DATA SCIENCE PROJECT

P496 Electric Motor Prediction

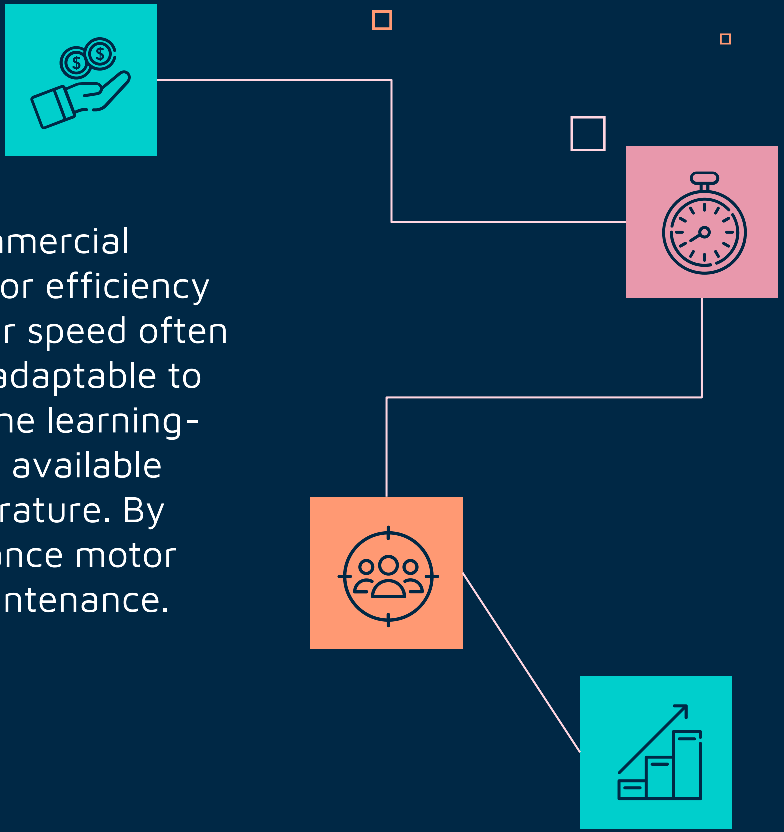
–GROUP 2 TEAM MEMBERS

Keerthi Mani Purna Chandrika
Neeraja B
Vaishnav P
Sajitha A
Aswanth TK
Sonam Akshay Sherkar
Jithin K S

CONTENTS

INTRODUCTION

Electric motors are widely used in industrial and commercial applications, where accurate speed estimation is crucial for efficiency and performance. Traditional methods for predicting motor speed often rely on physical models, which can be complex and less adaptable to varying conditions. This project aims to develop a machine learning-based predictive model to estimate motor speed using available attributes such as voltage, current, torque, and temperature. By leveraging data-driven techniques, the model can enhance motor control, improve efficiency, and support predictive maintenance.



PROBLEM STATEMENT

Electric motors are essential in various industrial and commercial applications, where accurate speed control is crucial for efficiency and performance. However, predicting motor speed based on operational parameters such as voltage, current, torque, and temperature remains a challenge due to the nonlinear and dynamic nature of motor behavior. This project aims to develop a predictive model using machine learning techniques to estimate motor speed based on available attributes. By leveraging historical and real-time data, the model will enable better control strategies, improve efficiency, and aid in predictive maintenance, ultimately enhancing the reliability and performance of electric motor systems.

DATASET DETAILS

The dataset comprises several sensor data collected from a permanent magnet synchronous motor (PMSM) deployed on a test bench. The PMSM represents a ccTest bench measurements were collected by the LEA department at Paderborn University. This dataset is mildly anonymized.

Comprehensive csv files containing all measurement sessions and features. Each row represents one snapshot of sensor data at a certain time step. Sample rate is 2 Hz (One row per 0.5 seconds). Distinctive sessions are identified with “profile_id”.

FEATURE SET

- ambient

Ambient temperature as measured by a thermal sensor located closely to the stator.

- coolant

Coolant temperature. The motor is water cooled. Measurement is taken at the outflow.

- u_d

Voltage d-component.

- u_q

Voltage q-component.

- motor_speed

Motor speed.

- torque

Torque induced by current.

- i_d
  Current d-component.

- i_q
  Current q-component.

- pm
  Permanent Magnet surface temperature representing the rotor temperature. This was measured with an infrared thermography unit.

- stator_yoke
  Stator yoke temperature is measured with a thermal sensor.

- stator_tooth
  Stator tooth temperature is measured with a thermal sensor.

- stator_winding
  Stator winding temperature measured with a thermal sensor.

- profile_id
  Each measurement session has a unique ID. Make sure not to try to estimate from one session onto the other as they are strongly independent.

