Energy Proceedings
Vol 45, 2024
ISSN 2004-2965

Security Concerns for Using Deep Learning Models in Predicting Hydrogen Production: A Comparative Study on Adversarial Attack

Chiagoziem C. Ukwuoma [1,2], Dongsheng Cai [1,2*], Chibueze D. Ukwuoma [3], Gladys W. Muoka [4], Chidera O. Ukwuoma [5], Olusola Bamisile [1,2], Qi Huang [1,2]

1 College of Nuclear Technology and Automation Engineering, Chengdu University of Technology, Sichuan P.R., 610059, China

2 Sichuan Engineering Technology Research Center for Industrial Internet Intelligent Monitoring and Application, Chengdu University of Technology, Sichuan P.R., 610059, China

3 Department of Physics, School of Engineering and Engineering Technology, Federal University of Technology Owerri, Nigeria

4 School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

5 Department of Prosthetics and Orthotics Technology, Federal University of Technology Owerri, Nigeria
(*Corresponding Author: caidongsheng@cdut.edu.cn)

ABSTRACT

In order to handle the rising global energy demand and lower carbon emissions, hydrogen, a clean and sustainable energy source, is essential. The creation of hydrogen is significant because it has the potential to transform the energy industry by providing a sustainable alternative to conventional fossil fuels. Deep learning has been a potent tool in recent years, exhibiting outstanding performance and dependability in a variety of domains, including the prediction of hydrogen generation. The optimization of hydrogen production methods to increase their effectiveness and reduce costs has shown promise. The susceptibility of deep learning models to adversarial attacks, which can reduce the precision and dependability of their predictions, is a growing worry. Adversarial attacks entail the purposeful alteration of input data to trick machine learning algorithms and provide false results. Such attacks may have far-reaching effects on hydrogen production prediction, thereby compromising the efficiency, economic feasibility, and safety of processes. To address these concerns, we conducted an extensive investigation into the susceptibility of deep learning models used for hydrogen production prediction to adversarial attacks using the co-gasification of biomass and plastics datasets. In the co-gasification of biomass and plastics dataset, the dependent variable was the quantity of hydrogen generated, and the independent variables included the gasification temperature, high-density polyethylene (HDPE) and rubber seed shell (RSS) particle size, and the quantity of plastic in the final product. The implemented adversarial attacks include the limited-memory broyden-fletcher-goldfarb-shanno (L-BFGS), fast gradient sign method (FGSM), basic iterative method, and projected gradient descent method (PGD). This study employed 4 machine learning regression models and a novel deep learning model based on Keras API to analyze the effect of the adversarial attack models under several perturbations including 0.1, 0.2, 0.4, 0.6 and 0.8. From the yielded result, it was evident that the FGSM and PGD adversarial attack has a significant influence on the employed model prediction results while the L-BFGS and the basic iterative method yielded results that will be addressed in our future works. Our research highlights the potential risks of relying on these models for decision-making in hydrogen production processes while also revealing the vulnerabilities of deep learning models in this crucial domain. We also highlight the significance of developing defense mechanisms and security protocols to protect the integrity of deep learning-based predictions in this crucial sector.

Keywords: machine learning, deep learning, predicting hydrogen production, co-gasification, adversarial attack

This is a paper for 15th International Conference on Applied Energy (ICAE2023), Dec. 3-7, 2023, Doha, Qatar.

NONMENCLATURE

Abbreviations	
HDPE	High-Density Polyethylene
RSS	Rubber Seed Shell
L-BFGS	Limited-Memory Broyden-Fletcher-Goldfarb-Shanno
FGSM	Fast Gradient Sign Method
PGD	Projected Gradient Descent Method
RFR	Random Forest Regressor
XGBoost	Extreme Gradient Boosting
SVR	Support Vector Regressor
KNN	K-Nearest Neighbor
RS	Renewable Sources
CO_2	Carbon Dioxide
CH_4	Methane
H_2	Hydrogen Gas
$Ni/CaFe_2O_4$	Nickel/Calcium Ferrite
RBF	Radial Basis Function
MLP	Multi-Layer Perceptron
ANN	Artificial Neural Network
SEE	Standard Error of Estimates
CNN	Convolutional Neural Networks
GA	Genetic Algorithms
DT	Decision Trees
MAE	Mean Absolute Error
MSE	Mean Square Error
RMSE	Root Mean Squared Error
RMSLE	Root Mean Squared Log Error
R^2	R-squared
Symbols	
$	Dollar
£	Pound
mm	Millimeter
C	Centigrade
vol %	Percentage in Volume
wt %	Weight in Volume
kg	Kilogram

1. INTRODUCTION

In recent times, there has been a growing emphasis on the quest for sustainable, cost-efficient, and long-lasting energy sources, driven by the escalating global demand for energy[1–3]. With the projected global population reaching 10 billion by 2050, energy consumption is expected to witness a significant upsurge, underscoring the imperative for sustainable solutions. Although fossil fuels have historically fueled global economic growth, their adverse environmental impact is undeniable. Consequently, the scientific community is actively exploring alternative methods of energy production that have minimal or no detrimental effects on the environment[4–7]. As previously mentioned, hydrogen is generated from substances containing hydrogen, such as carbohydrates or water. It's important to note that a substantial 96% of the world's hydrogen production traditionally relies on fossil fuels. Specifically, 30% is derived from naphtha reforming, 48% from natural gas steam reforming, and 18% from coal gasification[8]. However, these conventional methods of hydrogen production are closely tied to the environmental challenges currently facing our planet. Hence, environmentalists and the energy sector are vigorously working to develop more environmentally friendly approaches to producing hydrogen, particularly using RS.

Simultaneously, in the quest to convert carbon into sustainable energy sources like hydrogen and syngas, the utilization of plastics and biowastes holds the potential to reduce the environmental impact of industrial processes found in sectors such as iron, steel, and cement[9][10]. The co-gasification of mixtures containing plastic and biomass, achieved through dry and steam reforming of CO_2, results in the production of H_2, with factors such as feed composition and catalyst selection influencing the conversion of waste plastics into valuable fuel products[11–15]. Various variables, including temperature, the ratio of polymers to biomass, CO_2/CH_4 ratios, and the choice of catalyst, all contribute to the H_2 production process[9][16]–[18]. Waste polymers like polyethylene and polypropylene exhibit low moisture and ash contents but possess high volatile content, viscosity, and heating value. Among these materials, polypropylene emerges as the most efficient for hydrogen production. However, when compared to biomass, which contains substantial quantities of hydrogen-rich molecules such as cellulose, hemicellulose, and lignin, polymers require more energy for gasification and yield less hydrogen[11][19].

While the fossil fuel and renewable energy sectors have traditionally been the main players in the production of green hydrogen, a third contender has now entered the arena. Green hydrogen derived from organic waste has emerged as a significantly more cost-effective alternative to both fossil fuels and renewable energy sources, offering a carbon-negative solution. This form of green hydrogen, produced from diverse combinations of organic waste, has the potential to power mobile homes and remote hospitals that lack access to conventional electricity sources. In contrast,

the electrochemical method used to produce green hydrogen, which relies on substantial amounts of freshwater and renewable energy, is environmentally friendly and efficient in separating water into hydrogen and oxygen. The cost of producing green hydrogen from waste blends is estimated to be approximately $3 per kilogram, whereas utilizing solar or wind energy can cost roughly $11 to $16 per kilogram. Moreover, each tonne of dry waste can yield between 40 and 50 kg of green H_2, although this yield may vary between 30 kg and 120 kg depending on the moisture content in the waste blends.

