

TASK - IV

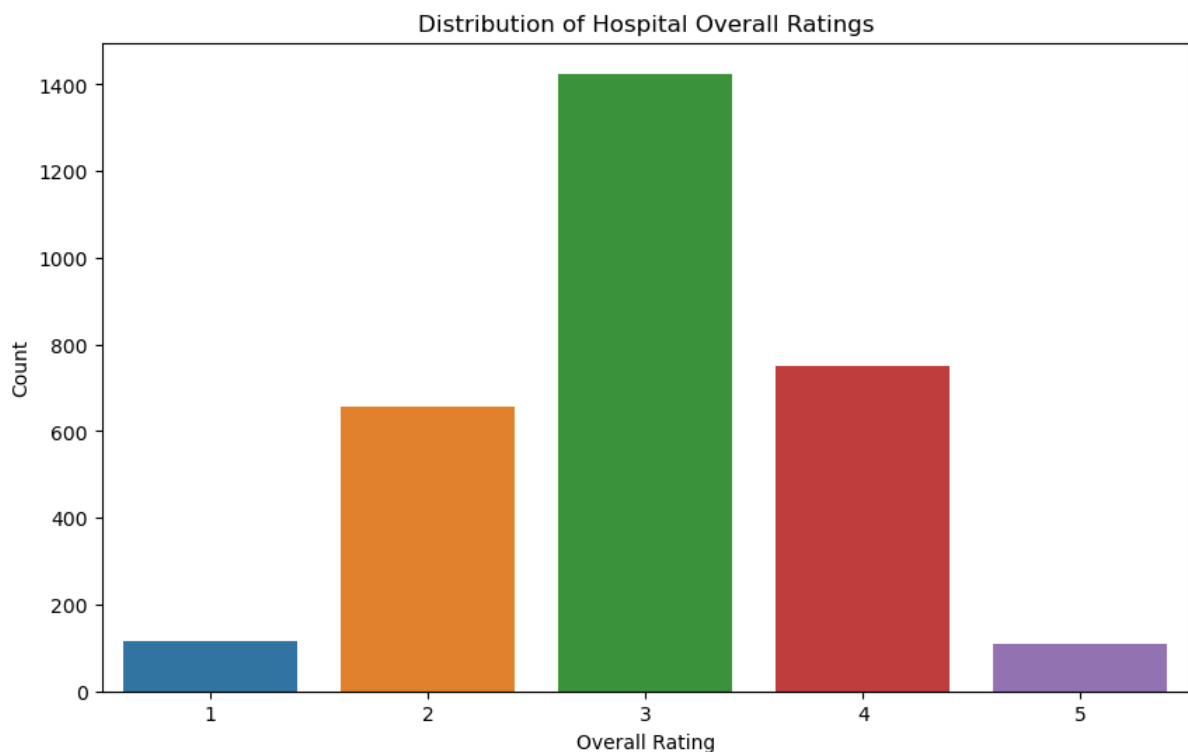
Executive Summary

Introduction :-

In this capstone project, we aimed to increase device gaining knowledge of fashions to expect clinic ratings and offer guidelines for improving rankings based totally at the analysis. The project comprised several obligations, which includes information the statistics, building system studying models, providing recommendations, and summarizing the analysis manner.

