



Internships Daily Diary (2025-26)

Week No: 4

Day: Friday

Date: 27-6-25

Today, I learned about Data processing or Data preprocessing in machine learning. Data preprocessing is a process of preparing the raw data & making suitable for a machine learning & steps for it that are :-

- i) Getting the dataset
- ii) Importing libraries
- iii) Importing dataset
- iv) Finding missing data
- v) Encoding categorical data
- vi) splitting dataset into training & test set
- vii) Feature scaling

I also learned about Encoding that's convert categorical data into numerical format so algorithms in ML can process it. There are three types of encoding

- i) Label encoding
- ii) One-Hot encoding
- iii) Ordinal encoding

2025/07/01 20:16

Internships Daily Diary (2025-26)

Week No: 4

Day: Thursday

Date: 26-6-25

Today, I solved task3 on EDA which stands for Exploratory Data Analysis. In this task first I import libraries which support to my code that's are numpy, pandas, matplotlib & seaborn. Then I import dataset using `read()`.

After importing dataset I run some basic functions like `info()`, `head()`, `tail()` then I start to handle data column by column. In this task I use count plot, histogram plot, box plot, bar plot, violin plot, heatmap plot & scatter plot. I also learn how to give interpretation on Jupyter notebook.



Internships Daily Diary (2025-26)

Week No: 4

Day: Wednesday

Date: 25-6-25

Today, I solved task 1 & task 2 on data cleaning both tasks are similar to done but datasets are different. First I import libraries which will support to code & help to done data cleaning the libraries are numpy, pandas, matplotlib, seaborn then I import dataset using read().

Then using methods like info(), describe(), drop_duplicates(inplace=True), etc. Then I handle one by one all columns which was need to clean like handling duplicates, filling missing values, standerdized columns, filling empty cells after all this I save cleaned data into new file.



Internships Daily Diary (2025-26)

Week No: 4

Day: Tuesday

Date: 24-6-25

Today, I learned all remaining plots in EDA that's one violin plot it made by Box plot + KDE plot, scatter plot it use to shows interconnection between two variables, pair plot it shows relationship between all feature in dataset, Heatmap plot used for correlation visualization &/or finding correlation between numerical variables, pie plot it's a part-to-whole relationship, swarm plot it is categorical scatter plot, strip plot, Joint plot & last one is Rug plot it used to show density within datapoints.

I also learned about colorbands like coolwarm, Diverging, cyclic, categorical & sequential.



Internships Daily Diary (2025-26)

Week No: 4

Day: Monday

Date: 23-6-25

Today, I learned about EDA & its type. EDA stands for Exploratory Data analysis. It visualizing data to understand its main feature, find patterns & discover how different parts of the data are connected. Its types are Univariate analysis, Bivariate analysis & multivariate analysis.

After theory part I studied some plots in EDA that are Line plot, Bar plot, Count plot, Histogram plot, KDE plot, Hist plot & Box plot I studied these plots with its example & also run it on Jupyter notebook.



Sadguru Balumama Shikshan Prasarak Mandal's
K.P.PATIL INSTITUTE OF TECHNOLOGY
(DTE CODE : 6814) (MSBTE Code : 1661)
Approved by AICTE, DTE Mumbai & Govt. of Maharashtra, Affiliated to MSBTE Mumbai

Internships Daily Diary(2025-26)
Department of Computer Engineering

Program Code

:CO5K

Name of the Student

: Rajeshwari Uttam Kumbhar

Name of the Mentor (Faculty)

: Miss. S.V. Mativade

Name of the Mentor (Industry)

: Mr. Abhijit P. Gatade

Enrollment Number

: 23213500243

Week	Day & Date	Discussion Topics/Activity	Details of Work Allotted Till Next Session /Corrections Suggested/Faculty Remarks	Signature of Industry Mentor
Week 4	Mon, Date 23-6-25	* EDA * Types of EDA * Plots	EDA using matplotlib & seaborn libraries, Types of visualization in EDA, Line, Bar, count etc plots	
	Tue, Date 24-6-25	* Plots	violin, scatter, pair, heatmap, Pie, swarm, strip, Joint, Rug plot	
	Wed, Date 25-6-25	* task 1 * task 2 on cleaning data	cleaning data using different functions & methods	
	Thu, Date 26-6-25	* task 3 on EDA	data cleaning using EDA, different types of plots	
	Fri, Date 27-6-25	* Data Processing in machine learning * encoding Types	definition of data Processing & encoding, label encoding, one hot, ordinal encoding	