Artificial intelligence techniques encompassing machine learning and deep learning algorithms have found valuable applications in tasks such as clustering, optimization, prediction, and classification within the domain of green hydrogen generation. These AI methodologies are instrumental in analyzing diverse data streams[11]. For instance, in Scotland, a real-time machine learning system is actively enhancing the production of green hydrogen by harnessing wind and tidal power[19]. To ensure the robustness and reliability of this system and facilitate data-driven decision-making, a cloud-based hydrogen management platform has been meticulously developed, integrating machine learning and optimization algorithms. This platform plays a pivotal role in identifying the most economically advantageous periods for hydrogen production and storage and has garnered £494,000 in support from the Department of Business, Energy, and Industrial Strategy. Consequently, the confidence and trust vested in decision-makers who employ machine learning models within specific domains are of utmost importance[20]. The enhancement of decision-making hinges on the ability to detect flaws and concealed biases within the operations of these models[21]. The utilization of artificial intelligence, particularly deep learning and machine learning models, in predicting hydrogen production brings to light the invaluable potential of these technologies. However, it is essential to acknowledge that as we delve deeper into AI's applications, particularly in critical domains like energy production, security concerns emerge as significant considerations.

Machine/Deep learning models are subject to adversarial attacks (see Fig. 1), according to recent research findings, which can introduce errors into these models both during the training and testing stages[22]. A technique for producing adverse instances is an adversarial attack. An example that is intended to induce a machine learning model to predict incorrectly even if it would appear to be legitimate to a person is called an adversarial example. Dalvi et al.[23] performed the earliest inquiry into this phenomenon in the context of spam filtering. They discovered that minor changes to spam emails' text might readily fool a linear classifier without materially altering the spam message's readability. Adversarial examples that target linear classifiers were introduced in this work. Following the groundbreaking work of Krishevsky et al.[24], who showed the amazing effectiveness of CNNs in a large-scale visual identification test, the interest in using deep learning models significantly increased. Szegedy et al.[25] described how deep neural networks, particularly in the computer vision field, are vulnerable to adversarial instances.

Fig. 1 Example of adversarial attack on numerical data

Motivation: One pressing issue revolves around the susceptibility of deep learning models to adversarial attacks, which can have a detrimental impact on their performance and consequently affect the quality of the decisions made based on their predictions. These adversarial attacks can manipulate the input data in subtle ways to mislead the AI model, potentially leading to incorrect predictions or compromised outcomes in the context of hydrogen production. In light of these security concerns, it becomes imperative to not only harness the power of AI for enhancing the efficiency and sustainability of hydrogen production but also to fortify these systems against potential threats. This dual focus on leveraging AI's capabilities while safeguarding against vulnerabilities will be crucial in ensuring the reliability and resilience of AI-driven decisions in this critical sector.

The aim of this study encompasses its examination of security concerns in hydrogen production prediction, its comparative study on adversarial attacks, its real-world application in the energy sector, and its potential contributions to mitigation strategies, all of which enhance the understanding and practical implementation of AI in this critical domain. The following highlights the major contribution of this manuscript;

- Exploring Security Implications in Hydrogen Production Prediction
- Conducts a comparative study using several machine learning models and a novel deep learning model for predicting hydrogen prediction.

- ❖ Conducts a comparative study on adversarial attacks specifically tailored to Machine learning models used in hydrogen production prediction.
- ❖ Finally, this study offers insights into potential mitigation strategies for securing deep learning models in hydrogen production prediction.

The structure of this paper comprises several sections. Section 2 delves into the studies related to the topic. The third section provides a detailed account of the comprehensive workflow and introduces the proposed model. Section 4 presents the outcomes of our experiments, engaging in discussions and identifying the limitations and future works. Finally, in Section 5, we conclude our work.

2. RELATED WORKS

The scientific investigation into predicting hydrogen production has yielded only a scant number of responses. The research community has conducted comprehensive inquiries into issues related to hydrogen production[26–29]. Furthermore, various methods and solutions for producing eco-friendly hydrogen via biological, chemical, or physical processes have been proposed. For instance, Nicolas et al.[30] probed the potential of generating eco-friendly hydrogen from bioethanol using nanocatalyst design. In the quest for eco-friendly hydrogen from seawater, Rafaeld'Amore-Domenech et al.[31] explored and compared four electrolysis techniques. Scarce literature addresses the prediction of hydrogen production. In Islamabad, Syed et al.[32] delved into a machine-learning algorithm to forecast hydrogen production from solar energy. Artificial intelligence, a tool of green technology, has the potential to facilitate the creation of eco-friendly hydrogen using diverse methods and resources. One widely used approach is methane drying and reforming, which employs machine and deep learning models to predict eco-friendly hydrogen generation based on various catalysts. Victor et al.[33] evaluated the Bayesian regularization algorithm, the Leven-Marquardt algorithm, and a scaled conjugate gradient algorithm as training algorithms for an ANN prediction model to estimate the quantities of CO and H_2 produced by the methane drying and reforming process. The empirical findings favored the Bayesian regularization technique, which exhibited the lowest SEE compared to the other methods. Hossain et al.[34] scrutinized the effectiveness of two ANN models in forecasting hydrogen-rich syngas generation from methane drying and reforming using advanced Ni/$CaFe_2O_4$ catalysts. The experiments' results were trained and validated using RBF and MLP neural network models, with the ANN-MLP-based approach outperforming the ANN-RBF-based approach in predicting hydrogen-rich syngas production.

To predict the overall hydrogen output from thermo-catalytic methane reforming, May et al.[35] assessed the performance of two deep learning models: one employing Bayesian regularization and another trained with the Levenberg-Marquardt method for a multilayer perceptron neural network. The experimental findings demonstrated that the Levenberg-Marquardt-trained neural network, configured as 7-16-1, outperformed the Bayesian regularization-trained network in forecasting green hydrogen production rates. Additionally, various ANN models were utilized, evaluated, and compared to forecast green hydrogen production[36]. Alternatively, a different approach involves producing green hydrogen through oxygen injection and hydrocarbon tanks submerged in water. In their work, Klemens et al.[37] introduced a data-centric AI system aimed at enhancing green hydrogen production within hydrocarbon reservoirs submerged in water. Their study represents a pioneering effort to improve oxygen injection techniques while optimizing hydrogen generation using an AI-based genetic optimization framework. Generating hydrogen from organic waste is considered one of the most prominent and cost-effective methods[37–40]. Nevertheless, the existing body of literature lacks an adequate number of AI models designed to strategize and enhance green hydrogen production from waste sources. Recent investigations[41][42] have concentrated on leveraging machine learning algorithms to maximize hydrogen generation from wastewater and sewage sludge. Hao-nan et al.[43] examined the application of five machine learning methods, including ANN, SVM, GA, DT, and RF to predict organic solid waste treatment outcomes. Their analysis was based on reviewing published papers from 2003 to 2020. It's worth noting that this study did not specifically address the application of these machine-learning methods for generating hydrogen from organic solid waste. This research holds significance due to the identified knowledge gap in this particular domain.

