2024-01-17

An Integrated Deep Learning Model with Genetic Algorithm (GA) for Optimal Syngas Production Using Dry Reforming of Methane (DRM)

Zarabian, Maryam

UNIVERSITY OF CALGARY

An Integrated Deep Learning Model with Genetic Algorithm (GA) for Optimal Syngas Production Using Dry Reforming of Methane (DRM)

by

Maryam Zarabian

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

CALGARY, ALBERTA
JANUARY, 2024

Abstract

Escalating energy demands and urbanization are substantial contributors to the exacerbation of the greenhouse effect. Predominantly, carbon dioxide (CO_2) emissions stem from fossil fuel derivatives. Additionally, significant methane (CH_4) emissions arise from natural gas reserves and biogas generated through the decomposition of organic materials. The dry reforming of methane (DRM) process not only holds promise in addressing environmental concerns associated with CO_2 and CH_4 but also serves as a recognized method for syngas production—a valuable industrial precursor for downstream chemical processes. Nevertheless, the commercial viability of this method is hindered by its high energy demand and side reactions disrupting the DRM reaction process. Utilizing a well-suited catalyst explicitly designed for DRM reaction to facilitate the process at lower temperatures without yielding undesirable by-products is regarded as a notable advancement. However, catalyst efficiency and thermodynamic elements, such as pressure, temperature, and feed ratio of reactants for DRM, are mutually influenced by each other. Hence, catalyst development and systematically controlling the thermodynamic aspects of the DRM reaction necessitate process modelling and optimization. This research employs an advanced deep learning model with a supervised greedy layer-wise pretraining approach as a viable alternative to mechanism-driven catalytic reaction process modelling, namely kinetics, to predict DRM reaction performance parameters. The thesis offers a pretraining approach with distinct advantages over the conventional deep learning methods. Notably, it demands less data, making it particularly valuable when experimental data collection is time-consuming and expensive. Furthermore, it conserves computational resources by significantly reducing the number of trainable parameters while demonstrating higher accuracy. The proposed deep learning model is exclusively trained using an experimental dataset. Prior to training, outliers within the experimental samples are identified and removed through the application of an unsupervised clustering algorithm, namely density-based spatial clustering of applications with noise (DBSCAN). Meanwhile, we investigate the causes of anomalies encountered during data acquisition. Subsequently, the process model combines with an evolutionary algorithm (EA), specifically a genetic algorithm (GA), to facilitate process optimization. Rather than exhaustively computing all potential values, the heuristic optimizer identifies feasible solutions by considering well-established DRM reaction constraints, ultimately providing optimal temperature and feed ratio values within a specified range approximately consistent with the experimental dataset.

Preface

This thesis is an original, unpublished, independent work by the author, Maryam Zarabian.

Acknowledgments

First and foremost, my greatest appreciation goes to my supervisor and, more importantly, my outstanding coach, Dr. Abraham O. Fapojuwo. His inspiring attitude, regular scheduling and gentle guidance stimulate me to work hard and be creative. I would like to acknowledge financial support from Carbonova Corp. and extend my acknowledgment to Carbonova co-founders Drs Mina Zarabian and Pedro Pereira-Almao for providing this opportunity. Special thanks to Carbonova employees Dr. Hector Guzman, Michael Bianchini and Alireza Zarabian, who contributed to my research by providing the experimental dataset as the primary source of this study and generously gave their time and insights. I would like to sincerely thank my examiners, Drs. Roberto Sousa, and Matthew Clarke. I am grateful for my family's endless love and support, either those living in Canada or Iran. I am proud to announce that this study is an extension of previous works at the University of Calgary conducted by my brilliant sister, Mina's Ph.D. research, establishing a novel eco-friendly chemical process, and my brother, Ali's M.Sc. research, designed a reactor for the process plant. I am so thankful to my dad, Naghi, who was my role model due to his hard-working personality, and my mom, Fereshteh, who trained me in commitment and unconditional love. Also, thanks to my beloved brother, Mohammadreza for standing beside me despite the long physical distance.

This work is dedicated to all women deprived of education, especially Afghan women.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Definition
AI	Artificial intelligence
ANN	Artificial neural networks
CFD	Computational fluid dynamics
CNF	Carbon nanofiber
CRZ	Carbonization reaction
DBSCAN	Density-based spatial clustering of applications with noise
DL	Deep learning
DNN	Deep neural networks
DRM	Dry reforming of methane
EA	Evolutionary algorithm
GA	Genetic algorithm
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MD	Methane decomposition
MSE	Mean squared error
ML	Machine learning
PCA	Principal component analysis
PSO	Particle swarm optimization
RMSE	Root mean squared error
RWGS	Reverse water gas shift
SRM	Steam reforming of methane
SSR	Software sensor for regression
TOS	Time on stream

List of Symbols

Symbol or abbreviation	Definition
ΔH	Enthalpy of reaction
E_a	Activation energy
kPa	Partial pressure
$F_{X_{(in)}}$	The inlet flow of gas X
$F_{X_{(out)}}$	The outlet flow of gas X
x	The vector of input/independent variables for a model
y	The vector of output/dependent variables for a model
z	Weighted sum
ϕ	Activation function
h	Predictive function
f	Objective function

Chapter 1

Introduction

1.1 Context and background

Artificial intelligence (AI) has already initiated a significant transformation in the chemical industry, and its influence is set to be even further in the coming years. AI offers outstanding opportunities in numerous applications ranging from predictive analytics and the optimization of process parameters to the enhancement of product quality. A recent survey conducted by McKinsey [1] revealed that industrial companies at the forefront of AI adoption outperformed their peers by a substantial margin, with a factor of 3.4. On a global scale, the survey projected that AI has the potential to generate approximately $13 trillion in total economic activity by the year 2030. Furthermore, there remains an untapped potential of approximately $1 trillion in value within the industrial sector. According to this survey, an increase of 10 to 15 percent in production and 4 to 5 in EBITA[1] was reported by operators applying AI in their industrial processing plants.

Figure 1.1 shows how AI improves a process plant through a series of actions. A process plant typically utilizes various sensors and instruments to capture process measurements such as temperature, pressure, and flow. This historical data is well-suited to leverage the advantages of AI, providing advanced process control and analytics. AI has the capability to identify patterns and reveal insights that are mostly invisible to humans. These insights can be given to the subject matter experts to integrate with the operating procedure and improve the process performance.

[1] Earnings Before Interest, Taxes, and Amortization

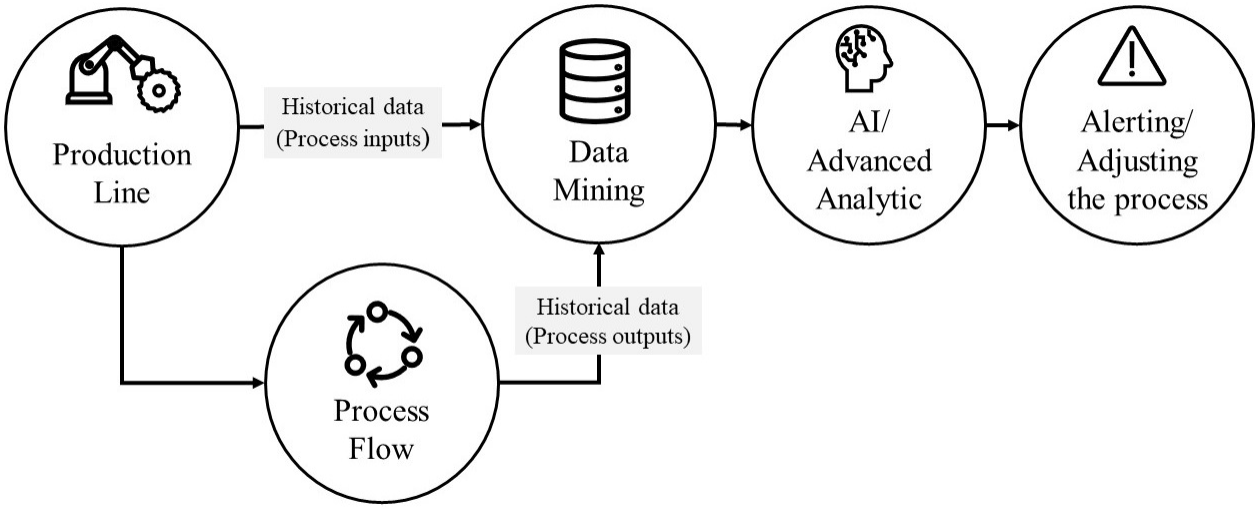


Figure 1.1: AI-assisted process plant.

1.2 Motivation

Carbonova technology transforms two primary sources of greenhouse gases (GHGs), carbon dioxide (CO_2) and methane (CH_4), into valuable carbon nanofibers. The Economist journal [2] described this innovative process magic as an energy-efficient and eco-friendly path. Carbon nanofiber is a remarkable substance with extraordinary thermal, electrical, mechanical, and chemical characteristics. Its versatility makes it valuable across a diverse range of sectors, including vehicles, construction materials, electronic devices, textiles, inks, protective coatings, and lubricants [3]. The developed synthesis path as shown in Figure 1.2 [4] encompasses two consecutive catalytic chemical reactions: syngas production via dry reforming of methane reaction (DRM) referred to Stage 1 (i.e., reforming reactor) and solid carbon nanofibers (CNFs) via carbonization reaction (CRZ) referred to Stage 2 (i.e., carbonization reactor).

In the literature, kinetic models often characterize the non-linear relationship governing the conversion process within the reactors. Nevertheless, these two consecutive processes exhibit complex kinetics and some physical limitations in measurement. Moreover, an accurate and fast prediction of the process outcomes under varying operational conditions poses a considerable challenge where an alternate kinetic model is required. Hence, the motivation for this study lies in leveraging the potential benefits of artificial intelligence (AI), particularly deep learning(DL), within Carbonova's two-stage conversion process. The primary goal of this study is to enhance process performance and product quality of the reforming reactor, known as Stage 1, aimed at improving the total solution.

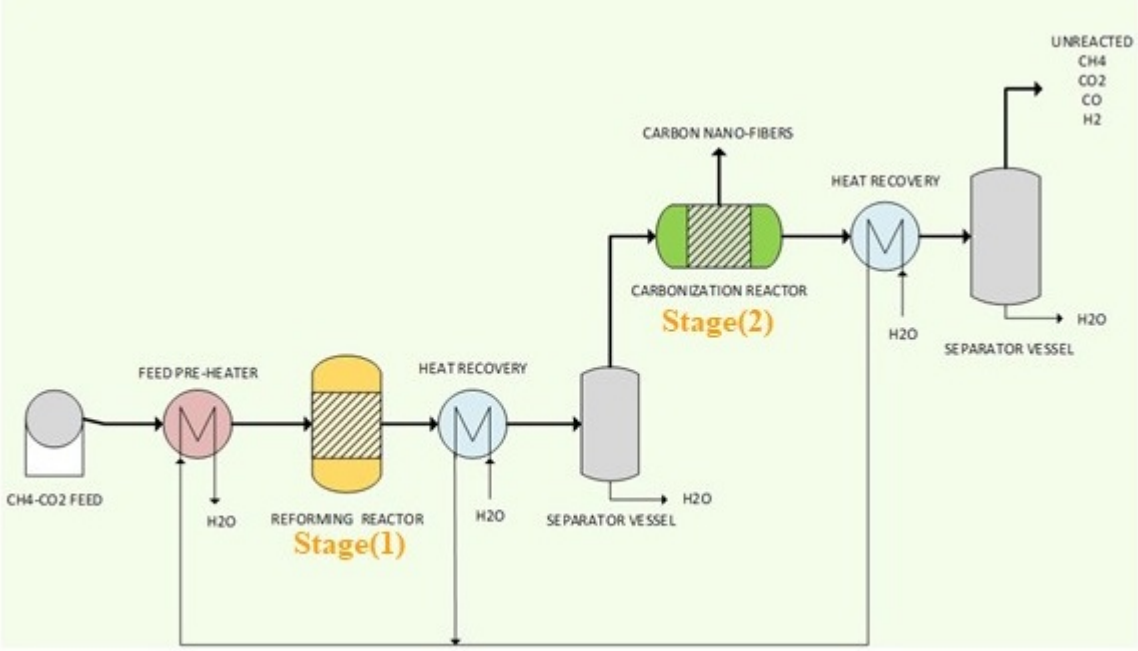


Figure 1.2: Carbonova chemical process (reprinted from Carbonova Corp. technical documents with permission [4]).

1.3 Problem statement and thesis objective

Syngas, a mixture of hydrogen (H_2) and carbon monoxide (CO), is a versatile and critical precursor in various industrial processes, which has gained significant importance in the global energy landscape. The traditional methods for syngas production primarily involve steam methane reforming (SMR) and partial oxidation of hydrocarbons, whereas dry reforming of methane (DRM) is a promising alternative which equally utilizes these two primary sources of greenhouse gases (GHGs) and produces syngas with an equal ratio of H_2 and CO denoted by [1:1] in accordance with the stoichiometric equation of the DRM reaction ($CH4 + CO2 \rightleftharpoons 2CO + 2H2$), well-suited syngas ratio for Carbonova technology. However, DRM presents several challenges, including complex reaction kinetics, catalyst deactivation, and requiring optimized various operating parameters for enhanced syngas yield. This study addresses the research problem of leveraging AI/DL capabilities for real-time adaptation of operational conditions such as pressure, temperature, and gas composition to optimize the DRM reaction process, specifically in the Carbonova Stage 1 context. In particular, the study answers the question: Which operational input conditions to Stage 1 maximize syngas production efficiently? The thesis objective aims to develop a deep learning model and optimization algorithms to tackle the foregoing research problem.

1.4 Contribution

This study investigates the experimental dataset in bench scale where collecting precisely labelled datasets with uniform distribution is expensive and often copes with physical limitations. Developing a deep learning model despite the lack of training samples is challenging, while previous works either implemented shallow learning or adopted synthetic simulated data to compensate for this shortage. This study contributes to deploying a deep learning model for the DRM reaction process, which benefits from the outstanding capabilities of pretraining for small datasets. Furthermore, it proposes a flexible optimizer for this catalytic reaction by integrating a deep neural network (DNN) and a genetic algorithm known as DNN+GA. The reaction at lower temperatures, such as the Carbonova temperature range of 500-586 °C, exhibits a multi-objective optimization problem tackled by this method. A recent open-source Python library, known as PyGAD, specifically published for genetic algorithms, facilitates the implementation of this optimizer without explicit programming.

1.5 Thesis outline

The thesis comprises five chapters to fulfill the research objectives. In Chapter 2, the literature review provides an introduction to the common applications of machine learning in chemical engineering as an interdisciplinary field. Then, it presents the industrial and environmental importance of dry reforming of methane, adding a discussion on the thermodynamics and kinetics of the reaction. The subsequence sections summarize conducted research involving AI to estimate the quality parameters of the DRM reaction and demonstrate the capabilities of these models to combine with evolutionary algorithms, offering a flexible optimizer that operates effectively under various operating conditions. Chapter 3 initially delves into the challenges of limited datasets for machine learning and how this research copes with them. Secondly, a process model utilizing deep learning and its predictive capabilities is presented. In Chapter 4, the AI-enabled process model provided in the previous chapter is integrated with a genetic algorithm to achieve the most efficient process under a defined condition. Despite the fact that this reaction can occur at around 300 °C, simultaneous reactions, specifically those causing carbon deposition, reduce the process efficiency at lower temperatures. However, scientists are encouraged to study the feasibility of this process at lower temperatures aimed at saving energy. This chapter initially delves into the process as a multi-objective optimization problem when the reaction occurs at a temperature ranging from 500-586 °C over a Ni/Al_2O_3 catalyst, and the optimizer handles the possibility of carbon deposition. In order to justify the method, in the following the same process is optimized under a different operating condition. Regarding the dataset

presented by Hossain et al. [5], in which the reaction serves at a temperature ranging from 700-800 °C over a $Ni/CaFe_2O_4$ catalyst, an ANN model is reproduced and combined with a genetic algorithm. However, due to less likelihood of side reactions and, consequently, less complexity, the optimizer tackles this new condition as a single-objective optimization problem while proposing a different optimal condition compared to the first problem, specifically a different gas composition, to reach maximum efficiency. Finally, in Chapter 5, the thesis concludes with a discussion on the engineering significance, thesis limitations and future work.

Chapter 2

Literature Review

2.1 Application of machine learning in chemical engineering

In chemical engineering, information collected from experiments, simulations, calculations, or open-source databases typically requires analysis to unfold the process concept. In order to analyze this information, engineers apply a standalone or a combination of two significant techniques known as mechanism-driven and data-driven models.

Mechanism-driven models, represented using chemical equations, elucidate the elementary steps or reactions that constitute the overall chemical process. These models provide a molecular-level perspective, exploring diverse pathways through which reactants transform into products based on experimental observations [6]. There is a close interrelation between reaction mechanisms and kinetics. Therefore, the kinetic measurements provide direct insight into the mechanism of a reaction. However, investigating the kinetics of heterogeneous reactions, especially catalytic ones, involves complex procedures [7]. In practice, factors such as heat and mass transfer introduce additional challenges in determining the reaction kinetics [8]. Furthermore, catalytic reactions exhibit unique mechanisms due to kinetics influenced by various catalysts [6].

In contrast, artificial intelligence (AI), specifically deep learning models, provides data-driven models able to extract the non-linear relationship between the process parameters and outputs, presenting process estimator models without accessing the reaction's mechanism. They are invaluable substitutes for mechanism-driven while their accuracy and agility facilitate process optimization.

The earliest attempts to employ computers as a problem solver, rather than a mere calculator, in chemical engineering traced back to [9] and [10]. They developed computer-aided models with the ability to design specific chemical processes. Since the 1990s, the significant rise in computing capacity led to the emergence of

advanced AI-enabled models, specifically deep learning models. This increase can be attributed to emerging integrated circuits and mass production of potent GPU cards thanks to the gaming industry. Furthermore, cloud platforms have democratized access to this computational power [11, p. 300]. The history of applying AI, particularly machine learning (ML), in chemical engineering was reviewed by [12], while the author highlighted the previous achievements and discussed future opportunities.