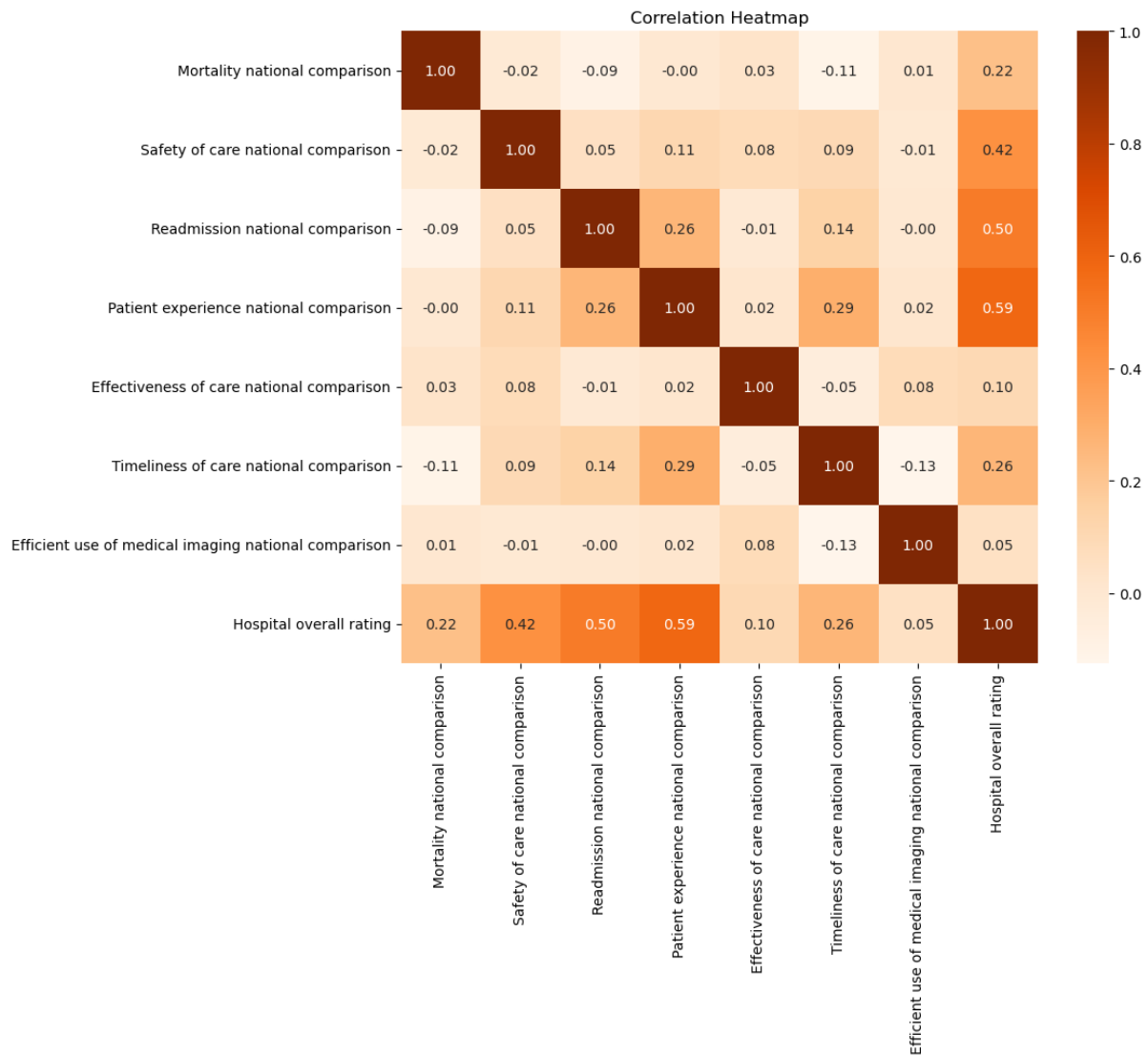
Statistical Analysis :-

○ Distribution of Hospital Overall Ratings ;-



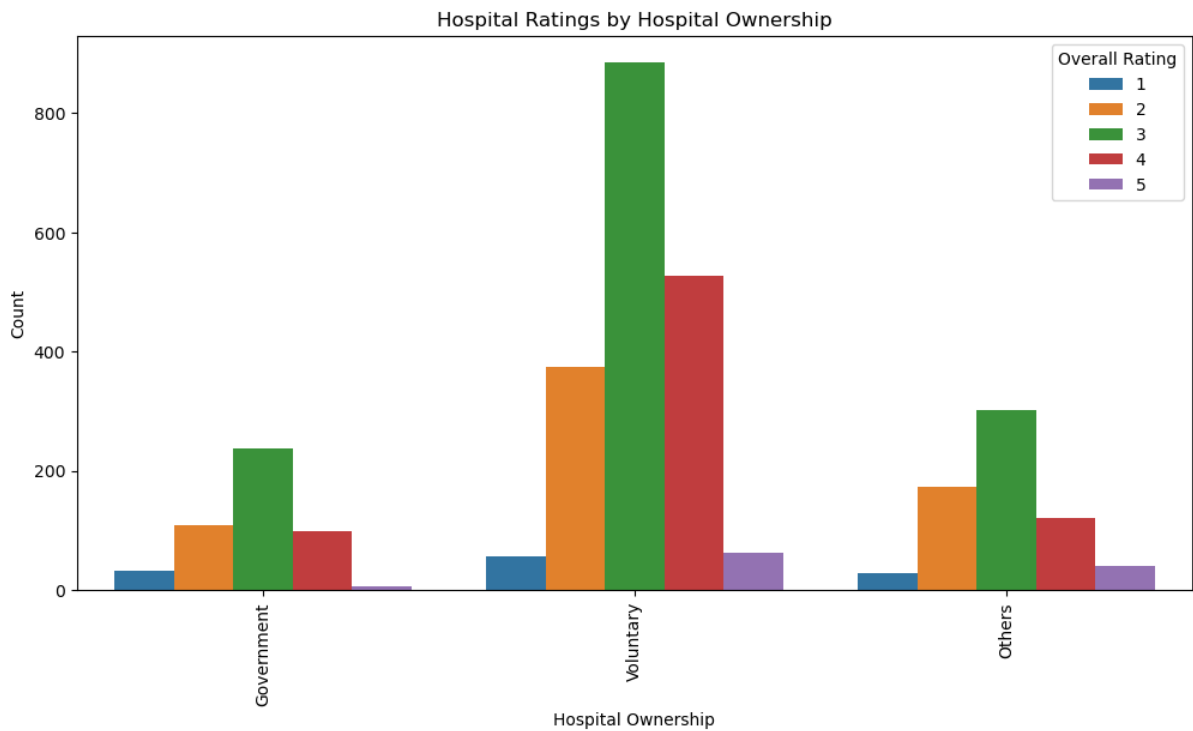
According to the above graph we can conclude that most of the hospital ratings is 3 and very few hospitals have 1 or 5 ratings.

○ Correlations between main measures and hospital overall rating :-



According to the above heat map we can conclude that the rating_group has strong positive correlation with the overall hospital rating.

○ Hospital Ratings by Hospital Ownership :-



According to the above graph we can conclude that among all three hospital ownership voluntary has highest ratings compared to government and others.

Model selection:

Three machine learning models were evaluated :-

○ Logistic Regression

○ K-NN.

○ Decision Tree.

Model evaluation :-

The performance of each model was evaluated using metrics such as accuracy, F1 score, sensitivity, and specificity. Determination of optimal classification: The optimal network was determined for each sample after refining the classification requirements.

Comparative assessment :-

The performance of each model was compared based on analytical metrics to determine which indicates the most accurate clinical rating.