3. MATERIAL AND METHODS

3.1 Adversarial Attack

The following Adversarial attacks were implemented in this paper including the L-BFGS, FGSM, Basic Iterative Method, and PGD [44]. The PGD method is one of the most effective adversarial Attack techniques

which produces adversarial samples quickly and simply and is mathematically expressed as;

$$x^{adv} = x + \epsilon sign\left(\nabla_x L(\theta, x, y_{true})\right) \quad (1)$$

x^{adv} is the perturbed adversarial sample, $\left(L(\cdot)\right)$ is the classification loss function, $\nabla_x L$ is the gradient concerning the unperturbed sample (x), is the DL model weights, and (y_{true}) is the true label, where is the magnitude of the perturbation that limits the amount of perturbation allowed in each pixel of an image. The L-BFGS creates an adversarial example using the least probable predicted class of a rained network for a certain data sample;

$$yLL = arg_y min\{p(y^{true}|x)\} \quad (2)$$

Where $arg_y min\{p(y^{true}|x)\}$ is the minimal probability that the provided data sample (x) has the true label y^{true}. PGD is an extension of FGSM and one of the most powerful first-order attack strategies. It repeatedly tries to create an ideal perturbation from a randomly chosen point inside an L-∞ ball, which establishes a region with a radius typically equal to epsilon around the original data point. Equation (3) depicts how the PGD iterates:

$$x^{t+1} = \prod_{x+s}\left(x^t + \alpha sign\left(\nabla_x l\left(\theta, x, y_{target}\right)\right)\right) \quad (3)$$

Where $Q(\cdot)$ is the projection function to project adversarial instances back onto the L-∞ ball after each iteration, x^t is the adversarial example at the $t-th$ iteration, α is the step size, and θ is the DL model weights.

3.2 Machine Learning Models

For the experiments, we selected four machine learning regression models including the SVR, XGBoost, RF and KNN. This study went further to develop a novel ANN model based on Keras API to support our study[45][46].

- ❖ RF: In the context of supervised learning, a RF is a potent machine learning method used for regression problems. It is an ensemble learning technique that integrates many decision tree regressors to provide precise regression predictions while reducing overfitting as shown in Fig. 2.
- ❖ XGBoost: is an efficient supervised machine learning approach for regression problems. Due to its remarkable performance and adaptability in handling complicated datasets, it is a preferred choice for numerous data-driven applications, from finance to healthcare. Decision trees are used to enhance prediction accuracy as illustrated in Fig. 3.

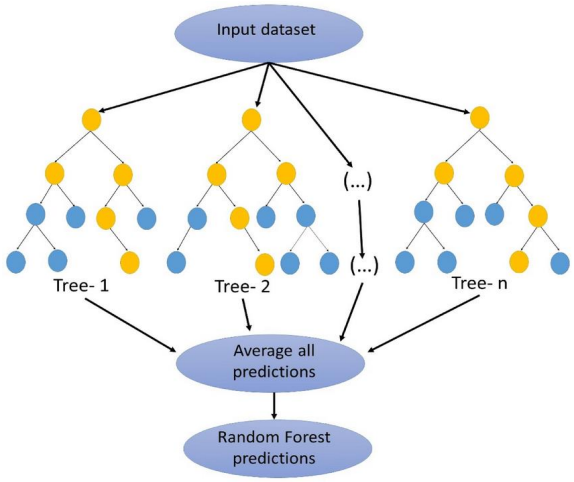

Fig. 2 Basic structure of the RL regressor

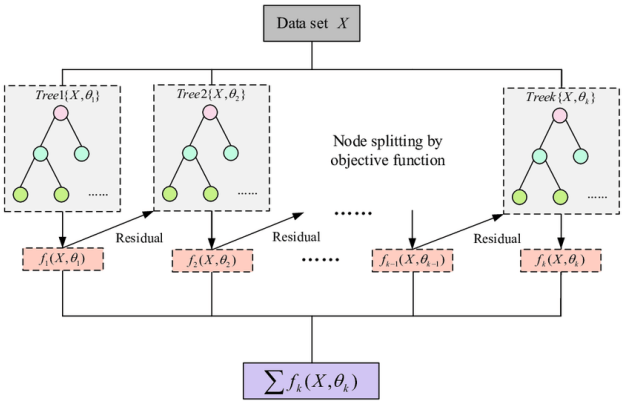

Fig. 3 Basic structure of the XGBoost regressor

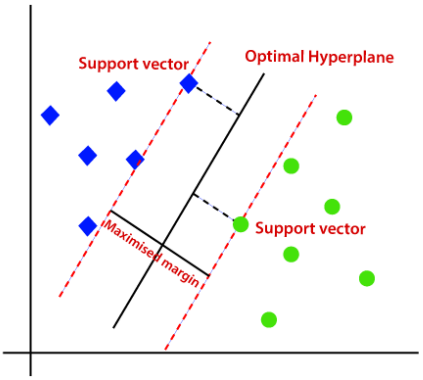

Fig. 4 Basic structure of the SVR

- ❖ SVR: SVM is a powerful supervised machine learning algorithm used for classification and regression tasks. It works by finding the optimal hyperplane that best separates data points belonging to different classes in a high-dimensional space. The key idea is to identify support vectors, which are the data points closest to the decision boundary, and use them to maximize the margin between classes as seen in Fig. 4.

Fig. 5 Basic structure of the KNN Regressor

- K-NN: To forecast or categorize specific data points, the k-NN model, an irregular supervised learning classifier, depends on closeness. It is adaptable and may be used for problems involving classification and regression. In classification, it chooses the class label that is most commonly present among a particular data point's closest neighbors based on a majority vote. To predict a classification in regression, the average of the KNN is calculated as seen in Fig 5.

3.3 Proposed Model

A novel deep learning model based on the Keras sequential model is proposed to further assess the effect of the adversarial attack in hydrogen production prediction as shown in Figures 6 and 7. This model is composed of interconnected nodes, or neurons, organized into layers. Information flows through these layers, starting with an input layer, passing through hidden layers, and concluding with an output layer. Each connection between neurons has a weight that adjusts during training, allowing the network to learn patterns and make predictions.

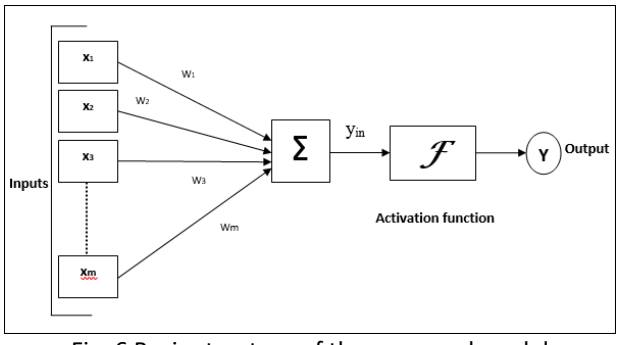

Fig. 6 Basic structure of the proposed model

Our architecture's input layer is made up of four features: gasification temperature, RSS particle size, HDPE, and the volume of plastic in the mixture. The input layer of the model has an architecture of [16, 3, 1] with 30 rows of data and 4 independent variables. The output layer does not get an activation function as it's a regression problem aimed at predicting numerical values directly. By promoting weight decay toward zero, the L2 regularization is used to prevent overfitting. The Adam optimizer, MSE loss, MAE, and other metrics are used to build the model. A batch size of 2, 3000 epochs, and a verbose output setting of 1 are used during training.