Machine learning as a subset of artificial intelligence, involves the development of statistical models and algorithms enabling computer systems to learn from data without being explicitly programmed [11, p. 4]. Machine learning systems are categorized based on the type and extent of supervision they receive into supervised, unsupervised, semi-supervised, self-supervised, and reinforcement learning [11, p. 9]. Two significant categories of supervised and unsupervised algorithms are used depending on whether the training set encompasses the desired solutions, known as labels, or not. Regression and classification are typical supervised learning tasks. The regression task predicts the continuous values or numeric data, whereas the classification categorizes the data points based on trained labels. Clustering, visualization algorithms, dimensionality reduction, and anomaly detection are the most common applications of unsupervised learning where no labelled data is available. Semi-supervised learning falls between these two paradigms and is typically used where the dataset is partially labelled and obtaining a fully labelled dataset is expensive or time-consuming. Regarding self-supervised algorithms, a fully labelled dataset is generated from the initially unlabeled dataset. Reinforcement learning (RL) deals with decision-making tasks. Agent, environment, and action are the main components of this paradigm. The agent continuously observes and interacts with the environment based on an award/penalty policy to achieve the predefined goal. Robotics, autonomous driving, game playing, recommendation systems and optimization are the most common applications of RL [11, pp. 10-17]. In the context of chemical engineering, soft sensors for regression (SSR) [13] represent a state-of-the-art technique for estimating the quality parameters of a process without full knowledge of the reaction kinetics. [14] categorized commonly used ANN architectures as estimators and enumerated a list of chemical process systems in which these techniques yielded satisfactory results. Moreover, recent research prominently employed deep learning, a subset of machine learning, in developing molecular systems. [15] designed a molecular system in which a recurrent neural network (RNN) model with Long Short-Term Memory (LSTM) architecture predicted chemical reactions. The research explored 11,000 elementary reactions while the model chained those elementary steps to predict the multi-step global reaction.

The literature survey by [16] enumerated the list of commonly used machine learning techniques and their relative application in chemical engineering. Data analysis tasks typically employ principal component analysis (PCA) and t-distributed stochastic neighbour embedding (t-SNE) algorithms in visualizing high-dimensional data, gaussian mixture modelling (GMM) and density-based spatial clustering of applications

with noise (DBSCAN) for clustering and diagnosing anomaly or novelty within samples, and LSTM for time series analysis. Classification and regression tasks commonly utilize decision trees, random forest (RF), support vector machine (SVM), and varied ensemble modelling. Moreover, artificial neural networks (ANN) are widely used in optimization, process control and monitoring tasks such as catalyst deactivation.

2.2 Dry reforming of methane (DRM)

2.2.1 The significance of dry reforming of methane (DRM), a catalytic reaction

Continuously increasing energy demand and subsequently increasing fossil fuel consumption drives carbon dioxide (CO_2) emissions. Additionally, natural gas reserves and biogas from municipal waste, sewage, agriculture waste, etc., are significant sources of methane (CH_4). Both mentioned gasses are the main contributors to the greenhouse effect, and scientists continually strive to mitigate their environmental impacts [17]. The most widely studied reforming technologies for conversion of CH4 to syngas i.e., a mixture of hydrogen (H_2) and carbon monoxide (CO) are steam (H_2O) reforming (SR), dry (CO_2) reforming (DRM), partial oxidation (PO), and autothermal reforming (AR) differing from oxidant used, final [H_2:CO] product ratio, kinetics, and energetics of the reaction [17]. Among them, DRM is the only known chemical reaction that equally utilizes these two main greenhouse gasses and converts them to an equivalent ratio of H_2 and CO [1:1], being a favourable feedstock for Fischer-Tropsch synthesis to produce liquid fuels [18]. DRM is a catalytic reaction whose performance is mainly dependent on catalyst efficiency. A catalyst is a compound that accelerates the rate of a chemical reaction while it is not consumed by the reaction. It offers an alternative reaction pathway, which requires less activation energy for the reactants to transform into products [6].

2.2.2 Thermodynamic of the reaction

DRM is a highly endothermic reaction requiring a high operating temperature to reach a large equilibrium conversion of methane and carbon dioxide to syngas.

$$\mathrm{CH_4 + CO_2 \rightleftharpoons 2CO + 2\,H_2} \qquad \Delta H_{298^\circ} = +247\,\mathrm{kJ\,mol^{-1}} \qquad \text{(DRM)}$$

Simultaneous side reactions, including reverse water gas shift (RWGS), Sabatier reaction (SR), steam reforming (SRM), methane decomposition (MD), and Boudouard reaction (BR), affect DRM equilibrium [19]:

$$CO_2 + H_2 \rightleftharpoons CO + H_2O \qquad \Delta H_{298^\circ} = +41.2\,\mathrm{kJ\,mol^{-1}} \qquad \text{(RWGS)}$$

$$CO_2 + 4\,H_2 \rightleftharpoons CH_4 + 2\,H_2O \qquad \Delta H_{298^\circ} = -165\,\mathrm{kJ\,mol^{-1}} \qquad \text{(SR)}$$

$$CH_4 + H_2O \rightleftharpoons CO + 3\,H_2 \qquad \Delta H_{298^\circ} = +206\,\mathrm{kJ\,mol^{-1}} \qquad \text{(SRM)}$$

$$CH_4 \rightleftharpoons C + 2\,H_2 \qquad \Delta H_{298^\circ} = +75\,\mathrm{kJ\,mol^{-1}} \qquad \text{(MD)}$$

$$2\,CO \rightleftharpoons C + CO_2 \qquad \Delta H_{298^\circ} = -171\,\mathrm{kJ\,mol^{-1}} \qquad \text{(BR)}$$

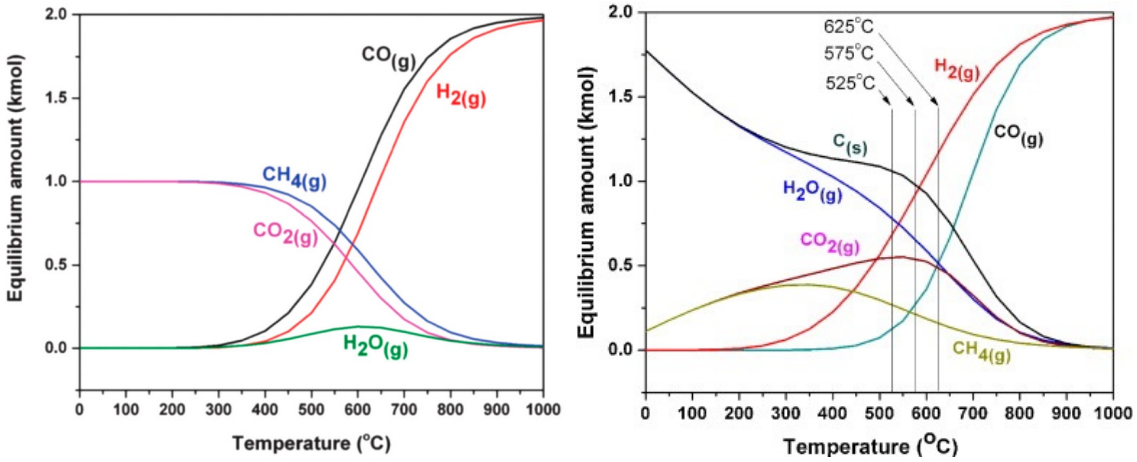
Since DRM is part of a complex network of reactions, the composition of DRM products, i.e., hydrogen (H_2) and carbon monoxide (CO), strongly depends on the feed ratio and conversion rate. According to the stoichiometric equation of the DRM reaction, a syngas ratio close to unity is expected. However, simultaneous side reactions, specifically RWGS, reduce the H_2/CO ratio to less than one by consuming hydrogen (H_2) molecules and transforming them into water (H_2O) molecules. Carbon deposition is another significant limitation of DRM due to side reactions. Solid carbon formed on the surface blocks the catalyst pores and disables the active sites. When catalyst activity drops, the primary reaction tends toward side reactions that ultimately disrupt the DRM reaction [17].

In addition to feed ratio, DRM is influenced by other thermodynamic components such as temperature and pressure. The thermodynamic simulations of DRM conducted by [20] and [21] demonstrated a substantial drop in both the yield of products and conversion rate by increasing the operating pressure.

While DRM is feasible at around 300°C, by-products are dominants during the possibility of side reactions. Depending on potential side reactions, whether carbon formation hinders CO formation or water formation consumes H_2 produced by DRM, the syngas ratio can be greater or lower than one. Even a near-unity syngas ratio is likely when side reactions equally impact H_2 and CO production. However, at high temperatures, particularly above 900°C, where side reactions are unlikely, DRM brings the syngas ratio close to one without the emergence of by-products [17].

Figure 2.1 illustrates the thermodynamic equilibrium of DRM reaction in a temperature ranging from 0 to 1000°C at a constant pressure of 1 atm. Figure 2.1(a) assumes no carbon formation is possible where H_2 and CO are the primary products, and H_2O is the only by-product [17]. The figure shows H_2 and CO formation starts above 300°C and continues at all temperatures with a constant H_2/CO ratio of 0.8-1 while a significant amount of H_2O appears through RWGS between 400-800°C. Figure 2.1(b) assumes carbon formation through MD and BR is also possible and significantly hinders CO formation, making the CO curve considerably far from the H_2 curve, which means the H_2/CO ratio increases in favour of H_2 [22]. However,

both Figure 2.1(a) and 2.1(b) show a syngas ratio close to one [1:1] at temperatures above 900°C, where no side reaction accompanies the DRM reaction [17].



(a) (b)

Figure 2.1: Thermodynamic equilibrium of DRM at temperatures ranging from 0-1000 °C, at a pressure of 1 atm and a feed ratio of [CO_2:CH_4]=1, calculated by HSC Chemistry 7.1 software. (a) assuming no carbon formation occurs (reprinted with permission from [17]). (b) assuming carbon formation occurs (reprinted with permission from [22]).

Thermal sintering is another severe issue that irreversibly deactivates the DRM catalyst at high temperatures due to metal particle growth and consequently blocks the pores and drops the catalyst's active sites [23]. Therefore, scientists face two significant challenges :

1. Developing a catalyst that is resilient enough against carbon deposition and thermal sintering and efficiently suppresses side reactions [24].
2. Running DRM in optimal operating conditions, including temperature, feed ratio and space velocity(SV) [1] [25].

2.2.3 The kinetic of the reaction

Unlike the thermodynamics of the DRM reaction, which is perfectly understandable, the kinetics and mechanism of the reaction remain a topic of debate. Figure 2.2 depicts an example of the kinetic scheme for DRM reaction on Ni-Rh/Al2O3 catalyst presented by [19] and by no means covers all available kinetics. The authors formulated the DRM reaction against 27 equations based on the biomolecular reactions strategy known as Langmuir–Hinshelwood kinetics [7] and fitted them with the experimental dataset. The research denoted the best-fitted formula in Equation 2.1 as the selected kinetic rate for the DRM reaction.

[1] The feed flow versus catalyst volume

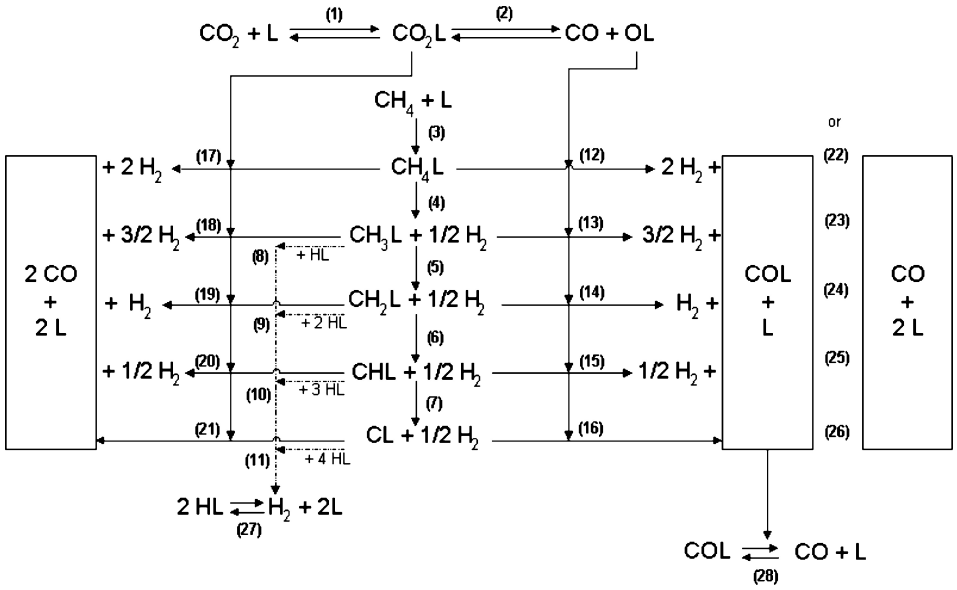


Figure 2.2: A sample of the kinetic scheme for DRM reaction (reprinted with permission from [19]).

$$r_{ref} = \frac{k_1 K_{CH4} K_{CO2} \left(P_{CH4} P_{CO2} / {P_{H2}}^{0.5} - {P_{H2}}^{1.5} P_{CO}{}^2 / K_{ref}\right)}{\left(1 + P_{CH4}/{P_{H2}}^{0.5} K_{CH4} + P_{CO2} K_{CO2}\right)^2} \tag{2.1}$$

Where

- r_{ref} is the reaction rate related to the DRM reaction.
- k_1 is the rate constant related to the first reforming step (the first elementary reaction).
- K_{CH4}, K_{CO2} and K_{ref} are the adsorption constant related to CH_4, CO_2 and the overall DRM reaction, respectively.
- P_{CH4}, P_{CO2}, P_{H2} and P_{CO} are the partial pressure relative to CH_4 and CO_2 as the reactants, and H_2 and CO as the products of the DRM reaction.

2.2.4 The criteria for a suitable DRM catalyst

The significant considerations in developing DRM catalysts center around the following properties [26, p. 179]:

- Activity: The catalysts should exhibit high activity by promoting the conversion of reactants into products. It involves enhancing the adsorption and activating reactant molecules on the catalyst surface.
- Selectivity: The catalysts should facilitate the desired reactions while minimizing unwanted

side reactions. The selectivity towards syngas formation and suppressing by-product formation is a crucial factor.

- Stability: The catalysts should maintain their activity over extended periods of operation. DRM is a highly demanding reaction that involves high temperatures and carbon deposition, so it is essential to ensure long-term performance.

Generally, catalysts are distinguished based on their composition as follows [27]:

- Active components: They provide a surface where reactant molecules are adsorbed and transformed into products. The nature and arrangement of active components significantly influence the catalyst's activity and selectivity in a specific reaction. In the literature, noble metals such as platinum (Pt), rhodium (Rh), and ruthenium (Ru) present high activity and selectivity toward DRM and prove more resilient against carbon deposition. However, non-noble metals are preferred due to their low cost, particularly Ni-based catalysts such as Ni/ZrO_2 and Ni/Al_2O_3 are widely used for DRM [17].
- Support: An active component is commonly deposited or supported on a suitable carrier material. The support serves as a structural framework, enhancing the catalyst's stability, surface area, and dispersion of active sites. Common support materials include ZrO_2, Al_2O_3, CeO_2, SiO_2, CaO, La_2O_3, MgO and TiO_2 [27].
- Promoters: The additional substances are incorporated into the catalyst formulation to improve its performance. They enhance the activity and selectivity or increase the catalyst's resistance against deactivation. Promoters play a crucial role in optimizing the catalyst for specific reactions. Fe, Mg, B, Co, Cu, Mn, Sn, V and Zr are common promoters for Ni-based DRM catalysts [28].

In order to achieve a comprehensive understanding of catalyst performance and design the most effective catalysts for specific applications, a combination of the following techniques is employed:

- Catalyst Characterization: Analyzing the physical and chemical characteristics of the catalyst that affect its performance using various techniques such as X-ray diffraction (XRD), X-ray photoelectron spectroscopy (XPS), scanning electron microscopy (SEM), transmission electron microscopy (TEM) [26, pp. 119-123].
- Reaction Kinetics: Studying the rate of product formation during a catalytic reaction under effective components, i.e., temperature and pressure [7].

- Product Analysis: Analyzing the products by various methods such as gas chromatography (GC) or mass spectrometry (MS) [26, p. 156].
- Stability Testing: Monitoring any changes in selectivity or activity of the catalyst as a function of time on stream while keeping other conditions constant [29].
- Modelling and Simulation: Employing data-driven models, e.g., AI-enabled models and computer simulations, to predict and understand the behaviour of the catalyst and optimize its performance (refer to section 2.1).

2.3 Application of deep learning to model dry reforming of methane (DRM) over varied catalysts

The industry typically offers datasets in a tabular format, representing the data as a vector of features. While deep learning has achieved remarkable success with homogeneous datasets like text, images, and audio, it faces serious challenges in the context of heterogeneous tabular datasets. In the literature [30] and [31], decision tree algorithms such as random forest (RF) and gradient-boosted decision tree (GBDT) are preferred as non-deep alternatives for this type of dataset. Nevertheless, software sensors prefer deep learning models due to their hierarchical abstracted representation, eliminating the workload of complicated feature engineering [32]. Their multimodal capability allows learning datasets containing images and numerical values simultaneously [30]. Additionally, the flexibility of deep neural networks (DNN) enables reusing pretrained layers in transfer learning, as well as various supervised and unsupervised pretraining methods where enough labelled data is not available [11, pp. 377–378]. In chemical engineering, we often deal with small datasets due to the expense of collecting experimental observations and implementing DNN with the assistance of pretraining is a common method studied by [33]. Table 2.1 summarizes the ANN application in modelling of DRM process on diverse catalysts. The previous works are mostly shallow learning, comprising only one hidden layer within their architecture, whereas deep learning models require at least two hidden layers [34]. The remaining presented only pure deep learning and never used the capabilities of pretraining for small datasets. For example, [5] utilized a tiny experimental dataset with only 27 samples and proposed a model that seriously suffered from overfitting. Some researchers preferred simulated datasets to provide enough training samples for deep learning models. For example, [35] implemented a deep learning model employing a simulated dataset for syngas production through steam reforming of methane (SRM). However, some process details are hidden from the simulator since they work based on assumptions, not the real world.

Table 2.1: Summary of ANN application in modelling of DRM process on diverse catalysts.