DATA EXPLORATION

```
print('Data Info:')
data.info()
```

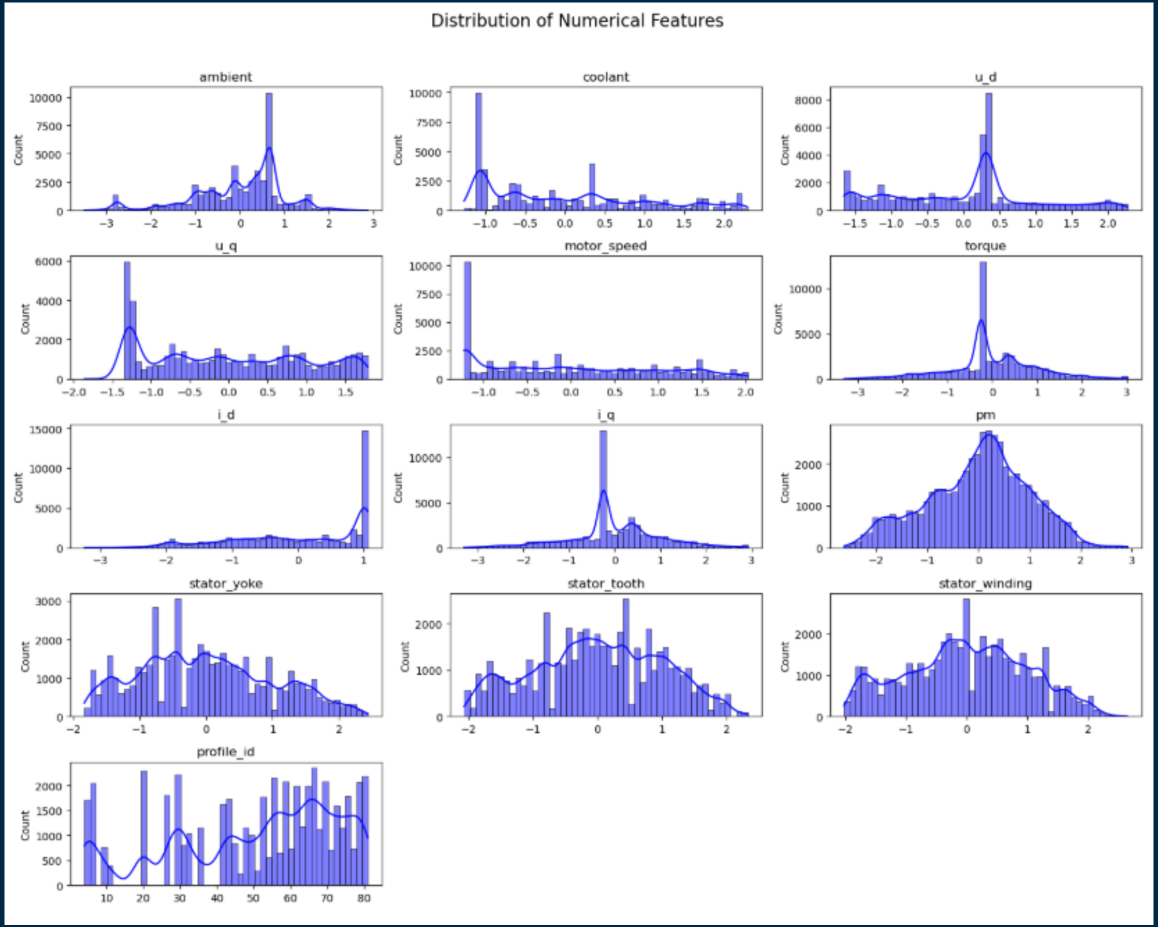
```
Data Info:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 998070 entries, 0 to 998069
Data columns (total 13 columns):
 #   Column          Non-Null Count   Dtype
---  ------          --------------   -----
 0   ambient         998070 non-null  float64
 1   coolant         998070 non-null  float64
 2   u_d             998070 non-null  float64
 3   u_q             998070 non-null  float64
 4   motor_speed     998070 non-null  float64
 5   torque          998070 non-null  float64
 6   i_d             998070 non-null  float64
 7   i_q             998070 non-null  float64
 8   pm              998070 non-null  float64
 9   stator_yoke     998070 non-null  float64
 10  stator_tooth    998070 non-null  float64
 11  stator_winding  998070 non-null  float64
 12  profile_id      998070 non-null  int64
dtypes: float64(12), int64(1)
memory usage: 99.0 MB
```

MISSING VALUES

```
missing_values=data.isnull().sum()
print('\nMissing Values in each column:')
print(missing_values)
```