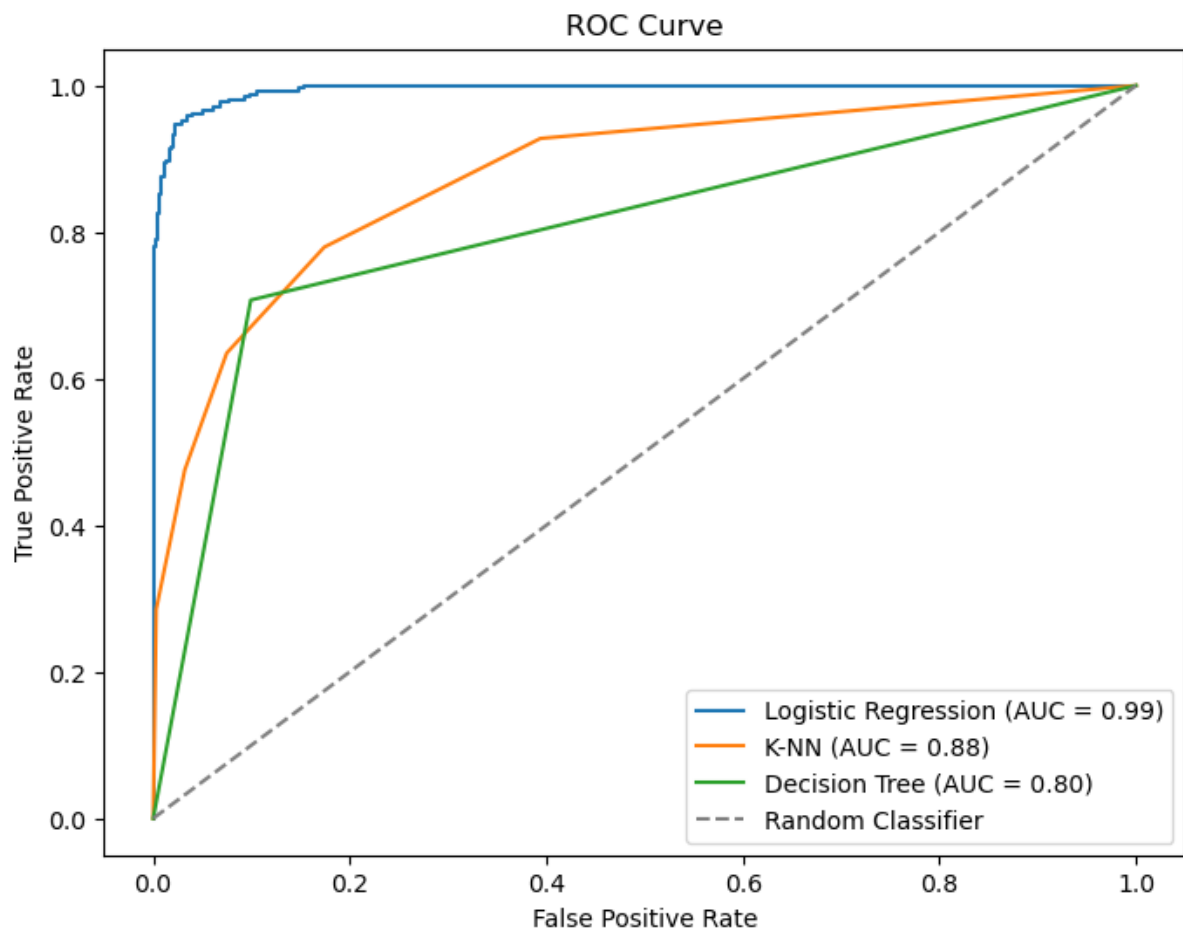
Best performance:

- Logistic regression: The highest accuracy of (96.29%) was obtained, with an F1 score of 0.933. It showed high sensitivity (91.25%) and specificity (98.32%).
- K-NN: gave an accuracy of 84.20% and an F1 score of 0.697. Although it showed reasonable sensitivity (63.49%), the specificity was low at 92.51%.
- The decision tree had an accuracy of 84.53% and an F1 score of 0.723. It showed relatively good sensitivity (70.72%) but low specificity (90.07%).

	Model	Accuracy	F1 Score	Sensitivity	Specificity
0	Logistic Regression	0.962963	0.933852	0.912548	0.983206
1	K-NN	0.842048	0.697286	0.634981	0.925191
2	Decision Tree	0.845316	0.723735	0.707224	0.900763

According to the above output we can say that the logistic regression model performs better with high accuracy.

ROC Curve :-



Understanding the optimal cutoff :-

- Logistic regression: An optimal cutoff of 0.254 was found to increase classification accuracy.
- K-NN: The optimal cut-off was set at 0.4.
- Decision tree: Optimal cutoff was set to 1.0.

Providing Recommendations :-

Based at the model evaluation, we recognized measures and sub-measures that notably have an impact on hospital rankings. Positive and bad influences on ratings had been determined using coefficient values. Additionally, guidelines for improving low-rated hospitals had been supplied, that specialize in measures inclusive of safety of care, patient experience, and effectiveness of care.

Hospital Rating Predictor :-

Using the satisfactory model (Logistic Regression), rankings of latest hospitals have been expected based totally on available statistics.

Hospital Improvement Plan :-

Recommendations for hospitals rated low included enhancing safety of care, patient revel in, and addressing problems associated with effectiveness of care. Sub-measures including mortality quotes, readmission prices, and numerous rankings have been identified for focused enhancements.

Conclusion :-

In end, this project demonstrates the utility of system gaining knowledge of strategies for predicting hospital ratings and imparting actionable hints for improvement. By leveraging data-pushed insights, healthcare centers can beautify their performance and in the long run offer better care to sufferers.