**BASIC INFORMATION**

**Title of Project:** Gas sensor temperature modulation

**Name:** Subhangi Dhasmana

**Email ID:** subhangi.dhasmana@gmail.com

**Contact Number:** 8287334826

**Google Drive Link:** https://drive.google.com/drive/u/0/folders/1agdjUnhYAEIqXhq1FnW8fWvg5VRsxer6

**GitHub Link:** https://github.com/SUBHANGI-DHASMANA/Gas-concentration-predictor

**Title**: **Gas sensor temperature modulation**

**Abstract:**

The Gas Sensor Temperature Modulation project aims to explore the effectiveness of different machine learning algorithms in predicting gas sensor readings based on temperature modulation. The project utilizes a carefully curated dataset consisting of temperature variations and corresponding gas sensor measurements. By evaluating the performance of five popular regression models, namely Linear Regression, RandomForestRegressor, DecisionTreeRegressor, Ridge, and Lasso, the project aims to identify the most suitable model for accurate prediction in this domain.

The dataset used in this project comprises temperature data along with corresponding gas sensor readings. Before model training, the dataset undergoes preprocessing steps such as handling missing values, removing outliers, and normalizing the features to ensure optimal model performance. Careful attention is paid to ensure the data quality and integrity throughout the process.

The Gas Sensor Temperature Modulation project involves training five different regression models on the preprocessed dataset. The Linear Regression model provides a baseline performance for comparison, while the RandomForestRegressor and DecisionTreeRegressor capture non-linear relationships. Additionally, the Ridge and Lasso models help address potential issues of overfitting and multicollinearity.

The Gas Sensor Temperature Modulation project contributes to the field of machine learning by comparing the performance of different regression models on the task of predicting gas sensor readings based on temperature modulation. The findings can be used to guide future research and real-world applications in areas such as environmental monitoring, industrial processes, and air quality assessment.

**Keywords:**

Regression,

Linear Regression,

Gas Concentration Time Series Prediction,

Random Forest regressor,

Decision tree regressor

**1. Introduction:**

The provided dataset for the Gas Sensor Temperature Modulation project consists of temperature variations and corresponding gas sensor measurements. This dataset serves as the foundation for training and evaluating machine learning models to predict gas sensor readings based on temperature modulation.

The dataset includes temperature data, which represents the varying temperatures recorded during the gas sensor measurements. Temperature modulation is a key factor that influences gas sensor responses, as changes in temperature can affect gas concentration levels. Therefore, capturing the relationship between temperature and gas sensor readings is crucial for accurate predictions.

The gas sensor measurements in the dataset represent the output responses of the gas sensors, which are influenced by the gas concentrations present in the environment. These measurements serve as the target variable for the prediction task. By analyzing the relationship between temperature and gas sensor readings, the models aim to accurately predict the gas sensor responses based on the temperature modulation patterns.

The Gas Sensor Temperature Modulation project utilizes five different machine learning regression models to predict gas sensor readings based on temperature modulation. Here is a brief description of each model:

**Linear Regression:**

Linear Regression is a simple and interpretable model that assumes a linear relationship between the input features (temperature) and the target variable (gas sensor readings). It provides a baseline performance for comparison and helps understand the overall trend between temperature and gas sensor responses.

**RandomForestRegressor**

RandomForestRegressor is an ensemble learning model that combines multiple decision trees to make predictions. It is effective at capturing non-linear relationships and interactions between features. The model creates a forest of decision trees and averages their predictions to generate the final output.

**DecisionTreeRegressor**

DecisionTreeRegressor is a single decision tree-based model that splits the data based on the input features and predicts the target variable at each leaf node. It captures non-linear relationships and can handle complex interactions. However, it may be prone to overfitting if not properly regularized.

**Ridge Regression** Ridge Regression is a regularized linear regression model that adds a penalty term to the loss function. It helps address the issue of multicollinearity, which can occur when there are high correlations between the input features. Ridge Regression encourages smaller coefficient values, reducing the impact of highly correlated features.

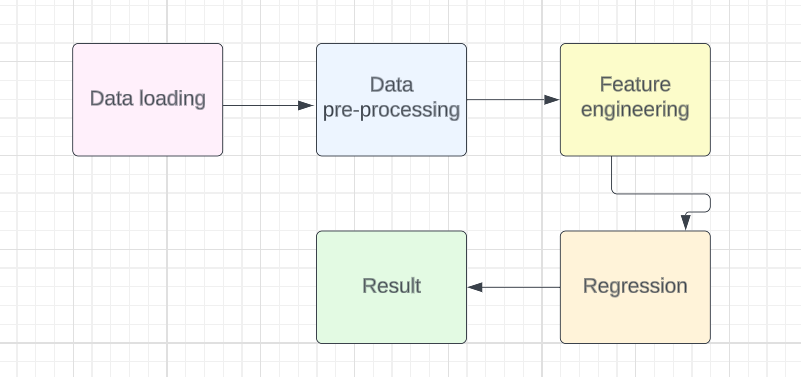
**Lasso Regression**

Lasso Regression is another regularized linear regression model that introduces a penalty term to the loss function. It not only addresses multicollinearity but also performs feature selection by encouraging sparse coefficient values. Lasso Regression can be useful in situations where only a subset of features has a significant impact on the target variable.

Each model is trained on the pre-processed dataset and evaluated using standard performance metrics such as mean squared error (MSE), mean absolute error (MAE), and coefficient of determination (R-squared). These metrics provide insights into the accuracy, precision, and goodness of fit of each model, allowing for a comparison of their predictive capabilities in the Gas Sensor Temperature Modulation project.

