**Show Renewal Prediction Models - Entertainment Sector**

Pooja Joshi & E.Code [E25007]

# Overview :

# This report outlines the process of developing a renewal prediction model, focusing on data preparation, model training, evaluation, and visualization of results. The analysis aims to understand factors influencing renewal rates and improve model accuracy through data balancing techniques.

# Objective:

· To build a predictive model for renewal status based on relevant features.

· To apply SMOTE for handling class imbalance in the dataset.

· To visualize model performance and class distribution post-SMOTE.

# Assigned Task(s) :

· Data loading and cleaning

· Feature engineering

· Applying SMOTE for class balancing

· Model training and evaluation

· Visualization of results

# Task Details :

### Task 1: Data Preparation and Model Training

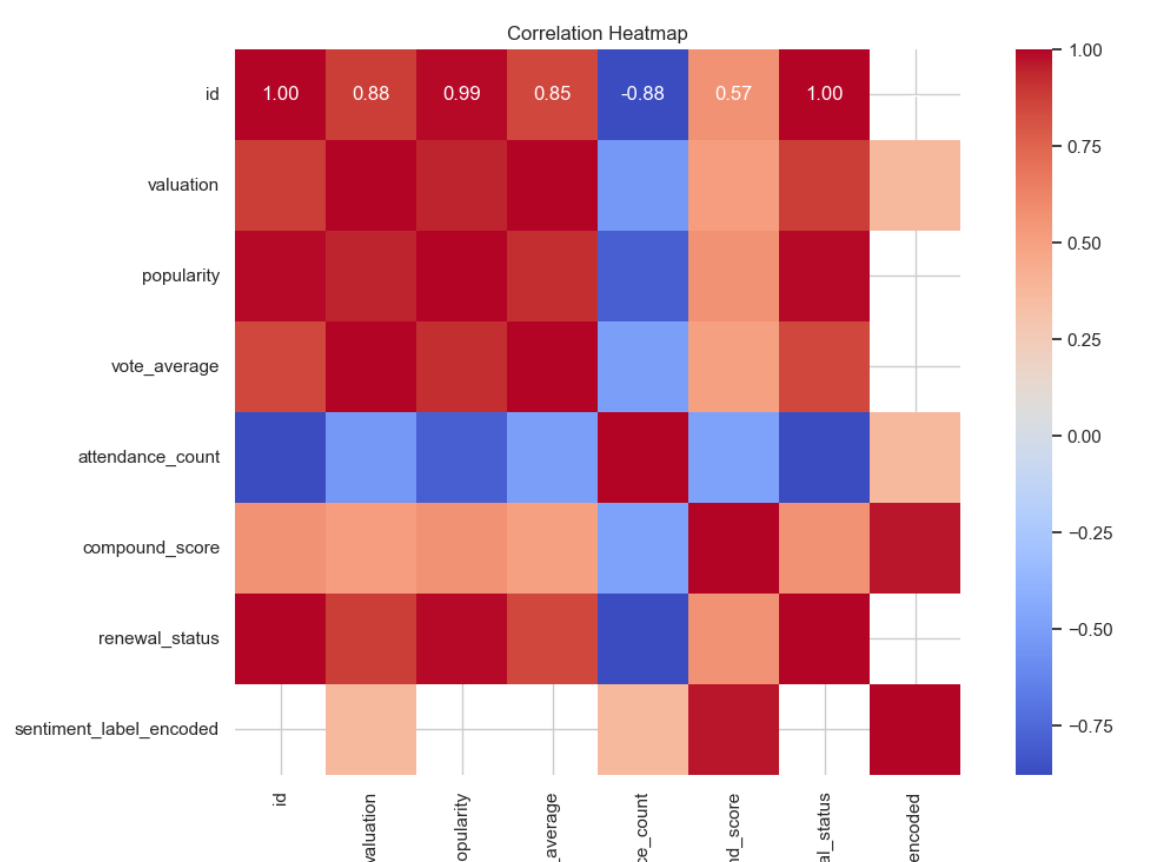
* **Status:** Completed
* **Details:**
  + Loaded the dataset and performed initial cleaning.
  + Engineered features such as popularity, vote average, and attendance count.
  + Applied SMOTE to balance the class distribution.