Related work	Study aim	ANN architecture
Ayodele et al. [37]	Syngas production from methane dry reforming.	ANN model (4:16:4) Input: CH_4 partial pressure, CO_2 partial pressure, feed ratio, reaction temperature Output: H_2 production rate, CO production rate, CH_4 conversion (%), CO_2 conversion (%)
Hossain et al. [5]	hydrogen-rich syngas production from methane dry reforming.	ANN model (3:12:4) Input: feed ratio, reaction temperature, metal loading in catalyst Output: CH_4 conversion (%), CO_2 conversion (%), H_2 yield, CO yield
Ayodele et al. [38]	CO-Rich Hydrogen Production Rate from Methane Dry Reforming.	ANN model (3:13:2) and (3:15:2) Input: CH_4 partial pressure, CO_2 partial pressure, reaction temperature Output: H_2 production rate, CO production rate
Alsaffar et al. [39]	Scavenging carbon deposition during hydrogen-rich	ANN model (3:16:1) and (3:20:1) Input: feed ratio, reaction temperature, N_2 flowrate Output: carbon deposition
Elmaz et al. [40]	Syngas Production from Methane Dry Reforming.	ANN model (4:9:9:4) Input: CH_4 partial pressure, CO_2 partial pressure, feed ratio, reaction temperature Output: H_2 production rate, CO production rate, CH_4 Conversion (%), CO_2 Conversion (%)
Ayodele et al. [41]	Carbon dioxide reforming of methane.	Various ANN architectures Input: calcination temperature, reduction temperature, reaction temperature, time on stream, Ni loading Output: CH_4 conversion (%), CO_2 conversion (%)

2.4 Combining the artificial neural network (ANN) with an evolutionary algorithm (EA) for process optimization

An estimator alone is invaluable; however, the main problem is getting to the most efficient process. Optimal operating conditions considering the process constraints lead to the most efficient process, where the optimization problem arises. [36] studied various methods to address an optimization problem related to a chemical process. The first step involves deriving the process model using different approaches, classified into mechanism-driven and data-driven. Next, the model as the basis for an objective function with the addition of various constraints serves to tackle the optimization problem through analytical or heuristic approaches. [37] applied a combination of artificial neural networks model with Box-Behnken design, for optimal syngas production through DRM reaction. Integrating software sensors with evolutionary algorithms such as particle swarm optimization (PSO) and genetic algorithm (GA) is a common technique to address optimization problems. Table 2.2 summarizes various optimization problems related to chemical engineering solved through this method.

Table 2.2: Summary of integrated ANN with EA

Related work	Contribution
Hao et al. [42]	The combustion process in a pulverized coal-burned utility boiler aimed at minimum nitrogen oxide (NOx) emissions. A three-layer ANN (29:31:1) combined with GA.
Nandi et al. [43]	Optimization of benzene isopropylation on H-beta catalytic process. Two distinct AI-enabled models, SVR and ANN with five hidden layers integrated with GA to obtain the maximum selectivity and yield.
Kana et al. [44]	Optimization of biogas production from saw dust and other co-substrates. A three-layer ANN(5:2:1) formed a fitness function for GA
Soleimani et al. [45]	Predicting the permeation flux and fouling resistance in commercial polyacrylonitrile (PAN) ultrafiltration (UF) membranes aimed at efficiently separating oil from industrial wastewater. Two distinct four-layer ANNs (4:8:1:1) independently investigated the non-linear relationship between operating conditions with permeation flux and fouling resistance. Those ANN models formed a multi-objective optimization problem (MOOP) solved by GA.
Vel´asco-Mejía et al. [46]	Optimization of the pharmaceutical crystallization process to improve the quality of the product. A four-layer ANN (9:6:10:1) discovered the process model followed as a fitness function for GA.
Shin et al. [47]	Submerged combustion vaporizer (SCV) aimed at minimum NOx in the flue gas of the process. GA combined with a three-layer ANN (4:17:1) as a fitness function.
Gbadago et al. [48]	Butadiene synthesis over a ferrite catalyst aimed at maximum selectivity, yield and conversion. The non-linear behaviour of the process was modelled with a five-layer DNN (3:5:5:5:3) using Python and a dataset simulated by CFD to combine with GA.
Khezri et al. [49]	Maximize the wax production rate of the large-scale gas to liquids (GTL) process. A four-layer ANN (4:7:15:1) integrated with GA.
Fang et al. [50]	Optimization of industrial propane dehydrogenation process. Four AI-enabled models, KNN, DT, SVM, and ANN, investigated through simulated dataset while finally DT integrated with PSO to optimize total profit (y1) and SVM integrated with PSO to optimized propylene yield (y2).
Alotaibi et al. [51]	Optimization of dry reforming of methane. Two distinct two-layer ANNs with (1:4:3:5) and (2:6:5:5) architectures adopted synthetic data simulated by CFD while a combination of ANN+GA optimized temperature, gas velocity and physical properties of the reactor.

2.5 Thesis work in the context of current state of the art

Applying machine learning in chemical engineering is a relatively new interdisciplinary topic that not only requires sophisticated knowledge from both computer science and chemistry fields but also demands detailed insight into the specific process to which it is applied. This search comprehensively studies the limitations of DRM reaction in syngas production and contributes to alleviating its commercialization obstacles. It presents a software sensor for regression (SSR), specifically employing deep learning on a relatively small experimental dataset for the first time. It employs a pretraining approach to cope with the shortage of training samples and simultaneously benefits from the advantages of deep learning for software sensors. The SSR estimates the quality parameters of DRM reaction and facilitates catalyst development and process optimization in conjunction with conventional kinetics models. Additionally, an integrated DNN+GA optimizer implemented for the first time in terms of DRM reaction presents a versatile methodology fitting with various DRM catalysts and operational conditions. The optimizer delves into process optimization as a multi-objective problem that aims explicitly to maximize gas consumption, i.e., CO_2 and CH_4 and achieve a near-unity syngas ratio. The constraint handling technique distinguishes feasible solutions and discards infeasible ones during optimization. These constraints are consistent with the DRM reaction borders in carbon deposition and catalyst deactivation issues. Furthermore, we employ the newly open-source library in Python known as PyGAD to implement the complicated details of the genetic algorithm in a few lines of programming codes.

Chapter 3

AI-enabled Process Models for Dry Reforming of Methane (DRM)

3.1 Introduction

This chapter explores the deployment of three distinct machine learning approaches: random forest (RF), pure deep learning, and pretrained deep learning in predicting the DRM reaction performance. These models are trained, validated, and tested comprehensively against an experimental dataset, and the superior model predicts the impact of temperature and feed ratio variations on DRM reaction performance. Section 3.2 provides insights into the dataset comprising the data acquisition process, preprocessing tasks, and the model variables. Section 3.3 offers a robust ensemble learning known as random forest due to its simplicity and high performance in the context of heterogeneous tabular datasets. Subsequently, sections 3.4 and 3.5 explore deep learning, leveraging distinct methods to optimize the network's parameters. The latter highlights how supervised greedy layer-wise pretraining facilitates deep learning and reduces demand for computational resources. With available models, section 3.6 discusses the evaluation metrics ensuring the reliability and generalizability of the aforementioned models as a regression task. Section 3.7 presents findings related to anomaly detection and root causes, demonstrating the performance of all models, and proposing the superior model along with its predictive capabilities. Finally, the chapter summary highlights the crucial aspects of this chapter's contribution and corresponding findings.

3.2 Dataset overview

3.2.1 Data acquisition and experimental setup

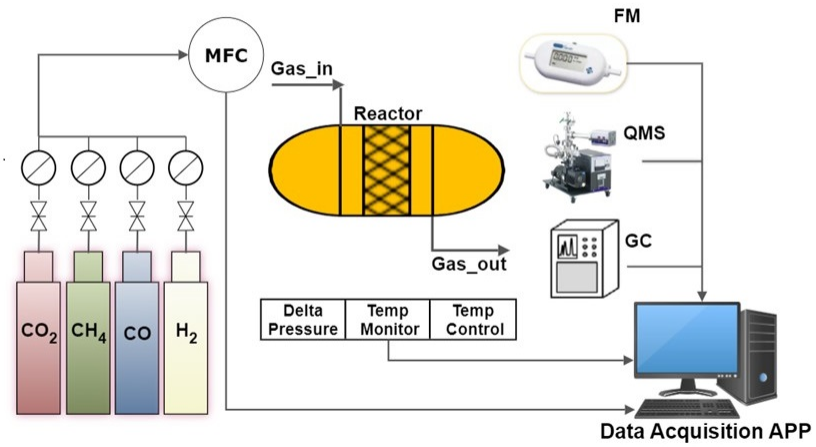


Figure 3.1: The experimental setup scheme for DRM reaction.

The experimental setup scheme for DRM reaction shown in Figure 3.1 measures gas components, monitors the thermodynamic elements, and collects historical data during the process [1]. Two primary gas components, methane (CH_4) and carbon dioxide (CO_2), are initially fed to the reactor. Furthermore, the experiment explores the recycling process by introducing external hydrogen (H_2) and carbon monoxide (CO) into the reactor at a specific stage. The experiment maintains the space velocity (SV) at a constant value for all samples by introducing argon (Ar), a gas component that does not contribute to the process. Meanwhile, it tracks catalyst stability over time on stream. The mass flow controller (MFC) and gas flow meter (FM) measure inlet and outlet flow, respectively. Regarding the inlet flow, MFC also controls the flow of each gas component. However, measuring the gas composition of the outlet flow is a more complex process, and this task is accomplished using two instruments: the Quadrupole Mass Spectrometer (QMS) and the Gas Chromatograph (GC).

3.2.2 Feature extraction

The process model requires DRM efficiency parameters, whereas the instruments in Figure 3.1 provide the experimental results in a simplified form. Equations (3.1),(3.2),(3.3),(3.4) and (3.5) are used to extract

[1]The calculations consider gauge pressure.

H_2/CO ratio, conversion rate and product yield percentages in which $F_{X_{(in)}}$ is the inlet flow rate (mol/min) and $F_{X_{(out)}}$ is the outlet flow rate (mol/min) for each denoted gas X. Additionally, values corresponding to the inlet and outlet flows are standardized as Std_Inlet_flow and Std_Outlet_flow based on quantities introduced into and exhausted from the reactor under steady-state conditions.

$$H_2/CO\ Ratio = \frac{F_{H_{2(\text{out})}}}{F_{CO_{(\text{out})}}} \times 100 \tag{3.1}$$

$$CH_4\ Conversion = \frac{F_{CH_{4(in)}} - F_{CH_{4(out)}}}{F_{CH_{4(in)}}} \times 100 \tag{3.2}$$

$$CO_2\ Conversion = \frac{F_{CO_{2(in)}} - F_{CO_{2(out)}}}{F_{CO_{2(in)}}} \times 100 \tag{3.3}$$

$$H_2\ Yield = \frac{F_{H_{2(\text{out})}}}{2 \times F_{CH_{4(\text{in})}}} \times 100 \tag{3.4}$$

$$CO\ Yield = \frac{F_{CO_{(out)}}}{F_{CH_{4(in)}} + F_{CO_{2(in)}}} \times 100 \tag{3.5}$$

3.2.3 Feature unit translation

In the literature, the process model for dry reforming of methane (DRM) considers the gas components introduced into the reactor on a partial pressure (kPa) rather than the milliliter per minute (ml/min) measurement typically used during data collection.

3.2.4 Anomaly detection and data cleaning

Errors are expected during data acquisition, and manually identifying the outliers is a time-consuming and often challenging task. An unsupervised clustering algorithm, namely density based spatial clustering of applications with noise (DBSCAN), is known as a robust outlier detection method. DBSCAN arranges each dense region as a cluster using only two hyperparameters (ε, min_samples). It initially defines a few randomly selected instances as core while any instances within a distance of ε (e.g., $\varepsilon = 0.5$) from these core samples are considered neighbours. A minimum number of samples (e.g., min_samples = 5) is essential to form a cluster. Instances neither qualified as core nor neighbour are treated as outliers with "-1" label. The algorithm identifies the best possible clusters through multiple iterations. In addition to removing outliers, we also discard all observations containing missing or invalid values [11, p. 279].

Table 3.1: Dataset Structure [52]

x	(Independent variables)	Description
x_1	TOS (h)	time on stream
x_2	Set temp (°C)	furnace temperature
x_3	Reaction temp (°C)	reaction temperature
x_4	Std_Inlet_flow (ml/min)	standard total inlet gas fed to the reactor (excluding Ar)
x_5	CH_4 (kPa)	methane partial pressure
x_6	CO_2 (kPa)	carbon dioxide partial pressure
x_7	CO (kPa)	carbon monoxide partial pressure
x_8	H_2 (kPa)	hydrogen partial pressure
y	(Dependent variables)	Description
y_1	Std_Outlet_flow (ml/min)	standard total outlet gas exhausted from the reactor (excluding Ar)
y_2	H_2/CO ratio (%)	syngas ratio
y_3	H_2 yield (%)	hydrogen moles in relation to reactants
y_4	CO yield (%)	carbon monoxide moles in relation to reactants
y_5	CH_4 Conversion (%)	reactant conversion rate
y_6	CO_2 Conversion (%)	reactant conversion rate

Table 3.2: Range of Variables [52]

Measures	Parameters	Min	Max
TOS (h)	time on stream	0	380
Set temp (°C)	furnace temperature	560	585
Reaction temp (°C)	reaction temperature	562	586
Std_Inlet_flow (ml/min)	standard total inlet gas fed to the reactor (excluding Ar)	196.81	257.73
CH_4 (kPa)	methane partial pressure	39.65	51.87
CO_2 (kPa)	carbon dioxide partial pressure	39.85	52.13
CO (kPa)	carbon monoxide partial pressure	0	12.25
H_2 (kPa)	hydrogen partial pressure	0	24.49

3.2.5 Dataset structure

The error-free dataset consists of 278 observations of DRM reactions collected during an ongoing process spanning approximately 380 hours [52]. The AI-enabled model tackles a regression problem in which the independent variables represented by the vector $\mathbf{x}$ determine the dependent variables denoted as vector $\mathbf{y}$, both presented in Table 3.1. The process model is depicted as $y = h(x)$, where $h(.)$ represents the predictive model for DRM performance parameters. The model describes the nonlinear relationship between input and output variables when each training instance is presented individually to the model.

Table 3.2 illustrates the range of input variables within the dataset, with distinct values representing furnace and reaction temperatures for each instance. These variations are monitored due to the endothermic nature of DRM reactions and some isolation issues, leading to gaps in some instances. According to the available dataset [52], all features employed by the process model are approximately in a similar range, and no feature scaling is required.

3.3 Random forest

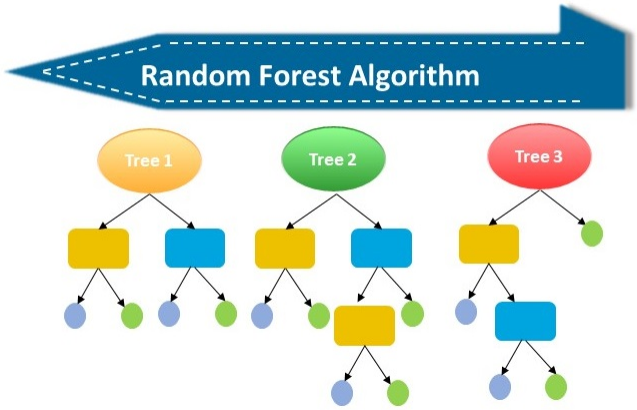


Figure 3.2: The Random Forest (RF) scheme.

A decision tree algorithm represents the predictor as a hierarchical and interpretable model. It arrives at a decision by recursively partitioning the dataset into subsets based on the features that optimally separate the data at each node. This recursive process continues until predefined stopping criteria are met, such as reaching a maximum tree depth or a minimum number of samples per leaf node [11, pp. 195-199]. Random forest (RF) employs an ensemble learning method in which multiple decision tree predictors are combined using either the bagging or, in some cases, the pasting technique [53]. Unlike traditional decision trees, which exhaustively explore all available features, RF selectively considers only a random subset of features when

searching for the optimal split at each node. This selective exploration not only accelerates the tree-growing process but also enhances the algorithm's robustness and its capacity to handle high-dimensional datasets. Furthermore, the ensemble approach fortifies RF against both overfitting and underfitting, particularly when dealing with minor outliers within the dataset. RF's adaptability in accommodating both continuous and categorical features, with minimal data preparation requirements, establishes it as a formidable machine learning algorithm for regression and classification tasks [11, p. 177].

3.4 Pure deep learning

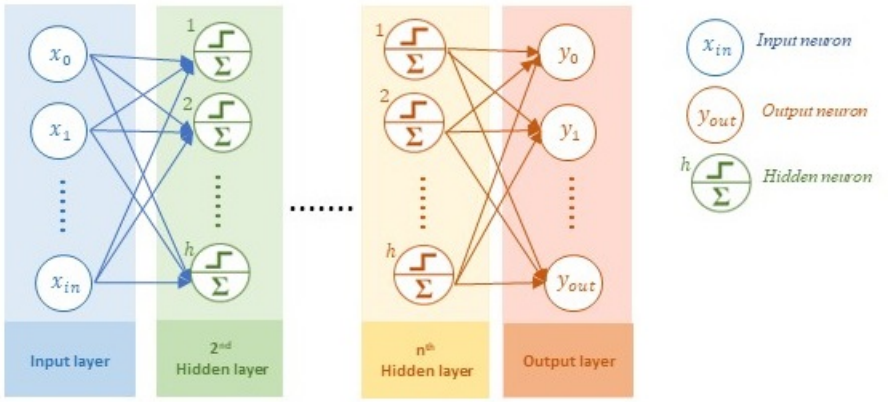


Figure 3.3: The Deep Neural Networks (DNN) scheme.

Deep learning models represent an extended version of artificial neural networks (ANNs) exclusively developed to tackle complex problems and handle extensive datasets. The foundational concept of ANN architecture, namely perceptron, draws its inspiration from neuroscience. It involves a weighted sum $z = \theta_1 x_1 + \theta_2 x_2 + \ldots + \theta_m x_m$, where $x = (x_1, \ldots, x_m)$ is the vector representing the input variables, $\theta_1, \theta_2, \ldots, \theta_m$ are the corresponding weights for each input variable, m denotes the total number of input variables and n is the number of output variables. This weighted sum is then subjected to a non-linear operation, known as an activation function (ϕ), to produce the output $h_\theta(x) = \phi(z)$. A multi-layer perceptron (MLP) comprises multiple perceptrons stacked together, significantly enhancing the model performance. The input variables, referred to as input neurons, constitute the initial perceptron as the input layer followed by a bias neuron (x_0=1). Subsequent intermediate layers, termed hidden layers, each houses an appropriate number of hidden neurons. Finally, the last perceptron, referred to as the output layer, calculates the output neurons.

Multiple hyperparameters within the ANN model define the learning framework while utilizing the back-propagation algorithm, a well-suited gradient-based algorithm, to minimize the network's prediction error [11, p. 309]. All instances within the training dataset are individually presented to the model, and

prediction occurs in a sequential manner. Meanwhile, the model strengthens the connection weights, helping reduce the network error, whereas it refines the connection weights responsible for incorrect predictions. The term pure deep learning, in this study, specifically refers to the standard deep learning approach.

