Application of Supervised Machine Learning Techniques in Fuel Cell Modelling

A Report submitted in Partial Fulfilment of the Requirement of the Award Degree of

BACHELOR OF TECHNOLOGY
in
ELECTRICAL & ELECTRONICS ENGINEERING
by

SANIYA BHARGAVA (201EE151)
CHITHRITHA V KUMAR (201EE217)

Under the guidance of
Dr. U. Vinatha
Associate Professor
Department of Electrical and Electronics Engineering

NITK Surathkal

DEPARTMENT OF ELECTRICAL & ELECTRONICS ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA

DECLARATION

We, hereby declare the report of the project work entitled "**Application of Supervised Machine Learning techniques in Fuel Cell Modelling**" which is being submitted to National Institute of Technology Karnataka, Surathkal for the award of the degree of Bachelor of Technology in Electrical and Electronics Engineering, is a bonafide report of the work carried out by us. The material contained in this report has not been submitted to any university or Institution for the award of any degree.

Roll No	**Name**	**Signature**
201EE151	SANIYA BHARGAVA	
201EE217	CHITHRITHA V KUMAR	

Department of Electrical and Electronics Engineering
Place: NITK Surathkal
Date : 28th November 2023

CERTIFICATE

We, hereby declare the report of the project work entitled "**Application of Supervised Machine Learning techniques in Fuel Cell Modelling**" which is being submitted by

S. no	Roll No	Name
1.	201EE151	SANIYA BHARGAVA
2.	201EE217	CHITHRITHA V KUMAR

As a record of the work carried out by them, it is accepted as the B.Tech Project work report submission in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology in Mining Engineering of National Institute of Technology Karnataka, Surathkal.

Guide(s):
Dr. U. Vinatha
Associate Professor
Department of Electrical and Electronics Engineering
NITK Surathkal

Chairman - DUGC
(signature with Date and Seal)

ACKNOWLEDGMENT

At the very outset, we wish to place on record our sincere thanks and gratitude to Dr. U. Vinatha, Associate Professor, Department of Electrical & Electronics Engineering, for giving us the invaluable opportunity to work on the project. We extend our heartfelt appreciation to Ajeya K, our mentor, whose guidance and encouragement played a pivotal role in shaping the success of this endeavor.

Ajeya K's unwavering support, insightful inputs, and commitment to our growth have been instrumental at every stage of the project. We are deeply grateful for the time and expertise shared, which greatly enriched our learning experience.

This project would not have been possible without the support of our peers and the resources provided by the Department of Electrical & Electronics Engineering. We express our gratitude to everyone who has been part of this journey, contributing to its success in various ways.

Thank you all for your unwavering support and belief in our capabilities.

Names (Roll.No.)
SANIYA BHARGAVA (201EE151)
CHITHRITHA V KUMAR (201EE217)

ABSTRACT

In recent years, machine learning (ML) has received growing attention and it has been used in a wide range of applications. This report explores the application of supervised machine learning techniques in the modeling of fuel cells, aiming to enhance the understanding and predictive capabilities of these complex energy conversion systems. Fuel cells play a crucial role in the transition towards sustainable energy solutions, and accurate modeling is essential for optimizing their performance and efficiency. The project leverages various supervised machine learning algorithms, particularly regression models, to analyze and predict the behavior of fuel cells under different operating conditions.

The intersection of electrical engineering and machine learning has sparked a revolutionary approach to understanding and optimizing fuel cell systems crucial for sustainable energy solutions. Leveraging supervised machine learning, particularly regression models, this innovative endeavor aims to decode the intricate behaviors of fuel cells across diverse operating conditions. By analyzing extensive datasets, these models unravel complex relationships between input parameters and fuel cell performance, providing predictive insights essential for system optimization. This integration promises real-time adaptability, empowering decision-making processes and elevating the efficiency and reliability of fuel cell technology. It signifies a transformative synergy between traditional engineering principles and cutting-edge machine learning, offering a more agile, data-driven path toward sustainable energy solutions.

TABLE OF CONTENTS

Polymer electrolyte membrane fuel cell	PEMFC
Artificial Neural Network	ANN
Support Vector Machine	SVM
Support Vector Machine Regressor	SVR
K Nearest Neighbours	KNN
Recurrent Neural Network	RNN
Rectified Linear Unit	ReLU
Mean Square Error	MAE
Mean Square Error	RMSE
Mean Absolute Error	MAE

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 - Introduction

1.1 General

Proton exchange membrane fuel cell (PEMFC) is considered essential for climate change mitigation, and a fast and accurate model is necessary for its control and operation in practical applications. In this study, various machine learning methods are used to develop data-based models for PEMFC performance attributes and internal states. Techniques such as Artificial Neural Network (ANN), Support Vector Machine Regressor (SVR) and K-Nearest Neighbours (KNN) are used to predict the cell voltage and Current Density for various operating conditions. Two different sets of data are considered in this study.

1.2 Fuel cells

A fuel cell uses the chemical energy of hydrogen or other fuels to cleanly and efficiently produce electricity. If hydrogen is the fuel, the only products are electricity, water, and heat. Fuel cells are unique in terms of the variety of their potential applications; they can use a wide range of fuels and feedstocks and can provide power for systems as large as a utility power station and as small as a laptop computer.

