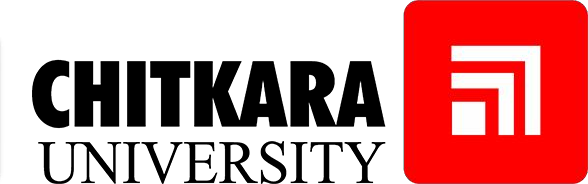
Artificial Intelligence and Machine Learning

Project Report Semester-IV (Batch-2022)

Case Study: Train data with MySQL

<https://drive.google.com/drive/folders/1Sm6rbmld5P4tQH47DaoeEhDHabuoD4LC?usp=sharing>



**Supervised By: Submitted By:**

Rajeev Bhardwaj Sparsh Mittal

2210990865

G-13

**Department of Computer Science and Engineering**

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

**Chitkara University, Punjab**

**DESCRIPTION ABOUT THE CASE-STUDY:-**

* Read dataset of train
* Display top 5 rows of dataset
* Display last 5 rows of dataset
* Shape of the dataset
* Getting information about our dataset
* Overall statistic about the dataset
* Data filtering
* Check null value in the dataset
* Drop the column
* Handle missing values
* Categorical data encoding
* Univariate analysis
* How many people survived and plot graph( for died)
* How many pclass were in 1st class, 2nd class, 3rd class plot graph
* Number of male and female
* Bivariant analysis
* Who has better chance of survival male or female
* Which passenger class has better chance of survival

**LIBRARY:**

* Pandas
* Seaborn
* Matplotlib

**METHODS:**

1. Read dataset: use read\_csv()
2. Top 5 rows: head()
3. Shape of dataset: use shape()
4. Getting information about the dataset: use .info()
5. Overall statistic of the dataset: use .describe()
6. Data filtering: use .unique()
7. Check null value in dataset: use .isnull()
8. Drop a column, Handle missing values: use .dropna()
9. Categorical data encoding: use .pd.get\_dummies()
10. Univariate analysis: use sns.boxplot()
11. Only find out survived people: use value\_counts()[1:]
12. To figure out that how many passenger in differ class: use sns.countplot()
13. Bivariant analysis: use sns.heatmap()
14. To find better one: use library plt
15. Plt.figure(figsize=()) - for size
16. Plt.xlabel() - for giving label to x-axis
17. Plt.ylabel() - for giving label to y-axis
18. Plt.title() - give title to graph
19. Plt.bar() – to make graph