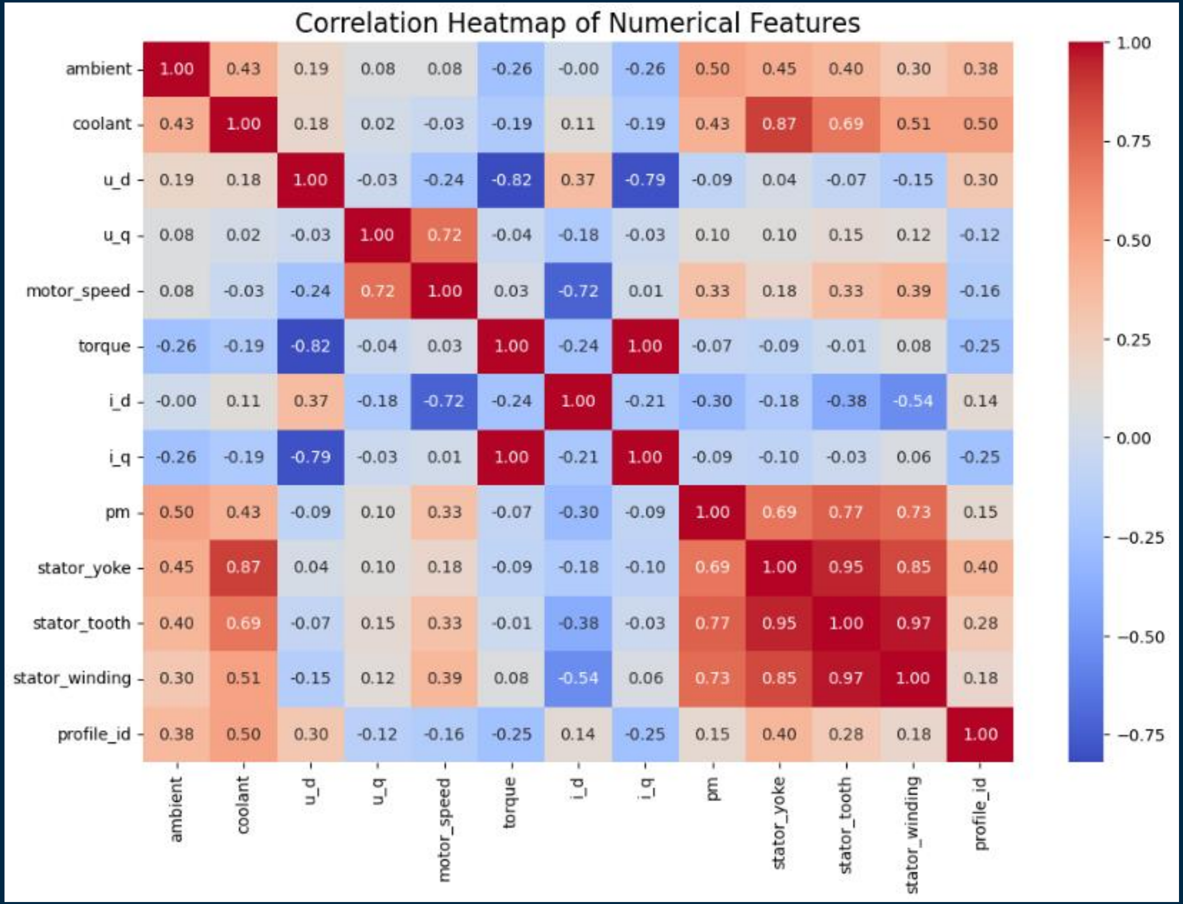
```
Missing Values in each column:
ambient           0
coolant           0
u_d               0
u_q               0
motor_speed       0
torque            0
i_d               0
i_q               0
pm                0
stator_yoke       0
stator_tooth      0
stator_winding    0
profile_id        0
dtype: int64
```

HISTOGRAM

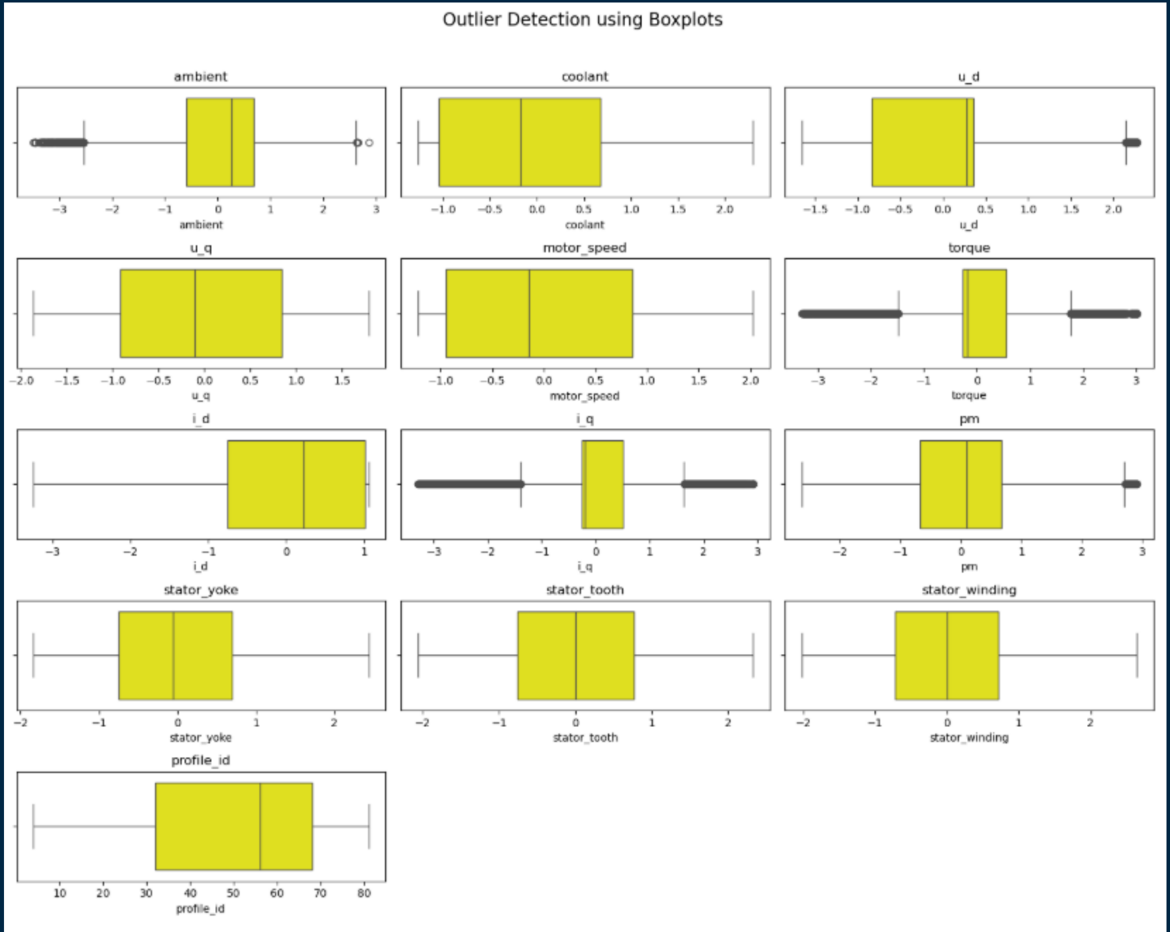


Distribution of Numerical Features
ambient
coolant
u_d
u_q
motor_speed
torque
i_d
i_q
pm
stator_yoke
stator_tooth
stator_winding
profile_id
Count

