|  |  |
| --- | --- |
| **Section** | **Description** |
| **Name** | Shiyas Rasi Pandarakkal |
| **MISIS ID** | M01087756 |
| **Dataset** | Bike Sharing |
| **Introduction** | * **Introduce the Business/Domain Scenario** What are the business /domain problems that company or organization facing? Briefly describe the realistic business/domain relevant problem. * **Define Your Chosen Indicators:** Clearly state the **one** business/domain indicators you will be investigating. For indicator, you must: Provide a precise definition (e.g., "Customer Churn Rate, defined as the percentage of customers who stop using our service within a three-month period."). * **Equation for indicator**: Include an equation for the indicator * **State the Importance of Indicators:** Explain, in general terms, why business indicator (or relevant KPI) is vital for any data-driven decision. Discuss how they help in monitoring performance, making strategic decisions, and achieving business/domain goals.   Why would a manager or stakeholder care about this metric? What kinds of decisions would it influence? (e.g., "Understanding churn is critical for developing customer retention strategies, which are more cost-effective than acquiring new customers.") |
| **Dataset** | * **Data Profile:** Present list that describes each attribute (column) in the dataset. Include the attribute name, a clear description of what it represents, and its data type (e.g., numerical, categorical, binary). |
| **Methodology** | * **Data Preprocessing:** Describe all the steps you took to clean and prepare the data. This includes, but is not limited to:   + Handling of missing values (e.g., imputation, removal).   + Encoding of categorical variables (e.g., one-hot encoding, label encoding).   + Feature scaling or normalization (e.g., StandardScaler, MinMaxScaler).   + Any feature engineering you performed (creating new features from existing ones).   + Features filtering using correlation is optional. * **Model Selection:**   + Identify the type of ML problem you are solving for the indicator (e.g., classification, regression, clustering).   + Select at least two appropriate machine learning algorithms for your problem. * **Model Training and Evaluation:**   + Use train-test split only.   + Explain the evaluation metrics you chose to measure model performance (e.g., Accuracy, Precision, Recall, F1-Score for classification; R-squared, MAE, RMSE for regression). |
| **Results and Analysis** | * **Presentation of Results:** Display the performance of your models in a clear and organized manner. Use either table or chart. Compare evaluation metrics across different models. * **Interpretation of Findings:** This is the most critical part of the section. Go beyond the numbers and explain what your results mean in the context of the business problem.   + What are the key drivers for your business/domain indicators? (e.g., "The model shows that customers with month-to-month contracts and higher monthly charges are significantly more likely to churn.").   + What actionable recommendations could you provide to the business/domain based on your findings? |