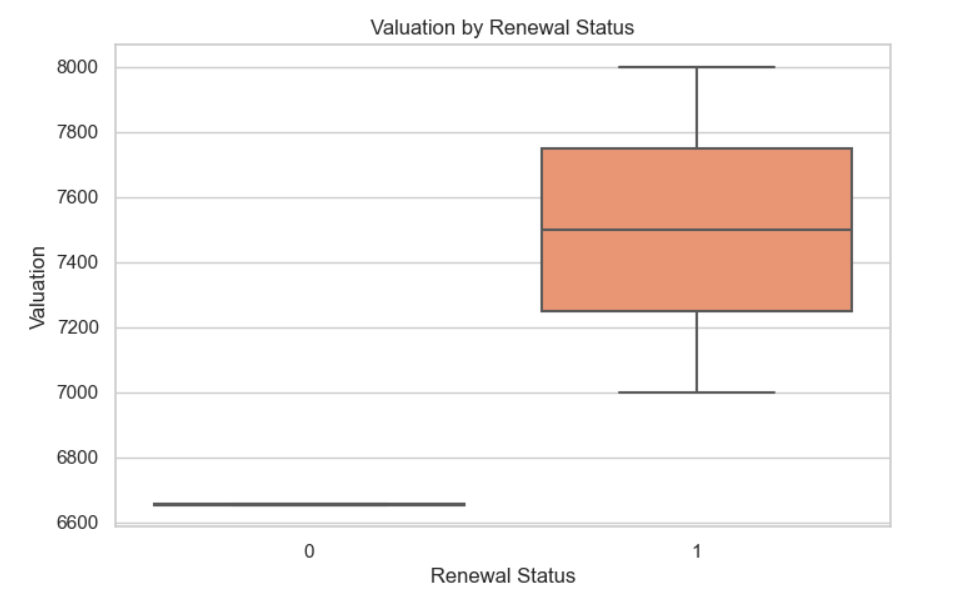
### Task 2: Model Evaluation

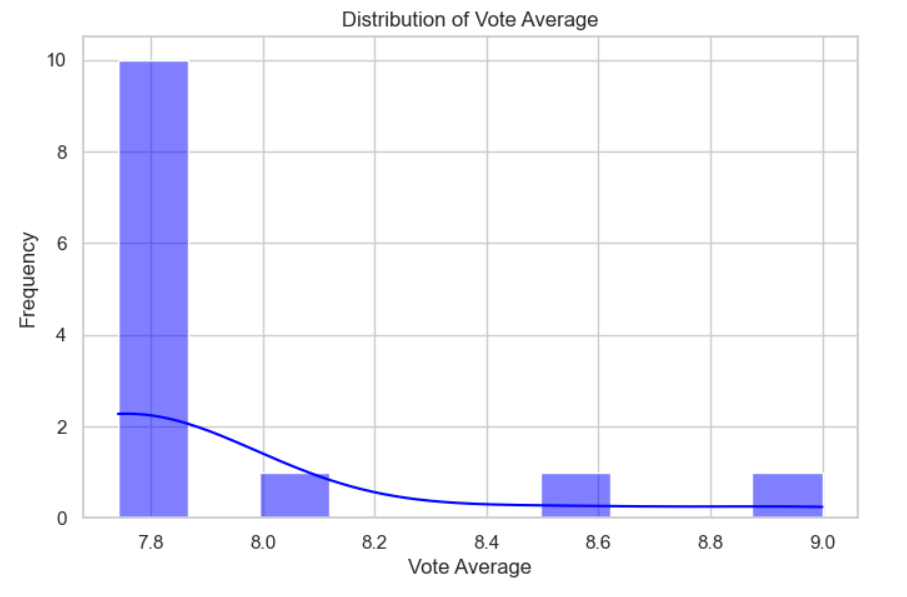
* **Status:** Completed
* **Details:**
  + Trained a Random Forest model on the resampled data.
  + Evaluated the model's performance using precision, recall, and F1-score metrics.

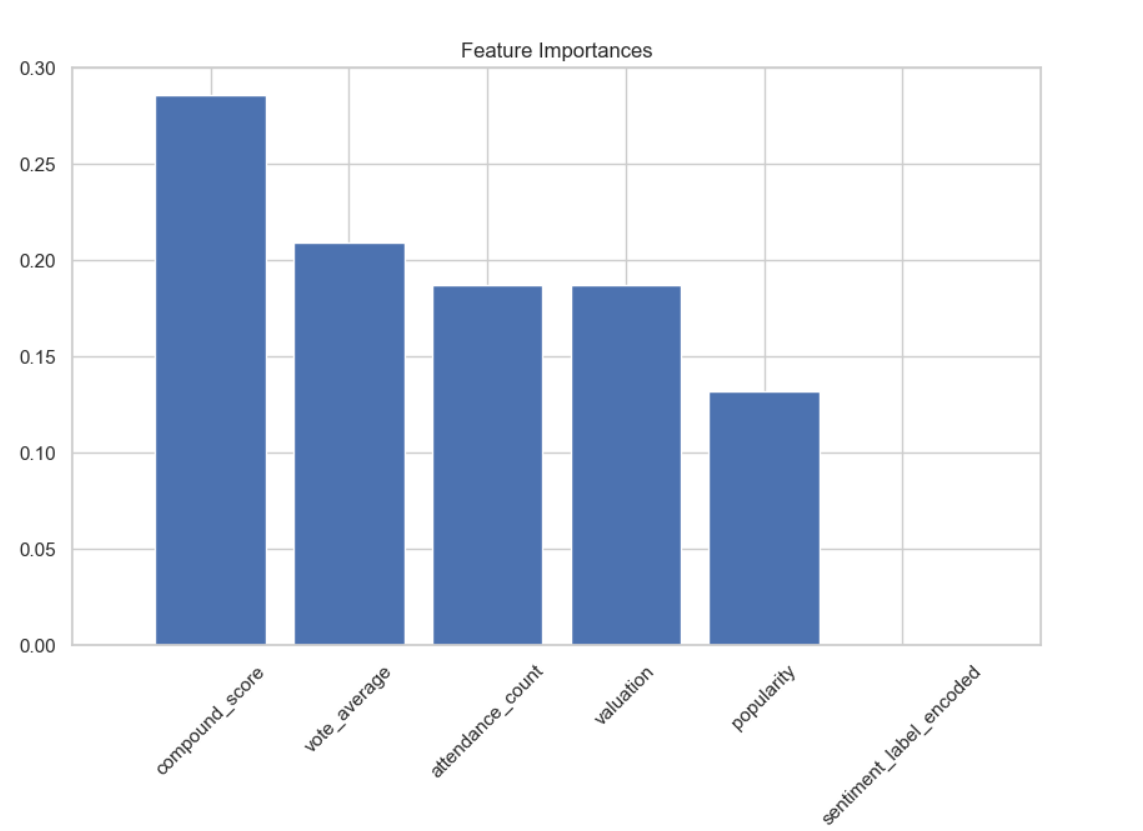
### Task 3: Visualization of Class Distribution and Model Performance