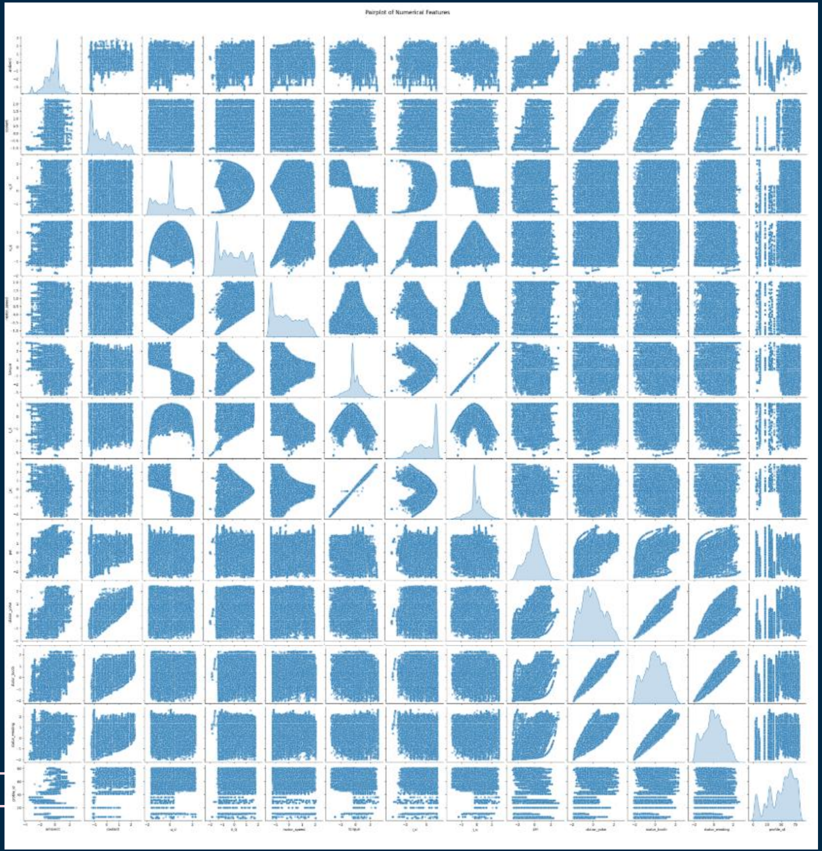
CORRELATION HEATMAP



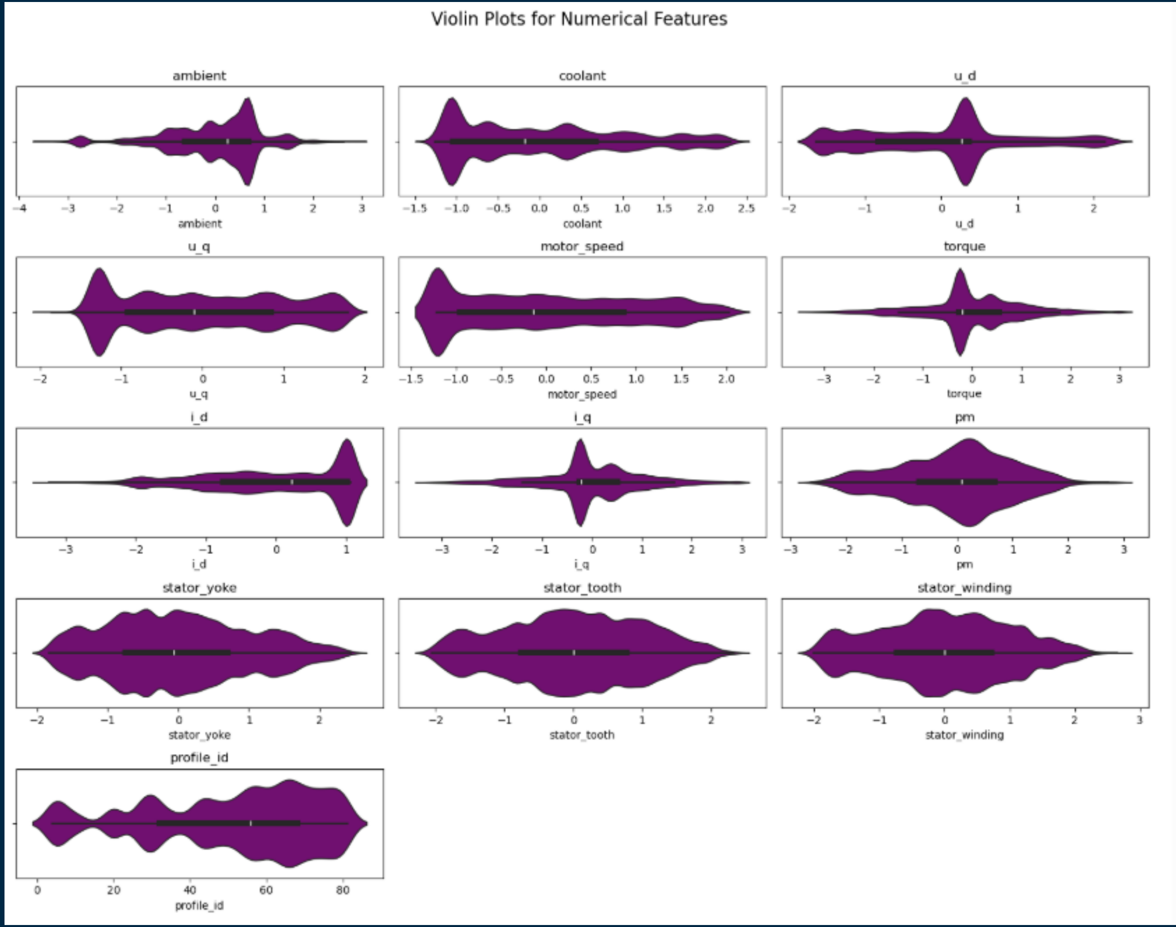
OUTLIER DETECTION



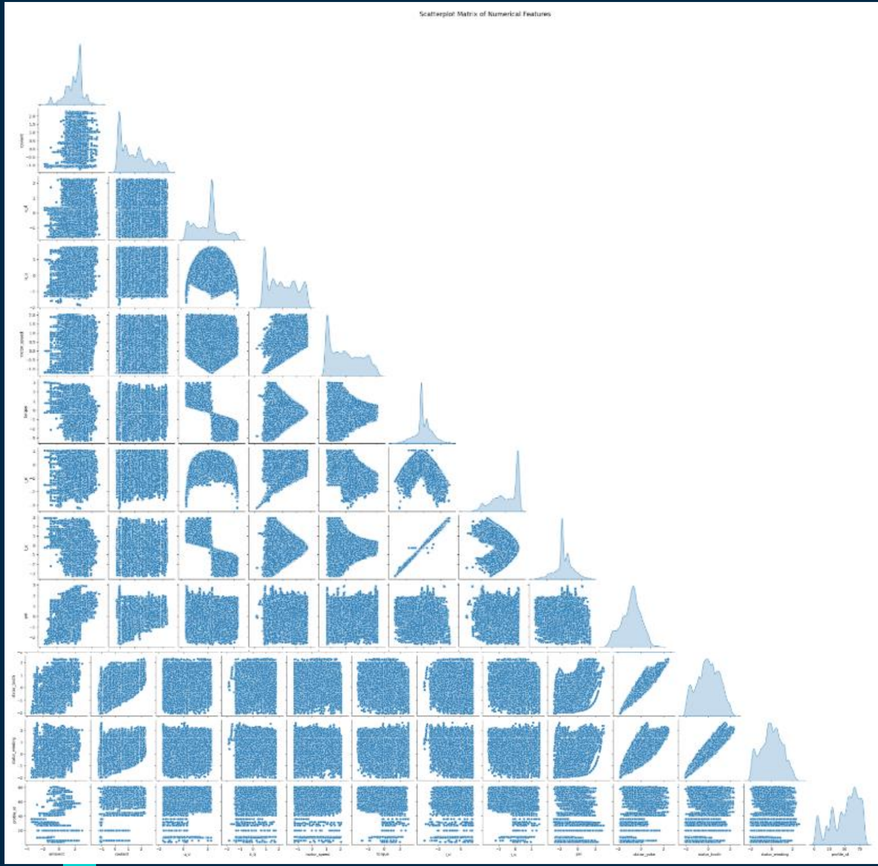
FEATURE RELATIONSHIP



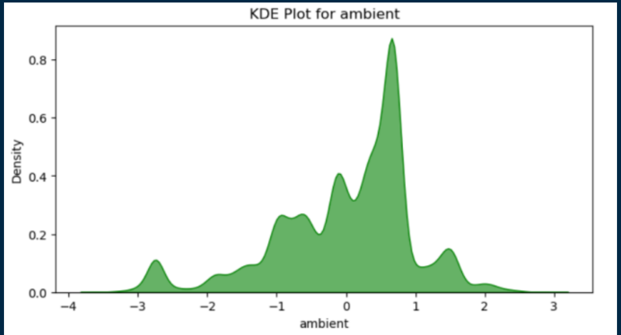
DISTRIBUTION COMPARISON

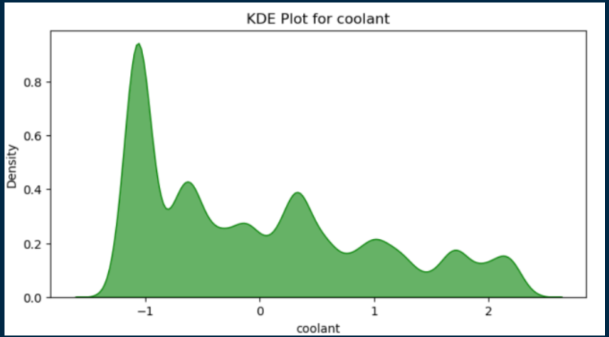


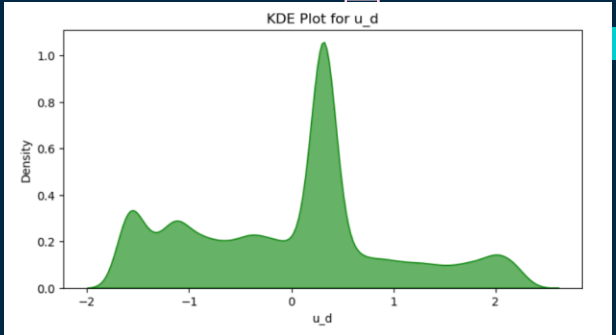
SCATTERPLOT

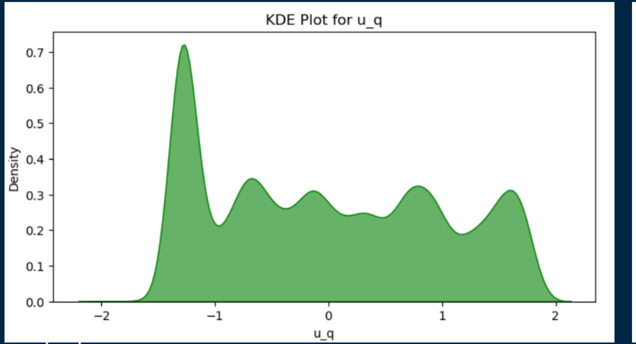


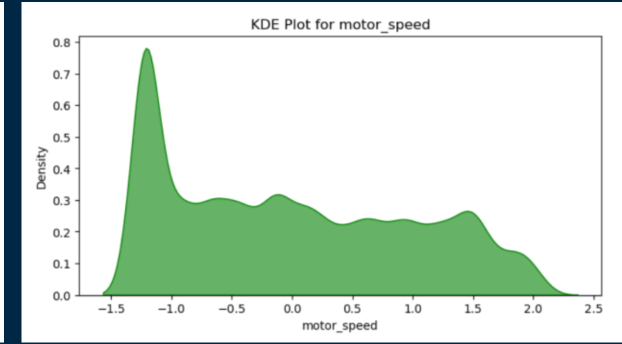
KDE PLOT

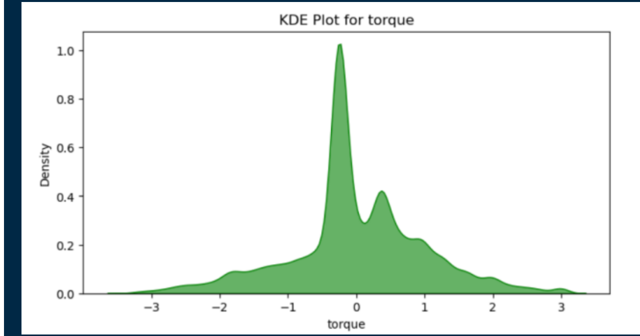

KDE Plot for ambient
Density
ambient

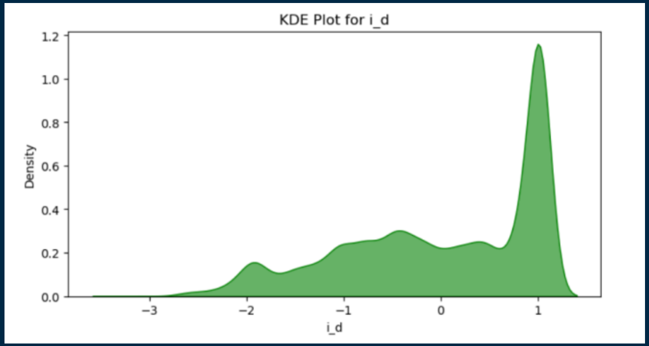

KDE Plot for coolant
Density
coolant

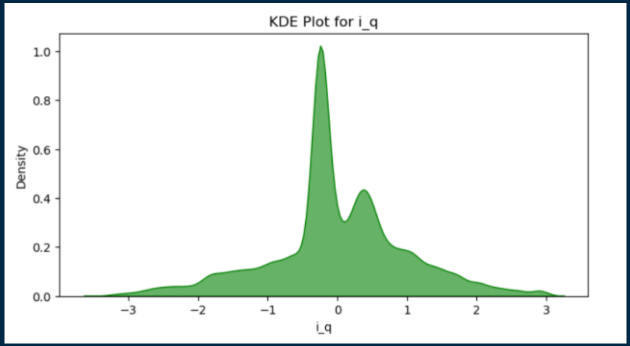

KDE Plot for u_d
Density
u_d

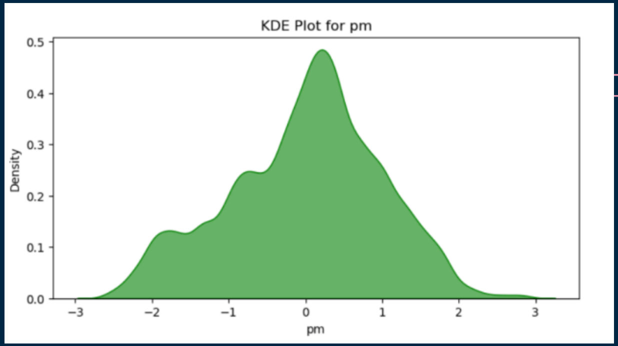

KDE Plot for u_q
Density
u_q

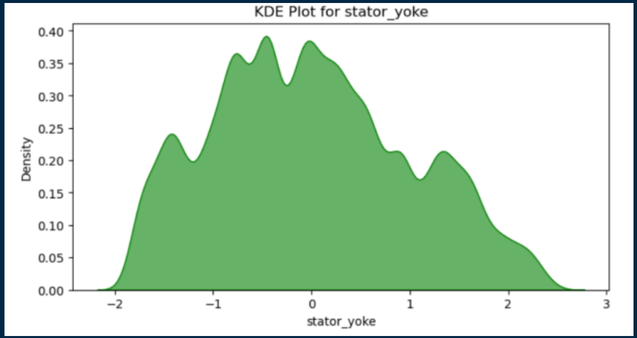

KDE Plot for motor_speed
Density
motor_speed

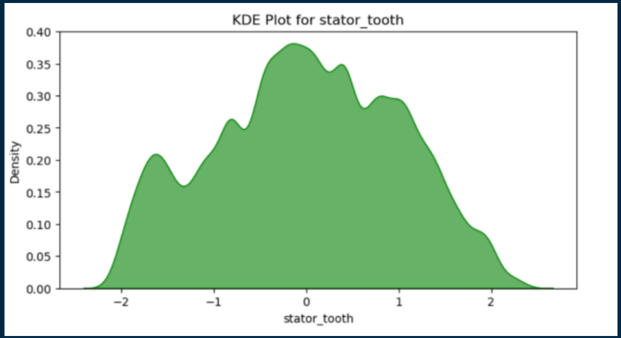

KDE Plot for torque
Density
torque