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense_1 (Dense)              (None, 16)                80
_________________________________________________________________
dense_2 (Dense)              (None, 3)                 51
_________________________________________________________________
dense_3 (Dense)              (None, 1)                 4
=================================================================
Total params: 135
Trainable params: 135
Non-trainable params: 0
_________________________________________________________________
```

Fig. 7 Proposed model summary

3.4 Dataset

The study's dataset was based on the research from Chin et al.[47]. In statistics, a sample size of 30 is typical. A population data set's confidence interval can be increased by a factor of 30 to support claims that the result is false[48]. The dataset consists of 30 experimental runs, with gasification temperature, RSS, and HDPE particle size, the volume of plastic in the mixture acting as independent variables, and the volume of hydrogen produced acting as the dependent variable (Table 1). A larger sample size, however, has a better likelihood of being representative of the population at hand. According to statisticians, a sample size of 30 is enough for the majority of distributions.

Table 1. Description of the hydrogen production data

Temperature (C)	RSS Particle Size (mm)	HDPE Particle Size (mm)	Percentage of Plastics in Mixture (wt%)	H2 (vol %)
800	0.25	0.25	10	46.676
700	0.125	0.375	20	50.123
600	0.5	0.25	30	47.751
800	0.5	0.25	10	45.952
500	0.375	0.375	20	44.781
700	0.375	0.625	20	43.031
600	0.5	0.25	10	45.324
900	0.375	0.375	20	49.23
800	0.5	0.5	30	44.355
600	0.5	0.5	30	44.208
700	0.375	0.375	0	44.466
700	0.375	0.375	40	46.603
700	0.625	0.375	20	43.072
800	0.25	0.5	30	47.396
700	0.375	0.375	20	39.98
800	0.25	0.25	10	46.338
700	0.375	0.375	20	38.569

700	0.375	0.375	20	49.868
800	0.25	0.25	30	46.545
700	0.375	0.375	20	38.612
600	0.5	0.5	10	41.032
700	0.375	0.375	20	38.625
600	0.25	0.5	30	47.123
700	0.375	0.375	20	38.621
600	0.25	0.25	10	48.634
800	0.5	0.25	30	48.475
600	0.25	0.5	10	48.132
700	0.375	0.375	20	39.262
600	0.25	0.25	30	46.502
800	0.5	0.5	10	41.93

3.5 Evaluation Metrics

This paper made use 5 evaluation metrics namely the MAE, MSE, RMSE, RMSLE, R^2. By dividing the total number of observations by the sum of all errors, the MAE determines the exact difference between the actual and anticipated values mathematically represented as

$$\mathrm{MAE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}\left(\mathrm{Y_i} - \widehat{\mathrm{Y}}_\mathrm{i}\right) \quad (4)$$

where n = number of samples, Y_i = observed values and $\hat{Y}_i$= predicted values. The squared variation in the actual and anticipated value is known as the mean squared error mathematically represented as;

$$\mathrm{MSE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}\left(\mathrm{Y_i} - \widehat{\mathrm{Y}}_\mathrm{i}\right)^2 \quad (5)$$

The RMSE corresponds to the square root of the average squared error, and its measurement unit aligns with that of the dependent variable.

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}\left(\mathrm{Y_i} - \widehat{\mathrm{Y}}_\mathrm{i}\right)^2} \quad (6)$$

$$\mathrm{RMSLE} = \sqrt{\frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}\left(\log(\mathrm{Y_i} + 1\,) - \log\left(\widehat{\mathrm{Y}}_\mathrm{i} + 1\right)\right)^2} \quad (7)$$

R^2 also known as the Coefficient of Determination or Fit Quality, measures how much better the performance of the regression line is than a simple mean line. It is dimensionless, analyzes model performance in every situation, and consistently produces numbers below one. it is mathematically represented below as;

$$\mathrm{R^2} = 1 - \frac{\text{sum squared regression (SSR)}}{\text{total sum of squares (SST)}} = 1 - \frac{\sum\left(\mathrm{Y_i} - \widehat{\mathrm{Y}}_\mathrm{i}\right)^2}{\sum(\mathrm{Y_i} - \overline{\mathrm{Y}})^2} \quad (8)$$

4. RESULTS AND ANALYSIS

For the machine learning models, we used the Grid search method to select their optimal hyperparameters for the training as depicted in Table 2. Furthermore, we used the boxplot to depict the inherent properties of the features of the dataset

Table 2. Training hyperparameter of the ml models

ML Model	Optimal Hyperparameter
RFR	Nos. of estimators = 30, random state = 100
XGB Regressor	Base score = 0.5, learning rate = 0.200, nos. of estimators = 50, max depth = 12, gamma = 0.7, alpha = 0.7, random state = 42
SVR	Kernel = 'rbf', random seed = 42
K-NN Regressor	Nos. neighbors = 4

Table 3. Descriptive statistics hydrogen production data

	Temperature (C)	RSS Particle Size (mm)	HDPE Particle Size (mm)	% of Plastics in Mixture (wt%)	H2 (vol %)
count	30.0000	30.0000	30.0000	30.0000	30.0000
mean	0.7778	0.6000	0.6000	0.5000	44.7072
std	0.1011	0.1819	0.1661	0.2274	3.6519
min	0.555556	0.2000	0.4000	0.0000	38.5690
25%	0.666667	0.4000	0.4000	0.2500	42.2053
50%	0.777778	0.6000	0.6000	0.5000	45.6380
75%	0.888889	0.8000	0.7500	0.7500	47.3278
max	1.000000	1.0000	1.0000	1.0000	50.1230

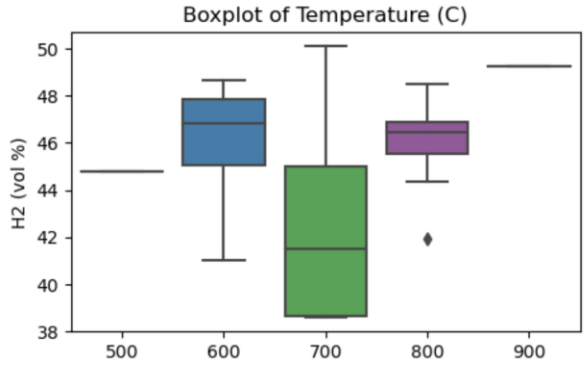

Fig. 8 Temperature vs. H_2

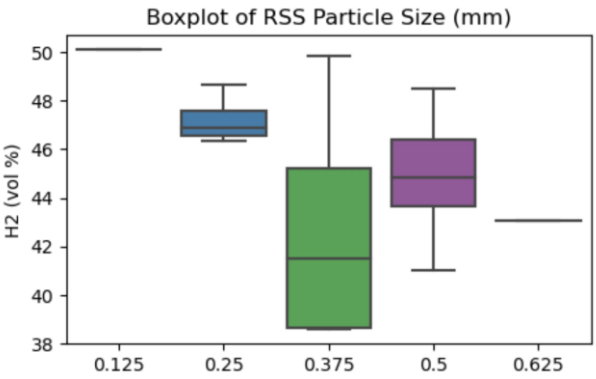

Fig. 9 RSS particle size vs. H_2

Table 3 shows the descriptive statistics of the employed dataset. To illustrate the data distribution, spot potential outliers, and gauge the range of values in each of the designated columns, we employed boxplots (Fig. 8 - Fig. 11) Boxplots offer a clear and succinct

representation of the most important statistical facts about a dataset, such as the median, quartiles, and any possible outliers. The boxplot displays the values in the chosen column's distribution on the x-axis and "H_2 (vol%)" on the y-axis, which may be important for making data-driven choices or seeing patterns and trends in the data.

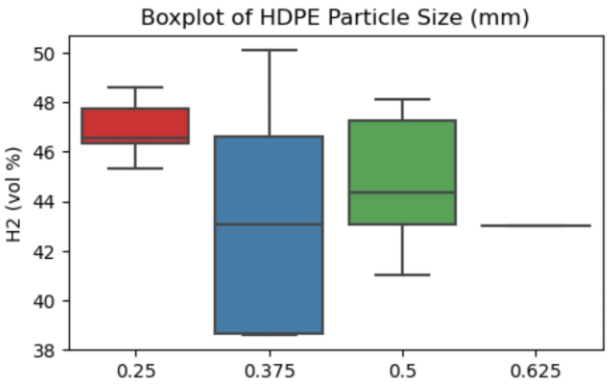

Fig. 10 HDPE particle size vs. H_2

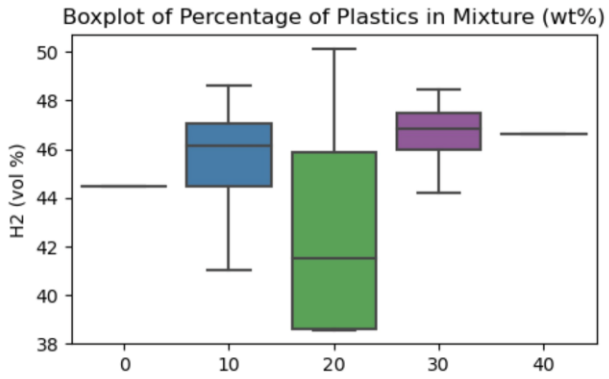

Fig. 11 % of plastics in mixture vs. H_2

