Explained Variance Ratio: [5.83938171e-01 2.50446927e-01 1.30956754e-01 1.87577082e-02

9.87786808e-03 3.29865585e-03 1.74587911e-03 9.70534500e-04

7.50214188e-06]

Cumulative Explained Variance Ratio: [0.58393817 0.8343851 0.96534185 0.98409956 0.99397743 0.99727608

0.99902196 0.9999925 1

Based on the provided explained variance ratios, here is the interpretation:

The first principal component explains approximately 58.39% of the total variance in the data.

The first two principal components combined explain around 83.44% of the total variance.

The first three principal components combined explain about 96.53% of the total variance.

The first four principal components combined explain approximately 98.41% of the total variance.

The first five principal components combined explain around 99.40% of the total variance.

The first six principal components combined explain about 99.73% of the total variance.

The first seven principal components combined explain approximately 99.90% of the total variance.

The first eight principal components combined explain around 99.99% of the total variance.

The explained variance ratio provides insights into the importance of each principal component in capturing the variability in the original feature data. By examining the cumulative explained variance ratio, you can determine how much of the total variance is explained as you include more principal components.

Regression fit

Mean Squared Error (MSE): The MSE value of 397.7437 indicates the average squared difference between the predicted and actual values. A lower MSE indicates better model performance, with values closer to zero indicating a more accurate model.

Mean Absolute Error (MAE): The MAE value of 13.232 indicates the average absolute difference between the predicted and actual values. Similar to MSE, a lower MAE indicates better model performance, with values closer to zero indicating a more accurate model.

Root Mean Squared Error (RMSE): The RMSE value of 19.9435 is the square root of the MSE. It provides a measure of the average prediction error in the same unit as the target variable. Lower RMSE values indicate better model performance.

R-squared (R²): The R-squared value of 0.8844 indicates the proportion of the variance in the target variable that is explained by the regression model. R-squared ranges from 0 to 1, with higher values indicating a better fit. In this case, the model explains approximately 88.44% of the variance in the target variable.

Adaboost

these are results comparing above Mean Squared Error (MSE): 784.55868481645

Mean Absolute Error (MAE): 20.9235013462186

Root Mean Squared Error (RMSE): 28.009974737875968

R-squared: 0.7719059100699602

Based on the provided results for the AdaBoost regression model, here is the comparison with the previous linear regression model:

Mean Squared Error (MSE): The AdaBoost model has a higher MSE value of 784.5587 compared to the linear regression model's MSE of 397.7437. A higher MSE indicates that the AdaBoost model's predictions have a larger average squared difference from the actual values.

Mean Absolute Error (MAE): The AdaBoost model has a higher MAE value of 20.9235 compared to the linear regression model's MAE of 13.232. A higher MAE indicates that the AdaBoost model's predictions have a larger average absolute difference from the actual values.

Root Mean Squared Error (RMSE): The AdaBoost model has a higher RMSE value of 28.009975 compared to the linear regression model's RMSE of 19.9435. A higher RMSE indicates that the AdaBoost model's predictions have a larger average prediction error in the same unit as the target variable.

R-squared (R²): The AdaBoost model has an R-squared value of 0.7719, which is lower than the linear regression model's R-squared of 0.8844. A higher R-squared indicates that the model explains a larger proportion of the variance in the target variable. Therefore, the linear regression model performs better in explaining the variance compared to the AdaBoost model.

XGBOOST

Mean Squared Error (MSE): 549.3717676649582

Mean Absolute Error (MAE): 17.30392768048885

Root Mean Squared Error (RMSE): 23.438681013763514

R-squared: 0.8402816056926166

Based on the provided results for the XGBoost regression model, here is a discussion comparing them with the previous models:

Mean Squared Error (MSE): The XGBoost model has an MSE of 549.3718, which is lower than the AdaBoost model's MSE of 784.5587. This indicates that the XGBoost model's predictions have a smaller average squared difference from the actual values compared to the AdaBoost model.

Mean Absolute Error (MAE): The XGBoost model has an MAE of 17.3039, which is lower than the AdaBoost model's MAE of 20.9235. A lower MAE indicates that the XGBoost model's predictions have a smaller average absolute difference from the actual values compared to the AdaBoost model.

Root Mean Squared Error (RMSE): The XGBoost model has an RMSE of 23.4387, which is lower than the AdaBoost model's RMSE of 28.009975. A lower RMSE indicates that the XGBoost model's predictions have a smaller average prediction error in the same unit as the target variable compared to the AdaBoost model.

R-squared (R²): The XGBoost model has an R-squared value of 0.8403, which is higher than the AdaBoost model's R-squared of 0.7719. A higher R-squared indicates that the XGBoost model explains a larger proportion of the variance in the target variable compared to the AdaBoost model.

In conclusion, the XGBoost model performs better than the AdaBoost model based on the provided metrics. The XGBoost model has lower MSE, MAE, and RMSE values, indicating smaller prediction errors. Additionally, the XGBoost model has a higher R-squared value, suggesting that it explains a larger proportion of the variance in the target variable.