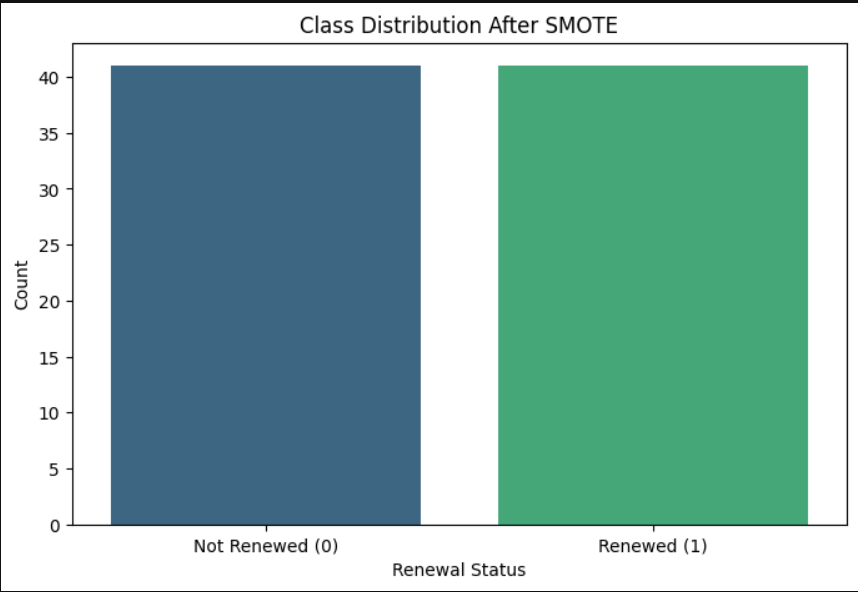
* **Status:** Completed
* **Details:**
  + Created visualizations to illustrate the class distribution before and after applying SMOTE, allowing for a clear understanding of the impact of data balancing on the renewal status prediction.
  + Generated performance metrics visualizations, including precision, recall, and F1-score, to assess the effectiveness of the model.











**Progress :**

· **Accomplishments:**

* Successfully balanced the dataset using SMOTE, achieving an equal class distribution of renewal statuses.
* Developed a functional predictive model with significant evaluation metrics.

· **Metrics:**

#### New Class Distribution After SMOTE

| **Renewal Status** | **Count** |
| --- | --- |
| 0 | 41 |
| 1 | 41 |

#### Model Evaluation Results

| **Metric** | **Precision** | **Recall** | **F1-Score** | **Support** |
| --- | --- | --- | --- | --- |
| **0** | 1.00 | 0.91 | 0.95 | 11 |
| **1** | 0.00 | 0.00 | 0.00 | 0 |
| **Accuracy** |  |  | 0.91 | 11 |

# Challenges and Solutions :

· **Challenges Faced:**

* Encountered class imbalance, with one class having significantly fewer samples.

· **Solutions Implemented:**

* Utilized SMOTE to generate synthetic samples, thus achieving a balanced dataset for training.

# Next Steps :

· **Upcoming Tasks:**

* Fine-tune the Random Forest model for better performance.
* Explore alternative models and techniques for comparison.

· **Goals:**

* · Aim for a higher precision and recall in the model evaluation metrics.
* Investigate additional features that could improve prediction accuracy.

# Conclusion :

* Summary: The report details the successful implementation of a renewal prediction model, addressing data imbalance and evaluating model performance.

# · **Acknowledgments**: Thank the audience for their time and attention.

# Instructions:

1. Use Google Docs. Single Column
2. TNR stands for Times New Roman: B - Bold
3. Use images as required with proper references
4. Use charts, tables as per your requirement.
5. Number of Pages: 2 to 8 for each task report.