3.5 Greedy layer-wise pretrained deep learning

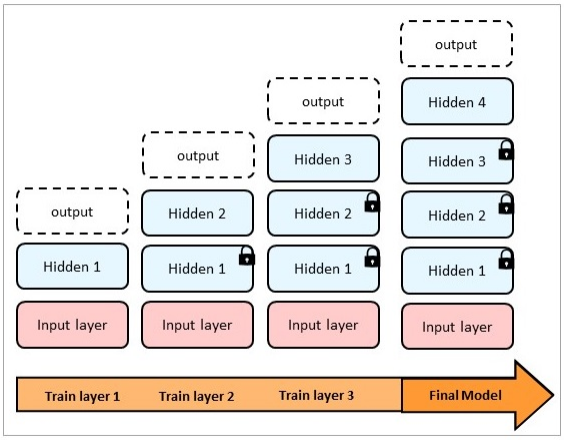


Figure 3.4: The supervised greedy-layer wise pretraining scheme.

Greedy algorithms are typically used to break down complex problems into smaller parts, solving each part independently to achieve an optimal solution. However, combining these individual optimal solutions does not always guarantee an overall optimal solution. Nevertheless, this method is advantageous in terms of saving computational resources. Therefore, instead of simultaneously optimizing all parameters within a neural network, a more efficient approach involves optimizing one hidden layer at a time. This technique, known as supervised greedy layer-wise pretraining [54, p. 323], provides an initial framework for the intermediate layers of a deep model. Moreover, reusing pretrained layers offers the possibility of obtaining a converged model with fewer instances. It is particularly invaluable when data acquisition is time-consuming and costly. As depicted in Figure 3.4 , the training process begins with a primary neural network involving a single hidden layer. Afterward, it drops the output layer and simultaneously locks all previously trained parameters. Then, an additional hidden layer is introduced into the model, while the output layer is returned as the final layer. The training process iterates this procedure multiple times with sequential performance evaluation to ensure meeting the aimed level of performance [55]. The pretrained deep model enhances convergence time and accuracy for the existing dataset compared to the corresponding pure deep model, primarily due to a significantly reduced parameter count (see the results and discussion section for more details).

3.6 Validation and testing the models

An AI-enabled model is evaluated based on its predictive accuracy. Reliable accuracy is presented when the model is evaluated against unseen data samples. It is accomplished by dividing the existing dataset into three distinct segments, as illustrated in Table 5. The largest segment is responsible for training the model, while the validation dataset helps fine-tune the model, and a test dataset simulates real-world scenarios. Moreover, a shuffling method reorders the chronological arrangement of the data samples.

Table 3.3: Dataset Segmentation

Segment	Percentage	Quantity
Training	64%	177
Validation	16%	45
Test	20%	56

This study evaluates the AI-enabled process models for DRM reaction using five distinct metrics including mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), coefficient of determination (R2), mean absolute percentage error (MAPE) as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\widehat{y}_i - y_i\right| \tag{3.6}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{3.7}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{3.8}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{N}(\hat{y}_i - \bar{y})^2} \tag{3.9}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\widehat{y}_i - y_i}{y_i}\right| \times 100 \tag{3.10}$$

where N, y_i, $\hat{y}_i$ and $\bar{y}$ are defined as follows:

- N: number of instances (samples) within the dataset.
- y_i: actual output value (ground truth) for the i^{th} instance within the dataset.
- $\hat{y}_i$: predicted output value for the i^{th} instance within the dataset.

- $\bar{y}$: mean of all output values within the dataset.

MAPE is frequently used to present the percentage error between the predicted and actual values. However, when expressing the accuracy of regression models, it is common to use the complement of MAPE (i.e., 100% - MAPE).

3.7 Results and discussion

3.7.1 Anomaly detection results

The DBSCAN algorithm explores the original dataset using two hyperparameters, namely ε (epsilon) and min_sample, to identify patterns and organize output values into multiple clusters, with the parameters and results listed in Table 3.4.

Table 3.4: DBSCAN algorithm hyperparameters/results

hyperparameters	values
ε	0.5
min_samples	5

results	values
cluster no.	20
outlier no.	153

The time index of outliers is invaluable for understanding their underlying causes. In order to visualize and analyze these outliers, we employ principal component analysis (PCA) to compress all six output values for each observation. Subsequently, we plot the compressed data over time, denoting outliers with red crosses. The visualization in Figure 3.5 provides valuable insights into the distribution of outliers over time, where the first dense region of outliers occurs during the early hours of the process, followed by another dense region emerging after approximately 200 hours.

3.7.2 Exploring anomaly causes

Anomaly detection not only improves the quality of experimental datasets but also facilitates the identification of potential error sources during the data acquisition process. It plays a crucial role in flagging irregularities that may arise when data is collected under non-steady-state conditions, as well as when issues related to catalyst performance and measuring instrument accuracy come into play. As depicted in Figure

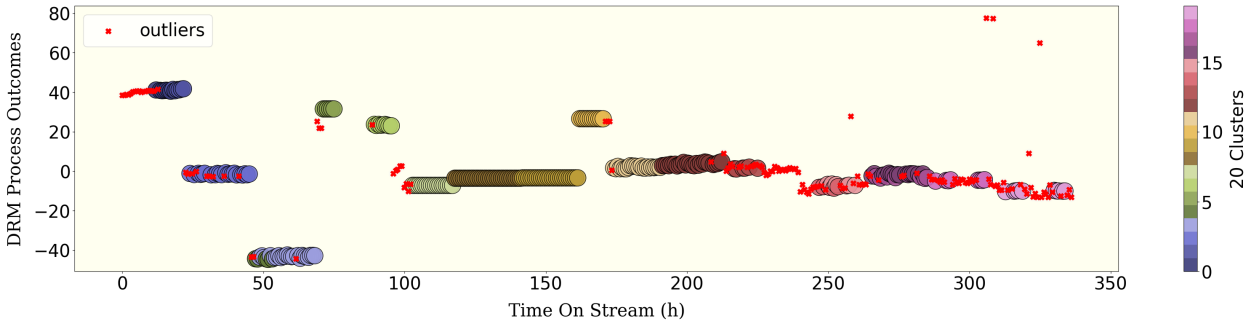


Figure 3.5: Projection of 6-dimensional DRM process outcomes to a 1-dimensional space plotted over time on stream to highlight the time index of outliers.

3.5, the initial hours of the process exhibit a high density of outliers, likely attributable to the unsteady state of the reaction. Subsequently, beyond this period, the process proceeds with minimal discernible errors, extending up to 200 hours. Table 3.5 presents a performance comparison of the random forest model using both the original and error-free datasets, highlighting the origins of the second outlier density. The predicted values, specifically those corresponding to y_1 (Std_Outlet_flow), demonstrate nearly comparable accuracy when utilizing either the original or the error-free dataset. In contrast, the application of the error-free dataset significantly enhances the predictive accuracy of other parameters, i.e., y_2 to y_6. The absence of anomaly for Std_Outlet_flow suggests that the reaction is following the favourable path, making it unlikely that catalyst-related issues arise in the middle of the process. Instead, the root causes are more likely associated with measuring instruments. Notably, the first parameter, Std_Outlet_flow, is measured independently by a flow meter (FM), whereas gas composition parameters y_2 to y_6 are measured using gas chromatograph (GS) and quadrupole mass spectrometer (QMS) instruments, as illustrated in Figure 3.1.

Table 3.5: Random Forest (RF) evaluation comparison

(100-MAPE) for RF model						
Dataset	Std_Outlet_flow (ml/min)	H_2/CO ratio (%)	H_2 yield (%)	CO yield (%)	CH_4 Conversion (%)	CO_2 Conversion (%)
Original dataset	**99.96%**	97.76%	98.08%	97.39%	82.71%	95.48%
Error-free dataset	**99.99%**	99.49%	99.60%	99.75%	98.27%	99.36%

3.7.3 Process model evaluation

3.7.3.1 Identifying hyperparameters for random forest:

We employ a cross-validation method to optimize hyperparameters and enhance model generalization performance. A hyperparameter grid comprising 200 candidates participates in conducting 10-fold cross-validation (known as k-folds CV), trying a total of 2000 model fits [11, pp. 91–93]. Table 3.6 presents the optimal hyperparameters along with their corresponding descriptions.

Table 3.6: Random Forest (RF) hyperparameters

Name	Description	Values
n_estimators	number of trees in the forest	86
min_samples_split	min number of samples placed in a node before the node is split	2
min_samples_leaf	min number of samples allowed in a leaf node	1
max_features	max number of features considered for splitting a node	Sqrt[2]
max_depth	max number of levels in each decision tree	100
bootstrap	sampling method with replacement	True

3.7.3.2 Identifying hyperparameters for deep learning:

In Figures 3.6(a) and 3.6(b), the MSE values for both the training and test datasets are illustrated in the context of proposed deep learning models as they introduce additional hidden layers into their architectures. The initial model, depicted in Figure 3.6(a), is a pure deep neural network, and it reaches its best performance when it incorporates an extensive architecture comprising nine hidden layers. In contrast, the supervised pretrained deep neural network model, presented in Figure 3.6(b), converges to its optimal performance with a more limited architecture of five hidden layers. Beyond this point, any further addition of layers does not yield remarkable performance enhancement. The remaining hyperparameters, as outlined in Table 3.7, are established through heuristic methods. Regarding both models, the empirical evidence suggests exponential linear unit (eLU) in Equation 3.11 as the preferred activation function to handle the non-linear operations for all layers within the deep neural networks [11, pp. 361–367].

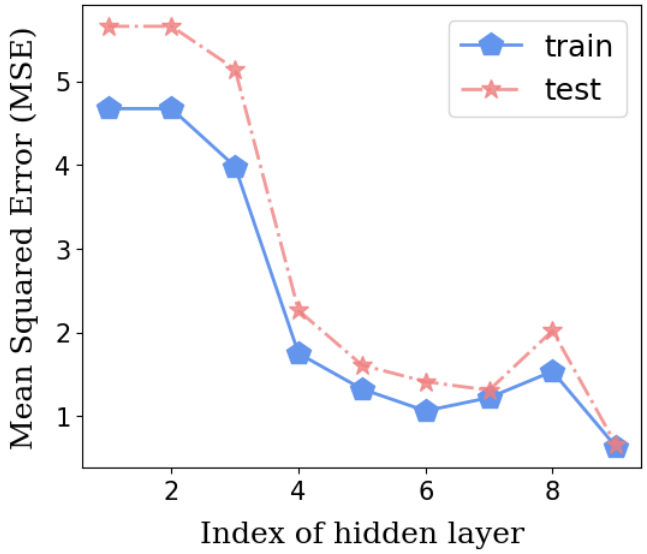
$$\mathrm{eLU}_{\alpha}(z) = \begin{cases} \alpha(e^{z} - 1) & \text{if } z < 0 \\ z & \text{if } z \geq 0 \end{cases} \tag{3.11}$$

where z and α are defined as follows:

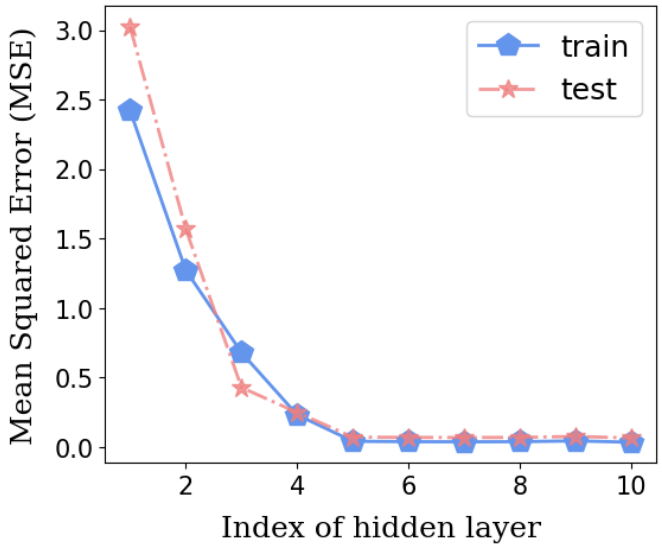
- z: the linear operation consisting of a weighted sum within each layer

[2] square root of the total number of features

- α: a constant controls the smoothness of the function when "z" has considerably negative values and the value of "α" is typically defined as one ($\alpha = 1$).



(a)



(b)

Figure 3.6: Incorporating additional layers versus MSE metric values for test and training dataset. (a) pure deep neural networks model. (b) supervised greedy layer-wise pretrained deep neural networks model.

Table 3.7: Deep learning hyperparameters

	Pure DNN	Pretrained DNN
Activation function	eLU	eLU
Number of hidden layers	9	5
Number of hidden neurons per layer	31	31
Optimizer	Nadam	Nadam
Learning rate (η)	0.001	0.001
Batch size	12	6
Epochs	5000	5000
Total parameters	9,399	5,431
Trainable parameters	9,399	1,184

Momentum, Nesterov Accelerated Gradient (NAG), AdaGrad, RMSProp, and Adam are popular variations of optimizer to adjust weights and bias in a deep neural network. In our research, the deep learning models are constructed using the Nadam optimizer. This powerful optimizer combines the adaptive learning rate feature from Adam with the momentum-adjusted gradient feature of NAG [11, pp. 379–386]. As illustrated in Table 3.7, the pretrained deep learning model demonstrates efficient convergence by utilizing only 1,184 trainable parameters out of a total of 5,431. This streamlined parameter utilization significantly conserves computational resources, which is a crucial aspect of this study.

3.7.3.3 Performance evaluation result

Tables 3.8, 3.9 and 3.10 present the evaluated metrics, respectively demonstrating the predictive accuracy of random forest, pure deep learning model and pretrained deep learning model as the proposed AI-enabled process models corresponded to the existing experimental dataset for the DRM reaction process. Notably, the R-squared values for the random forest model in Table 3.8 and the supervised pretrained deep neural network model in Table 3.10 are identical, both outperforming the pure deep neural network model in Table 3.9, which exhibits lower R-squared values. Random forest and pretrained deep learning models exhibit comparable performance with the existing dataset. However, this study favours pretrained deep learning models due to the advantages of deep learning for software sensors (refer to section 2.3)

Table 3.8: Metrics evaluated the test dataset for the random forest model

Random Forest model						
	Flow-out (ml/min)	H_2/CO ratio (%)	H_2 yield (%)	CO yield (%)	CH_4 Conversion (%)	CO_2 Conversion (%)
MAE	3.69E-02	4.14E-01	7.18E-02	5.14E-02	1.09E-01	1.34E-01
MSE	1.10E-02	4.78E-01	1.00E-02	7.00E-03	3.00E-02	3.70E-02
RMSE	1.05E-01	6.91E-01	1.00E-01	8.40E-02	1.73E-01	1.92E-01
R2	1.00E+00	9.99E-01	9.99E-01	9.99E-01	9.99E-01	9.99E-01
(100-MAPE)	99.99%	99.49%	99.60%	99.75%	98.27%	99.36%

Table 3.9: Metrics evaluated test dataset for the pure deep neural networks model

Pure deep neural networks model						
	Flow-out (ml/min)	H_2/CO ratio (%)	H_2 yield (%)	CO yield (%)	CH_4 Conversion (%)	CO_2 Conversion (%)
MAE	7.01E-01	9.82E-01	2.36E-01	3.15E-01	4.65E-01	3.78E-01
MSE	9.97E-01	2.33E+00	1.06E-01	2.53E-01	4.56E-01	2.94E-01
RMSE	9.98E-01	1.53E+00	3.26E-01	5.03E-01	6.75E-01	5.42E-01
R2	9.95E-01	9.93E-01	9.89E-01	9.78E-01	9.89E-01	9.93E-01
(100-MAPE)	99.71%	98.86%	98.66%	98.37%	91.08%	98.14%

Table 3.10: Metrics evaluated test dataset for the supervised pretrained deep neural networks model

Supervised pretrained deep neural networks model						
	Flow-out (ml/min)	H_2/CO ratio (%)	H_2 yield (%)	CO yield (%)	CH_4 Conversion (%)	CO_2 Conversion (%)
MAE	2.11E-01	4.00E-01	7.80E-02	6.94E-02	1.27E-01	1.36E-01
MSE	7.70E-02	2.59E-01	9.00E-03	1.00E-02	3.10E-02	3.40E-02
RMSE	2.77E-01	5.09E-01	9.50E-02	1.00E-01	1.76E-01	1.84E-01
R2	1.00E+00	9.99E-01	9.99E-01	9.99E-01	9.99E-01	9.99E-01
(100-MAPE)	99.91%	99.53%	99.55%	99.68%	97.77%	99.37%

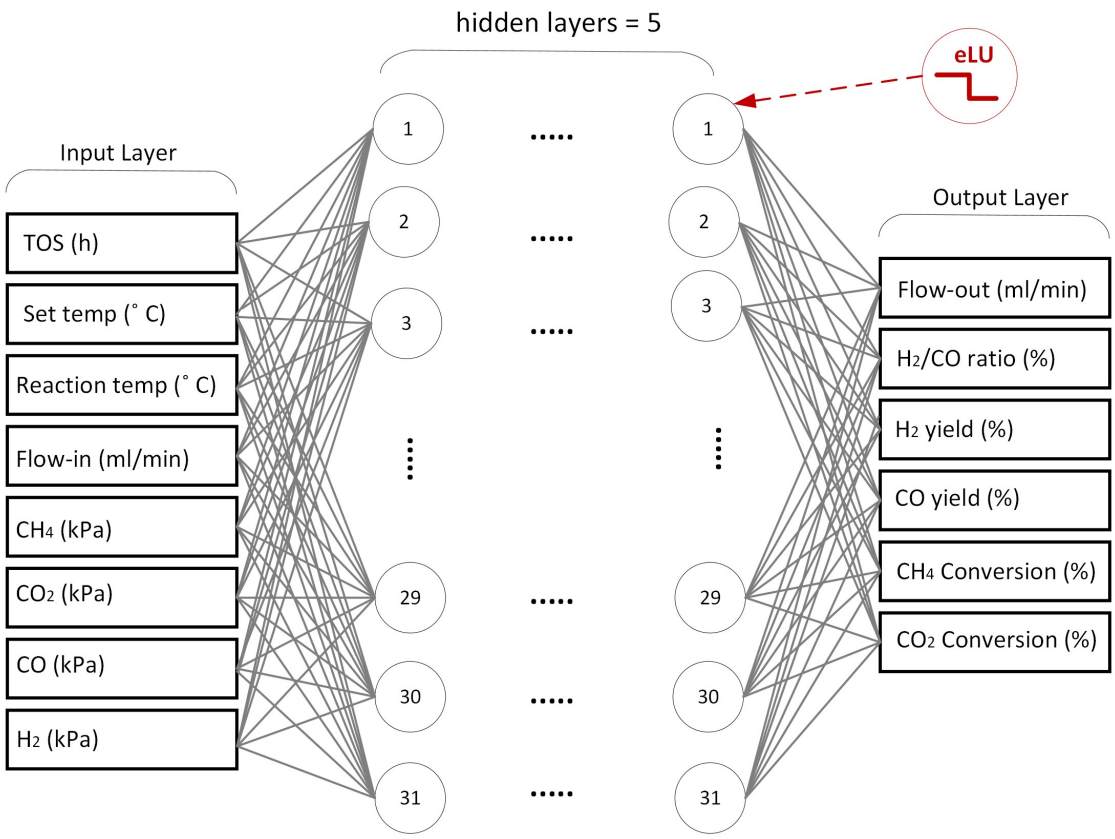
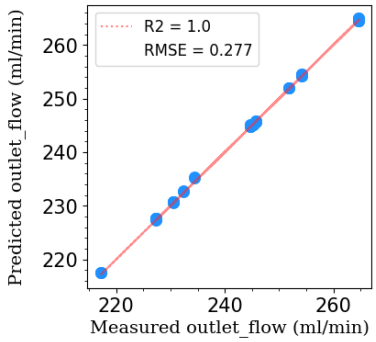
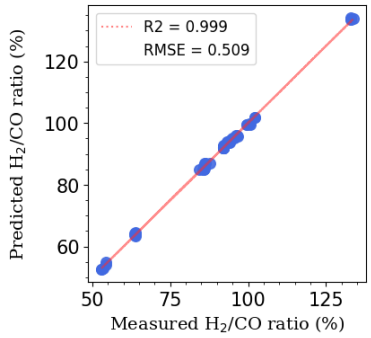


Figure 3.7: The architecture of DNN model for estimating the quality parameters of the DRM reaction using greedy-layer wised pretraining.

