

Model Reporting

We utilized the Random Forest Regressor to analyze the Zomato dataset. The model was trained on relevant features from the dataset and evaluated using metrics such as Mean Squared Error (MSE) and R-squared (R^2). Results demonstrated the model's effectiveness in predicting restaurant ratings/prices (or any target variable analyzed). Key findings and performance metrics are summarized below:

- Dataset: Zomato dataset from Kaggle.
- Model: Random Forest Regressor.
- Evaluation Metrics: MSE, R^2 .

✓ The libraries We've imported cover all the essential tools for preprocessing, model building, and evaluation.

1. Load the dataset and inspect for missing values or outliers.
2. Encode categorical variables (e.g., city or restaurant type in Zomato dataset).
3. Scale numerical features (like ratings, prices).
4. Split the data into training and test sets.
5. Train the model using the RandomForestRegressor.
6. Tune hyperparameters with RandomizedSearchCV.
7. Evaluate using mean_squared_error and r2_score.
8. Optionally, visualize results (e.g., residual plots or feature importances).

✓ We applied a Random Forest Regressor to predict target variables in the Zomato dataset without using scaling methods. The model was evaluated on test data, yielding the following performance metrics:

- MSE: 35.26
- R^2 : 0.49

These results indicate moderate predictive accuracy, with the model explaining approximately 49% of the variance in the target variable. Further improvements can be achieved through feature engineering, hyperparameter tuning, or exploring alternative models.

✓ Using a scaled version of the Zomato dataset, the Random Forest Regressor was evaluated for predictive performance. Key metrics are as follows:

- MSE: 0.51
- R^2 : 0.49

Scaling resulted in a normalized MSE, but the R^2 value remained consistent, explaining approximately 49% of the variance in the target variable. This suggests that scaling had minimal impact on overall model performance. Further tuning or alternative methods could be explored for improvement.

✓ The Random Forest Regressor was validated using 5-fold cross-validation with scaled data. The MSE across folds was:

- Fold MSEs: [0.5076, 0.4986, 0.5056, 0.5120, 0.4845]
- Average MSE: 0.5017

This validation indicates consistent model performance across different subsets, with minimal variation in MSE, confirming the model's reliability on the dataset.