Refer Unit 2 of course curriculum
- https://www.w3resource.com/python-exercises/pandas/python-pandas-data-frame-exercise-1.php
- https://www.w3schools.com/python/pandas/pandas_intro.asp
- https://www.geeksforgeeks.org/introduction-to-pandas-in-python/

H. Resources/Equipment Required

Sr. No.	Instrument/Equipment with Broad Specifications
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM Python versions: 2.7.X, 3.6.X
2	Python IDEs and Code Editors (jupyter, spyder, google colab) Open Source: Anaconda Navigator

I. Safety and necessary Precautions followed

- Read the experiment thoroughly before starting and ensure that you understand all the steps and concepts involved from underpinning theory.
- Keep the workspace clean and organized, free from clutter and unnecessary materials.
- Use the software according to its intended purpose and instructions.
- Ensure that all the necessary equipment and software are in good working condition.
- Never eat or drink in the lab, as it can cause contamination and create safety hazards.
- If any accidents or injuries occur, immediately notify the instructor and seek medical attention if necessary.

J. Source code and Output:

- to create a dataframe from a dictionary and display it

Output:

- to sort the DataFrame first by 'name' in ascending order

Output:

- to delete the one specific column from the DataFrame

Output:

- to write a DataFrame to CSV file using tab separator

Output:

K. Practical related Quiz.

1. What are the significant features of the pandas Library?
2. Mention the different types of Data Structures in Pandas?
3. How can we sort the DataFrame?

L. References / Suggestions (lab manual designer should give)

https://www.w3resource.com/python-exercises/pandas/python-pandas-data-frame-exercise-16.php
https://www.interviewbit.com/pandas-interview-questions/
https://www.javatpoint.com/python-pandas
https://pandas.pydata.org/docs/getting_started/tutorials.html
https://www.kaggle.com/learn/pandas
https://www.educative.io/answers/how-to-delete-a-column-in-pandas

M. Assessment-Rubrics

Criteria	Total Marks	Exceptional (5 - Marks)	Satisfactory (4 to 3 - Marks)	Developing (2 - Marks)	Limited (1 -Mark)
Engagement	5	Performed practical him/her self	Performed practical with others help	Watched other students performing practical but not tried him/her self	Present in practical session but not attentively participated in performance
Accuracy	5	Accurately done	1-2 errors/mistakes found	3-5 errors/mistakes identified	More than 5 errors/mistakes committed
Documentation	5	No errors, Program is well Executed and Documented Properly.	Complete write-up and output tables but presentation is poor	Some of the commands missing with missing outputs	Poor write-up and diagram or missing content
Understanding & Explanation	5	Fully understood the performance & can explain perfectly	Understood the performance but cannot explain	Partially understood the performance &can give little explanation	Partially understood and cannot give explanation
Time	5	Completed thework within 1 week	Work is submitted later than 1 week but by the end of 2ndweek	Work done after 2nd week but before the end of3rd week	Work submittedafter 3 week time
Total Marks:			**Signature with Date:**		

Date: ……

Practical No.8: Write a Pandas program to create a line plot of the opening, closing stock prices of given company between two specific dates.

A. Objective

Getting familiarized with python machine learning libraries.

B. Expected Program Outcomes (POs)

PO1, PO2, PO3, PO4, PO6, PO7

C. Expected Skills to be developed based on competency:

- Programming skills.
- Debugging skills.

D. Expected Course Outcomes(Cos)

CO -2

E. Practical Outcome(PRo)

To plot various graphs from the CSV file in Python

F. Expected Affective domain Outcome(ADos)

- Handle tools /components/equipment carefully with safety and necessary precaution.
- In software ethics Environment sustainability and environment consciousness whenever suitable.

