

Case Study Rubric: Predicting Parking Ticket Appeal Success in Charlottesville

Goal:

The goal of this case study is for you to preprocess a real-world dataset, train three classification models to predict the success of an appeal, and then evaluate and compare all three models' performances. You will experience the typical workflow of a time-series analysis project, all the way from raw, uncleaned data to building a working predictive model. The project more broadly explores how predictive modeling can (and cannot) explain human decision processes such as appeal approvals.

Tasks Overview:

1. Data Preprocessing

- a. Import and clean the Charlottesville Parking Tickets dataset.
- b. Encode categorical variables and scale numeric features.

2. Model Training

- a. Train three classification models:
 - i. Logistic Regression (linear baseline)
 - ii. Random Forest
 - iii. XGBoost

3. Model Evaluation

- a. Compare performance metrics, including Accuracy, F1-Score, and ROC-AUC
- b. Extract and visualize feature importances or coefficients

4. Draw Conclusions

- a. Interpret results to explain any patterns that correlate with appeal success

Evaluation Criteria:

Category	Description	Points
Data Preprocessing	Dataset loaded, clean, encoded, and split properly	20
Model Training	All models trained and cross-validated	25
Evaluation and Comparison	Meaningful metrics and visualization included	25

Drawing Insights and Conclusions	Discusses correlation strength, fairness implications, data limitations, and more	20
Presentation Quality	Clear organization, visuals, and a readable report	10
Total		100

Deliverables:

- Cleaned dataset
- Jupyter notebook with full model training code
- Evaluation metrics reported (Accuracy, F1-Score, and ROC-AUC)
- Final reflection report
- Code and PDF report uploaded to GitHub repository