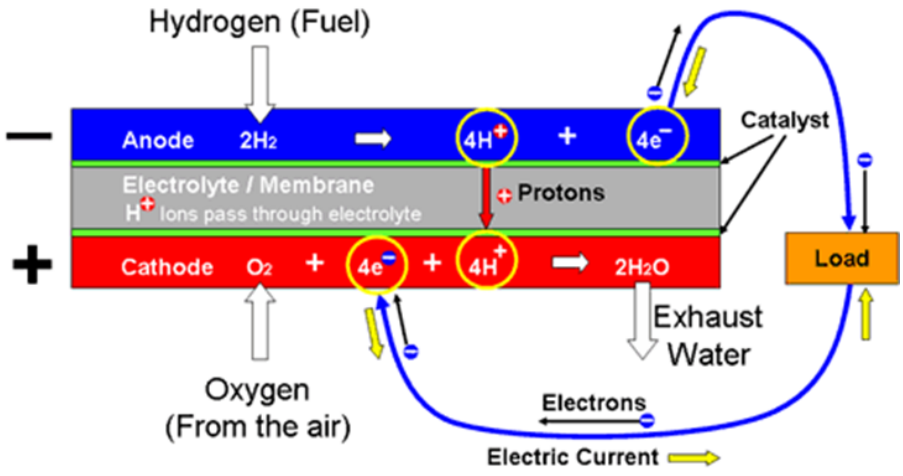

Fig. 1 Fuel Cell electrochemistry

1.2.1 How does a fuel cell work

Fuel cells operate on a fascinating principle, resembling batteries but distinguished by their continuous operation without the need for recharging. Their functionality is sustained as long as a constant supply of fuel is maintained. A fuel cell consists of two electrodes—a negative electrode (or anode) and a positive electrode (or cathode)—sandwiched around an electrolyte. A fuel, such as hydrogen, is fed to the anode, and air is fed to the cathode. In a hydrogen fuel cell, a catalyst at the anode separates hydrogen molecules into protons and electrons, which take different paths to the cathode. The electrons go through an external circuit, creating a flow of electricity. The protons migrate through the electrolyte to the cathode, where they unite with oxygen and the electrons to produce water and heat.

1.2.2 Fuel Cell Classification and Types

Fuel cells, revolutionary devices in the realm of energy conversion, are classified based on the type of electrolyte they employ. This classification profoundly influences electrochemical reactions, catalyst requirements, operating temperatures, fuel compatibility, and application suitability. Various fuel cell types, each with distinct advantages, limitations, and potential applications, are under development. Various kinds of Fuel cells are

1. *Polymer Electrolyte Membrane (PEM) Fuel Cells*:Also known as Proton Exchange Membrane (PEM) fuel cells, these exhibit high power density, low weight, and volume advantages. Employing a solid polymer electrolyte and porous carbon electrodes with platinum or platinum alloy catalysts, PEM fuel cells operate at around 80°C, allowing quick startup and enhanced durability.Primarily used in transportation applications such as cars, buses, and heavy-duty trucks.
2. *Direct Methanol Fuel Cells (DMFCs)*:Description: Unlike most fuel cells powered by hydrogen, DMFCs use pure methanol directly mixed with water as the fuel. They offer advantages in energy density, ease of transportation, and are commonly used in portable applications like cell phones and laptops.
3. *Alkaline Fuel Cells (AFCs)*: Developed early, AFCs use a potassium hydroxide solution as the electrolyte. Recent advancements include AFCs with a polymer membrane electrolyte, closely resembling PEM fuel cells. Challenges include susceptibility to CO2 and liquid electrolyte issues.
4. *Phosphoric Acid Fuel Cells (PAFCs)*: Utilizing liquid phosphoric acid as an electrolyte, PAFCs are a mature technology, commonly used for stationary power generation. They are tolerant of impurities in fossil fuels and efficient in co-generating electricity and heat.
5. *Molten Carbonate Fuel Cells (MCFCs)*:High-temperature fuel cells using a molten carbonate salt mixture as an electrolyte. Non-precious metals can be used as catalysts, reducing costs. Internal reforming capability and improved efficiency make them suitable for various applications like Natural gas and coal-based power plants.

6. *Solid Oxide Fuel Cells (SOFCs)*:Utilizing a hard, non-porous ceramic compound as the electrolyte, SOFCs operate at high temperatures, eliminating the need for precious-metal catalysts. They are sulfur-resistant and can use a variety of fuels, but challenges include slow startup and high operating temperatures.

1.3 Machine Learning and its Applications

Machine learning represents a powerful methodology for enhancing the performance and efficiency of fuel cells, integral devices converting chemical energy into electrical power. This transformative approach involves training models on existing data to create fitting models, enabling highly accurate predictions for nonlinear problems with computational efficiency. In the domain of fuel cell applications, machine learning plays a pivotal role in various aspects. It optimizes operating parameters by analyzing real-time sensor data, such as temperature, humidity, cell voltage and pressure, leading to improved efficiency, extended lifespan, and reduced maintenance costs. Predictive maintenance capabilities allow machine learning to foresee potential component failures, enabling proactive measures and minimizing downtime. Additionally, machine learning excels in fault detection and diagnostics, swiftly identifying anomalies and deviations in fuel cell systems for prompt corrective actions. The technology aids in fuel cell design by analyzing materials properties, suggesting optimal combinations for enhanced component performance. Machine learning also dynamically controls fuel cell operation, adjusting parameters in real-time to optimize performance based on changing conditions. Moreover, it contributes to energy management in hybrid systems, monitors fuel quality, conducts data analysis for research purposes, aids in emission reduction, and enhances user experience in applications like fuel cell-powered vehicles. This multifaceted integration of machine learning showcases its versatility and profound impact on advancing fuel cell technology across diverse domains. In this report we will be using Supervised Machine Learning Regression models to predict Current Density and Cell Voltage of fuel cells using parameters as input features such as relative humidity, losses, power, etc.

1.4 Objectives

In pursuit of accurate fuel cell modeling, the focus is on obtaining precision through a specified set of parameters. The key lies in identifying the most suitable machine learning technique to ensure the precise representation of fuel cell characteristics. The goal is to refine modeling approaches to achieve optimal accuracy in capturing the intricate details of fuel cell behavior. This involves a strategic selection of parameters and a nuanced understanding of the machine learning tools employed. Ultimately, the aim is to enhance the reliability and fidelity of fuel cell models through the application of tailored machine learning methodologies.

CHAPTER 2 - Model Development

2.1 Data acquisition and description

Dataset-1: It serves as a crucial component in this project, offering insights into the behavior of a fuel cell system under varying conditions. This dataset is derived from simulation mode, encompassing five essential input features: activation drop loss, ohmic drop loss, concentration drop loss, power, and efficiency. The simulation results provide a comprehensive understanding of the relationships between these inputs and the output feature, current density. Obtained through meticulous simulation processes, this dataset forms the foundation for training and evaluating our models. Figure 2 represents the relation of each input parameter with the output feature.