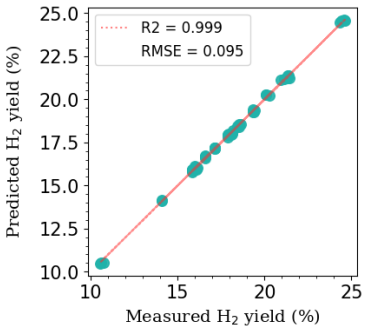
As a result, the superior DRM process model proposed in this research is a deep learning model utilizing the supervised greedy layer-wise pretraining approach. It consists of five hidden layers with 31 hidden neurons for each layer, as shown in Figure 3.7. This pretraining approach explicitly facilitates the training process and reduces the training time through lower parameters while having the outstanding ability to handle small datasets [11, p. 378]. Since R-squared is recognized as a highly informative metric for assessing model accuracy [56], Figure 3.8 presents parity plots illustrating the predicted outcomes of the DRM process model against ground truth values obtained from experiments. These plots correspond to previously unseen data and offer valuable insights into the capabilities of the proposed model.



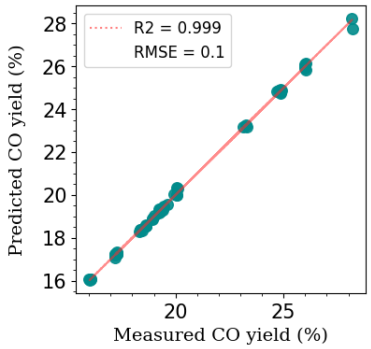
(a) predicted Outlet_flow vs. measured Outlet_flow.



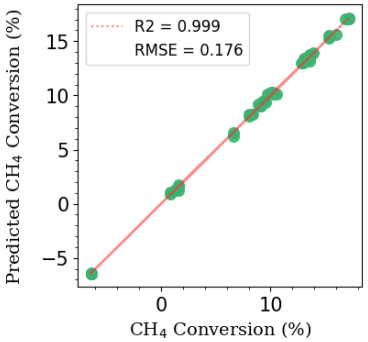
(b) predicted H_2/CO ratio vs. measured H_2/CO ratio.



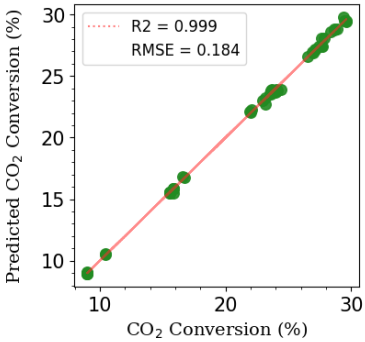
(c) predicted H_2 yield vs. measured H_2 yield.



(d) predicted CO yield vs. measured CO yield.



(e) predicted CH_4 conversion vs. measured CH_4 conversion.



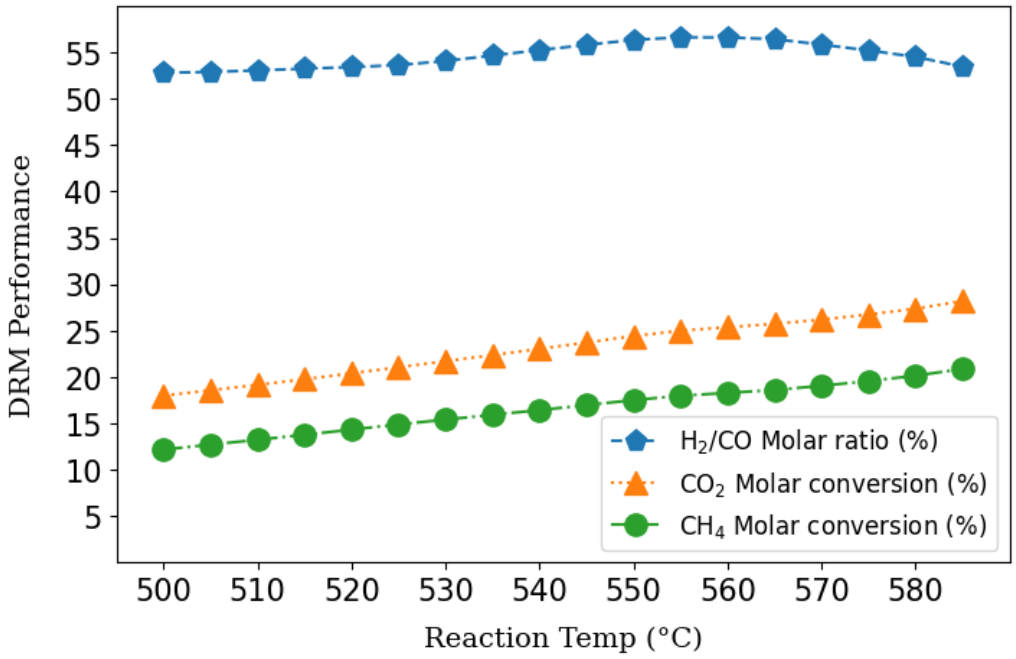
(f) predicted CO_2 conversion vs. measured CO_2 conversion.

Figure 3.8: Parity plots comparing predicted values versus experimental values.

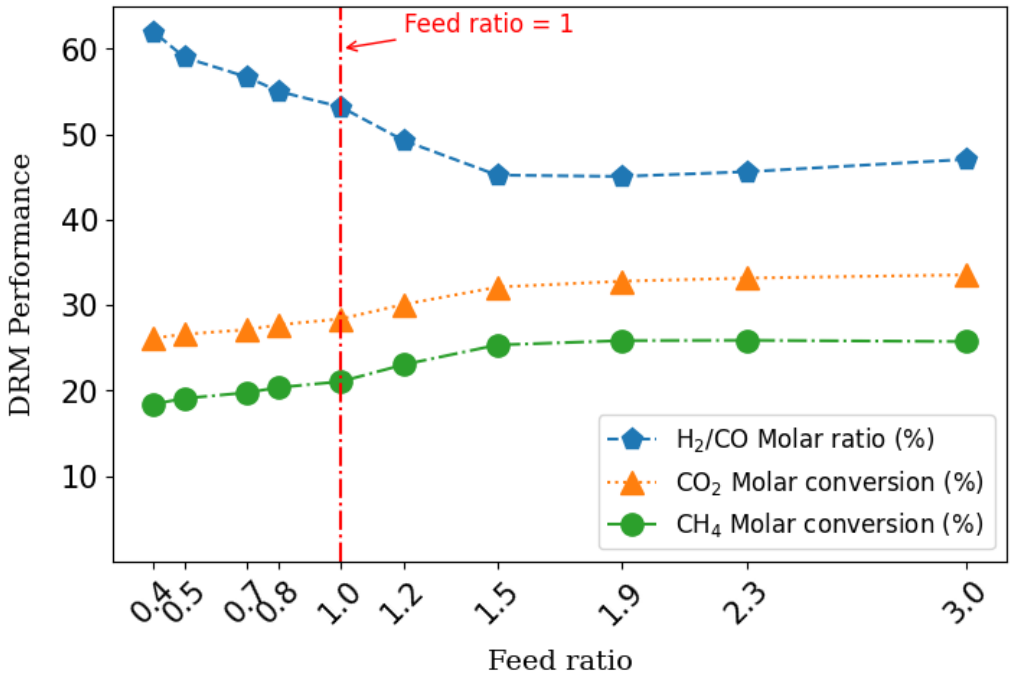
3.7.4 Predicting the impact of temperature and feed ratio on DRM reaction performance

Figure 3.9(a) demonstrates the impact of temperature ranging from 500 °C to 586 °C on DRM performance, as projected by the superior deep learning model. This analysis maintains a consistent feed flow of 210 ml/min and a feed ratio of one, i.e., [CO_2:CH_4]=1 for each temperature point, which is consistent with the existing experimental dataset. The graph exhibits a notable enhancement in conversion rates with rising temperatures, while the syngas ratio displays a gradual fluctuation of 52% to 56%. DRM is characterized as a highly endothermic reaction, yielding favourable results at temperatures exceeding 900°C. Furthermore, elevating the temperature in the DRM reaction with an activation energy greater than zero (Ea > 0), according to the Arrhenius statement [26, p. 14] accelerates the reaction rate. As a result, higher molar conversion of CH_4 and CO_2 are expected [37]. Previous works [57], [58] and [25] also presented enhancement in conversion rates with rising temperatures, irrespective of catalyst variety and quantity. A feed ratio of one, denoted as [CO_2:CH_4] = 1, theoretically results in identical molar conversions of CH_4 and CO_2 and a syngas ratio of one, denoted as [H_2:CO] = 1, in accordance with the stoichiometric equation of the DRM reaction ($CH_4 + CO_2 \rightleftharpoons 2CO + 2\,H_2$). However, in practice, the reaction deviates from the primary path through side reactions. Additionally, byproducts impact the catalyst activity, leading to unequal conversion rates and a less than one syngas ratio [17].

In Figure 3.9(b), the superior deep learning model provides an illustration of how the feed ratio [CO_2:CH_4], ranging from 0.4 to 3.0, influences the DRM reaction performance. This analysis maintains a consistent feed flow of 210 ml/min and temperature at 586 °C for each feed ratio point, which is consistent with the parameters within the experimental dataset. Furthermore, the vertical dashed red line marks the point where the input gas composition is balanced, signifying [CO_2:CH_4] = 1. When the feed ratio is below the balance, i.e., [CO_2:CH_4] < 1, the model predicts a high syngas ratio with a low conversion rate. Conversely, a lower syngas ratio and higher conversion rate are expected above the balance. However, once the feed ratio surpasses 1.5, the figure indicates no further improvement in conversion rate and significant change in syngas ratio for [CO_2:CH_4] > 1.5. The feed ratio directly influences the rate of side reaction interference. [25] comprehensively studied the influence of feed ratio on DRM performance. According to their finding, the feed ratio lower than the balance elicits carbon deposition due to CH_4-rich condition and above the balance has the potential for water formation due to RWGS reaction.



(a) Influences of reaction temperature (500-586 °C)



(b) Influence of feed ratio [CO_2:CH_4]

Figure 3.9: Quality parameters of the process predicted by deep learning model for DRM reaction with constant feed flow of 210 ml/min.

3.8 Chapter summary

This chapter studies three AI-enabled process models for the DRM reaction, ultimately proposing a deep learning model with the supervised greedy layer-wise pretraining approach as the superior process model. The study assesses the performance of all introduced process models and notably highlights the enhanced capabilities of this particular approach. The DRM reaction typically serves as an initial stage for downstream processes while investigating the quality of the final product, likely leading to the creation of a visual dataset. Therefore, deep learning models offer a distinct advantage as they can handle a mixed numerical and visual dataset. They also provide reusing capabilities through transfer learning, making them particularly valuable when the process model requires extending or alternating the input or output variables. The chapter further explores the impact of temperature and feed ratio variations using the proposed deep learning model and compares the findings with results obtained from related experimental works.

Chapter 4

Optimal Syngas Production via Integrated Deep Neural Network (DNN) and Genetic Algorithm (GA)

4.1 Introduction

A software sensor using deep learning has the ability to estimate the quality parameters of the DRM reaction. However, we must adjust the input parameters to achieve the optimal process. Figure 4.1 illustrates how an optimizer fine-tunes the input parameters in an iterative procedure based on the history of output observations. The optimizer employs the AI-enabled model as an objective function. Meanwhile, it evaluates the feasibility of the suggested solutions using the constraints. The optimizer continues to meet the specified conditions or reach the maximum iteration. Deep learning models are inherently black box models [11, p. 131], and analytically optimizing the objective functions using a deep learning model is impossible. In contrast, evolutionary algorithms (EA) have the ability to stochastically and heuristically explore the entire search space and employ machine learning models as an objective function. They quickly obtain optimal parameters, often suboptimal but close to the global optimum. Evolutionary algorithms (EAs) commonly follow a population-based approach, beginning with an initial set of randomly generated input variables as the first population. This population is then assessed against an objective function, also known as a fitness function. Subsequently, the population is iteratively improved to either maximize or minimize the objective function within specified iterations or to meet predetermined objective criteria [59, pp. 3–6]. EAs categorize

the search space of an optimization problem into feasible and infeasible solutions imposed by constraints. These constraints often incorporate into the objective function in a format of penalty functions [59, pp. 8–9]. The evolutionary algorithms (EAs) enhance their solutions by diversifying their exploration of various areas within the search space and intensifying their focus on the most promising solutions identified thus far [59, p. 11].

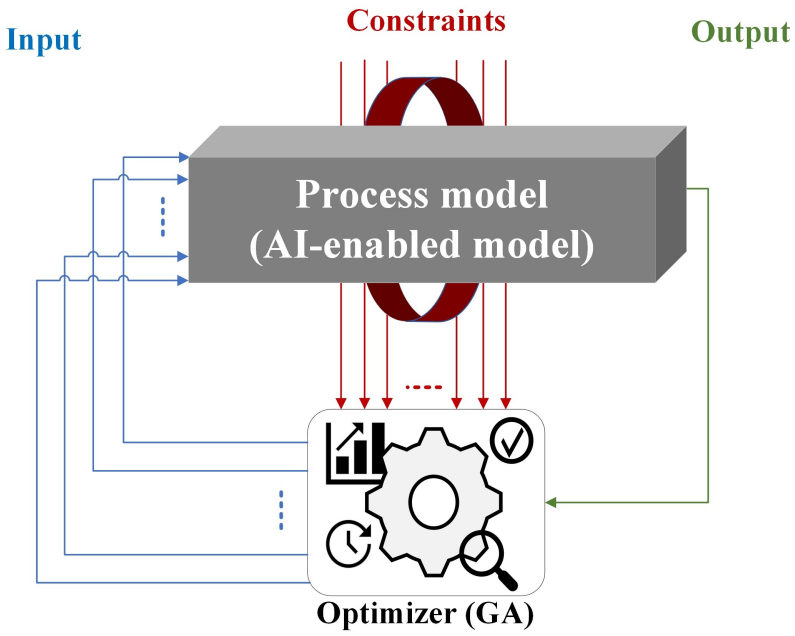


Figure 4.1: An optimizer scheme.

This study specifically employs a genetic algorithm (GA) as an evolutionary algorithm to determine the optimal reaction temperature and feed gas composition to reach an efficient syngas production via DRM reaction. Dry reforming of methane is a challenging process that exposes varied outcomes at various operational conditions. Specifically, different temperature ranges, pressure ranges and catalyst quality significantly change the reaction mechanism (refer to section 2.2.3). Furthermore, the reactor design [29] remarkably influences the results. For example, in a temperature range between 500-600 °C, which is the studied range of our primary dataset, RWGS reaction is unavoidable due to the close thermodynamic equilibrium of DRM and RWGS (refer to section 2.2.2). However, at higher temperatures, 700-800 °C DRM competes with another category of side reactions. Due to this variety, researchers typically study this reaction in a specific assumption. However, their research also presents totally different experimental or simulated datasets based on their assumptions. While mechanism-driven models require a new kinetics model for each tiny change in assumption, the AI-enable models as a data-driven approach only require retraining the models, which is

commonly possible by retransfer learning in the context of deep learning models. In some cases, manipulating the objective function is also required depending on the dataset and objectives. An optimization problem can be either a multi-objective function when the objectives are non-dominated or a simple single-objective function when there is no trade-off among objectives, and they intensify each other. Integrating GA with an AI-enabled process estimator offers outstanding versatility to meet the variety of outcomes resulting from various reaction mechanisms and optimize the process for various catalysts, temperature ranges and other operating conditions.

4.2 An optimization problem for syngas production via DRM over Ni/Al_2O_3 catalyst

The primary dataset of this study comes from dry reforming of methane over Ni/Al_2O_3 catalyst in a temperature range of 500-586 °C shown with other assumed parameter ranges in Table 3.2. The details of the catalyst synthesis are out of the scope of this thesis. However, different catalyst compositions and their corresponding characterizations significantly influence the reaction mechanism and raise a distinct optimization problem.

4.2.1 DNN process model

The experimental dataset is exposed to an instance of deep neural networks, while the deep learning model copes with the smallness of the dataset through pretraining. The deployed model works as a regressor model to estimate the quality parameters of the process. In Chapter 3, we comprehensively discussed the predictive accuracy of the model with relative input and output variables. In this chapter, the deployed model serves as an objective function and combines with a genetic algorithm for optimal syngas production via DRM reaction.

4.2.2 DNN-GA process optimization

Genetic algorithm (GA) is a population-based evolutionary algorithm inspired by Charles Darwin's hypothesis "survival of the fittest" [60, p. 36]. It is a heuristic optimization algorithm to solve complex problems where the fittest solutions in each generation, namely selected individuals, are improved for producing the next generation. In this study, a combination of a genetic algorithm and a deep learning model works as an optimizer for the dry reforming of methane (DRM) process. Considering the experimental dataset, two operating variables of the process, i.e., feed gas composition [CO_2:CH_4] and temperature (°C), which significantly

influence DRM performance, are optimized to reach the maximum syngas molar ratio and molar conversion simultaneously. The optimizer considers the problem's search space close to the dataset with which the deep learning model was previously trained. Considering this search space, an independent evaluation of each of these objectives, i.e., maximum syngas molar ratio and maximum molar conversion, shows that these objectives are in conflict with each other, which means that we can reach the maximum syngas ratio at the expense of the lower conversion and vice versa. A problem is considered a multi-objective optimization problem (MOOP) only when objectives are in conflict with each other [60, p. 16]. We first tackle the problem as a single objective optimization problem to show the conflict between objectives. Secondly, consider the problem as a multi-objective problem, presenting the correct version of the solution. The results and discussion section (4.4) discusses more details and evidence relative to these non-dominated objectives. In the following, we delve into optimizing the DRM process in the assumed search space as a multi-objective problem.