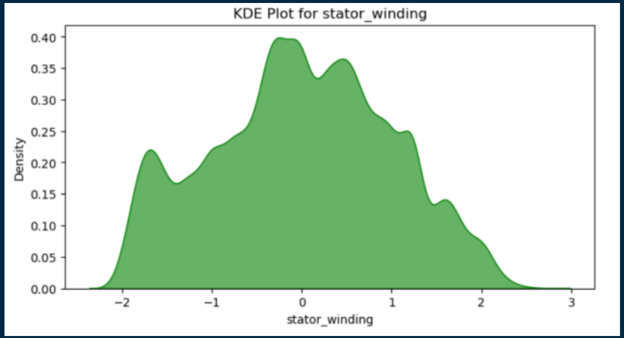

KDE Plot for i_d
Density
i_d

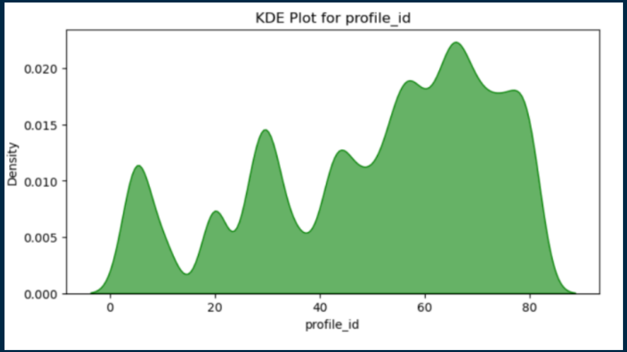

KDE Plot for i_q
Density
i_q


KDE Plot for pm
Density
pm


KDE Plot for stator_yoke
Density
stator_yoke


KDE Plot for stator_tooth
Density
stator_tooth


KDE Plot for stator_winding
Density
stator_winding


KDE Plot for profile_id
Density
profile_id

FINDINGS

Standardization Check

- The mean of most columns is close to zero.
- The standard deviation is close to 1.
- This confirms that the data is standardized.

Multicollinearity Check

High correlations exist between:

```
i_q and torque (0.996)

stator_tooth and stator_yoke (0.95)

stator_winding and stator_tooth (0.966)
```

These features carry redundant information and may need removal or transformation.

Unnecessary Columns

profile_id: Likely categorical and may not be useful in a regression model.

One of each highly correlated pair should be removed to reduce multicollinearity.