Dataset-2: Complementing the simulated data, Dataset-2 is sourced from internet repositories, adding real-world diversity to our project. This dataset comprises four distinct input features: current density, relative humidity, pressure, and power density. Acquired from external sources, the data represents empirical observations in fuel cell environments. The output feature of this dataset is the cell voltage of a fuel cell, a critical parameter influencing overall system performance. By incorporating both simulated and real-world datasets, our project aims to develop models that can generalize well across different operational scenarios, enhancing the robustness and applicability of our findings.

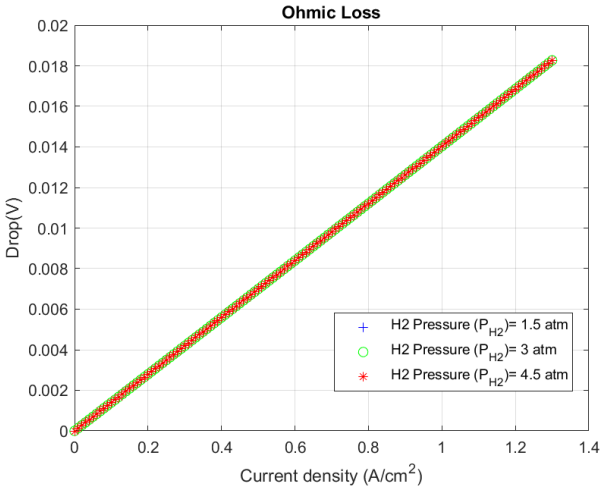

Fig. 2(a) Ohmic Drop Loss (V) Vs Current Density (A/cm^2)

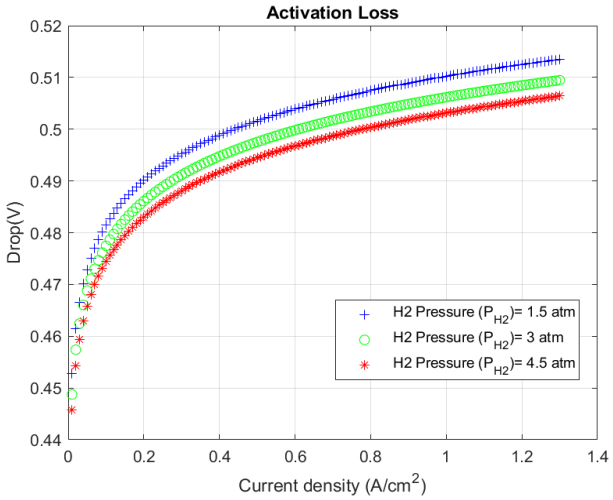

Fig. 2(b) Activation Drop Loss (V) Vs Current Density (A/cm^2)

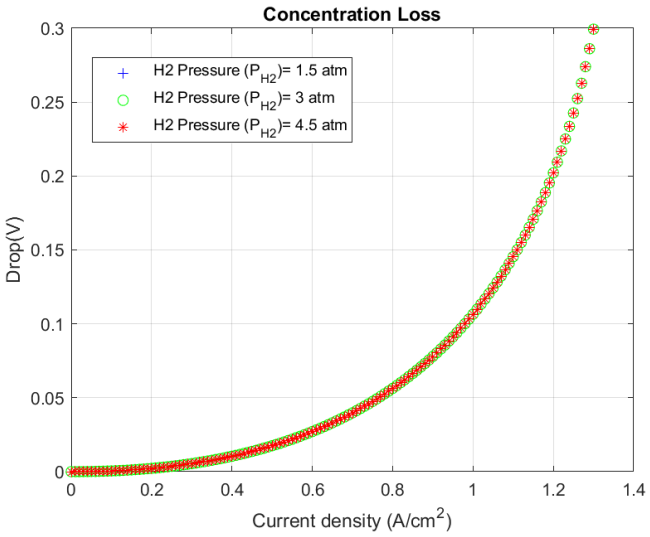

Fig. 2(c) Concentration Drop Loss (V) Vs Current Density (A/cm^2)

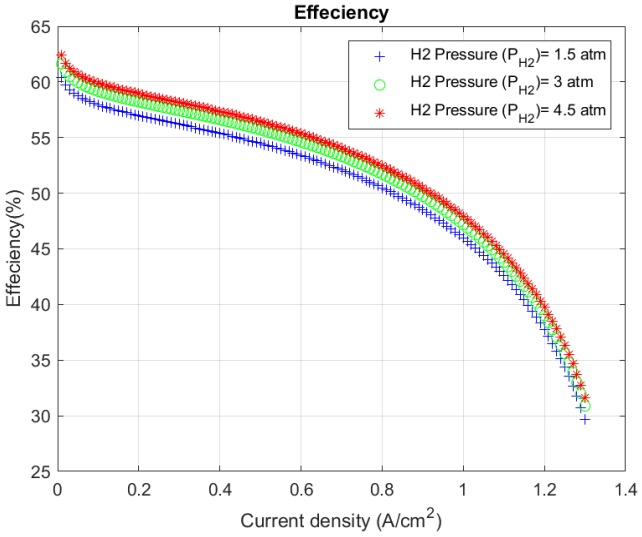

Fig. 2(d) Efficiency(%) Vs Current Density (A/cm^2)