Based on the provided results for the Random Forest classifier, here is a discussion comparing them with the previous models:

Accuracy: The Random Forest classifier has an accuracy of 0.0826, which is not a commonly used metric for regression models. It is more commonly used for classification tasks. Therefore, the accuracy value might not provide meaningful information for a regression model.

Mean Absolute Error (MAE): The Random Forest classifier has an MAE of 19.585, which is higher than both the AdaBoost model's MAE of 17.304 and the linear regression model's MAE of 13.232. A higher MAE indicates that the Random Forest classifier's predictions have a larger average absolute difference from the actual values compared to the other models.

Root Mean Squared Error (RMSE): The Random Forest classifier has an RMSE of 29.1335, which is higher than both the AdaBoost model's RMSE of 23.4387 and the linear regression model's RMSE of 19.9435. A higher RMSE indicates that the Random Forest classifier's predictions have a larger average prediction error in the same unit as the target variable compared to the other models.

R-squared (R²): The Random Forest classifier has an R-squared value of 0.7532, which is lower than both the AdaBoost model's R-squared of 0.8403 and the linear regression model's R-squared of 0.8844. A higher R-squared indicates that the model explains a larger proportion of the variance in the target variable. Therefore, the Random Forest classifier performs worse than both the AdaBoost and linear regression models in explaining the variance.

In conclusion, based on the provided metrics, the Random Forest classifier performs worse than both the AdaBoost and linear regression models. It has higher MAE, RMSE values, and a lower R-squared value, indicating larger prediction errors and lower explanatory power.

Bagging

baggiging results Mean Absolute Error (MAE): 19.36560595802302

Root Mean Squared Error (RMSE): 28.105312638547435

R-squared: 0.7703505341384163

Mean Squared Error (MSE): 789.9085985104942

Root Mean Squared Error (RMSE): 28.105312638547435

Mean Absolute Error (MAE): 19.36560595802302

R-squared (R²): 0.7703505341384163

Mean Absolute Error (MAE): The Bagging classifier has an MAE of 19.3656, which is higher than the AdaBoost model's MAE of 17.304, but lower than the Random Forest classifier's MAE of 19.585. A higher MAE indicates that the Bagging classifier's predictions have a larger average absolute difference from the actual values compared to the AdaBoost model, but a smaller average absolute difference compared to the Random Forest classifier.

Root Mean Squared Error (RMSE): The Bagging classifier has an RMSE of 28.1053, which is higher than the AdaBoost model's RMSE of 23.4387, but lower than the Random Forest classifier's RMSE of 29.1335. A higher RMSE indicates that the Bagging classifier's predictions have a larger average prediction error in the same unit as the target variable compared to the AdaBoost model, but a smaller average prediction error compared to the Random Forest classifier.

R-squared (R²): The Bagging classifier has an R-squared value of 0.7704, which is higher than the AdaBoost model's R-squared of 0.8403, but lower than the Random Forest classifier's R-squared of 0.7532. A higher R-squared indicates that the model explains a larger proportion of the variance in the target variable. Therefore, the Bagging classifier performs better than the Random Forest classifier in explaining the variance but worse than the AdaBoost model.

In conclusion, based on the provided metrics, the Bagging classifier performs better than the Random Forest classifier in terms of MAE and RMSE but worse in terms of R-squared. It performs worse than the AdaBoost model in terms of MAE, RMSE, and R-squared.

Boosting

Accuracy: 0.07582938388625593

Mean Squared Error (MSE): 871.2173324306026

Root Mean Squared Error (RMSE): 29.51639091133268

Mean Absolute Error (MAE): 19.947190250507788

R-squared (R²): 0.7467117139637218

Based on the provided results for the boosting (AdaBoost) model, here is a discussion of the metrics:

Accuracy: The accuracy value of 0.0758 is not typically used as a performance metric for regression models. It is more commonly used for classification tasks. Therefore, the accuracy value might not provide meaningful information for a regression model.

Mean Squared Error (MSE): The boosting model has an MSE of 871.2173, which is higher than the previous models. A higher MSE indicates that the boosting model's predictions have a larger average squared difference from the actual values compared to the other models.

Root Mean Squared Error (RMSE): The RMSE of 29.5164 is higher than the previous models. A higher RMSE indicates that the boosting model's predictions have a larger average prediction error in the same unit as the target variable compared to the other models.

Mean Absolute Error (MAE): The MAE of 19.9472 is higher than the previous models. A higher MAE indicates that the boosting model's predictions have a larger average absolute difference from the actual values compared to the other models.

R-squared (R²): The boosting model has an R-squared value of 0.7467, which is lower than the previous models. A higher R-squared indicates that the model explains a larger proportion of the variance in the target variable. Therefore, the boosting model performs worse in explaining the variance compared to the previous models.

In conclusion, based on the provided metrics, the boosting (AdaBoost) model performs worse than the previous models in terms of MSE, RMSE, MAE, and R-squared. It has higher prediction errors and explains a smaller proportion of the variance in the target variable.