G. Prerequisite Theory

- Refer Unit 2 of course curriculum
- https://www.w3resource.com/python-exercises/pandas/plotting/pandas-plotting-exercise-2.php
- https://www.w3schools.com/python/matplotlib_plotting.asp
- https://www.geeksforgeeks.org/graph-plotting-in-python-set-1/
- https://www.geeksforgeeks.org/introduction-to-pandas-in-python/

H. Resources/Equipment Required

Sr. No.	Instrument/Equipment with Broad Specifications
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM Python versions: 2.7.X, 3.6.X
2	Python IDEs and Code Editors (jupyter, spyder, google colab) Open Source: Anaconda Navigator

I. Safety and necessary Precautions followed

- Read the experiment thoroughly before starting and ensure that you understand all the steps and concepts involved from underpinning theory.
- Keep the workspace clean and organized, free from clutter and unnecessary materials.
- Use the software according to its intended purpose and instructions.
- Ensure that all the necessary equipment and software are in good working condition.
- Never eat or drink in the lab, as it can cause contamination and create safety hazards.
- If any accidents or injuries occur, immediately notify the instructor and seek medical attention if necessary.

J. Source code and Output:

Write a Pandas program to create a line plot of the opening, closing stock prices of given company between two specific dates.

Output:

K. Practical related Quiz.

1. What is the difference between plt.show() and plt.savefig() in Matplotlib?
2. How can you set the font size of a plot in Matplotlib?
3. What is the difference between a scatter plot and a line plot in Matplotlib?
4. What is the purpose of the plt.grid() function in Matplotlib?
5. How can you change the color of a plot in Matplotlib?
6. Why matplotlib is used in machine learning?

L. References / Suggestions (lab manual designer should give)

https://www.w3resource.com/python-exercises/pandas/plotting/pandas-plotting-exercise-2.php
https://www.interviewbit.com/pandas-interview-questions/
https://www.javatpoint.com/python-pandas
https://testbook.com/interview/matplotlib-interview-questions
https://climbtheladder.com/matplotlib-interview-questions/

M. Assessment-Rubrics

Criteria	Total Marks	Exceptional (5 - Marks)	Satisfactory (4 to 3 - Marks)	Developing (2 - Marks)	Limited (1 -Mark)
Engagement	5	Performed practical him/her self	Performed practical with others help	Watched other students performing practical but not tried him/her self	Present in practical session but not attentively participated in performance
Accuracy	5	Accurately done	1-2 errors/mistakes found	3-5 errors/mistakes identified	More than 5 errors/mistakes committed
Documentation	5	No errors, Program is well Executed and Documented Properly.	Complete write-up and output tables but presentation is poor	Some of the commands missing with missing outputs	Poor write-up and diagram or missing content
Understanding & Explanation	5	Fully understood the performance & can explain perfectly	Understood the performance but cannot explain	Partially understood the performance &can give little explanation	Partially understood and cannot give explanation
Time	5	Completed thework within 1 week	Work is submitted later than 1 week but by the end of 2ndweek	Work done after 2nd week but before the end of3rd week	Work submittedafter 3 week time
Total Marks:			**Signature with Date:**		

Date: ……

Practical No.9: Write a Pandas program to create a plot of Open, High, Low, Close, Adjusted Closing prices and Volume of given company between two specific dates.

A. Objective

Getting familiarized with python machine learning libraries.

B. Expected Program Outcomes (POs)

PO1, PO2, PO3, PO4, PO6, PO7

C. Expected Skills to be developed based on competency:

- Programming skills.
- Debugging skills.

D. Expected Course Outcomes(Cos)

CO -2

E. Practical Outcome(PRo)

To plot various graphs from the CSV file in Python

F. Expected Affective domain Outcome(ADos)

- Handle tools /components/equipment carefully with safety and necessary precaution.
- In software ethics Environment sustainability and environment consciousness whenever suitable.