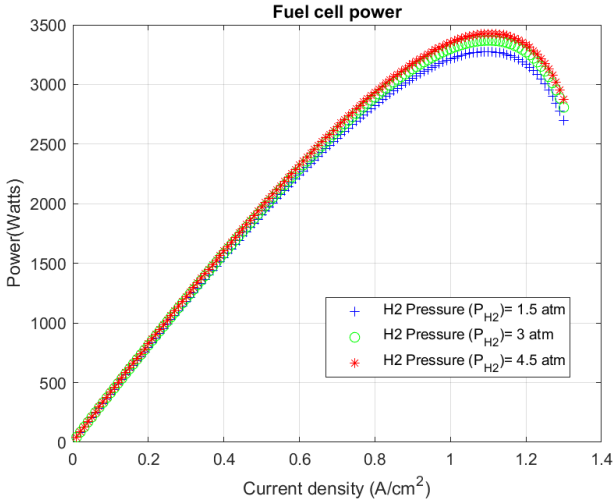

Fig. 2(e) Power(W) Vs Current Density (A/cm^2)

2.1.1 Normalization of data

Min-Max Scaling, also known as Min-Max normalization or feature scaling, is a preprocessing technique employed to rescale numerical features within a specific range. This transformation ensures that all features have a consistent scale, preventing certain features from dominating others due to differences in their magnitudes.The Min-Max Scaling process transforms the original feature values into a new scale, typically ranging from 0 to 1.

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

2.1.2 Data division

A substantial 80% of the data is entrusted to the model during the training phase, allowing it to absorb the underlying patterns and intricacies. The remaining 20% is reserved for a critical mission – to rigorously test the model's mettle on previously unseen data.

2.1.3 Loss Functions

In the context of our regression analysis, we employ two essential metrics to evaluate the performance of our predictive model: Mean Absolute Error (MAE) and Mean Squared Error (MSE).

1. *Mean Absolute Error* - It measures the average absolute difference between the predicted values and the actual values. The formula for MAE is as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \widehat{y_i}\right|$$

2. *Mean Squared Error* - It measures the average of the squared differences between the predicted values and the actual values. The formula for MSE is as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y_i}\right)^2$$

R-squared (Coefficient of Determination) is a statistical metric commonly used to evaluate the goodness of fit of a regression model. It provides a measure of how well the independent variable(s) in the model explain the variability observed in the dependent variable. R-squared is a value between 0 (indicates that the model does not explain any of the variability in the dependent variable.)and 1(indicates that the model explains all of the variability in the dependent variable).

$$R_2 = 1 - \frac{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y_i}\right)^2}{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \overline{y_i}\right)^2}$$

where,

- n = number of data points in the dataset
- y_i = The actual (true) value for the *i*-th data point.
- $\widehat{y_i}$ = The predicted value for the *i*-th data point
- $\overline{y_i}$ = The mean of the actual values

2.2 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is a combination of neurons which are the basic processing units that regulate the computational activity based on the designated activation function and associated inputs.This architecture is a clone of the biological neuron in terms of functionality and with the help of a backpropagation algorithm, it can solve and model regressions accurately.

From a mathematical perspective, the following operation can be described as follows, internally all the inputs received by a neuron (x_1, x_2, x_3, ….., x_r) are multiplied by their corresponding weights (w_1, w_2, w_3, ……, w_r) and summed up together. Later this summation of products is assigned a threshold also referred to as bias (w_0) where the result of this operation is passed to an activation function. A schematic representation of the neuron is shown in Fig. 3 and the mathematical representation of the neuron is shown in the following equation:

$$y(t) = \Phi\left(\sum_{i=1}^{r}\left(w_i x_i - w_0\right)\right)$$

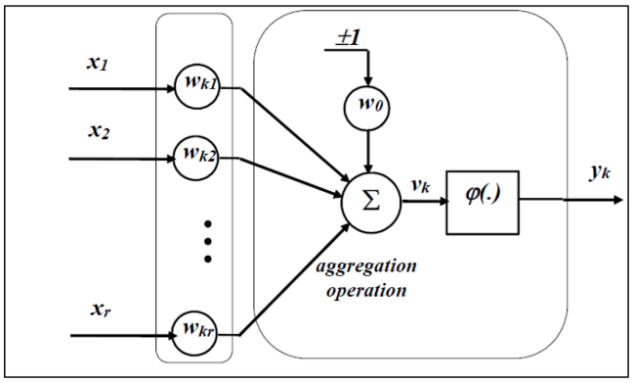

Fig. 3 Schematic of a Neuron

where $y(t)$ is the output of the neuron and Φ is the activation function, the threshold w_0 determines whether an aspect of a neuron generates an output or if it is inactive. e. This entire set of these weights and biases are considered hyperparameters which are calibrated depending on the activation function and associated error at each neuron. ANN here is used as a supervised machine learning tool where relevant inputs also referred to as feature vectors along with their corresponding outputs are provided to the model for training.

2.2.1 Initial ANN parameters

Table 1 Initial ANN parameters considered for Dataset 1

Input features	5
Output feature	1
Hidden Layers	2
Nodes / Neurons	Hidden layer - 1 → 64 Hidden layer - 2 → 32 Final layer → 1
Activation Function	ReLu
Epochs	100
Loss function	MAE / MSE
Optimizer	Adam optimizer

2.2.2 Input and output features

Table 2 Input and Output features of Dataset 1

Input Features	1. Activation drop loss (V) 2. Concentration drop loss (V) 3. Ohmic drop loss (V) 4. Power (W) 5. Efficiency (%)
Output Feature	1. Current Density (A/cm^2)

2.2.3 Activation Function

The activation functions incorporate the non-linearity to the neural networks depending on the sensitivity of the input feature vectors. There are various nonlinear functions such as Sigmoid (Logistic function), Hyperbolic Tangent (Tanh) and Rectified Linear Unit (ReLU). ReLU is the activation function that is considered in this study. It is mostly implemented in deep learning and the ReLU function is composed of two linear pieces to account for non-linearities. Unlike sigmoid and Tanh, the output of ReLU does not have a maximum value helping it to address the vanishing gradient issue and is computationally easy as there is no exponential operation [37]. After considering the initial results and literature it is

decided to opt for the ReLU function to save computational expense as the activation function for the neural network considering the linear and nonlinear attributes.

