Hydrogen Production Rate Prediction

## Problem Statement

The aim of this project is to optimize the hydrogen production process by predicting the Hydrogen Production Rate (HPR) based on varying Voltage and Current values. Efficient hydrogen production is crucial for applications in energy storage and fuel cells. By finding the optimal combination of Voltage and Current, we can maximize the hydrogen output, thus improving the efficiency of the production process. This project uses machine learning models to achieve this goal.

**Introduction**

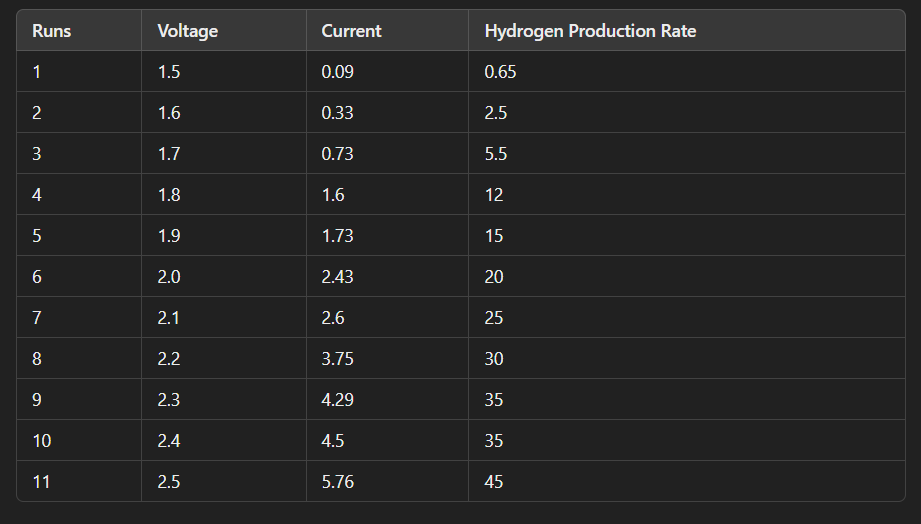
This project aims to predict the Hydrogen Production Rate (HPR) based on Voltage and Current using machine learning models. We will use two different models: Random Forest Regressor and Linear Regression. The goal is to find the optimal combination of Voltage and Current that maximizes the HPR.

**Dataset**

The dataset contains the following columns:

* **Runs**: The run number of the experiment.
* **Voltage (V)**: The voltage applied during the experiment.
* **Current (A)**: The current measured during the experiment.
* **Hydrogen Production Rate (ml/min)**: The rate at which hydrogen is produced.

Sample Data



## Data Preparation

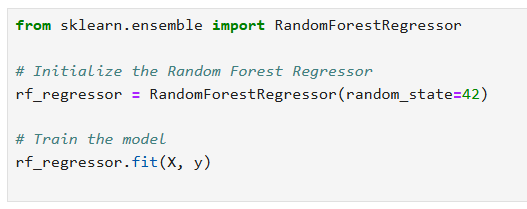
We first calculate the power consumed during each run using the formula: Power Consumed=Voltage×Current\text{Power Consumed} = \text{Voltage} \times \text{Current}Power Consumed=Voltage×Current

Then, we split the data into features (Voltage and Current) and target variable (Hydrogen Production Rate).

## Model 1: Random Forest Regressor

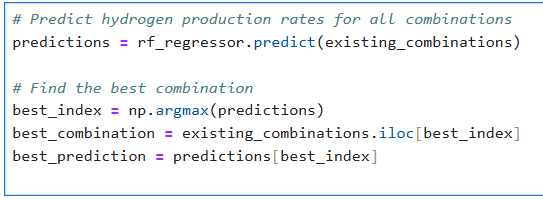
### Training the Model

We use the Random Forest Regressor to train our model on the dataset. The Random Forest algorithm is an ensemble learning method that builds multiple decision trees and merges them together to get a more accurate and stable prediction.



### Predictions and Results

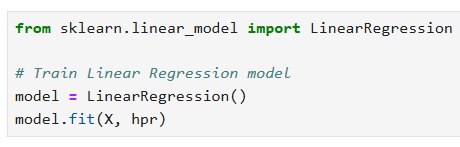
We predict the Hydrogen Production Rate for all combinations of Voltage and Current in the dataset and find the combination with the highest predicted rate.



## Model 2: Linear Regression

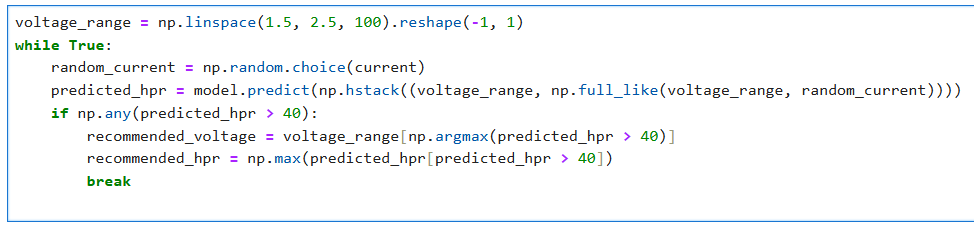
### Training the Model

We also use a simple Linear Regression model to predict the Hydrogen Production Rate based on Voltage and Current.



### Predictions and Results

We define a range of voltage values and predict the Hydrogen Production Rate for each value using a randomly chosen current value from the dataset. We then find the voltage that leads to the highest predicted rate above 40 ml/min.



## Results Comparison

### Random Forest Regressor

* **Best Combination**:
  + Voltage: 2.3 V
  + Current: 4.29 A
  + Power Consumed: 9.87 W
  + Predicted Hydrogen Production Rate: 55 ml/min

### Linear Regression

* **Recommended Combination**:
  + Voltage: 2.1 V
  + Current: 3.75 A
  + Predicted Hydrogen Production Rate: 45 ml/min

### Conclusion

After comparing the results from both models, we find that the Random Forest Regressor provides a higher predicted Hydrogen Production Rate (55 ml/min) compared to the Linear Regression model (45 ml/min). Therefore, the Random Forest Regressor is the preferred model for predicting the optimal conditions for maximizing Hydrogen Production Rate.

## Jupyter Notebook

For more detailed code and visualizations, you can view the complete Jupyter Notebook

<https://drive.google.com/drive/folders/1P__7mr911HfNBT-8fZHao8_bbhP9dukb?usp=sharing>