EDA SUMMARY

1. Data contains the following numerical columns:
['ambient', 'coolant', 'u_d', 'u_q', 'motor_speed', 'torque', 'i_d', 'pm', 'stator_tooth']

2. Correlation heatmap shows relationships between numerical variables.
3. Outliers were detected,visualized using boxplots and removed using IQR Method.
4. Pairplots,violin plots and scatter plots reveal feature distributions and relationships.
5. KDE plots provide additional insights into feature distributions.

MODEL BUILDING

We used various machine learning models for predicting motor speed.The models used are,

Model Name	Accuracy
Linear Regression	91.80%
Lasso	91.80%
Ridge Regression	91.80%
Decision Tree	99.96%
Random Forest Regression	99.98%
Lightgbm	99.97%
ANN Regression	99.96%
KNN	99.92%

Found that Random Forest is the best model for this dataset.So we are moving forward with Randomset.

```
Best Hyperparameters: {'positive': False, 'n_jobs': 1, 'fit_intercept': True, 'copy_X': True}
Best Score: 0.9177849684485843
R² Score: 0.9180255240953006
Mean Absolute Error: 0.21275086731260015
Mean Squared Error: 0.08214825278035429
```

Linear Regression

```
Best lasso parameters: {'max_iter': 1000, 'alpha': 0.001}
Best lasso R2: 0.9177359286556456
R² Score: 0.9180255240953006
Mean Absolute Error: 0.21275086731260015
Mean Squared Error: 0.08214825278035429
```

Lasso Regression

```
Best Ridge parameters: {'max_iter': 3000, 'alpha': 0.1}
Best Ridge R2: 0.9177849684489006
R² Score: 0.9180255240953006
Mean Absolute Error: 0.21275086731260015
Mean Squared Error: 0.08214825278035429
```

Ridge Regression

```
Best parameters:  {'min_samples_split': 5, 'min_samples_leaf': 4, 'max_depth': 30, 'criterion': 'squared_error'}
Best cross-validation accuracy:  0.998946105358332
R² Score: 0.9990691596680058
Mean Absolute Error: 0.012103684415361674
Mean Squared Error: 0.0009328136111501858
```