4.1 Result Analysis

In this section, we present the results obtained from our analysis.

Table 4. Results

Model	MAE	MSE	RMSE	R2	MSLE	RMSLE
RFR	2.516	8.946	2.991	0.285	0.004	0.066
XGBoost	2.880	12.266	3.502	0.020	0.006	0.077
SVR	3.361	12.809	3.579	-0.024	0..006	0.080
KNN	3.266	11.564	3.401	0.076	0.006	0.075
Proposed Model	1.775	5.488	2.323	0.459	0.003	0.053

From Table 4, RFR has the lowest MAE, MSE, and RMSE among the models, indicating it performs the best in terms of accuracy and error metrics. However, its R^2 is relatively low, suggesting that it doesn't explain a significant portion of the variance in the data. XGBoost has higher errors (MAE, MSE, RMSE) and a very low R^2, indicating poorer performance compared to Random Forest. SVR performs slightly worse than XGBoost in terms of error metrics and has a negative R^2, suggesting it doesn't fit the data well. KNN falls in between RFR and XGBoost in terms of error metrics and R^2.

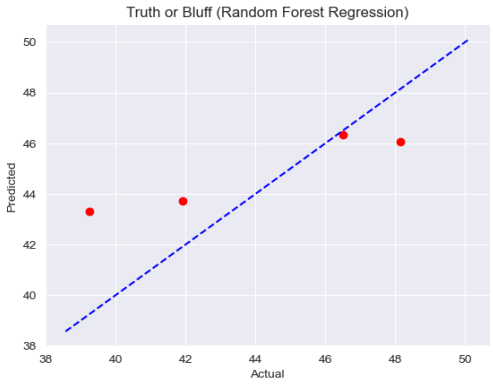

Fig. 12 RRF prediction

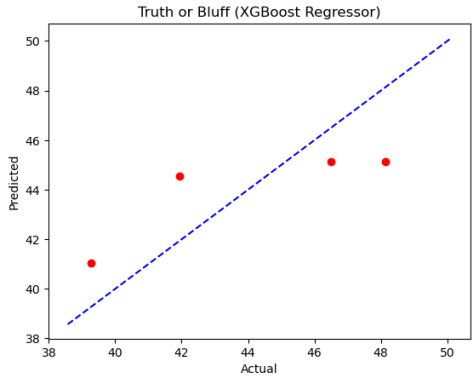

Fig. 13 XGBoost regessor prediction

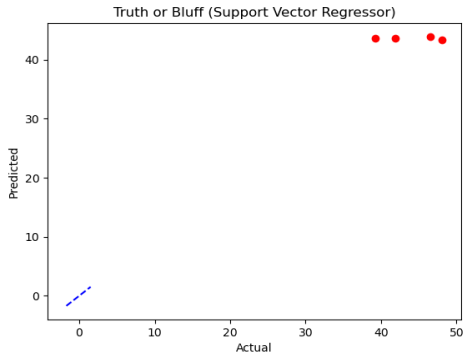

Fig. 14 SVR prediction

The Proposed Model outperforms all the other models in terms of MAE, MSE, RMSE, and R^2. It has the lowest error values and the highest R^2, indicating it provides the most accurate predictions and explains a significant portion of the data's variance. It also has the lowest MSLE and RMSLE, suggesting that it handles the data's wide range and potential skewness well.Fig. 12 – Fig. 15 illustrates the machine learning model prediction

vs. the actual result. The proposed model result is shown in Figure 16.

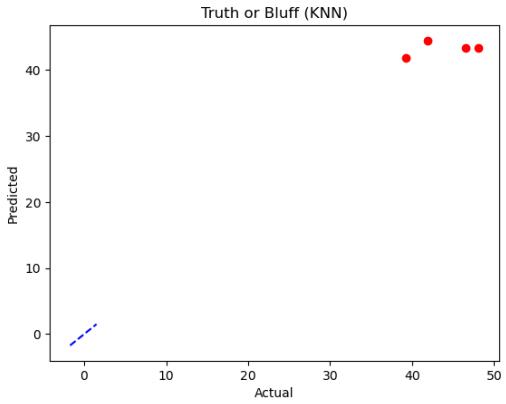

Fig. 15 K-NN regessor prediction

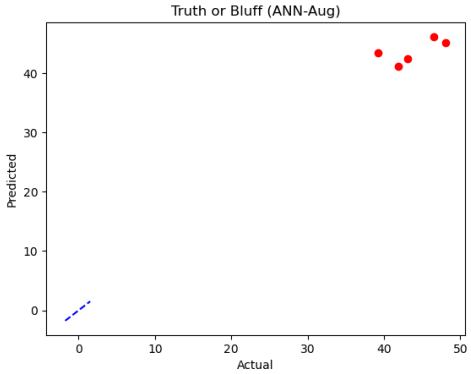

Fig. 16 Proposed model prediction

4.2 Adversarial Attack Analysis

In this section, explore the effects of the L-BFGS, FGSM, Basic Iterative Method, and PGD on the hydrogen production prediction. These adversarial attacks are all white box attacks because they rely on access to the model's internal information, such as gradients or model parameters. However, they can be adapted to black-box scenarios with some modifications, but their effectiveness might be reduced. We used an epsilon range of 0.1, 0.2, 0.4, 0.6, and 0.8 in our experiments. Table 5 to Table 9 demonstrate the achieved results. from the obtained results, it was evident that the FGSM and the PGD Adversarial Attack had a significant effect on the predicted results of the implemented models.

Table 5. Effects of the FGSM attack

Model	Perturbation	MAE	MSE	RMSE	R2	MSLE	RMSLE
RF	**-**	**2.516**	**8.946**	**2.991**	**0.285**	**0.004**	**0.066**
	0.1	2.530	9.653	2.991	0.229	0.005	0.069
	0.2	3.283	15.718	2.991	-0.256	0.008	0.089
	0.4	3.612	15.566	2.991	-0.244	0.008	0.089
	0.6	3.361	14.366	2.991	-0.148	0.007	0.086
	0.8	3.361	14.412	2.991	-0.152	0.007	0.086
XGB	**-**	**2.880**	**12.266**	**3.502**	**0.020**	**0.006**	**0.077**
	0.1	2.880	12.266	3.502	0.020	0.006	0.077
	0.2	3.479	15.562	3.502	-0.243	0.008	0.087
	0.4	3.730	16.455	3.502	-0.315	0.008	0.091
	0.6	3.356	13.536	3.502	-0.082	0.007	0.081
	0.8	3.361	13.557	3.502	-0.083	0.007	0.083
SVM	**-**	**3.361**	**12.809**	**3.579**	**-0.024**	**0..006**	**0.080**
	0.1	3.375	12.802	3.579	-0.023	0.006	0.080
	0.2	3.384	12.819	3.579	-0.024	0.006	0.080
	0.4	3.390	12.985	3.579	-0.038	0.007	0.081
	0.6	3.391	13.343	3.579	-0.066	0.007	0.083
	0.8	3.390	13.831	3.579	-0.105	0.007	0.084
KNN	**-**	**3.266**	**11.564**	**3.401**	**0.076**	**0.006**	**0.075**
	0.1	4.027	18.834	3.401	-0.505	0.009	0.096
	0.2	4.693	23.963	3.401	-0.915	0.012	0.110
	0.4	3.140	12.854	3.401	-0.027	0.007	0.081
	0.6	2.755	11.798	3.401	0.057	0.006	0.078
	0.8	3.361	13.698	3.401	-0.095	0.007	0.083
P. Model	**-**	**1.775**	**5.488**	**2.323**	**0.459**	**0.003**	**0.053**
	0.1	2.056	6.454	2.342	0.363	0.003	0.057
	0.2	2.267	7.704	2.343	0.240	0.004	0.061
	0.4	3.464	13.605	2.343	-0.342	0.007	0.082
	0.6	4.612	23.277	2.343	-1.296	0.011	0.107
	0.8	4.737	27.057	2.343	-1.669	0.013	0.114

Table 6. Effects of the PGD adversarial attack

Model	Perturbation	MAE	MSE	RMSE	R2	MSLE	RMSLE
RL	**-**	**2.516**	**8.946**	**2.991**	**0.285**	**0.004**	**0.066**
	0.1	2.530	9.653	2.991	0.229	0.005	0.069
	0.2	3.283	15.718	2.991	-0.256	0.008	0.089
	0.4	3.612	15.566	2.991	-0.244	0.008	0.089
	0.6	3.361	14.366	2.991	-0.148	0.007	0.086
	0.8	3.361	14.412	2.991	-0.152	0.007	0.086
XGB	**-**	**2.880**	**12.266**	**3.502**	**0.020**	**0.006**	**0.077**
	0.1	2.880	12.266	3.502	0.020	0.006	0.077
	0.2	3.479	15.562	3.502	-0.243	0.008	0.087
	0.4	3.730	16.455	3.502	-0.315	0.008	0.091
	0.6	3.356	13.536	3.502	-0.082	0.007	0.083
	0.8	3.361	13.557	3.502	-0.083	0.007	0.083
SVM	**-**	**3.361**	**12.809**	**3.579**	**-0.024**	**0..00 6**	**0.080**
	0.1	3.375	12.802	3.579	-0.023	0.006	0.080
	0.2	3.384	12.819	3.579	-0.024	0.006	0.080
	0.4	3.390	12.985	3.579	-0.038	0.007	0.081
	0.6	3.391	13.343	3.579	-0.66	0.007	0.083
	0.8	3.390	13.831	3.579	-0.105	0.007	0.084
KNN	-	3.266	11.564	3.401	0.076	0.006	0.075
	0.1	4.027	18.834	3.401	-0.505	0.009	0.096
	0.2	4.693	23.963	3,401	-0.915	0.012	0.110
	0.4	3.140	12.854	3.401	-0.027	0.007	0.081
	0.6	2.755	11.798	3.401	0.057	0.006	0.078
	0.8	3.361	13.698	3.401	-0.095	0.007	0.083
P. Model	-	1.775	5.488	2.323	0.459	0.003	0.053
	0.1	2.056	6.454	2.343	0.363	0..003	0.057
	0.2	2.267	7.704	2.343	0.240	0.004	0.061
	0.4	3.464	13.606	2.343	-0.342	0.007	0.082

	0.6	4.612	23.277	2.343	-1.296	0.011	0.107
	0.8	5.134	34.627	2.343	-2.416	0.016	0.128

Table 7. Effects of the L-BFGS adversarial attack

Model	Perturbation	MAE	MSE	RMSE	R2	MSLE	RMSLE
RL	-	**2.516**	**8.946**	**2.991**	**0.285**	**0.004**	**0.066**
	0.1	2.516	8.946	2.991	0.285	0.004	0.066
	0.2	2.516	8.946	2.991	0.285	0.004	0.066
	0.4	2.516	8.946	2.991	0.285	0.004	0.066
	0.6	2.516	8.946	2.991	0.285	0.004	0.066
	0.8	2.516	8.946	2.991	0.285	0.004	0.066
XGB	-	**2.880**	**12.266**	**3.502**	**0.020**	**0.006**	**0.077**
	0.1	2.880	12.266	3.502	0.020	0.006	0.077