2.3 Support Vector Machine (SVM)

Support Vector Machine (SVM) is one of the most used techniques to resolve classification problems. The SVM helps us to find a line/hyperplane (in multidimensional space) that separates two classes. A decision boundary serves as a demarcation line to split the classes from each other. Unlike Support Vector Machine classifiers that aim to find a hyperplane to separate different classes, SVM Regressors are designed to predict a continuous outcome.

2.3.1 Initial SVR parameters

Table 3 Initial SVR parameters considered for Dataset 1

Input features	Dataset 1 → 5
Output feature	1
Epsilon	0.1
C	1.0
Kernel	rbf - radial basis function
Degree of polynomial	3

2.3.2 Input and output features

Table 4 Input and Output features of Dataset 1

Input features - Dataset 1	1. Activation drop loss (V) 2. Concentration drop loss (V) 3. Ohmic drop loss (V) 4. Power (W) 5. Efficiency (%)
Output Features - Dataset 1	1. Current Density (A/cm^2)

2.4 Cascaded Models

Cascaded regression models refer to a series of regression models that are applied sequentially to refine predictions or estimates progressively. Each subsequent model in the cascade aims to correct or enhance the output of the previous one, resulting in a more accurate and fine-tuned final prediction.

2.4.1 Input and output features

Table 5 Input and Output features

Features	**Dataset 1**	**Dataset 2**
Input features	1. Activation drop loss (V) 2. Concentration drop loss (V) 3. Ohmic drop loss (V) 4. Power (W) 5. Efficiency (%)	1. Current Density (A/m^2) 2. Power_Density (mW/cm2) 3. Pressure (psig) 4. Relative_Humidity(%)
Output features	1. Current Density (A/m^2)	1. Cell Voltage (V)
Datapoints	131 points	784 points

2.4.2 Cascading Neural Networks (NN) into Support Vector Machines (SVM)

In the implemented architecture, a cascaded framework has been established by integrating a Neural Network (NN) and a Support Vector Machine (SVM). This cascaded structure is designed such that the output of the Neural Network, which takes multiple input feature vectors (5 input features → Dataset 1 and 4 input features → Dataset 2), serves as the input to the Support Vector Machine, allowing for a sequential fine-tuning of predictions and predicting a single output. The Neural Network, recognized for its capacity to capture intricate patterns and relationships in data, acts as the initial model in the cascade. Its output, representing a learned feature representation, is subsequently fed into the Support Vector Machine.

2.4.3 Cascading Support Vector Machines (SVM) into Neural Networks (NN)

This experimental framework involves the application of a cascaded architecture, integrating Support Vector Machines (SVM) into Neural Networks (NN), to model and analyze two distinct datasets. Dataset 1 comprises five input features, while Dataset 2 is characterized by four input features and one output feature.

In the cascaded model, the Support Vector Regressor serves as the initial stage, capturing intricate patterns and relationships in the data, leveraging the five input features in Dataset

1. Subsequently, the output of the Support Vector Regressor, representing a learned feature representation, seamlessly transitions to the next stage—Neural network. This sequential flow enables a fine-tuning process wherein the discriminative capabilities of the SVM are applied to enhance predictions further. For Dataset 2, with its unique input-output structure, the cascaded SVM-NN architecture is adept at handling the four input features and predicting the single output feature.

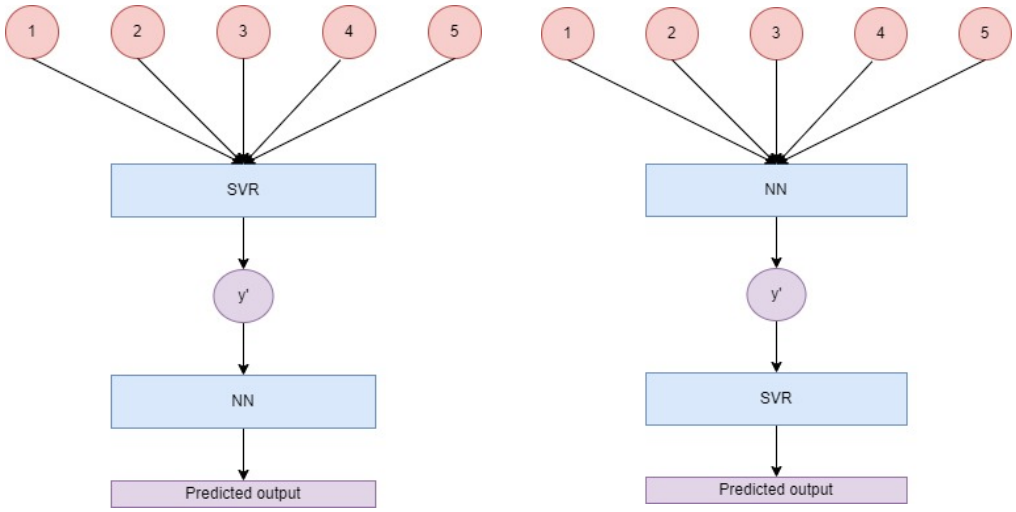

Fig.4 Representation of Cascaded Models

2.5 K-Nearest Neighbours

K-Nearest Neighbors (KNN) Regression is a supervised machine learning algorithm used for predicting continuous values, making it suitable for regression tasks.In KNN regression, the prediction for a new data point is based on the values of its 'k' nearest neighbors in the feature space.Neighbors are determined using a distance metric, commonly Euclidean distance, which measures the geometric distance between data points.For each new data point, the algorithm identifies its 'k' nearest neighbors from the training dataset.The predicted value for the new data point is often the average (or weighted average) of the target values of these 'k' neighbors.The choice of 'k' is a crucial parameter that impacts the model's performance. A small 'k' can lead to high sensitivity to noise, while a large 'k' may smooth out patterns in the data. In this study we have used k = 5.

CHAPTER 3 - Results

3.1 Output plots of the Neural network model

The figures below represent the output of the Neural network model that is trained with Dataset - 1.