G. Prerequisite Theory

- Refer Unit 2 of course curriculum
- https://www.w3resource.com/python-exercises/pandas/plotting/pandas-plotting-exercise-13.php
- https://www.w3schools.com/python/matplotlib_plotting.asp
- https://www.geeksforgeeks.org/graph-plotting-in-python-set-1/
- https://www.geeksforgeeks.org/introduction-to-pandas-in-python/

H. Resources/Equipment Required

Sr. No.	Instrument/Equipment with Broad Specifications
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM Python versions: 2.7.X, 3.6.X
2	Python IDEs and Code Editors (jupyter, spyder, google colab) Open Source: Anaconda Navigator

I. Safety and necessary Precautions followed

- Read the experiment thoroughly before starting and ensure that you understand all the steps and concepts involved from underpinning theory.
- Keep the workspace clean and organized, free from clutter and unnecessary materials.
- Use the software according to its intended purpose and instructions.
- Ensure that all the necessary equipment and software are in good working condition.
- Never eat or drink in the lab, as it can cause contamination and create safety hazards.
- If any accidents or injuries occur, immediately notify the instructor and seek medical attention if necessary.

J. Source code and Output:

Write a Pandas program to create a plot of Open, High, Low, Close, Adjusted Closing prices and Volume of given company between two specific dates

Output:

K. Practical related Quiz.

1. What is Matplotlib?
2. What is the difference between pyplot and pylab?
3. What are some alternatives to Matplotlib?
4. What are some of the features provided by Matplotlib?
5. What are some of the major components of any graphs or plot?

L. References / Suggestions (lab manual designer should give)

https://www.w3resource.com/python-exercises/pandas/plotting/
https://www.interviewbit.com/pandas-interview-questions/
https://www.javatpoint.com/python-pandas
https://testbook.com/interview/matplotlib-interview-questions
https://climbtheladder.com/matplotlib-interview-questions/

M. Assessment-Rubrics

Criteria	Total Marks	Exceptional (5 - Marks)	Satisfactory (4 to 3 - Marks)	Developing (2 - Marks)	Limited (1 -Mark)
Engagement	5	Performed practical him/her self	Performed practical with others help	Watched other students performing practical but not tried him/her self	Present in practical session but not attentively participated in performance
Accuracy	5	Accurately done	1-2 errors/mistakes found	3-5 errors/mistakes identified	More than 5 errors/mistakes committed
Documentation	5	No errors, Program is well Executed and Documented Properly.	Complete write-up and output tables but presentation is poor	Some of the commands missing with missing outputs	Poor write-up and diagram or missing content
Understanding & Explanation	5	Fully understood the performance & can explain perfectly	Understood the performance but cannot explain	Partially understood the performance &can give little explanation	Partially understood and cannot give explanation
Time	5	Completed thework within 1 week	Work is submitted later than 1 week but by the end of 2ndweek	Work done after 2nd week but before the end of3rd week	Work submittedafter 3 week time
Total Marks:			**Signature with Date:**		

Date: ……

Practical No.10: Write a Pandas program to implement following operation

- to find and drop the missing values from the given dataset
- to remove the duplicates from the given dataset.

A. Objective

Getting familiarized with python machine learning libraries.

B. Expected Program Outcomes (POs)

PO1, PO2, PO3, PO4, PO6, PO7

C. Expected Skills to be developed based on competency:

- Programming skills.
- Debugging skills.

D. Expected Course Outcomes(Cos)

CO -3

E. Practical Outcome(PRo)

Understand data pre-processing using Python

F. Expected Affective domain Outcome(ADos)

- Handle tools /components/equipment carefully with safety and necessary precaution.
- In software ethics Environment sustainability and environment consciousness whenever suitable.