**2. Proposed Methodology [ Pictorial Diagram and Explanation of each sub modules]**

**Diagram / Flow Chart**

****

**a. Datasets**

Dataset is related to Gas sensor temperature modulation. It comprises the various features such as

| **Time (s)** | **CO (ppm)** | **Humidity (%r.h.)** | **Temperature (C)** | **Flow rate (mL/min)** | **Heater voltage (V)** | **R (MOhm)** |
| --- | --- | --- | --- | --- | --- | --- |

. It has only float values in all features.

**b. Pre processing**

All features are float values. There are no NULL values in the rows.

Shape of data frame = (295719, 20)

**c. Feature Scaling**

For desirable model training we have to remove features which are not in use.

So, we removed CO, Humidity, Temperature, Flow rate and Heater columns.

**d. Train Test Split**

We divide the data in Training and Testing data so that we can check the score and test our data.

**e. Model Training and Testing with the help of Regression**

Models used are - Linear regression, Random forest regressor, Decision Tree regressor, Ridge and Lasso

**f. Performance Measure**

There are various Performance metrics that can help to measure performance like Confusion Matrix, Classification Report, F1 Score and Accuracy Score, etc.

**3. Result & Discussion**

Then we apply the various Classification models after splitting the dataset into training and testing data -

Linear Regression

Random forest regressor

Decision Tree regressor

Ridge

Lasso

After applying the model we calculate the scores.

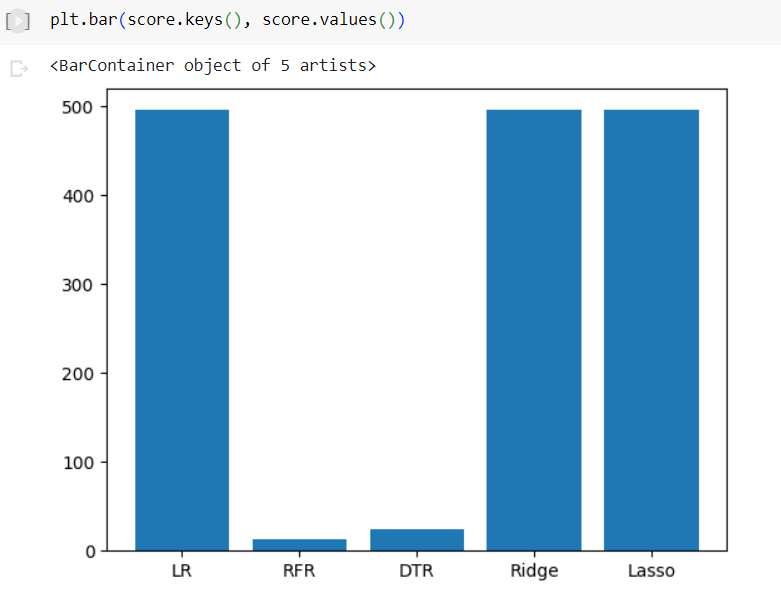
LR :- 495.42470887462537

RFR :- 12.525798828229167

DTR :- 23.98265944208694

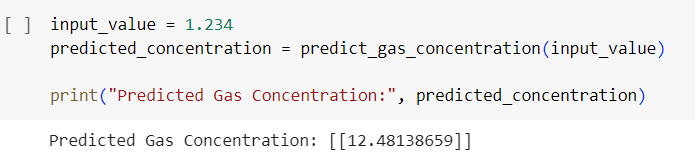
Ridge :- 495.42470887462537

Lasso :- 495.4247115869228



We can see the best accuracy is shown by Logistic Regression, Ridge and Lasso on our data.

After that I have created predict\_gas\_concentration function which takes time as input and give concentration as output



4. **Conclusion**

From the Gas concentration temperature modulation Project we could conclude that the Best Model that fits the data is **LINEAR REGRESSION**.

This model is perfect to predict concentration using time series.

In conclusion, the Gas Sensor Temperature Modulation project aimed to predict gas sensor readings based on temperature modulation using five different machine learning regression models. The project utilized a carefully curated dataset comprising temperature variations and corresponding gas sensor measurements.

Through the evaluation of Linear Regression, RandomForestRegressor, DecisionTreeRegressor, Ridge, and Lasso models, valuable insights were gained regarding the effectiveness of each model in capturing the relationship between temperature and gas sensor responses. The models were trained on a preprocessed dataset, which underwent steps such as handling missing values, removing outliers, and normalizing features to ensure optimal model performance.

The project's findings demonstrated that each model had its strengths and characteristics. Linear Regression provided a baseline understanding of the overall trend between temperature and gas sensor readings. RandomForestRegressor and DecisionTreeRegressor proved effective in capturing non-linear relationships and interactions. Ridge and Lasso Regression models addressed issues such as multicollinearity and feature selection.

By accurately predicting gas sensor readings based on temperature modulation, this project enables improved decision-making and enhanced gas monitoring and control systems. The insights gained from the project can aid in developing advanced guidance systems and personalized recommendations for industries and organizations involved in gas monitoring and management.

**References**

1. <https://www.sciencedirect.com/science/article/abs/pii/S0925400516302817>
2. <https://www.sciencedirect.com/science/article/abs/pii/S0925400516302817>
3. <https://www.sciencedirect.com/science/article/abs/pii/S0925400521000770>
4. <https://www.sciencedirect.com/science/article/abs/pii/S0921510717303574>
5. <https://www.sciencedirect.com/science/article/abs/pii/S0925400599002415>
6. <https://s3-us-west-2.amazonaws.com/ieeeshutpages/xplore/xplore-shut-page.html>
7. <https://www.mdpi.com/2079-9292/7/4/54>