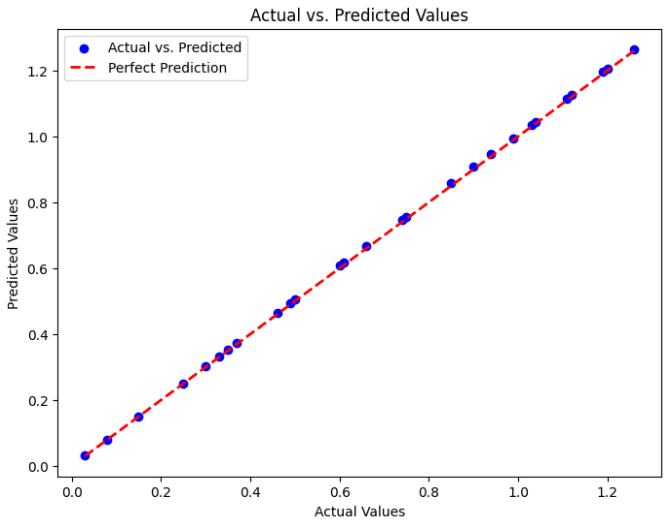

Fig. 5(a) Predicted Vs. Actual Current Density values for Dataset-1

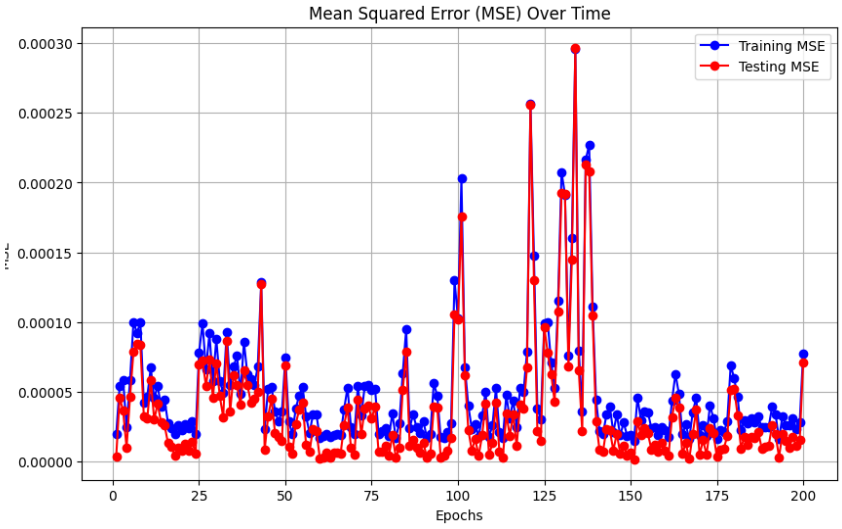

Fig.5(b) X-axis→Epochs;Y-axis→MSE for Dataset-1

3.2 Output plot of the SVR model

The figures below represent the output of the Support Vector Regressor model that is trained with Dataset - 1.

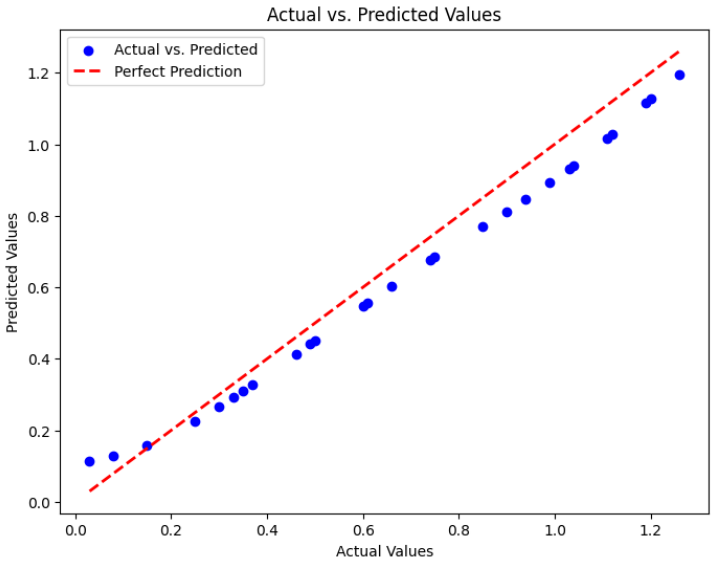

Fig.6 Predicted Vs. Actual Current Density values for Dataset-1

3.3 Output plots of the Cascaded models

The actual and predicted values of the cascaded methods and K-nearest neighbours is plotted as shown in figures 7-9.