	0.2	2.880	12.266	3.502	0.020	0.006	0.077
	0.4	2.880	12.266	3.502	0.020	0.006	0.077
	0.6	2.880	12.266	3.502	0.020	0.006	0.077
	0.8	2.880	12.266	3.502	0.020	0.006	0.077
SVM	-	**3.361**	**12.809**	**3.579**	**-0.024**	**0..006**	**0.080**
	0.1	3.361	12.809	3.579	-0.024	0..006	0.080
	0.2	3.361	12.809	3.579	-0.024	0..006	0.080
	0.4	3.361	12.809	3.579	-0.024	0..006	0.080
	0.6	3.361	12.809	3.579	-0.024	0..006	0.080
	0.8	3.361	12.809	3.579	-0.024	0..006	0.080
KNN	-	**3.266**	**11.564**	**3.401**	**0.076**	**0.006**	**0.075**
	0.1	3.266	11.564	3.401	0.076	0.006	0.075
	0.2	3.266	11.564	3.401	0.076	0.006	0.075
	0.4	3.266	11.564	3.401	0.076	0.006	0.075
	0.6	3.266	11.564	3.401	0.076	0.006	0.075
	0.8	3.266	11.564	3.401	0.076	0.006	0.075
P. Model	-	**1.775**	**5.488**	**2.323**	**0.459**	**0.003**	**0.053**
	0.1	1.775	5.488	2.323	0.459	0.003	0.053
	0.2	1.775	5.488	2.323	0.459	0.003	0.053
	0.4	1.775	5.488	2.323	0.459	0.003	0.053
	0.6	1.775	5.488	2.323	0.459	0.003	0.053
	0.8	1.775	5.488	2.323	0.459	0.003	0.053

Table 8. Effects of the basic iterative method adversarial attack

Model	Perturbation	MAE	MSE	RMSE	R2	MSLE	RMSLE
RL	-	**2.516**	**8.946**	**2.991**	**0.285**	**0.004**	**0.066**
	0.1	2.516	8.946	2.991	0.285	0.004	0.066
	0.2	2.516	8.946	2.991	0.285	0.004	0.066
	0.4	2.516	8.946	2.991	0.285	0.004	0.066
	0.6	2.516	8.946	2.991	0.285	0.004	0.066
	0.8	2.516	8.946	2.991	0.285	0.004	0.066
XGB	-	**2.880**	**12.266**	**3.502**	**0.020**	**0.006**	**0.077**
	0.1	2.880	12.266	3.502	0.020	0.006	0.077
	0.2	2.880	12.266	3.502	0.020	0.006	0.077
	0.4	2.880	12.266	3.502	0.020	0.006	0.077
	0.6	2.880	12.266	3.502	0.020	0.006	0.077
	0.8	2.880	12.266	3.502	0.020	0.006	0.077
SVM	-	**3.361**	**12.809**	**3.579**	**-0.024**	**0..006**	**0.080**
	0.1	3.361	12.809	3.579	-0.024	0..006	0.080
	0.2	3.361	12.809	3.579	-0.024	0..006	0.080
	0.4	3.361	12.809	3.579	-0.024	0..006	0.080
	0.6	3.361	12.809	3.579	-0.024	0..006	0.080
	0.8	3.361	12.809	3.579	-0.024	0..006	0.080
KNN	-	**3.266**	**11.564**	**3.401**	**0.076**	**0.006**	**0.075**
	0.1	3.266	11.564	3.401	0.076	0.006	0.075
	0.2	3.266	11.564	3.401	0.076	0.006	0.075
	0.4	3.266	11.564	3.401	0.076	0.006	0.075
	0.6	3.266	11.564	3.401	0.076	0.006	0.075
	0.8	3.266	11.564	3.401	0.076	0.006	0.075
P. Model	-	**1.775**	**5.488**	**2.323**	**0.459**	**0.003**	**0.053**
	0.1	1.775	5.488	2.323	0.459	0.003	0.053
	0.2	1.775	5.488	2.323	0.459	0.003	0.053
	0.4	1.775	5.488	2.323	0.459	0.003	0.053
	0.6	1.775	5.488	2.323	0.459	0.003	0.053
	0.8	1.775	5.488	2.323	0.459	0.003	0.053

4.3 Result Discussions

As seen in Table 5, the results indicate that all models tested are vulnerable to FGSM adversarial attacks, as evidenced by the degradation in performance metrics with increasing perturbation. The severity of the vulnerability varies among models, with some models (e.g., Proposed Model) being more sensitive than others. For the Random Forest Regression model, the MAE, MSE, and RMSE increase slightly with higher perturbation values, indicating that the model's performance degrades as the perturbation increases. R^2 values are negative, suggesting that the model's predictions are worse than simply using the mean of the target values. Similar to Random Forest, XGBoost shows a decrease in performance with increasing perturbation. The R^2 values are also negative, indicating poor predictive performance. The SVR model also exhibits a decline in performance as perturbation increases. R^2 values are negative. KNN performs relatively better, with higher R^2 values compared to the previous models. However, it still shows a decrease in performance as perturbation increases. The proposed model initially performs well with low perturbation but experiences a significant drop in R^2 values as perturbation increases. This suggests that the proposed model is sensitive to adversarial attacks.

Table 6 indicates that all tested models are vulnerable to PGD adversarial attacks. As the perturbation level increases, the models' predictive performance deteriorates, as reflected in higher error metrics and negative R^2 values. For the RFR, the MAE, MSE, and RMSE increase as the perturbation level (epsilon) increases, indicating that the model's performance degrades with stronger attacks. R^2 values are mostly negative, suggesting poor predictive performance. This indicates that the RFR is vulnerable to PGD attacks. Similar to RFR, XGBoost exhibits a decrease in performance as the perturbation level increases. R^2 values are also negative, indicating that the model's

predictions deteriorate under stronger attacks. SVR follows a similar pattern, with a decrease in performance as the perturbation level increases. R^2 values remain negative. KNN initially performs reasonably well with low perturbation but shows a significant drop in R^2 values as the attack strength increases. This suggests that KNN is also vulnerable to PGD attacks. The proposed model follows a similar pattern as the other models, with a decrease in performance as perturbation increases. The R^2 values are negative for stronger attacks.

Table 7 and Table 8 shows that the L-BFGS and the Basic Iterative Method of Adversarial attack had no influence on the predicted results and yielded the same result on all perturbation. This indicates an unusual scenario and can be caused by several factors such as data entry error, Incorrect Implementation, the dataset (If the dataset used for the experiments is highly structured or has some unique characteristics), lack of diversity in models and data, etc. which will be looked into our future work. In conclusion, the recorded results highlight the importance of implementing robustness techniques and defenses against adversarial attacks in machine learning models to mitigate their susceptibility to such attacks.

4.4 Limitations and Future Works