G. Prerequisite Theory

- Refer Unit 3 of course curriculum
- https://www.w3resource.com/pandas/dataframe/dataframe-dropna.php
- https://www.kdnuggets.com/2020/07/easy-guide-data-preprocessing-python.html
- https://www.geeksforgeeks.org/data-preprocessing-machine-learning-python/
- https://builtin.com/machine-learning/how-to-preprocess-data-python

H. Resources/Equipment Required

Sr. No.	Instrument/Equipment with Broad Specifications
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM Python versions: 2.7.X, 3.6.X
2	Python IDEs and Code Editors (jupyter, spyder, google colab) Open Source: Anaconda Navigator

I. Safety and necessary Precautions followed

- Read the experiment thoroughly before starting and ensure that you understand all the steps and concepts involved from underpinning theory.
- Keep the workspace clean and organized, free from clutter and unnecessary materials.
- Use the software according to its intended purpose and instructions.
- Ensure that all the necessary equipment and software are in good working condition.
- Never eat or drink in the lab, as it can cause contamination and create safety hazards.
- If any accidents or injuries occur, immediately notify the instructor and seek medical attention if necessary.

J. Source code and Output:

- to find and drop the missing values from the given dataset

Output:

- to remove the duplicates from the given dataset.

Output:

K. Practical related Quiz.

1. What is Data Preprocessing?
2. What are the missing values? and How do you handle missing values?
3. What are some common problems that occur during data processing? How can they be fixed?
4. Describe types of Data.
5. Explain confusion matrix.

L. References / Suggestions (lab manual designer should give)

https://www.javatpoint.com/data-preprocessing-machine-learning
https://climbtheladder.com/data-preprocessing-interview-questions/
https://avinashnavlani.medium.com/data-science-interview-questions-part-5-data-preprocessing-588d2c66ddd2
https://www.mlstack.cafe/interview-questions/data-processing
https://alvintoh.gitbook.io/machine-learning-a-z-hands-on-python-r-in-data-sci/section-2-part-1-data-preprocessing/quiz-1-data-preprocessing

M. Assessment-Rubrics

Criteria	Total Marks	Exceptional (5 - Marks)	Satisfactory (4 to 3 - Marks)	Developing (2 - Marks)	Limited (1 -Mark)
Engagement	5	Performed practical him/her self	Performed practical with others help	Watched other students performing practical but not tried him/her self	Present in practical session but not attentively participated in performance
Accuracy	5	Accurately done	1-2 errors/mistakes found	3-5 errors/mistakes identified	More than 5 errors/mistakes committed
Documentation	5	No errors, Program is well Executed and Documented Properly.	Complete write-up and output tables but presentation is poor	Some of the commands missing with missing outputs	Poor write-up and diagram or missing content
Understanding & Explanation	5	Fully understood the performance & can explain perfectly	Understood the performance but cannot explain	Partially understood the performance &can give little explanation	Partially understood and cannot give explanation
Time	5	Completed thework within 1 week	Work is submitted later than 1 week but by the end of 2ndweek	Work done after 2nd week but before the end of3rd week	Work submittedafter 3 week time
Total Marks:			**Signature with Date:**		

Date: ……

Practical No.11: Write a Pandas program to filter all columns where all entries present, check which rows and columns has a NaN and finally drop rows with any NaNs from the given dataset.

A. Objective

Getting familiarized with python machine learning libraries.

B. Expected Program Outcomes (POs)

PO1, PO2, PO3, PO4, PO6, PO7

C. Expected Skills to be developed based on competency:

- Programming skills.
- Debugging skills.

D. Expected Course Outcomes(Cos)

CO -3

E. Practical Outcome(PRo)

Understand data pre-processing using Python

F. Expected Affective domain Outcome(ADos)

- Handle tools /components/equipment carefully with safety and necessary precaution.
- In software ethics Environment sustainability and environment consciousness whenever suitable.