Decision Tree

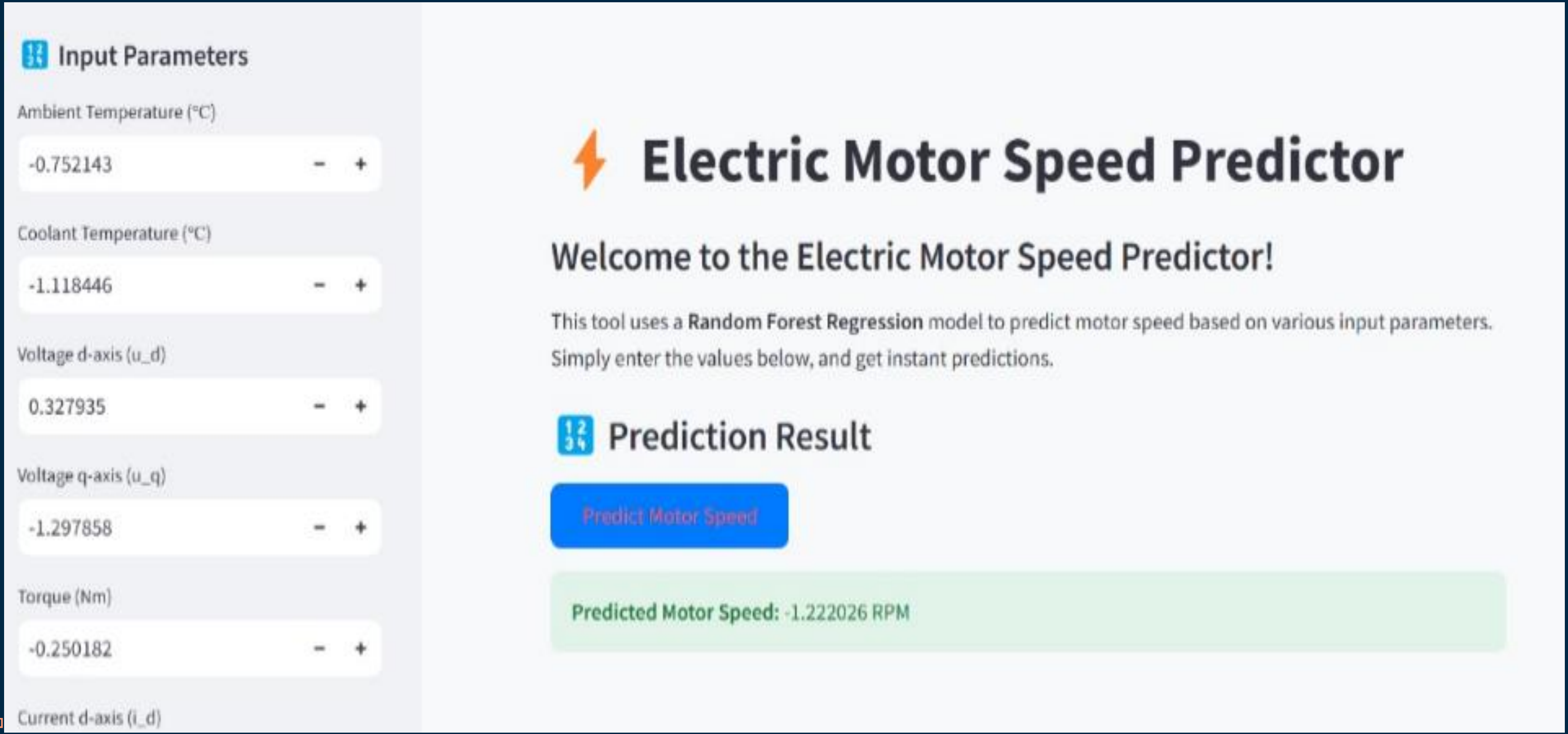
```
R² Score: 0.9998401118931888
Mean Absolute Error: 0.003913363138344772
Mean Squared Error: 0.0001602270520176478
```

Random Forest Regression

```
Best Parameters: {'num_leaves': 200, 'max_depth': 15, 'learning_rate': 0.1, 'lambda_l1': 1.0}
Best RMSE: 0.017043296583392563
R² Score: 0.9997628885775379
Mean Absolute Error: 0.007137565866361852
Mean Squared Error: 0.0002376140726068108
```

Lightgbm Tree

DEPLOYMENT


Input Parameters
Ambient Temperature (°C)
-0.752143
Coolant Temperature (°C)
-1.118446
Voltage d-axis (u_d)
0.327935
Voltage q-axis (u_q)
-1.297858
Torque (Nm)
-0.250182
Current d-axis (i_d)
Electric Motor Speed Predictor
Welcome to the Electric Motor Speed Predictor!
This tool uses a Random Forest Regression model to predict motor speed based on various input parameters.
Simply enter the values below, and get instant predictions.
Prediction Result
Predict Motor Speed
Predicted Motor Speed: -1.222026 RPM

CONCLUSION

In conclusion, predicting motor speed using machine learning offers a data-driven approach to enhance efficiency, reliability, and control in electric motor systems. By leveraging key attributes, the model can provide accurate speed estimations, enabling better performance monitoring and predictive maintenance. This approach can lead to reduced downtime, optimized energy consumption, and improved overall operational efficiency.

Do you have any questions?

THANKS