

---

# LGBM & XGBM

---

- Started with importing the basic libraries such as, **NumPy, Pandas, Matplotlib, Seaborn & Warnings**
- Loaded the Datasets and called out the first few entries of the dataset.
- Copied the Datasets into other data frame to perform **Exploratory Data Analysis**.
  - o Dropped the unusual and not used for model attributes from the datasets
  - o Shifted the **Target Variable** to the last
  - o Checked for null values and the number of null values per attribute
  - o Got some visual representation of the missing values through **Heat Map**
  - o Created a table see the number of total entries in the datasets
  - o Filled the null values in the **Age** attribute with the **Mean Method**
  - o Filled the null values in the **Embarked** attribute with the **Mode Method**
  - o Filled the null values in the **Fare** attribute with the **Median Method**
  - o And after tat checked if the null values were cleared or not
- Got on with the visualizations
  - o Went in with **Histograms**
  - o Then with the **Density Plots**
  - o After that we **Label Encoded** the categorical variables in the attributes
  - o And plotted the **Pie Chart** for the survived attribute and **Bar Chart** for that same attribute
  - o And pie chart and bar chart for the embarked data also
  - o Plotted the **Box Plot** for the visualization of the outlier's present
  - o At last ended with **Heat Map** of the **Correlation Matrix** to identify the relation between the attributes
  - o And all these were done for both train and test data
- Made **Log Transformation** of the data with log logic and added them to the dataset
- And removed the original attributes and done this for both the datasets
- Built the models for both **LGBM & XGBM** and before the data was split
- Done with **The Evaluation Metrics** for both the model and had them plotted for better understanding and tabulated them
- Applied the **Hyper Parameter Tuning** for both the models and observed the best parameters for the both the models and analysed the **Strengths and Weaknesses Of Both The Models Through Evaluation Metrics**