Lubricant Performance Prediction: Data Preprocessing
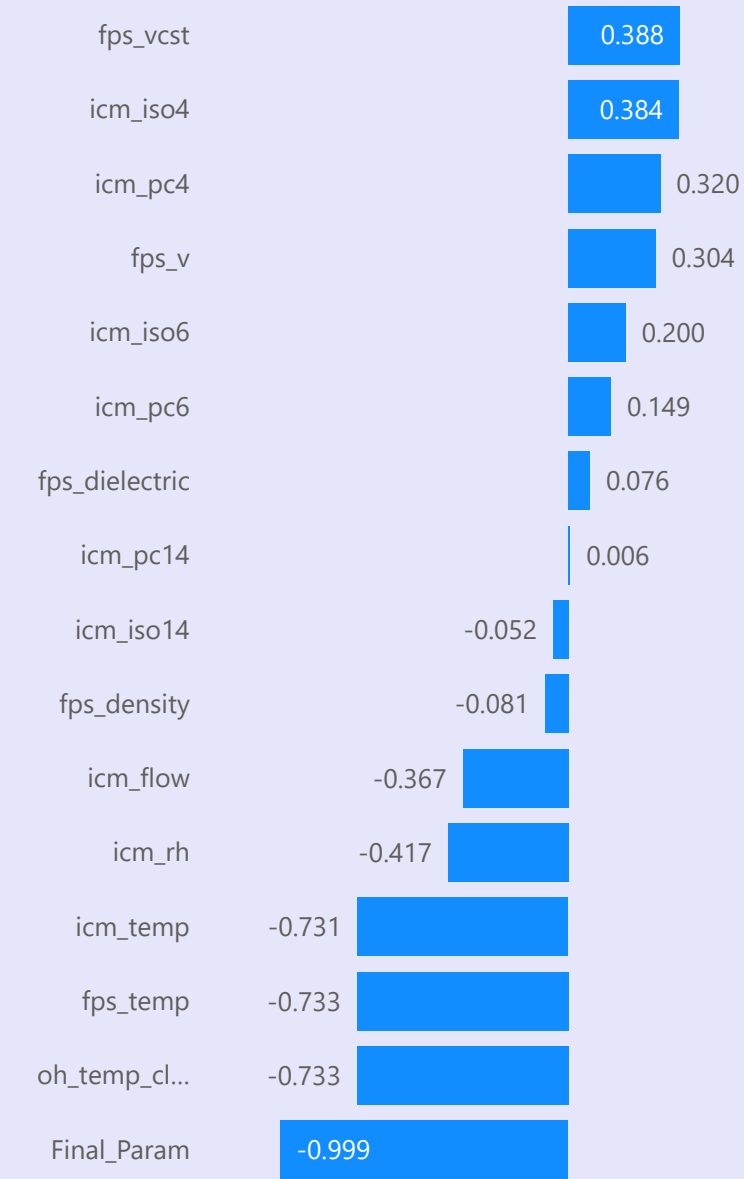
Missing values in Dataset

Feature	Count
icm_iso4	59392
icm_rh	13072
fps_densi...	13068
icm_flow	13068
icm_temp	13060
oh_temp...	2697
fps_temp	3

Data Preprocessing

100%

Stage	Value
Total data	258645
Cleaned Data	245535
Empty Rows	13101
Imputed	2700
Extreme Outliers	9