4.2.2.1 The objective functions

MOOP with n objectives is expressed with an aggregated fitness function $f(\mathbf{x})$ that invariably entails a trade-off between these objectives, denoted in Equation 4.1.

$$max f(\mathbf{x}) = (f_1(\mathbf{x}),\ f_2(\mathbf{x}),\ \ldots,\ f_n(\mathbf{x})) \tag{4.1}$$

Where

- $f_i(\mathbf{x})$ is the fitness function of the i^{th} objective, and $i \in \{1, n\}$.
- $\mathbf{x}$ is the vector of decision variables (i.e., the input variables for a deep learning model), defined as an m-dimensional vector $\mathbf{x} = (x_1, x_2, \ldots, x_m)$, where m is the number of decision variables.
- The optimal solution is a vector $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_m^*)$ that yields the maximum objective function $f(\mathbf{x}^*) = (f_1(\mathbf{x}^*), f_2(\mathbf{x}^*), \ldots, f_n(\mathbf{x}^*))$.

In order to maximize $f(\mathbf{x})$, the MOOP searches a continuous domain $x_{(j)}$ for the j^{th} element of $\mathbf{x}$, restricted as follows:

$$x_{(j)} \in \left[x_{(j)_{min}}, x_{(j)_{max}}\right],\quad for\ all\quad j \in \{1, m\} \tag{4.2}$$

A deep learning model as a regressor, denoted by $y = h_{\boldsymbol{\theta}}(\mathbf{x})$, predicts the quality parameters of the DRM reaction, where it takes an 8-dimensional vector $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8)$ as the operating conditions of the process and calculates the process efficiency, represented by a 6-dimensional vector $\mathbf{y} = (y_1, y_2, y_3, y_4, y_5, y_6)$ according to the Table 3.1 in section 3.2.5. The objective functions for the optimization problem are described as follows:

$$f_1(\mathbf{x}) = f_1(h_\theta(\mathbf{x})) = y_2 \tag{4.3}$$

$$f_2(\mathbf{x}) = f_2(h_\theta(\mathbf{x})) = y_5 + y_6 \tag{4.4}$$

$$f_3(\mathbf{x}) = f_3(h_\theta(\mathbf{x})) = \left(\frac{y_1}{x_4}\right) \tag{4.5}$$

Where

- $f_1(\mathbf{x})$ calculates the syngas molar ratio.
- $f_2(\mathbf{x})$ calculates the molar conversion.
- The amount of carbon deposited is a significant objective in estimating the efficiency of the DRM reaction. When it is unavailable, the ratio of "Std_Outlet_flow" over "Std_Inlet_flow" defined by $f_3(\mathbf{x})$ is an essential component of our MOOP. We can restrict mass production, i.e., coke or water, by maximizing the $f_3(\mathbf{x})$.

A deep learning model takes eight input variables into account, according to Table 3.1 in section 3.2.5. However, we assume that the "set temperature" is equivalent to the "reaction temperature", the "time on stream (TOS)" is set to a value of "0", and neither external H_2 nor external CO gases are fed to the reactor to simplify the problem. Additionally, the "Std_Inlet_flow" is determined by the quantity of fed gasses, i.e., x_5 and x_6, unit transformation also applied to x_5 and x_6 according to the preprocessing and scaling in sections 3.2.2 and 3.2.3. Therefore, the search space of our optimization problem, which includes three major non-zero decision variables- reaction temperature (°C), CH_4 feed flow (ml/min), and CO_2 feed flow (ml/min), comprises random values uniformly distributed between $x_{(j)_{\min}}$ and $x_{(j)_{\max}}$, according to Equation 4.6 as follows:

$$
\begin{cases}
\text{Time on stream (h)} & \longrightarrow\ x_1 \in [0,0] \\
\text{Set temperature (°C)} & \longrightarrow\ x_2 = x_3 \\
\text{Reaction temperature (°C)} & \longrightarrow\ x_3 \in [500,586] \\
\text{Std\_Inlet\_flow (ml/min)} & \longrightarrow\ x_4 \text{ as a function of } x_5 \text{ and } x_6 \\
CH_4 \text{ feed flow (ml/min)} & \longrightarrow\ x_5 \in [50,149] \\
CO_2 \text{ feed flow (ml/min)} & \longrightarrow\ x_6 \in [50,149] \\
\text{CO (ml/min)} & \longrightarrow\ x_7 \in [0,0] \\
H_2 \text{ (ml/min)} & \longrightarrow\ x_8 \in [0,0]
\end{cases}
\tag{4.6}
$$

In order to achieve the desired result in the assumed search space, the DRM reaction process should adhere to specific constraints as follows:

- A syngas molar ratio of between 50% and 100%.
- An outlet-to-inlet flow ratio of between 1 and 2.
- A conversion gap of lower than 10 for these two reactants.

We discuss more details of these boundaries in the results and discussion section. The pseudocode implementing the constraint-handling technique (CHT) [61] for our MOOP through a penalty function is as follows:

Algorithm 1: Constraint handling technique.

Input : Values of y_1, y_2, y_5, y_6, x_4
Output: $f(\mathbf{x})$

```
if y2 < 50 or y2 > 100 then
    f(x) = 0;
end if
if y1/x4 < 1 or y1/x4 > 2 then
    f(x) = 0;
end if
if |y5 − y6| > 10 then
    f(x) = 0;
end if
```

Weighted-based genetic algorithm (WBGA) is a common method to solve the MOOP proposed by Hajela and Lin [62], in which scaled fitness functions $f_i'(\mathbf{x})$ combine through a weighted sum as shown in Equation 4.7 and ultimately result in a single-objective problem prior to optimization. The weighting coefficients

control the contribution of individual fitness functions based on their relative importance [63]. In essence, the MOOP is converted to a single objective optimization problem (SOOP) as follows:

$$\max f(\mathbf{x}) = w_1\ f'_1(\mathbf{x}) + \ w_2\ f'_2(\mathbf{x}) + \ldots + \ w_n\ f'_n(\mathbf{x}) \tag{4.7}$$

Where

- $f'_i(\mathbf{x})$ is the scaled fitness function for $f_i(\mathbf{x})$.
- w_i is the weighting coefficient of the i^{th} objective function, $i = 1, 2, \ldots, n$.
- $\sum_{i=1}^{n} w_i = 1$

The relative importance for $f_1(\mathbf{x})$, $f_2(\mathbf{x})$ and $f_3(\mathbf{x})$ in Equation 4.3, 4.4 and 4.5 yields to weighting coefficients w = 0.04, 0.48, 0.48 where w = (w_1, w_2, w_3) according to the following preference model in Equation 4.8:

$$\begin{cases} f_1(\mathbf{x}) \ll\ f_2(\mathbf{x}),\ where\ f_1(\mathbf{x})\ is\ much\ less\ important\ than\ f_2(x) \\ f_2(\mathbf{x}) \approx\ f_3(\mathbf{x}), where\ f_2(\mathbf{x})\ and\ f_3(\mathbf{x})\ \ are\ equally\ importamt \end{cases} \tag{4.8}$$

Since the magnitude of each fitness function is different, the WBGA applies scaled function $f'_i(\mathbf{x})$ in Equation 4.9 as follows:

$$f'_i(\mathbf{x}) = \ \lambda_i\ .f_i(\mathbf{x})\ \ ,\ where\ \lambda_i\ is\ the\ scaling\ factor\ for\ f_i(\mathbf{x}) \tag{4.9}$$

The acceptable range of values for functions $f_1(\mathbf{x})$, $f_2(\mathbf{x})$ and $f_3(\mathbf{x})$ respectively related to Equations 4.3, 4.4 and 4.5 are as follows:

- $0 < f_1(\mathbf{x}) < 100$:

  According to the available experimental dataset, without introducing external H_2 to the reactor, H_2 selectivity is consistently lower than CO selectivity; however, a near-unity ratio, denoted as 100%, is possible.

- $0 < f_2(\mathbf{x}) < 100$:

  The thermodynamic equilibrium conversion[1] of CH_4 and CO_2 at a pressure of 1 atm, a temperature range of 500-600 °C, and an equal gas composition [CO_2:CH_4] = 1, assumed

[1]Calculated by Gibbs free energy minimization algorithm on HSC Chemistry 7.1 software

for no carbon formation, cannot exceed 50%, [17] and [22]. As a result, the cumulative conversion of both reactants, i.e., $f_2(\mathbf{x})$ remains within the range [0, 100]. The available dataset follows the aforementioned concept.

- $0 < f_3(\mathbf{x}) < 2$:

  According to the stoichiometric equation of the DRM reaction (refer to section 2.2.2), two moles of the reactants convert to four moles of the products. Assuming the process produces no byproduct, and all reactants fed to the reactor directly convert to products through the DRM reaction, the outlet flow would be twice the inlet flow.

The scaled fitness function $f_i'(\mathbf{x})$, which assumes values approximately ranging from 0 to 100 for $f_1(\mathbf{x})$, $f_2(\mathbf{x})$ and $f_3(\mathbf{x})$, respectively represented in Equations 4.3, 4.4 and 4.5, is subjected to a scaling factor as illustrated in Equation 4.10:

$$\begin{cases} f_1'(\mathbf{x}) = f_1(\mathbf{x}), where\ \lambda_1 = 1 \\ f_2'(\mathbf{x}) = f_2(\mathbf{x}), where\ \lambda_2 = 1 \\ f_3'(\mathbf{x}) = f_3(\mathbf{x}) \times 50, where\ \lambda_3 = 50 \end{cases} \tag{4.10}$$

The SOOP fitness function $f(\mathbf{x})$ is given as follows:

$$f(\mathbf{x}) = f(h_\theta(\mathbf{x})) = (y_2) \times 0.04 + (y_5 + y_6) \times 0.48 + \left(\left(\frac{y_1}{x_4}\right) \times 50\right) \times 0.48 \tag{4.11}$$

4.2.2.2 The genetic algorithm implementation

The flowchart in Figure 4.2 depicts the steps during the genetic algorithm (GA):

1. Step 1. Randomly initialize the first population of the solutions consisting of M=10 solutions: $S_k \quad for\ k \in \{1, M\}$, $where\ M = 10$

2. Step 2. Calculate the objective functions $f(\mathbf{x})$ in Equation 4.11, where n = 3. This step proceeds by scoring the solutions based on the objective criteria, e.g., maximizing the objective.

3. Step 3. Select the fittest solutions, e.g., two or more solutions with the highest scores and send them to the parent mating pool.

4. Step 4. Generate offspring solutions through crossover and mutation operations on parents within the mating pool. The GA applies crossover, e.g., single-point crossover as an exploiting tool and mutation as an exploring tool to keep the intensity and diversity of the offspring.

5. Step 5. Iterate the above steps 2 to 4 until the maximum number of generations or acceptable results is reached

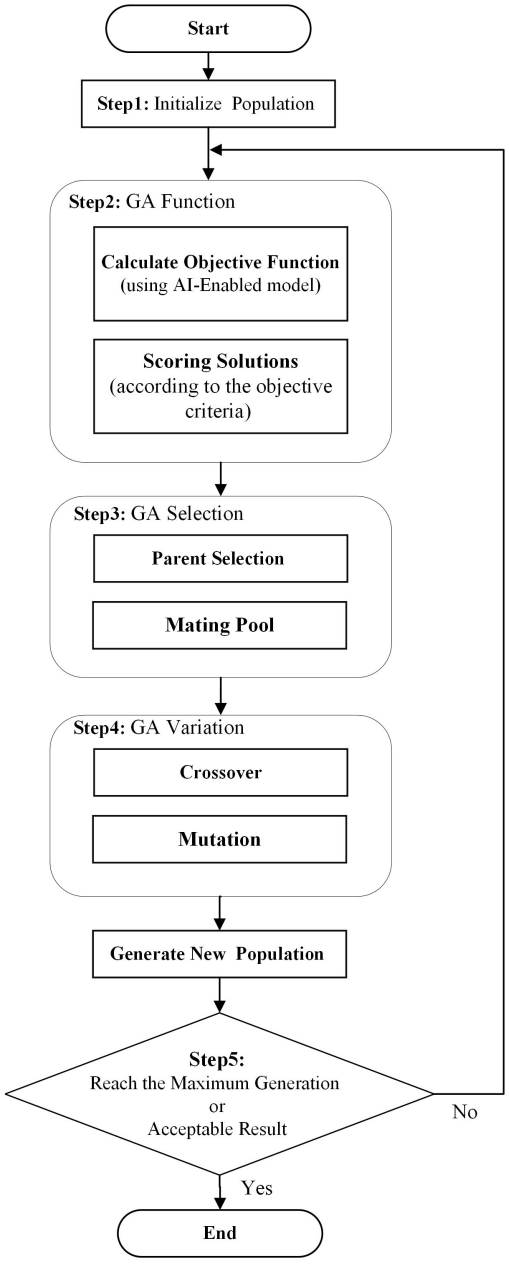


Figure 4.2: The genetic algorithm flowchart.

The size of the parent mating pool (e.g., 4) and the population (e.g., 10) are determined as a function of the genes (i.e., 3) and generation number (e.g., 200). A larger population and parent mating pool result in enhanced diversity of the solutions by exploring a broader search space, whereas it increases the convergence time. In contrast, a smaller population and parent mating pool are essential to speed up the convergence. Plotting the rate of new solutions versus generation and fitness progress versus generation, similar to Figure 4.4, presents valuable insights into choosing appropriate values for those hyperparameters. In order to solve the DRM optimization problem through GA, we employ the PyGAD 2.19 Python library [64]. The library applies the GA class in which various attributes, methods, functions, etc., build the genetic algorithm. For example, an instance of the GA class for the DRM optimization problem initializes as follows:

```
ga_instance = pygad.GA(num_generations=200,
                       sol_per_pop=10,
                       num_parents_mating=4,
                       num_genes=3,
                       gene_space=[
                           np.arange(500, 587).tolist(),
                           np.arange(50, 150).tolist(),
                           np.arange(50, 150).tolist(),
                       ],
                       gene_type=[int,int,int],
                       mutation_type="adaptive",
                       mutation_percent_genes=[50, 20],
                       fitness_func=fitness_function)
```

where

- num_generations: The number of iterations (i.e., generations) the GA goes through to find the best solution.
- sol_per_pop: The number of solutions (i.e., individuals) within the population for each iteration of the GA.
- num_genes: The number of decision variables (i.e., genes) that form each solution within the population.
- num_parents_mating: The number of selected solutions (i.e., mating pool size) that contribute to the next generation of population.

- gene_space: This parameter receives the search space of the decision variables as an array.
- gene_type: This parameter assigns the data type of each decision variable.
- mutation_type: This parameter assigns the mutation operation, which defaults to "random". Other types include, for example, "adaptive".
- fitness_func: This parameter accepts the fitness function $f(\mathbf{x})$.
- mutation_percent_genes: In general GA, a constant value is assigned to a mutation probability, and all solutions have an equal likelihood of mutation. Marsili et al. [65] introduced a more efficient search by adaptive mutation in which mutation probability is a function of fitness. Adaptive GA increases the mutation probability of low-quality solutions while decreasing this probability for high-quality solutions. This approach reduces the chance of disrupting a high fitness solution and enhances the exploratory role of a low-fitness solution. PyGAD implements this strategy by utilizing the mutation_type attribute in the GA class. It adjusts the mutation rate based on the quality of solutions, employing a lower rate (e.g., 20) for solutions of high quality and a higher rate (e.g., 50) for solutions of lower quality, according to the following pseudo-code:

Algorithm 2: Adaptive mutation strategy.

Input : All solutions in each population
Output: Mutation rates for each solution

Compute fitness average, $f(\mathbf{x})_{\text{avg}}$ regarding the population;
for *each solution in population* **do**
    Compute fitness $f(\mathbf{x})$ for the individual;
    if $f(\mathbf{x}) < f(\mathbf{x})_{avg}$ **then**
        Set the mutation rate to high;
    end if
    else if $f(\mathbf{x}) \geq f(\mathbf{x})_{avg}$ **then**
        Set the mutation rate to low;
    end if
end for

4.3 An optimization problem for syngas production via DRM over $Ni/CaFe_2O_4$ catalyst

Catalytic reactions can reveal various mechanisms caused by changes in the chemical and physical characteristics of the catalyst. Additionally, each catalyst composition can tolerate a different range of thermodynamic parameters. As discussed in section 2.2.2, the reaction temperature plays a crucial role in the DRM reaction process, significantly impacting the conversion rates and syngas ratios. Thus, scientists often conduct separate studies of the DRM reaction using various catalysts and different temperature ranges, leading to the publication of various datasets. In order to justify the use of ANN-GA optimization for the DRM reaction, this section focuses on optimizing syngas production through the DRM reaction over a different catalyst composition, namely $Ni/CaFe_2O_4$, developed and published by Hossain et al. [5]. The flexibility of ANN-GA allows for the optimization of varied datasets stemming from different DRM reaction mechanisms. The parameters and objectives for optimization can also be adjusted.

4.3.1 Reproducing the existing ANN process model

This section discusses the reproduction of an artificial neural network (ANN) model, originally presented by Hossain et al. [5] for the DRM reaction over a $Ni/CaFe_2O_4$ catalyst. In the cited study, the researchers maintained a consistent inlet flow of 100 ml/min and a gas hourly space velocity (GHSV) of 30,000 h^{-1} STP[2] for all the samples. However, they evaluated the catalyst's activity with different metal loadings: 5 wt%, 10 wt%, and 15 wt%, with a feed ratio [CH_4:CO_2] of 0.4, 0.7, and 1.00, at reaction temperatures of 700 °C, 750 °C, and 800 °C under atmospheric conditions. These systematic changes provided a dataset comprising 27 samples. Table 4.1 provides a comprehensive overview of the dataset, presenting the independent and dependent variables. They partitioned the dataset, allocating 70% of the data for training purposes and reserving the remaining 30% for testing and validation, distributing this portion equally between testing and validation. Following multiple training iterations, they identified 12 hidden neurons as the optimized configuration for the ANN model with a single hidden layer. It is noteworthy that they applied the sigmoid function in Equation 4.12 as the activation function.

$$\sigma(z) = \frac{1}{1 + e^{-z}} \tag{4.12}$$

[2]Standard temperature and pressure

Table 4.1: Dataset Structure [5]

x	(Independent variables)	Description
x_1	Feed ratio	gas composition fed to the reactor [CH_4:CO_2]
x_2	Reaction temp (°C)	reaction temperature
x_3	Metal loading (%)	(weight of metal/total weight of catalyst) $\times$ 100%
y	(Dependent variables)	Description
y_1	CH_4 Conversion (%)	reactant conversion rate
y_2	CO_2 Conversion (%)	reactant conversion rate
y_3	H_2 yield (%)	hydrogen moles in relation to reactants
y_4	CO yield (%)	carbon monoxide moles in relation to reactants

Since the sigmoid activation function saturates at 0 or 1 [11, p. 358] outputting values between 0 and 1, the four dependent variables outlined in Table 4.1 are presented to the ANN model in their raw values rather than as percentages, as they originally fall within the range of 0 to 1. Furthermore, the temperature values (x_2), which are originally between 700 and 800, are normalized between 0.7 and 0.8 to align with the range of other features. Figure 4.3 presents the architecture of the ANN model as identified by the authors.