From the recorded results we can see that the experimented dataset is few and for deep learning models' optimal performance, a large dataset for training is needed. Hence data augmentation will be looked into in our next study. Secondly, Table 7 and Table 8 shows that the L-BFGS and the basic iterative method of adversarial attack had no influence on the predicted results and yielded the same result on all perturbation which is unusual. Further study will include analyzing the reason why the results are the same. Lastly, the adversarial attacks implemented are all white box attacks, this study will further look into the effect of the black box attack on machine learning models for predicting hydrogen production.

5. CONCLUSION

Our investigation revealed that FGSM and PGD adversarial attacks had a significant impact on the predictive performance of the employed machine learning regression models. These attacks resulted in a degradation of performance metrics, with increasing perturbation levels. The severity of vulnerability varied among models, with the Proposed Model being particularly sensitive to adversarial attacks. For the RFR, as perturbation increased, the MAE, MSE, and RMSE all increased, indicating deteriorating performance. Negative R^2 values suggested that the model's predictions became worse than using the mean of the target values. XGBoost and SVM models displayed similar patterns of vulnerability, with negative R^2 values indicating poor predictive performance under stronger attacks. KNN initially performed better but still exhibited sensitivity to PGD attacks.

Our findings underscore the importance of developing defense mechanisms and security protocols to protect the integrity of deep learning-based predictions in the critical domain of hydrogen production. Furthermore, they highlight the potential risks associated with relying on these models for decision-making in hydrogen production processes. Robustness techniques must be implemented to mitigate the susceptibility of these models to adversarial attacks. An intriguing aspect of our study was the identical results obtained from the L-BFGS and Basic Iterative Method attacks across all perturbation levels and models. This unusual scenario calls for further investigation into potential factors, such as data entry errors, incorrect implementations, dataset characteristics, and model diversity, which may have contributed to this unexpected outcome.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the support of the National Natural Science Foundation of China (NFSC, Grant No. 52007025) and the Science and Technology Support Program of Sichuan Province (2022JDRC0025).

DECLARATION OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. All authors read and approved the final manuscript.

REFERENCE

[1] Agyekum E B, Velkin V I. Optimization and techno-economic assessment of concentrated solar power (csp) in south-western africa: a case study on ghana[J]. Sustainable Energy Technologies and Assessments, 2020, 40:100763.

[2] Gyamfi B A, Adebayo T S, Bekun F V, et al. Beyond environmental kuznets curve and policy implications to promote sustainable development in mediterranean[J]. Energy Reports, 2021, 7:6119–6129.

[3] Agyekum E B. Energy poverty in energy rich ghana: a swot analytical approach for the development of ghana's renewable energy[J]. Sustainable Energy Technologies and Assessments, 2020, 40:100760.

[4] Tarhan C, Çil M A. A study on hydrogen, the clean energy

of the future: hydrogen storage methods[J]. Journal of Energy Storage, 2021, 40:102676.
[5] Zhao Y, Ramzan M, Adebayo T S, et al. Role of renewable energy consumption and technological innovation to achieve carbon neutrality in spain: fresh insights from wavelet coherence and spectral causality approaches[J]. Frontiers in Environmental Science, 2021, 9:769067.
[6] Yaqoob S J, Motahhir S, Agyekum E B. A new model for a photovoltaic panel using proteus software tool under arbitrary environmental conditions[J]. Journal of Cleaner Production, 2022, 333:130074.
[7] Adebayo T S, Awosusi A A, Oladipupo S D, et al. Dominance of fossil fuels in japan's national energy mix and implications for environmental sustainability[J]. International Journal of Environmental Research and Public Health, 2021, 18(14):7347.
[8] da Silva Veras T, Mozer T S, da Silva César A, et al. Hydrogen: trends, production and characterization of the main process worldwide[J]. International journal of hydrogen energy, 2017, 42(4):2018–2033.
[9] Devasahayam S. Decarbonising the portland and other cements—via simultaneous feedstock recycling and carbon conversions sans external catalysts[J]. Polymers, 2021, 13(15):2462.
[10] Devasahayam S, Raju G B, Hussain C M. Utilization and recycling of end of life plastics for sustainable and clean industrial processes including the iron and steel industry[J]. Materials Science for Energy Technologies, 2019, 2(3):634–646.
[11] Block C, Ephraim A, Weiss-Hortala E, et al. Co-pyrogasification of plastics and biomass, a review[J]. Waste and Biomass Valorization, 2019, 10:483–509.
[12] You S, Ok Y S, Tsang D C W, et al. Towards practical application of gasification: a critical review from syngas and biochar perspectives[J]. Critical Reviews in Environmental Science and Technology, 2018, 48(22-24):1165–1213.
[13] Devasahayam S. Opportunities for simultaneous energy/materials conversion of carbon dioxide and plastics in metallurgical processes[J]. Sustainable Materials and Technologies, 2019, 22:e00119.
[14] Sepe A M, Li J, Paul M C. Assessing biomass steam gasification technologies using a multi-purpose model[J]. Energy Conversion and Management, 2016, 129:216–226.
[15] Sterner M. Bioenergy and renewable power methane in integrated 100% renewable energy systems. limiting global warming by transforming energy systems: limiting global warming by transforming energy systems[M]. kassel university press GmbH, 2009.
[16] Saad J M, Williams P T. Manipulating the h2/co ratio from dry reforming of simulated mixed waste plastics by the addition of steam[J]. Fuel Processing Technology, 2017, 156:331–338.