G. Prerequisite Theory

- Refer Unit 3 of course curriculum
- https://www.w3resource.com/python-exercises/pandas/filter/pandas-filter-exercise-25.php
- https://www.kdnuggets.com/2020/07/easy-guide-data-preprocessing-python.html
- https://www.geeksforgeeks.org/data-preprocessing-machine-learning-python/
- https://builtin.com/machine-learning/how-to-preprocess-data-python

H. Resources/Equipment Required

Sr. No.	Instrument/Equipment with Broad Specifications
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM Python versions: 2.7.X, 3.6.X
2	Python IDEs and Code Editors (jupyter, spyder, google colab) Open Source: Anaconda Navigator

I. Safety and necessary Precautions followed

- Read the experiment thoroughly before starting and ensure that you understand all the steps and concepts involved from underpinning theory.
- Keep the workspace clean and organized, free from clutter and unnecessary materials.
- Use the software according to its intended purpose and instructions.
- Ensure that all the necessary equipment and software are in good working condition.
- Never eat or drink in the lab, as it can cause contamination and create safety hazards.
- If any accidents or injuries occur, immediately notify the instructor and seek medical attention if necessary.

J. Source code and Output:

Write a Pandas program to filter all columns where all entries present, check which rows and columns has a NaN and finally drop rows with any NaNs from the given dataset.

Output:

K. Practical related Quiz.

1. Pandas is an open-source _____ Library?
   A. Ruby
   B. Javascript
   C. Java
   D. Python

2. All pandas data structures are ___ mutable but not always _____mutable.
   A. size, value
   B. semantic, size
   C. value, size
   D. none of the mentioned

3. Pandas key data structure is called?
   A. Keyframe
   B. DataFrame
   C. Statistics
   D. Econometrics

4. Which of the following is correct Features of DataFrame?
   A. Potentially columns are of different types
   B. Can Perform Arithmetic operations on rows and columns
   C. Labeled axes (rows and columns)
   D. All of the above

5. Which of the following thing can be data in Pandas?
   A. a python dict
   B. an ndarray
   C. a scalar value
   D. All of the above

6. PANDAS stands for _____
   A. Panel Data
   B. Panel Dashboard
   C. Panel Data analyst
   D. Panel Data

L. References / Suggestions (lab manual designer should give)

https://www.javatpoint.com/data-preprocessing-machine-learning
https://climbtheladder.com/data-preprocessing-interview-questions/
https://avinashnavlani.medium.com/data-science-interview-questions-part-5-data-preprocessing-588d2c66ddd2
https://www.mlstack.cafe/interview-questions/data-processing
https://alvintoh.gitbook.io/machine-learning-a-z-hands-on-python-r-in-data-sci/section-2-part-1-data-preprocessing/quiz-1-data-preprocessing

M. Assessment-Rubrics

Criteria	Total Marks	Exceptional (5 - Marks)	Satisfactory (4 to 3 - Marks)	Developing (2 - Marks)	Limited (1 -Mark)
Engagement	5	Performed practical him/her self	Performed practical with others help	Watched other students performing practical but not tried him/her self	Present in practical session but not attentively participated in performance
Accuracy	5	Accurately done	1-2 errors/mistakes found	3-5 errors/mistakes identified	More than 5 errors/mistakes committed
Documentation	5	No errors, Program is well Executed and Documented Properly.	Complete write-up and output tables but presentation is poor	Some of the commands missing with missing outputs	Poor write-up and diagram or missing content
Understanding & Explanation	5	Fully understood the performance & can explain perfectly	Understood the performance but cannot explain	Partially understood the performance &can give little explanation	Partially understood and cannot give explanation
Time	5	Completed thework within 1 week	Work is submitted later than 1 week but by the end of 2ndweek	Work done after 2nd week but before the end of3rd week	Work submittedafter 3 week time
Total Marks:			**Signature with Date:**		

Date: ……

Practical No.12: Write a Python program using Scikit-learn to print the keys, number of rows-columns, feature names and the description of the given data.

A. Objective

Getting familiarized with python machine learning libraries.

B. Expected Program Outcomes (POs)

PO1, PO2, PO3, PO4, PO6, PO7

C. Expected Skills to be developed based on competency:

- Programming skills.
- Debugging skills.