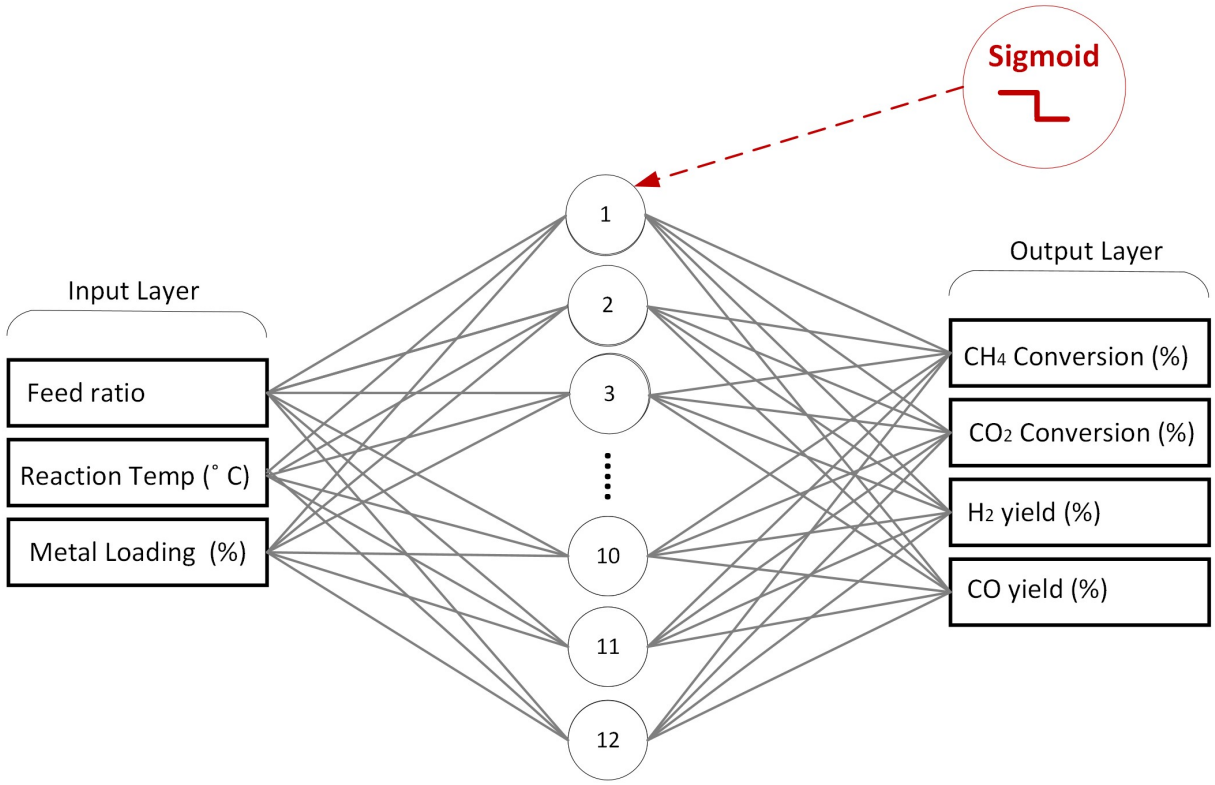


Figure 4.3: The architecture of ANN model for estimating the quality parameters of the DRM reaction over $Ni/CaFe_2O_4$ catalyst [5].

As discussed in section 3.4, a multi-layer perceptron (MLP) performs a back-propagation algorithm to fine-tune the connection weights aimed at reducing the prediction error. The authors did not identify the remaining hyperparameters within the paper. However, the training regulations, which encompass the batch size, optimizer, and the number of epochs, in addition to those mentioned earlier, are detailed in Table 4.2.

Table 4.2: The ANN model hyperparameters

	ANN
Activation function	Sigmoid
Number of hidden layers	1
Number of hidden neurons per layer	12
Optimizer	Nadam
Learning rate (η)	0.001
Batch size	10
Epochs	5000
Total parameters	100
Trainable parameters	100

The predictive accuracy of the ANN model as outlined in Table 4.3 is assessed for each dependent variable listed in Table 4.1 using the metrics discussed in Section 3.6 . These assessments are performed when the model predicts unseen samples in the test dataset to demonstrate its generalization capability.

Table 4.3: Metrics evaluated test dataset for the artificial neural networks model

Artificial neural networks model				
	CH_4 Conversion (%)	CO_2 Conversion (%)	H_2 yield (%)	CO yield (%)
MAE	6.39E-02	7.39E-02	3.80E-02	4.00E-02
MSE	5.00E-03	7.00E-03	3.00E-03	2.00E-03
RMSE	7.10E-02	8.40E-02	5.50E-02	4.50E-02
R2	7.75E-01	7.07E-01	5.50E-01	6.70E-01
Accuracy (100-MAPE)	77.63%	68.73%	72.91%	74.75%

4.3.2 ANN-GA process optimization

Similar to Section 4.2, a genetic algorithm employs a derived version of the ANN model[3] discussed in Section 4.3.1 as the objective function. The ANN model as a regressor, denoted by $y = h_{\boldsymbol{\theta}}(\mathbf{x})$, predicts the quality parameters of the DRM reaction, where it takes a 3-dimensional vector $\mathbf{x} = (x_1,\ x_2,\ x_3)$ as the operating conditions of the process and calculates its efficiency represented by a 4-dimensional vector $\mathbf{y} = \big(y_1,\ y_2, y_{3,},\ y_4\big)$ according to dataset structure presented in Table 4.1.

[3]An artificial neural network (ANN) requires at least two hidden layers to name as a deep neural network (DNN) [34].

The objective functions for the optimization problem are described as follows:

$$f_1(\mathbf{x}) = f_1(h_\theta(\mathbf{x})) = y_1 + y_2 \tag{4.13}$$

$$f_2(\mathbf{x}) = f_2(h_\theta(\mathbf{x})) = y_3 + y_4 \tag{4.14}$$

where

- $f_1(\mathbf{x})$ calculates the molar conversion rate.
- $f_2(\mathbf{x})$ calculates the yield of products.

The search space of the optimization problem comprises random values uniformly distributed between $x_{(j)_{\min}}$ and $x_{(j)_{\max}}$, according to Equation 4.15 as follows:

$$\begin{cases} \text{Feed ratio} & \longrightarrow\ x_1 \in [0.1, 1.0] \\ \text{Reaction temperature (°C)} & \longrightarrow\ x_2 \in [700, 800] \\ \text{Metal loading (\%wt)} & \longrightarrow\ x_3 \in [5, 15] \end{cases} \tag{4.15}$$

The optimizer seeks the defined search space for optimal temperature (°C), feed ratio and metal loading (%wt) in the catalyst to meet the maximum values of $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ as the objective functions. The results and discussion section demonstrate that these objectives dominate each other and make a single-objective optimization problem as follows:

$$f(\mathbf{x}) = f(h_\theta(\mathbf{x})) = y_1 + y_2 + y_3 + y_4 \tag{4.16}$$

where

- The optimal solution is a vector $\mathbf{x}^* = \left(x_1^*,\ x_2^*\ ,\ x_3^*\right)$ that similarly yields the maximum objective function for $f_1(\mathbf{x})$, $f_2(\mathbf{x})$ and $f(\mathbf{x})$

The GA is implemented using the open source PyGAD 2.19 Python library similar to Section 4.2.2.2 as follows:

```
ga_instance = pygad.GA(num_generations=50,
                       num_parents_mating=8,
                       sol_per_pop=16,
                       num_genes=3,
                       gene_space=[np.arange(0.1,1.1,0.1).tolist(),
                                   np.arange(0.7,0.80,0.01).tolist(),
                                   np.arange(5,16,1).tolist()
                                   ],
                       gene_type=[float,float,int],
                       mutation_type="adaptive",
                       mutation_percent_genes=[50,20],
                       random_seed=2,
                       fitness_func=fitness_function)
```

As demonstrated in the GA class, a larger parent mating pool (e.g., 8) and population (e.g., 16), compared to the previous problem in Section 4.2.2.2, are utilized to delay rapid convergence. However, due to the lower complexity of the problem, GA converges faster than the previous problem related to Ni/Al_2O_3 catalyst. Figure 4.6 in the results and discussion section details the convergence progress and exploration rate of $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ separately and jointly as $f(\mathbf{x})$.

4.4 Results and discussion

4.4.1 Optimal operating conditions for DRM over Ni/Al_2O_3 catalyst

In this section, we first examine the functions $f_1(\mathbf{x})$, $f_2(\mathbf{x})$ and $f_3(\mathbf{x})$ separately and discuss the validity of their corresponding solutions as well as any conflicts between these objectives. Secondly, we propose a comprehensive multi-objective solution that incorporates all the aforementioned functions.

4.4.1.1 Single-objective optimization problem

4.4.1.1.1 Maximum syngas molar ratio

The optimization of the maximum syngas ratio represents a single-objective problem, with $f_1(\mathbf{x})$ as the fitness function in Equation 4.3. The GA initializes a random population and evolves it over 200 generations. Figure 4.4(a) shows that fitness values progressively improve through these generations. The number of

new solutions explored in each generation, as depicted in Figure 4.4(e), helps determine whether further improvement in $f_1(\mathbf{x})$ is feasible. The GA successfully achieves the maximum syngas ratio by optimizing the values outlined in Table 4.4. The results reveal that a high syngas ratio is attained within the assumed search space, but this comes at the cost of low conversion rates, particularly in methane conversion. In comparison to the feed flow rate illustrated in Figure 3.9 (210 ml/min), the GA recommends a similar feed gas composition [CO_2:CH_4] = 1, but with a higher feed flow rate (298 ml/min). However, this increase in feed flow leads to higher space velocity (refer to section 2.2.2), which negatively affects the DRM conversion and the yield of products, as supported by the findings in [25] and [29]. Furthermore, the research conducted by [66] and [67] demonstrate that DRM, in the presence of Ni (i.e., Ni/Al_2O_3 catalyst), has the potential for CO_2 methanation ($CO_2 + 4\,H_2 \rightleftharpoons CH_4 + 2\,H_2O$), which significantly reduces the CH_4 conversion rate (4.25%). (refer to simultaneous reactions in section 2.2.2)

4.4.1.1.2 Maximum molar conversion rate

The second objective focuses on maximizing the conversion rate without regard to the syngas ratio. Figure 4.4(b) illustrates the evolution of the fitness function $f_2(\mathbf{x})$ over 200 generations, and Figure 4.4(f) displays the count of newly explored solutions for each generation. Table 4.4 presents the optimal values of the input variables and their relative outcomes. Research from [25] has demonstrated the potential of methane decomposition (MD) ($CH_4 \rightleftharpoons C + 2\,H_2$) in methane-rich conditions when a feed ratio is lower than one (i.e., [CO_2:CH_4] < 1), resulting in substantial carbon deposition and a decline in catalyst activity in the initial stages. Our AI-enabled model predicts high conversion rates at the cost of low syngas ratio, as indicated in Table 4.4. However, it is essential to note that the model does not offer a detailed understanding of side reactions because the dataset lacks information on carbon deposition or water formation. Therefore, the ratio of outlet-to-inlet flow presented as $f_3(\mathbf{x})$, essentially defines the boundaries of the DRM problem in the optimization, and it must be included as the third fitness function in Equation 4.11. (refer to simultaneous reactions in section 2.2.2)

4.4.1.1.3 Maximum outlet-to-inlet flow ratio

Examining the outlet-to-inlet flow ratio, denoted as $f_3(\mathbf{x})$, on its own does not yield accurate results. As displayed in Table 4.4, the GA suggests equal gas compositions (i.e., [CO_2:CH_4] = 1) at the maximum temperature within the defined search space. The progress of the fitness function values and exploration rate in each generation is depicted in Figures 4.4(c) and 4.4(g), respectively. However, the final solution presents a negative methane conversion rate, indicating the production of methane rather than its consumption. This issue arises from suggesting relatively low feed flow (i.e., 93.72 ml/min) and, consequently, low space

velocity, which significantly increases the methanation reaction (i.e., similar to Section 4.4.1.1.1). As a result, the objective of maximizing the outlet-to-inlet flow ratio contradicts the maximum syngas ratio and the conversion rate, especially in the case of methane conversion.

4.4.1.2 Multi-objective optimization problem

Since all aforementioned objectives denoted as $f_1(\mathbf{x})$, $f_2(\mathbf{x})$, and $f_3(\mathbf{x})$ conflict, a multi-objective function incorporating all these objectives denoted as $f(\mathbf{x})$ proposes the optimal temperature and gas compositions [CO_2:CH_4] to attain the most efficient DRM reaction within the defined search space. The multi-objective optimization problem (MOOP) encompasses three constraints aimed at balancing trade-offs among these objectives as follows:

i) 50% < Syngas ratio < 100%

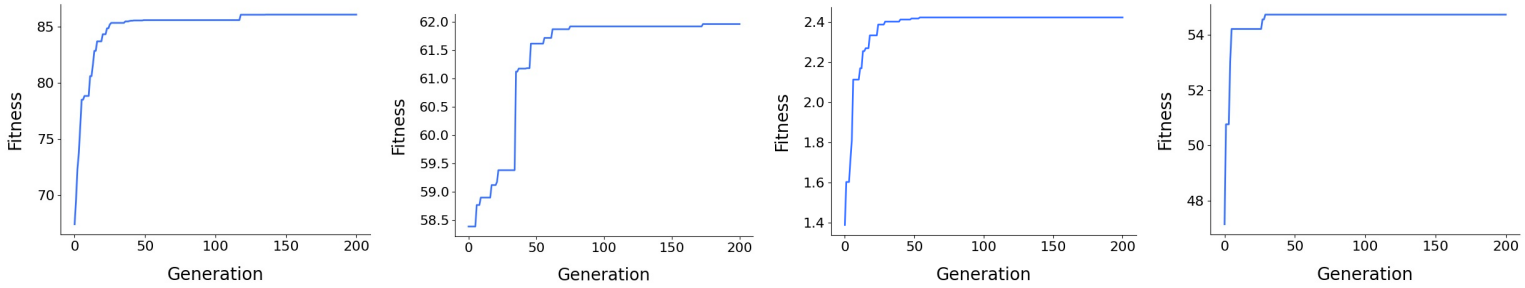
Considering that the RWGS reaction significantly influences the DRM reaction in the temperature range of 500-586 °C (refer to simultaneous reactions in section 2.2.2), it invariably results in a syngas ratio of lower than one, (i.e., [H_2:CO] < 1). Hence, we establish a 50% syngas ratio as the minimum objective and a 100% syngas ratio as the maximum, although achieving the latter is rare.

ii) 1 < Std_outlet_flow / Std_inlet_flow) < 2

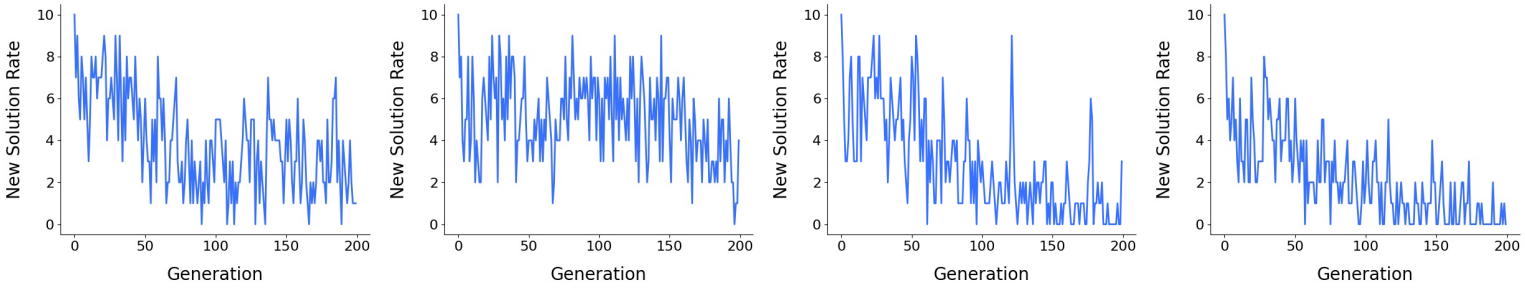
In accordance with the stoichiometric equation of the DRM reaction (as detailed in section 2.2.2), two moles of reactants convert into four moles of products. Consequently, an outlet-to-inlet flow ratio of two is expected for an efficient process. In practice, a ratio lower than one explicitly indicates carbon deposition, while a ratio greater than two signifies methanation (i.e., producing methane). This constraint confines these boundaries and offers rough insights into the presence of by-products.

iii) | CH_4 conversion – CO_2 conversion | < 10

A close to equilibrium DRM reaction consumes an equal amount of CH_4 and CO_2. Assuming no carbon deposition occurs during the process, the conversion gap primarily relates to the RWGS reaction [19]. This side reaction leads to an increase in CO_2 molar conversion while reducing the H_2 yield. The formation of water due to the RWGS reaction represents one of the significant challenges in the DRM process. This constraint confines the amount of simultaneous RWGS reaction.



(a) $f_1(\mathbf{x})$ fitness vs. generation. (b) $f_2(\mathbf{x})$ fitness vs. generation. (c) $f_3(\mathbf{x})$ fitness vs. generation. (d) $f(\mathbf{x})$ fitness vs. generation.



(e) No. new solutions for $f_1(\mathbf{x})$ vs. generation. (f) No. new solutions for $f_2(\mathbf{x})$ vs. generation. (g) No. new solutions for $f_3(\mathbf{x})$ vs. generation. (h) No. new solutions for $f(\mathbf{x})$ vs. generation.