3.3.1 Cascaded NN into SVR

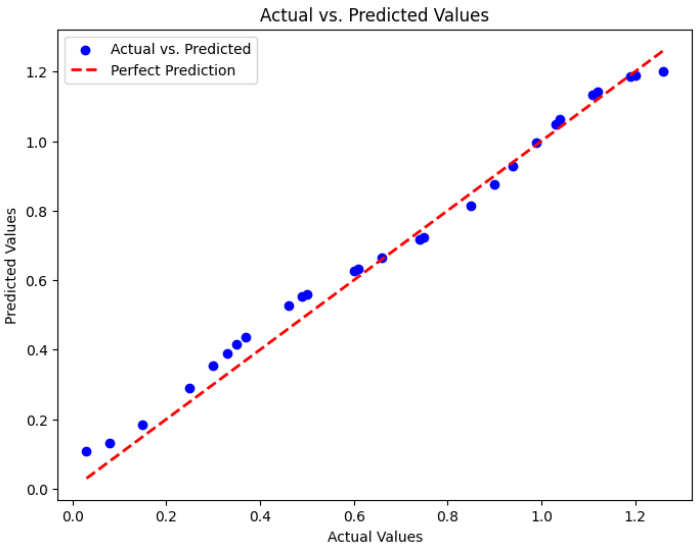

Fig.7(a) Predicted Vs. Actual Current Density values for Dataset-1

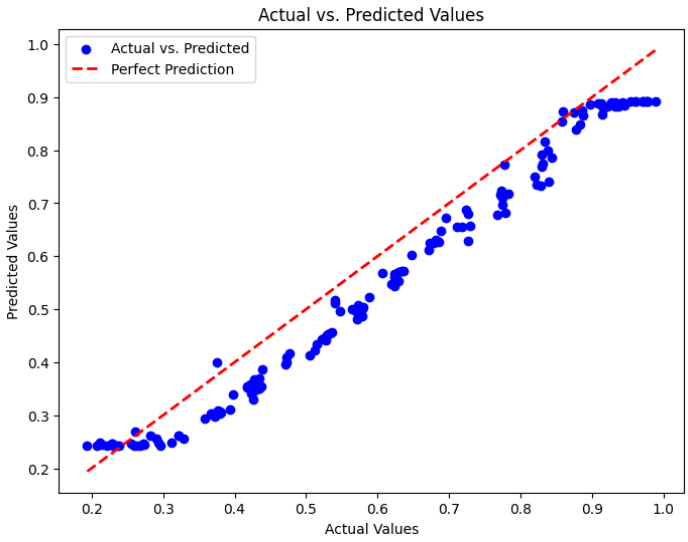

Fig.7(b) Predicted Vs. Actual Cell Voltage values for Dataset-2

3.3.2 Cascaded SVR into NN

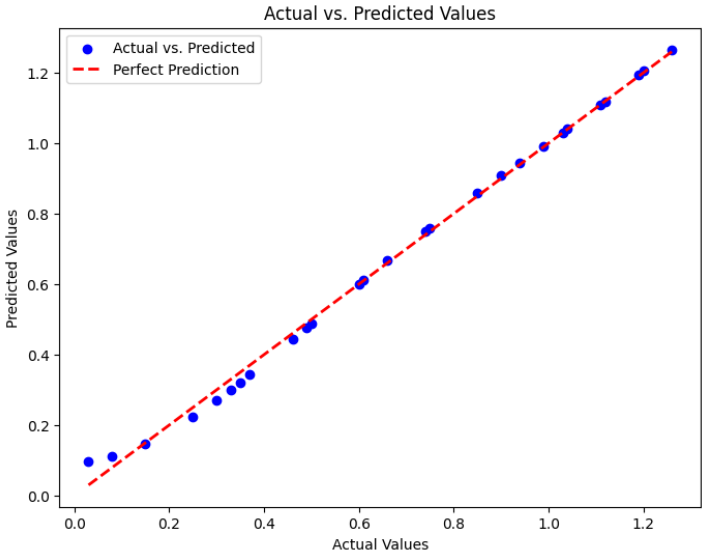

Fig.8(a) Predicted Vs. Actual Current Density values for Dataset-1

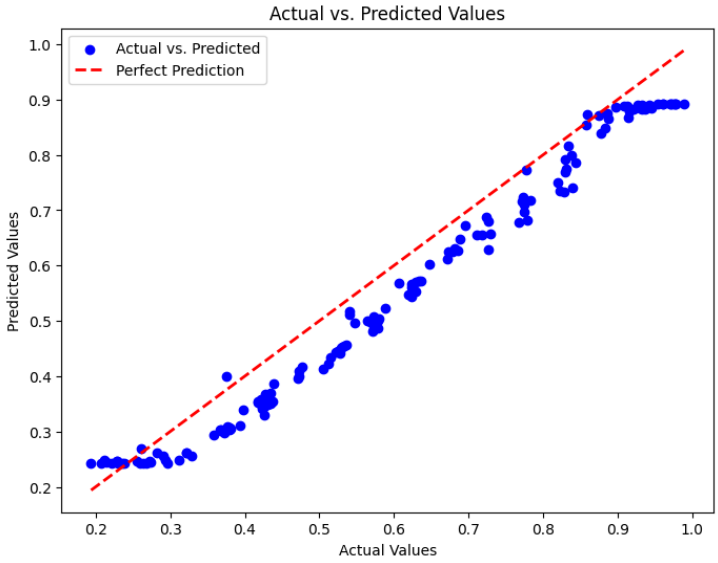

Fig.8(b) Predicted Vs. Actual Cell Voltage values for Dataset-2

3.4 Output plot of the K - Nearest neighbours model

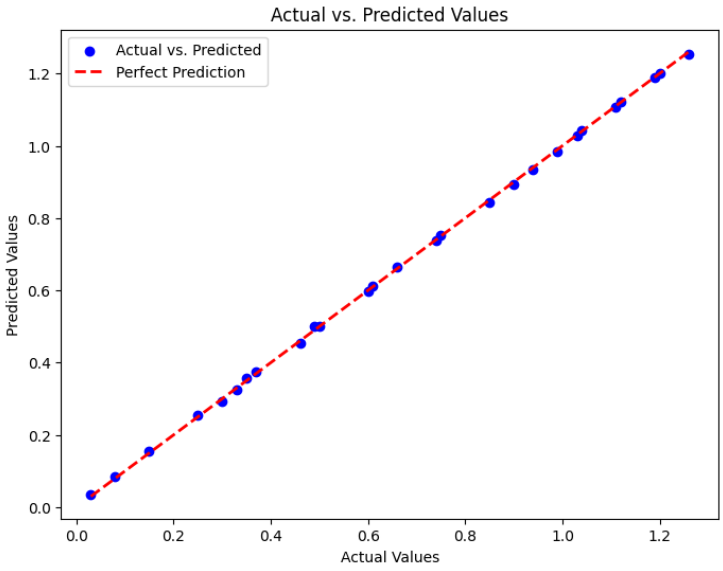

Fig.9(a) Predicted Vs. Actual Current Density values for Dataset-1

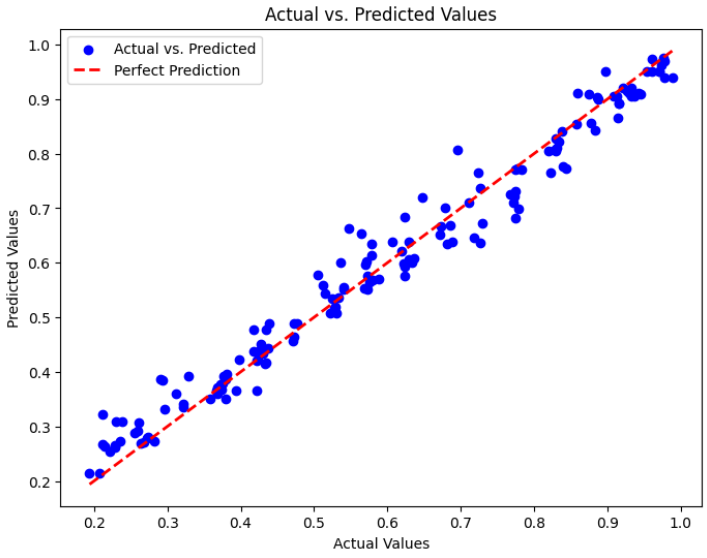

Fig.9(b) Predicted Vs. Actual Cell Voltage values for Dataset-2

3.5 Comparison of Different Models

Here, we have used the following three metrics to evaluate the performance of our model(s).