D. Expected Course Outcomes(Cos)

CO -3

E. Practical Outcome(PRo)

Understand data pre-processing using Python

F. Expected Affective domain Outcome(ADos)

- Handle tools /components/equipment carefully with safety and necessary precaution.
- In software ethics Environment sustainability and environment consciousness whenever suitable.

G. Prerequisite Theory

- Refer Unit 3 of course curriculum
- https://www.w3resource.com/machine-learning/scikit-learn/iris/python-machine-learning-scikit-learn-iris-basic-exercise-2.php
- https://www.kdnuggets.com/2020/07/easy-guide-data-preprocessing-python.html
- https://www.geeksforgeeks.org/data-preprocessing-machine-learning-python/
- https://builtin.com/machine-learning/how-to-preprocess-data-python

H. Resources/Equipment Required

Sr. No.	Instrument/Equipment with Broad Specifications
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM Python versions: 2.7.X, 3.6.X
2	Python IDEs and Code Editors (jupyter, spyder, google colab) Open Source: Anaconda Navigator

I. Safety and necessary Precautions followed

- Read the experiment thoroughly before starting and ensure that you understand all the steps and concepts involved from underpinning theory.
- Keep the workspace clean and organized, free from clutter and unnecessary materials.
- Use the software according to its intended purpose and instructions.
- Ensure that all the necessary equipment and software are in good working condition.
- Never eat or drink in the lab, as it can cause contamination and create safety hazards.
- If any accidents or injuries occur, immediately notify the instructor and seek medical attention if necessary.

J. Source code and Output:

Write a Python program using Scikit-learn to print the keys, number of rows-columns, feature names and the description of the given data.

Output:

K. Practical related Quiz.

1. Define Dimensionality reduction.
2. Describe Feature subset selection.
3. How to split data into 3 sets (train, validation and test)?

L. References / Suggestions (lab manual designer should give)

https://www.w3resource.com/machine-learning/scikit-learn/iris/python-machine-learning-scikit-learn-iris-basic-exercise-1.php
https://www.javatpoint.com/data-preprocessing-machine-learning
https://climbtheladder.com/data-preprocessing-interview-questions/
https://avinashnavlani.medium.com/data-science-interview-questions-part-5-data-preprocessing-588d2c66ddd2
https://www.mlstack.cafe/interview-questions/data-processing

M. Assessment-Rubrics

Criteria	Total Marks	Exceptional (5 - Marks)	Satisfactory (4 to 3 - Marks)	Developing (2 - Marks)	Limited (1 -Mark)
Engagement	5	Performed practical him/her self	Performed practical with others help	Watched other students performing practical but not tried him/her self	Present in practical session but not attentively participated in performance
Accuracy	5	Accurately done	1-2 errors/mistakes found	3-5 errors/mistakes identified	More than 5 errors/mistakes committed
Documentation	5	No errors, Program is well Executed and Documented Properly.	Complete write-up and output tables but presentation is poor	Some of the commands missing with missing outputs	Poor write-up and diagram or missing content
Understanding & Explanation	5	Fully understood the performance & can explain perfectly	Understood the performance but cannot explain	Partially understood the performance &can give little explanation	Partially understood and cannot give explanation
Time	5	Completed thework within 1 week	Work is submitted later than 1 week but by the end of 2ndweek	Work done after 2nd week but before the end of3rd week	Work submittedafter 3 week time
Total Marks:			**Signature with Date:**		

Date: ……

Practical No.13: Write a Python program to implement K-Nearest Neighbour supervised machine learning algorithm for given dataset.

A. Objective

Understand supervised machine learning concept.

B. Expected Program Outcomes (POs)

PO1, PO2, PO3, PO4, PO5, PO6, PO7

C. Expected Skills to be developed based on competency:

- Programming skills.
- Debugging skills.

