## **Artificial Intelligence and Machine Learning**

Project Report

Semester-IV (Batch-2022)

**Case Study**: - SQL Worksheet

https://drive.google.com/file/d/1g01OoUidT6ZWI5xMxQDxUP2GzEwDDM31/view?usp=sharing

A red and white sign

Description automatically generated with low confidence

|  |  |
| --- | --- |
| **Supervised By:**  **Rajeev Bhardwaj** | **Submitted By:**  **Vanshika**  **Roll Number: 2210990945**  **Group: 14** |

**Department of Computer Science and Engineering**

## **Chitkara University Institute of Engineering & Technology,**

## **Chitkara University, Punjab**

**Description about Case Study: -**

* Read Sql File
* Display Top 5 Rows Of The Dataset
* Check Last 3 Rows Of The Dataset
* Find Shape Of Our Dataset(Number Of Rows & Number Of Columns)
* Get Information About Our Dataset Like Total Number Rows,Total Number Of Columns,Datatype Of Each Column And Memory Requirement
* Get Overall Statistics About The Dataframe
* Data Filtering
* Check Null Values In The Dataset
* Drop The Column
* Handle Missing Values
* Categorical Data Encoding
* What Is Univariate Analysis ?
* How Many People Survived And How Many Died Plot On Graph
* How Many Passengers Were In First Class,Second Class And Third Class Plot Those Figures On Graph
* Number Of Male And Female Passengers
* Bivariant Analysis
* Who Has Better Chance Of Survival Male Or Female
* Which Passenger Class Has Better Chance Of Survival (First,Second,Or Third Class) ?
* Feature Engineering

**Library: -**

* Pandas
* mysql.connector
* MatPlotLib
* matplotlib.pyplot
* Seaborn

**Methods: -**

1. sql.connect():

Description: Establishes a connection to a SQL database allowing interaction and execution of SQL queries from python.

2. head():

Description: Display the first few rows of the dataframe.

3. tail():

Description : Display the last few rows of the dataframe.

4. shape():

Description: Returns the shape (number of rows and columns) of the dataframe.

5. info():

Description: Provide information about the dataframes such as rows, columns, data types and missing values.

6. describe():

Description: It generates descriptive statistics of the numerical columns of the dataframe.

7. filter():

Description: It is used to select or filter specific columns from dataframes based on their labels or column names.

8. isnull():

Description: Returns true or false for each value in the dataframe, indicating if the value is missing or not.

9. drop():

Description: Remove specific rows or columns from the dataframes.

10. sum():

Description: Calculates the sum of values along a specified axis in a dataframe.

11. get\_dummies():

Description: Used to convert categorical variables into dummy/indicator variables.

12. plt.figure():

Description: Used to create a new figure or modify the properties of an existing figure.

13. plt.title():

Description: Sets the title of the current axis or a subplot in a figure

14. plt.xlabel():

Description: Sets the labels for the x-axis of the current axis or a subplot in a figure.

15. plt.ylabel():

Description: Sets the labels for the y-axis of the current axis or a subplot in a figure

16. plt.grid():

Description: Adds grid lines to the current axis or a subplot in a figure.

17. plt.show():

Description: Displays the current figure that has been created or modified.

18. sns.histplot():

Description: Create a histogram to visualize the distribution of numerical data.

19. add\_axes():

Description: Adds an axis to the current figure at the specified position and with the specified size.

20. axis():

Description: Used to set or retrieve the current axis limits for the x- and y-axes in a plot

21. pie():

Description: Create a pie chart representing the proportions of categorical data as slices of a circle.

22. dropna():

Description: Used to remove missing values (NaN) from a dataframe or a series.

23. groupby():

Description: Used to group data from a dataframe according to specified criteria and perform operation on the resulting groups.

24. subplots():

Description: Creates a grid of subplots within a single figure allowing multiple plots to be displayed together.

25. set\_title():

Description: Sets the title of the current axis or the subplot in a figure.