0%
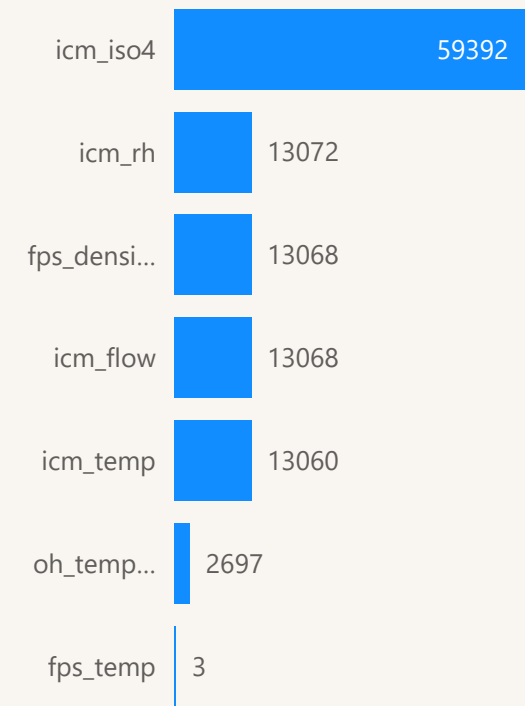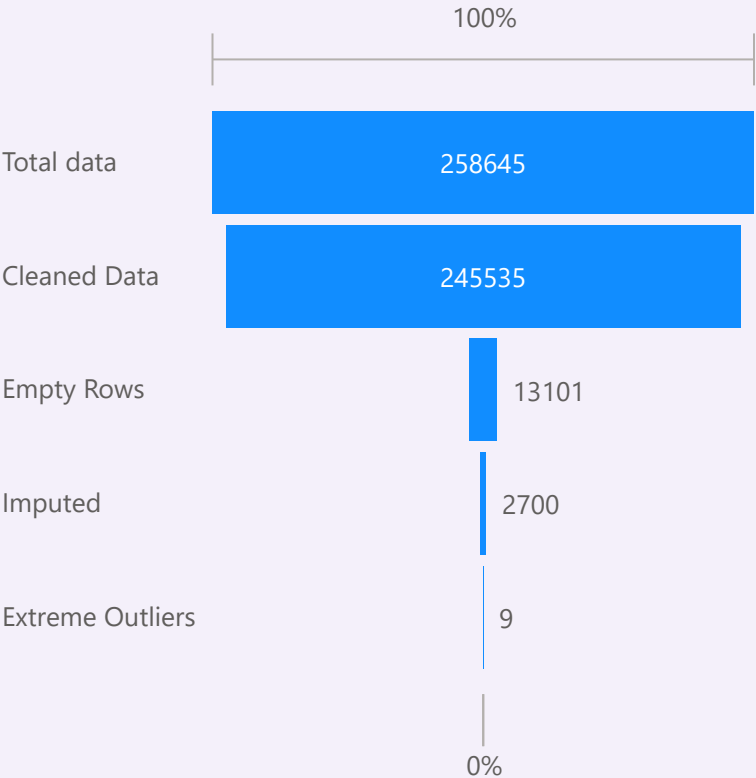
Correlation with Target Variable (oh_od)

Feature	Correlation
fps_vcst	0.388
icm_iso4	0.384
icm_pc4	0.320
fps_v	0.304
icm_iso6	0.200
icm_pc6	0.149
fps_dielectric	0.076
icm_pc14	0.006
icm_iso14	-0.052
fps_density	-0.081
icm_flow	-0.367
icm_rh	-0.417
icm_temp	-0.731
fps_temp	-0.733
oh_temp_cl...	-0.733
Final_Param	-0.999

Features

- [ ] Cleaned Feature
- [ ] Neural network
- [ ] Regression Feature

Final_Param
fps_density
fps_dielectric
fps_temp
fps_v
fps_vcst
icm_iso14
icm_iso4
icm_iso6
icm_pc14
icm_pc4
icm_pc6
icm_rh
icm_temp

Objective: Achieving 95% accuracy in predicting oil degradation (oh_od) by solving the linear regression problem, ensuring precise and reliable predictions.

FInal Param = oh_parama + oh_paramb + oh_paramc

This combination enhances the model's performance and accuracy.

- Data Split: The dataset was divided into 60% training, 20% validation, and 20% test sets.
- Feature Scaling: StandardScaler() was applied to normalize features with a mean of 0 and a standard deviation of 1.
- Final Datasets: The processed datasets were saved for model training and evaluation.

The model achieved 98% accuracy in predicting oil degradation, leveraging optimized features, proper data splitting, and scaling techniques for highly reliable and precise predictions.