D. Expected Course Outcomes(Cos)

CO - 4

E. Practical Outcome(PRo)

Apply K-Nearest Neighbour supervised machine learning algorithm for given dataset

F. Expected Affective domain Outcome(ADos)

- Handle tools /components/equipment carefully with safety and necessary precaution.
- In software ethics Environment sustainability and environment consciousness whenever suitable.

G. Prerequisite Theory

- Refer Unit 4 of course curriculum
- https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning
- https://www.geeksforgeeks.org/k-nearest-neighbor-algorithm-in-python/
- https://www.analyticsvidhya.com/blog/2021/01/a-quick-introduction-to-k-nearest-neighbor-knn-classification-using-python/
- https://realpython.com/knn-python/

H. Resources/Equipment Required

Sr. No.	Instrument/Equipment with Broad Specifications
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM Python versions: 2.7.X, 3.6.X
2	Python IDEs and Code Editors (jupyter, spyder, google colab) Open Source: Anaconda Navigator

I. Safety and necessary Precautions followed

- Read the experiment thoroughly before starting and ensure that you understand all the steps and concepts involved from underpinning theory.
- Keep the workspace clean and organized, free from clutter and unnecessary materials.
- Use the software according to its intended purpose and instructions.
- Ensure that all the necessary equipment and software are in good working condition.
- Never eat or drink in the lab, as it can cause contamination and create safety hazards.
- If any accidents or injuries occur, immediately notify the instructor and seek medical attention if necessary.

J. Source code and Output:

Write a Python program to implement K-Nearest Neighbour supervised machine learning algorithm for given dataset.

Output:

K. Practical related Quiz.

1. Define steps in Supervised Machine learning
2. List types of Supervised Learning
3. Write advantage and disadvantage of supervised machine learning.

L. References / Suggestions (lab manual designer should give)

https://www.w3schools.com/python/python_ml_knn.asp
https://www.digitalocean.com/community/tutorials/k-nearest-neighbors-knn-in-python
https://scikit-learn.org
/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html
https://www.datacamp.com/tutorial/k-nearest-neighbor-classification-scikit-learn
https://towardsdatascience.com/knn-using-scikit-learn-c6bed765be75
https://github.com/chingisooinar/KNN-python-implementation/blob/main/k_nearest_neighbors_from_scratch.ipynb

M. Assessment-Rubrics

Criteria	Total Marks	Exceptional (5 - Marks)	Satisfactory (4 to 3 - Marks)	Developing (2 - Marks)	Limited (1 -Mark)
Engagement	5	Performed practical him/her self	Performed practical with others help	Watched other students performing practical but not tried him/her self	Present in practical session but not attentively participated in performance
Accuracy	5	Accurately done	1-2 errors/mistakes found	3-5 errors/mistakes identified	More than 5 errors/mistakes committed
Documentation	5	No errors, Program is well Executed and Documented Properly.	Complete write-up and output tables but presentation is poor	Some of the commands missing with missing outputs	Poor write-up and diagram or missing content
Understanding & Explanation	5	Fully understood the performance & can explain perfectly	Understood the performance but cannot explain	Partially understood the performance &can give little explanation	Partially understood and cannot give explanation
Time	5	Completed thework within 1 week	Work is submitted later than 1 week but by the end of 2ndweek	Work done after 2nd week but before the end of3rd week	Work submittedafter 3 week time
Total Marks:			**Signature with Date:**		

Date: ……

Practical No.14: Write a Python program to implement a machine learning algorithm for given dataset. (It is recommended to assign different machine learning algorithms group wise – micro project)

A. Objective

Understand different machine learning algorithms.

B. Expected Program Outcomes (POs)

PO1, PO2, PO3, PO4, PO5, PO6, PO7

C. Expected Skills to be developed based on competency:

- Programming skills.
- Debugging skills.