Mean Squared Error (MSE) :
MSE is a measure of the average squared difference between the predicted values and the actual values.

Mean Absolute Error (MAE) :
MAE measures the average absolute differences between the predicted values and the actual values.

R-Squared :
R-squared represents the proportion of the variance in the dependent variable that's predictable from the independent variables.

Table 6 Comparison of different models

		MSE		MAE		R-Squared	
Models	Dataset	Training dataset	Testing dataset	Training dataset	Testing dataset	Training dataset	Testing dataset
Neural Network	1	0.000043	0.000026	0.004605	0.004955	0.999705	0.999805
Support Vector Regressor	1	0.004413	0.004633	0.060836	0.063486	0.969632	0.965213
Cascaded NN into SVR	1	0.002050	0.001826	0.038825	0.036575	0.985892	0.986286
	2	0.003634	0.003814	0.054659	0.056618	0.928640	0.931759
Cascaded SVR into NN	1	0.000466	0.000392	0.014474	0.013344	0.996796	0.997058
	2	0.002120	0.001964	0.036733	0.035534	0.958371	0.964861
K-Nearest Neighbours	1	0.000053	0.000027	0.005058	0.004593	0.999633	0.999795
	2	0.001064	0.001526	0.023290	0.029489	0.979108	0.972700

CHAPTER 4 - Conclusion

In this study, artificial neural network (ANN), Support Vector Machine Regressor (SVR) and K-Nearest neighbors (KNN) have been used as the machine learning methods to develop data-based models for the performance attributes and internal states of proton exchange membrane fuel cell (PEMFC). PEMFC operating conditions such as cell current, Activation drop loss, ohmic drop loss, concentration drop loss, efficiency and power are used as input feature parameters, while the output parameters include the predicted cell voltage and Current Density for various operating conditions.The accuracy of the data-based models developed are evaluated under various conditions. 80% of the data points are used for training, while the remaining 20% of data points are used for testing. Cascaded Regression data-based models are developed using NN and SVM.It is demonstrated that KNN clearly shows an advantage in comparison with Cascaded models. In this comprehensive evaluation of regression models applied to the two datasets, we employed three key performance metrics—Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-Squared. The models under consideration include Neural Network (NN), Support Vector Regressor (SVR), Cascaded NN into SVR, Cascaded SVR into NN, and K-Nearest Neighbors (KNN).The Neural Network exhibited exceptional performance on both training and testing datasets, as indicated by the minimal MSE and MAE values and near-perfect R-Squared scores. This underscores the NN's ability to capture intricate patterns in the data and generalize well to unseen instances.SVR showcased strong predictive capabilities, albeit with slightly higher errors compared to the NN. The R-Squared values, indicating the proportion of variance explained, remained notably high, signifying the SVR's effectiveness in capturing the underlying relationships in the data.The cascaded models demonstrated consistent performance, leveraging the strengths of both Neural Networks and Support Vector Regressors. The results suggest that this sequential fine-tuning approach contributes to enhanced predictive accuracy, especially evident in Dataset-1.KNN demonstrated competitive performance, particularly on Dataset-1, with low MSE and MAE values and impressive R-Squared scores. However, it exhibited slightly higher errors on Dataset-2, indicating a potential sensitivity to variations in input feature characteristics.The cascaded models, incorporating both NN and SVR, proved effective in achieving a balance between abstraction and discrimination, resulting in robust predictive capabilities.

REFERENCES

1. https://www.mpoweruk.com/hydrogen_fuel.htm

2. https://www.energy.gov/eere/fuelcells/types-fuel-cells

3. Arun Saco, P. Shanmuga Sundari, Karthikeyan J and Anand Paul. "An Optimized Data Analysis on a Real-Time Application of PEM Fuel Cell Design by Using Machine Learning Algorithms." *Algorithms*, vol. 15, no. 10, 2022, https://www.mdpi.com/1999-4893/15/10/346.
4. Danqi Su, Jiayang Zheng, Junjie Ma, Zizhe Dong, Zhangjie Chen and Yanzhou Qin. "Application of Machine Learning in Fuel Cell Research." *Energies*, vol. 16, no. 11, 2023, https://www.mdpi.com/1996-1073/16/11/4390.
5. Mohamed Derbeli,Cristian Napole andOscar Barambones. "Machine Learning Approach for Modeling and Control of a Commercial Heliocentris FC50 PEM Fuel Cell System." *Mathematics*, vol. 9, 2021.
6. Mojgan Fayyazi,Paramjotsingh Sardar,Sumit Infent Thomas,Roonak Daghigh,Ali Jamali,Thomas Esch,Hans Kemper,Reza Langari andHamid Khayyam. "Artificial Intelligence/Machine Learning in Energy Management Systems, Control, and Optimization of Hydrogen Fuel Cell Vehicles." *Sustainability*, vol. 15, no. 6, 2023, https://www.mdpi.com/2071-1050/15/6/5249.
7. Adithya Legala, Jian Zhao, Xianguo Li. "Machine learning modeling for proton exchange membrane fuel cell performance." Energy and AI, vol. 10, 2022, https://doi.org/10.1016/j.egyai.2022.100183.