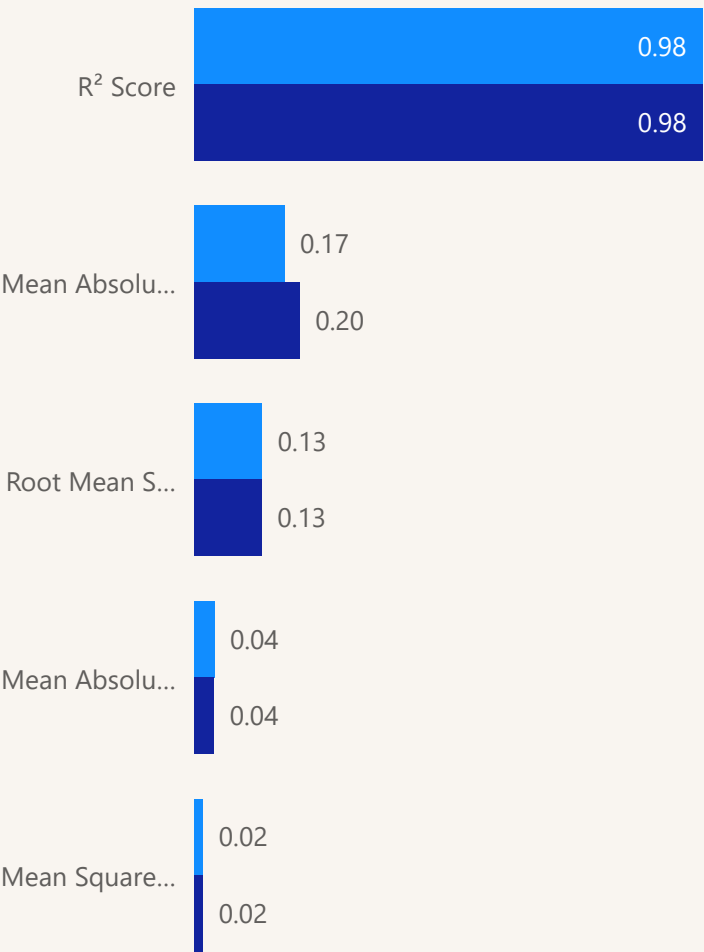# Lubricant Performance Prediction: Model Insights

Gradient Boosting Regression | Linear Regression | Neural Network | **Random Forest Regression**

Test Vs Validation Result

● Average of Test Value ● Average of Validation Value

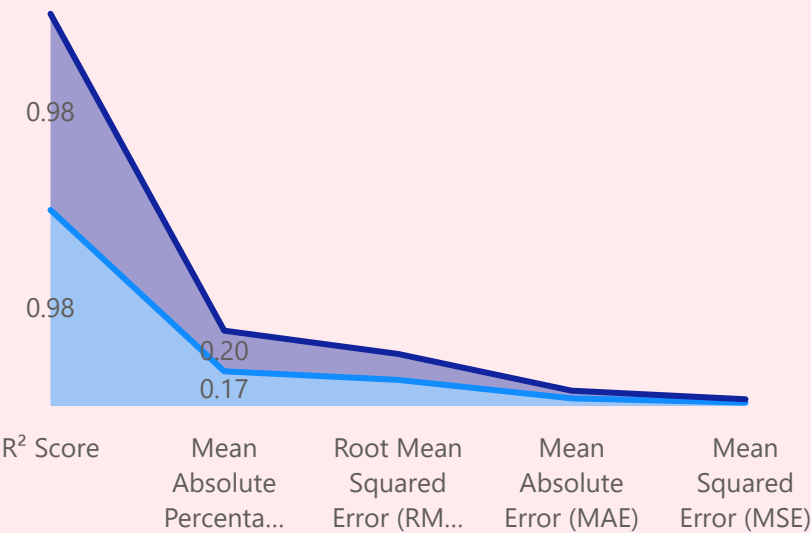|	Test	Validation
R² Score	0.98	0.98
Mean Absolu...	0.17	0.20
Root Mean S...	0.13	0.13
Mean Absolu...	0.04	0.04
Mean Square...	0.02	0.02

98.27%
Test Accuracy

98.31%
Validation Accuracy

High-Performing
Overall performance

Test Vs Validation Performance

● Average of Test Value ● Average of Validation Value

0.98
0.98
0.20
0.17

R² Score | Mean Absolute Percenta... | Root Mean Squared Error (RM... | Mean Absolute Error (MAE) | Mean Squared Error (MSE)

Feature Importance

Feature	Importance
icm_temp	0.018
fps_temp	0.011
fps_vcst	0.004
icm_pc4	0.003
icm_pc6	0.002
fps_v	0.002
icm_pc14	0.001
icm_iso14	0.000
icm_iso6	0.000
icm_iso4	0.000