D. Expected Course Outcomes(Cos)

CO - 5

E. Practical Outcome(PRo)

Apply machine learning algorithm for given dataset

F. Expected Affective domain Outcome(ADos)

- Handle tools /components/equipment carefully with safety and necessary precaution.
- In software ethics Environment sustainability and environment consciousness whenever suitable.

G. Prerequisite Theory

- Refer Unit 5 of course curriculum
- https://www.geeksforgeeks.org/machine-learning-with-python/
- https://www.javatpoint.com/machine-learning-algorithms
- https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/
- https://www.coursera.org/articles/machine-learning-algorithms

H. Resources/Equipment Required

Sr. No.	Instrument/Equipment with Broad Specifications
1	Computer system with operating system: Windows 7 or higher Ver., macOS, and Linux, with 4GB or higher RAM Python versions: 2.7.X, 3.6.X
2	Python IDEs and Code Editors (jupyter, spyder, google colab) Open Source: Anaconda Navigator

I. Safety and necessary Precautions followed

- Read the experiment thoroughly before starting and ensure that you understand all the steps and concepts involved from underpinning theory.
- Keep the workspace clean and organized, free from clutter and unnecessary materials.
- Use the software according to its intended purpose and instructions.
- Ensure that all the necessary equipment and software are in good working condition.
- Never eat or drink in the lab, as it can cause contamination and create safety hazards.
- If any accidents or injuries occur, immediately notify the instructor and seek medical attention if necessary.

J. Source code and Output:

Write a Python program to implement a machine learning algorithm for given dataset. (It is recommended to assign different machine learning algorithms group wise – micro project)

Output:

K. Practical related Quiz.

1. Define real world examples of unsupervised Learning
2. List types of Supervised Learning
3. Write advantage and disadvantage of un supervised machine learning.
4. Differentiate Supervised and Unsupervised Learning

L. References / Suggestions (lab manual designer should give)

https://www.w3schools.com/python/python_ml_knn.asp
https://data-flair.training/blogs/machine-learning-algorithms-in-python/
https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/
https://github.com/topics/machine-learning-projects
https://www.datacamp.com/tutorial/k-nearest-neighbor-classification-scikit-learn
https://towardsdatascience.com/knn-using-scikit-learn-c6bed765be75

M. Assessment-Rubrics

Criteria	Total Marks	Exceptional (5 - Marks)	Satisfactory (4 to 3 - Marks)	Developing (2 - Marks)	Limited (1 -Mark)
Engagement	5	Performed practical him/her self	Performed practical with others help	Watched other students performing practical but not tried him/her self	Present in practical session but not attentively participated in performance
Accuracy	5	Accurately done	1-2 errors/mistakes found	3-5 errors/mistakes identified	More than 5 errors/mistakes committed
Documentation	5	No errors, Program is well Executed and Documented Properly.	Complete write-up and output tables but presentation is poor	Some of the commands missing with missing outputs	Poor write-up and diagram or missing content
Understanding & Explanation	5	Fully understood the performance & can explain perfectly	Understood the performance but cannot explain	Partially understood the performance &can give little explanation	Partially understood and cannot give explanation
Time	5	Completed thework within 1 week	Work is submitted later than 1 week but by the end of 2ndweek	Work done after 2nd week but before the end of3rd week	Work submittedafter 3 week time
Total Marks:			**Signature with Date:**		

Introduction to Machine Learning

4350702

Lab manuals are prepared by

Miss. Kumundrini B. Prajapati
Lecturer
Government Polytechnic Gandhinagar

Ms. Jasmine J. Karagthala
Lecturer
Government Polytechnic for Girls, Ahmedabad

Branch Coordinator

Mr. B. H. Kantevala

Head of the Department

Computer Engineering

Government Polytechnic Ahmedabad

Committee Chairman

Mr. R. D. Raghani

(HOD-EC)

Principal (I/C)

Government Polytechnic, Gandhinagar