**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| Team Member’s Name, Email and Contribution: |
| Name :Sneha Owandkar  Email : sowandkar21[@gmail.com](mailto:autirupalivikas@gmail.com)  Contribution : I have prepare this project individually, understand the data and prepare questions individually. I have complete the data understanding part & library codes.         Also created a ppt . |
| Please paste the GitHub Repo link. |
| Github Link:- |
| Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words) |
| Data set name – Telecom Churn Analysis  Components: Customer churn, Customer retention, Customer relationship management (CRM), Data mining techniques, Telecom industry.  Shape –  3333 rows × 21 columns  Columns - 'State', 'Account length', 'Area code', 'International plan', 'Voice mail plan', 'Number vmail messages', 'Total day minutes', 'Total day calls', 'Total day charge', 'Total eve minutes', 'Total eve calls', 'Total eve charge', 'Total night minutes', 'Total night calls', 'Total night charge', 'Total intl minutes', 'Total intl calls', 'Total intl charge', 'Customer service calls', 'Churn'  Problem Statement :  Orange S.A., formerly France Télécom S.A., is a French multinational telecommunications corporation. The Orange Telecom's Churn Dataset, consists of cleaned customer activity data (features), along with a churn label specifying whether a customer canceled the subscription. Explore and analyze the data to discover key factors responsible for customer churn and come up with ways/recommendations to ensure customer retention.  Approaches :    1. Understand the data.                 Data understanding focuses on the comprehension of the information available in the project. In this step we basically check on the kind of variables provided with the dataset, dtype of the columns, shape of the data frame.    2. Basic cleaning.              Our dataset contains numbers of null values which might tend to disturb our accuracy hence we dropped them at the beginning of our project in order to get a better result.  Pandas isnull() and notnull() methods are used to check and manage NULL values in a data frame    Conclusion :    Some conclusions drawn from the analysis are as follows.     * **Electronic check mediums are the highest churners**. * Contract Type – Monthly customers are more likely to churn because of no contract terms, as they are free-to-go customers. * No Online security, No Tech Support category are high churners. * Non-senior Citizens are high churners. |