[17] Devasahayam S. Catalytic actions of mgco3/mgo system for efficient carbon reforming processes[J]. Sustainable Materials and Technologies, 2019, 22:e00122.
[18] Devasahayam S, Strezov V. Thermal decomposition of magnesium carbonate with biomass and plastic wastes for simultaneous production of hydrogen and carbon avoidance[J]. Journal of Cleaner Production, 2018, 174:1089–1095.
[19] Wang Z, Burra K G, Lei T, et al. Co-pyrolysis of waste plastic and solid biomass for synergistic production of biofuels and chemicals-a review[J]. Progress in Energy and Combustion Science, 2021, 84:100899.
[20] Guleria P, Naga Srinivasu P, Ahmed S, et al. Xai framework for cardiovascular disease prediction using classification techniques[J]. Electronics, 2022, 11(24):4086.
[21] Ravegnini G, Ferioli M, Morganti A G, et al. Radiomics and artificial intelligence in uterine sarcomas: a systematic review[J]. Journal of Personalized Medicine, 2021, 11(11):1179.
[22] Biggio B, Roli F. Wild patterns: ten years after the rise of adversarial machine learning[C]. In: *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*. 2018: 2154–2156.
[23] Dalvi N, Domingos P, Mausam, et al. Adversarial classification[C]. In: *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. 2004: 99–108.
[24] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Advances in neural information processing systems, 2012, 25.
[25] Szegedy C, Zaremba W, Sutskever I, et al. Intriguing properties of neural networks[J]. 2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings, 2013.
[26] Vidas L, Castro R. Recent developments on hydrogen production technologies: state-of-the-art review with a focus on green-electrolysis[J]. Applied Sciences, 2021, 11(23):11363.
[27] Ishaq H, Dincer I, Crawford C. A review on hydrogen production and utilization: challenges and opportunities[J]. International Journal of Hydrogen Energy, 2022, 47(62):26238–26264.
[28] Dincer I, Acar C. Review and evaluation of hydrogen production methods for better sustainability[J]. International journal of hydrogen energy, 2015, 40(34):11094–11111.
[29] Faizollahzadeh Ardabili S, Najafi B, Shamshirband S, et al. Computational intelligence approach for modeling hydrogen production: a review. engineering applications of computational fluid mechanics 12 (1): 438--458. 2018.
[30] Bion N, Duprez D, Epron F. Design of nanocatalysts for green hydrogen production from bioethanol[J]. ChemSusChem, 2012, 5(1):76–84.
[31] d'Amore-Domenech R, Santiago O, Leo T J. Multicriteria analysis of seawater electrolysis technologies

for green hydrogen production at sea[J]. Renewable and Sustainable Energy Reviews, 2020, 133:110166.
[32] Haider S A, Sajid M, Iqbal S. Forecasting hydrogen production potential in islamabad from solar energy using water electrolysis[J]. International Journal of Hydrogen Energy, 2021, 46(2):1671–1681.
[33] Ayodele B V, Mustapa S I, Alsaffar M A, et al. Artificial intelligence modelling approach for the prediction of co-rich hydrogen production rate from methane dry reforming[J]. Catalysts, 2019, 9(9):738.
[34] Hossain M A, Ayodele B V, Cheng C K, et al. Artificial neural network modeling of hydrogen-rich syngas production from methane dry reforming over novel ni/cafe2o4 catalysts[J]. International Journal of Hydrogen Energy, 2016, 41(26):11119–11130.
[35] Alsaffar M A, Ghany M A R A, Ali J M, et al. Artificial neural network modeling of thermo-catalytic methane decomposition for hydrogen production[J]. Topics in Catalysis, 2021, 64:456–464.
[36] Abdelkareem M A, Soudan B, Mahmoud M S, et al. Progress of artificial neural networks applications in hydrogen production[J]. Chemical Engineering Research and Design, 2022, 182:66–86.
[37] Katterbauer K, Qasim A, Marsala A, et al. A data driven artificial intelligence framework for hydrogen production optimization in waterflooded hydrocarbon reservoir[C]. In: *Abu Dhabi International Petroleum Exhibition and Conference*. 2021: D041S123R003.
[38] Santhosh J, Sarkar O, Mohan S V. Green hydrogen-compressed natural gas (bio-h-cng) production from food waste: organic load influence on hydrogen and methane fusion[J]. Bioresource Technology, 2021, 340:125643.
[39] Soler Crespo V E, Linares Hurtado J I, Arenas Pinilla E M, et al. Hydrogen from municipal solid waste as a tool to compensate unavoidable ghg emissions[J]. 2022.
[40] Rezaeitavabe F, Saadat S, Talebbeydokhti N, et al. Enhancing bio-hydrogen production from food waste in single-stage hybrid dark-photo fermentation by addition of two waste materials (exhausted resin and biochar)[J]. Biomass and Bioenergy, 2020, 143:105846.
[41] Hosseinzadeh A, Zhou J L, Altaee A, et al. Machine learning modeling and analysis of biohydrogen production from wastewater by dark fermentation process[J]. Bioresource technology, 2022, 343:126111.
[42] Haq Z U, Ullah H, Khan M N A, et al. Hydrogen production optimization from sewage sludge supercritical gasification process using machine learning methods integrated with genetic algorithm[J]. Chemical Engineering Research and Design, 2022, 184:614–626.
[43] Guo H, Wu S, Tian Y, et al. Application of machine learning methods for the prediction of organic solid waste treatment and recycling processes: a review[J]. Bioresource technology, 2021, 319:124114.
[44] Akhtar N, Mian A. Threat of adversarial attacks on deep learning in computer vision: a survey[J]. Ieee Access, 2018, 6:14410–14430.
[45] Zhou Z-H. Machine learning[J]. 2021.
[46] Mahesh B. Machine learning algorithms-a review[J]. International Journal of Science and Research (IJSR).[Internet], 2020, 9(1):381–386.
[47] Chin B L F, Yusup S, Al Shoaibi A, et al. Optimization study of catalytic co-gasification of rubber seed shell and high density polyethylene waste for hydrogen production using response surface methodology[J]. Advances in bioprocess technology, 2015, :209–223.
[48] Chang H J, Huang K, Wu C. Determination of sample size in using central limit theorem for weibull distribution[J]. International journal of information and management sciences, 2006, 17(3):31.