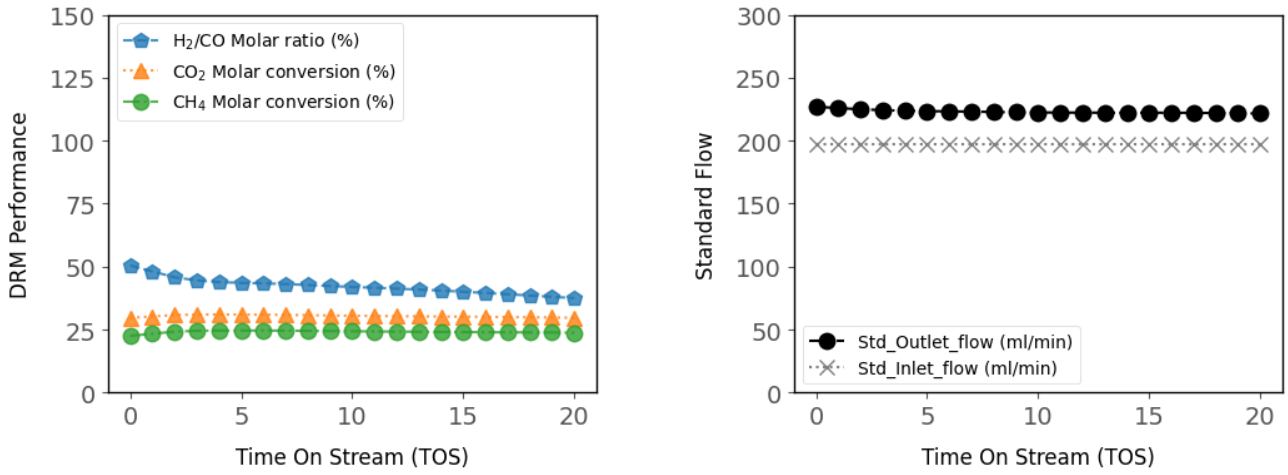
Figure 4.4: Fitness function values and exploration rate for each generation.

Table 4.4: The optimal reaction temperature/feed ratio for each objective function

Function		Operating Condition		Feed Ratio (CO_2/CH_4)	DRM Performance	
$f_1(x)$	maximum syngas ratio	Reaction Temp (°C):	569	1.00	H_2/CO molar ratio (%)	86.07
		CO_2 (ml/min):	149		CO_2 molar conversion(%)	15.46
		CH_4 (ml/min) :	149		CH_4 molar conversion(%)	4.25
$f_2(x)$	maximum molar conversion rate	Reaction Temp (°C):	586	0.66	H_2/CO molar ratio (%)	19.13
		CO_2 (ml/min):	71		CO_2 molar conversion(%)	35.46
		CH_4 (ml/min) :	108		CH_4 molar conversion(%)	26.50
$f_3(x)$	maximum outlet-to-inlet flow ratio	Reaction Temp (°C):	586	1.00	H_2/CO molar ratio (%)	2.81
		CO_2 (ml/min):	50		CO_2 molar conversion(%)	21.91
		CH_4 (ml/min) :	50		CH_4 molar conversion(%)	-10.38
$f(x)$	MOOP through the SOOP fitness function	Reaction Temp (°C):	579	1.50	H_2/CO molar ratio (%)	50.47
		CO_2 (ml/min):	126		CO_2 molar conversion(%)	29.66
		CH_4 (ml/min) :	84		CH_4 molar conversion(%)	22.51

In the following, the MOOP is converted to the SOOP fitness function $f(\mathbf{x})$ given by Equation 4.11 through the WBGA method. Figure 4.4(d) shows that the solution converges after approximately 30 generations and stabilizes in the succeeding ones. Concurrently, Figure 4.4(h) provides insight into the exploration rate within the $f(\mathbf{x})$ solutions. According to Table 4.4, the most effective DRM reaction occurs at a feed ratio of $[CO_2{:}CH_4] = 1.5$ and a reaction temperature of 579 °C. Yoo et al. [25] studied a feed ratio of greater than one $[CO_2{:}CH_4] > 1$ over the Ni/Al_2O_3 catalyst at a temperature varying from 700 °C to 850 °C (i.e., a higher temperature range than the defined search space in our problem), leading to simultaneous of significant RWGS reaction. However, the MOOP restricts the RWGS reaction through the constraint handling method while suggesting 579 °C as the optimum temperature.

Catalyst stability is a significant concern in the DRM reaction process, with carbon deposition notably reducing catalyst activity. The analyzed dataset includes a "time on stream" attribute, allowing the deep learning model to predict catalyst stability. Under the optimal conditions suggested by DNN-GA, specifically with a feed ratio of $[CO_2{:}CH_4] = 1.5$ at a temperature of 579 °C, Figure 4.5 presents the quality parameters of the DRM reaction over a twenty-hour period. Figure 4.5(a) displays the syngas ratio and conversion rate, while Figure 4.5(b) compares the Std_outlet_flow with the constant Std_inlet_flow. The findings show relatively constant conditions, which come from the stability of the catalyst.

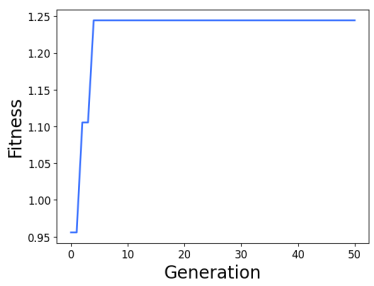


(a) (b)

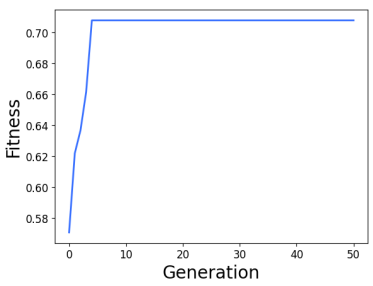
Figure 4.5: Catalyst stability prediction over a twenty-hour period. (a) DRM performance at optimal operating conditions. (b) a comparison of Std_outlet_flow with the constant Std_inlet_flow at optimal operating conditions.

4.4.2 Optimal operating conditions for DRM over Ni/CaFe2O4 catalyst

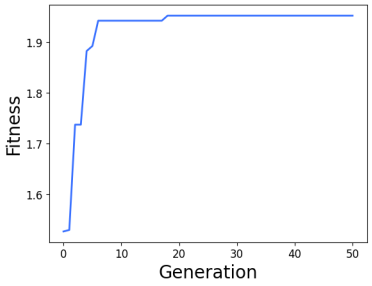
This section presents the results obtained for optimal syngas production via the DRM reaction over a $Ni/CaFe_2O_4$ catalyst and compares them with those previously obtained over a Ni/Al_2O_3 catalyst. The datasets presented to the AI-enabled models explicitly cover different temperature ranges, with DRM over $Ni/CaFe_2O_4$ occurring between 700-800 °C and DRM over Ni/Al_2O_3 approximately between 500-600 °C. Furthermore, variations in thermodynamic parameters such as space velocity and pressure, depending on catalyst characterization and reactor design, are present in these two datasets. As discussed in Section 3.7.4, higher temperatures significantly enhance the conversion of reactants into syngas. Therefore, we typically expect a more efficient process in the 700-800 °C range, which is evident in the dataset for $Ni/CaFe_2O_4$. However, at lower temperatures, due to the thermodynamic proximity of DRM to other reactions like the RWGS, simultaneous reactions become more likely, making the process more complex. Consequently, a similar optimization method utilizing a genetic algorithm and an objective function derived from an AI-enabled model reveals different results for those datasets. These differences stem from distinct datasets collected under various operating conditions, leading to different AI-enabled models and search spaces for the optimizer. The optimizer initially maximizes the objective function $f_1(\mathbf{x})$ for the maximum molar conversion rate and $f_2(\mathbf{x})$ for the maximum product yield separately. Subsequently, $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ are optimized jointly as an objective function referred to as $f(\mathbf{x})$. Figures 4.6(a) and 4.6(b) demonstrate the convergence of $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ in early generations. Increasing the population size and parent mating pool, as discussed in Section 4.3.2, does not significantly delay early convergence. The exploration rates of the objectives are roughly similar, as shown in Figures 4.6(d) and 4.6(e). Figure 4.6(c) illustrates the fitness values related to $f(\mathbf{x})$, which require more evolutionary iterations and converge after approximately 20 generations due to more complexity. Nevertheless, Table 4.5 displays the same optimal values for $f_1(\mathbf{x})$, $f_2(\mathbf{x})$ and $f(\mathbf{x})$ as distinct objectives, demonstrating $f(\mathbf{x})$ as a single-objective optimization problem. In essence, optimizing either $f_1(\mathbf{x})$ or $f_2(\mathbf{x})$ leads to the optimization of the joint objectives, i.e., $f(\mathbf{x})$. This single-objective problem arises due to the lower complexity of the DRM reaction process at a temperature range between 700-800 °C, resulting in less synchrony with other reactions during the process. In conclusion, the optimizer selects the maximum temperature within this range (800 °C) and equal gas composition [CO_2:CH_4] with the total feed flow of 100 (ml/min) for DRM over $Ni/CaFe_2O_4$, unlike DRM over Ni/Al_2O_3 selected an intermediate temperature of the specified range (579 °C) and a gas composition [CO_2:CH_4] of 1.5 with the total feed flow of 210(ml/min). It is evident that the DRM reaction over the $Ni/CaFe_2O_4$ catalyst at temperatures between 700-800 °C operates close to equilibrium and requires a gas composition consistent with the stoichiometric equation of the DRM (refer to Section 2.2.2) for optimal performance.



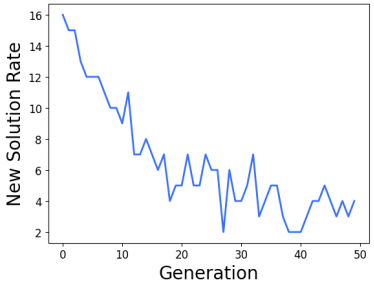
(a) $f_1(\mathbf{x})$ fitness vs. generation.



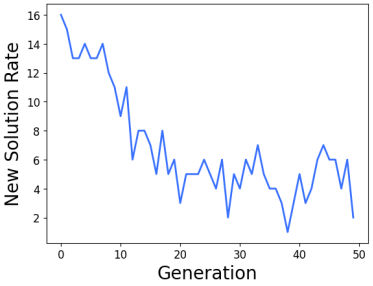
(b) $f_2(\mathbf{x})$ fitness vs. generation.



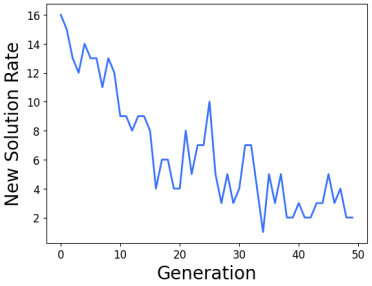
(c) $f(\mathbf{x})$ fitness vs. generation.



(d) No. new solutions for $f_1(\mathbf{x})$ vs. generation.



(e) No. new solutions for $f_2(\mathbf{x})$ vs. generation.



(f) No. new solutions for $f_3(\mathbf{x})$ vs. generation.

Figure 4.6: Fitness function values and exploration rate for each generation.

Table 4.5: The optimal reaction temperature/feed ratio for each objective function

Function		Operating Condition		Feed Ratio (CO_2/CH_4)	DRM Performance	
$f_1(x)$	maximum molar conversion rate	Reaction Temp (°C):	800	1.00	H_2 yield	35.60
		Metal loading (%wt)	5		CO yield	35.18
		CO_2 (ml/min):	50		CO_2 molar conversion(%)	61.20
		CH_4 (ml/min) :	50		CH_4 molar conversion(%)	63.24
$f_2(x)$	maximum yield of products	Reaction Temp (°C):	800	1.00	H_2 yield	35.60
		Metal loading (%wt)	5		CO yield	35.18
		CO_2 (ml/min):	50		CO_2 molar conversion(%)	61.20
		CH_4 (ml/min) :	50		CH_4 molar conversion(%)	63.24
$f(x)$	SOOP fitness function	Reaction Temp (°C):	800	1.00	H_2 yield	35.60
	(maximum molar conversion and yield of products)	Metal loading (%wt)	5		CO yield	35.18
		CO_2 (ml/min):	50		CO_2 molar conversion(%)	61.20
		CH_4 (ml/min) :	50		CH_4 molar conversion(%)	63.24

4.5 Chapter summary

In this chapter, the software sensor for regression discussed in the previous chapter, also known as an AI-enabled process model, is employed with the aim of optimizing the process. We combine this software sensor with a genetic algorithm to demonstrate the process optimization. Instead of manually calculating all possible process inputs and their corresponding outputs, we leverage a genetic algorithm to determine the optimal inputs rapidly. Genetic algorithms use an objective function to maximize or minimize objectives across various iterations. In this study, our objective function is primarily derived from the AI-enabled process model. DRM is a catalytic reaction that yields diverse outcomes over various catalysts differing in chemical and physical characteristics and under different thermodynamic parameters, especially temperature. This optimization method combines an optimizer with a data-driven model, offering the flexibility to model various datasets with unique attributes and adjust the objective function as needed. For the first step, the optimizer suggests a feed ratio [CO_2:CH_4] of 1.5 at 579 °C in the context of defined search space for an efficient DRM reaction process over Ni/AL2O3. For the second step, the optimizer suggests a feed ratio [CO_2:CH_4] of one at 800 °C in the context of defined search space for an efficient DRM reaction process over Ni/$CaFe_2O_4$. The DRM reaction optimization problem can be single-objective or multi-objective, depending on the dataset and operating conditions. The optimizer employs the AI-enabled model trained with the provided dataset as the objective function, and the ultimate results significantly depend on the dataset. It is noteworthy that the higher the temperature range, the closer to the equilibrium the DRM reaction process operates. Therefore, considering the DRM reaction at higher temperatures over a catalyst with high selectivity and activity, the optimizer probably suggests an equal gas composition according to the stoichiometric equation of the reaction (refer to Section 2.2.2). However, scientists attempt to provide efficient operating conditions at lower temperatures aimed at saving energy and low-cost catalysts to save financial resources where the reaction reveals complicated behaviour.

Chapter 5

Conclusions and Future Work

This chapter provides an overview of the thesis, outlining the research findings and the engineering importance of the thesis. It also enumerates potential future directions that should be considered by stakeholders who wish to employ software sensors in their process plants.

5.1 Thesis summary and conclusions

The thesis aims to address a gap in the existing literature concerning the deployment of a deep learning model for a chemical reaction process known as dry reforming of methane (DRM). This research introduces a software sensor for syngas production via DRM by leveraging experimental data and proposing a predictive model to estimate the quality parameters of the process. Both supervised and unsupervised learning assist in implementation.

Unsupervised learning initially identifies the outliers and discards them from the primary dataset. Subsequently, ensemble and deep learning predict the outcome. Ultimately, evaluation metrics, applied to samples never seen before by the models, are used to compare the performance and generalization capabilities of the developed models. Software sensors prefer deep learning models due to their inherent advantages. Nonetheless, these models are remarkably data-demanding, and conventional deep learning is unsuccessful in bench scale with limited training samples. Therefore, the outstanding capabilities of pretraining for small datasets tackle this issue and facilitate deep learning.

Meanwhile, results from developed models illustrate the root cause of outliers and specifically demonstrate which instruments generate those errors during data collection. Since the available dataset encompasses time series, the time index of error occurrence is traceable, and it is visualized with the assistance of a dimensionality reduction algorithm known as principal component analysis (PCA).

The primary objective of this thesis is to propose a real-time adaptive optimizer for an efficient version of the process. Operators typically apply various ranges of thermodynamic parameters, such as temperature and flow, for the DRM reaction process based on physical limitation and catalyst resiliency. While a software sensor utilizing deep learning has the ability to adapt to those various ranges, a genetic algorithm integrated with the software sensor proposes the optimal thermodynamic values within the range according to defined objectives. The optimizer proposes distinct optimal values for two datasets serving at various thermodynamic parameter ranges and operating setups. In particular, it proposes a gas composition of $[CO_2{:}CH_4]$=1.5 for the primary dataset received from Carbonova Corp. In contrast, the preferred gas composition for the second dataset received from previous work in the literature is $[CO_2{:}CH_4]$=1

5.2 Engineering significance of thesis findings

Dry reforming of methane is an environmentally and industrially important chemical process yet to be commercialized due to high energy demand and catalyst deactivation as its significant barriers. It deserves more effort and has been under development in terms of both catalyst improvement and process optimization over the last decades. This research contributes to the advancement of this development by proposing a software sensor for regression (SSR), especially utilizing deep learning, presenting advanced analytic capabilities as follows:

- Predictive analysis: The SSR provides predictive analysis by estimating the quality parameters of the DRM reaction process.
- Prescriptive analysis: An integration of the SSR with various optimization algorithms provides prescriptive analysis to improve the process. A total solution not only recommends optimal values for the thermodynamic parameters of the process but can also be used for developing and optimizing the catalysts applied to the process.
- Diagnostic analysis: The SSR serves as an estimator, and additionally, it functions as a monitoring tool for various applications, such as reporting errors, alerting accidents, and forecasting maintenance.
- Descriptive analysis: An integration of the SSR with a dimensionality reduction algorithm is employed for visualization and descriptive analysis.

Unlike previous works utilizing shallow learning, this study implements SSR using deep learning with the following advantages:

- Multi-dimensional support: A process plant inherently generates complex data structures as multi-dimensional tensors collected by various sensors and instruments. The hierarchical architecture of neural networks in deep learning has the ability to accept a substantial number of attributes without limitation, either as independent or dependent variables and unfold the non-linear relationship between those variables.
- Multi-modal support: Data extracted from various sensors in a process plant consists of various formats such as numeric, images and audio. Deep learning involves developing models that can effectively process and integrate distinct types of information, referred to as multiple modalities.
- Transfer learning: One of the features that outperforms deep learning from other machine learning models is the capability to reuse a model for new problems with similar tasks. Thus, retraining the model is available when operators add or remove specific attributes from the existing dataset or collect a new dataset under different operating conditions. This feature accelerates the deployment of new models and demands less training data.
- Less feature engineering demand: Raw data generated by a process plant is often required to transform into a more suitable format for machine learning algorithms using various tasks such as feature scaling, handling outliers and encoding categorical variables to improve the model performance. However, a hierarchical abstracted architecture of deep learning demands less feature engineering effort, making it well-suited for training the complex data structure of a process plant and the deployment of software sensors.

5.3 Thesis limitations and suggestions for future work

Accessing error-free datasets with comprehensive attributes is the main concern of machine learning engineers in developing software sensors for process plants. Though deep learning reduces feature engineering workload, it consistently plays a crucial role. The thesis outlines the following recommendations for future work, which significantly centers around data acquisition:

- Developing a software sensor for a process plant is an interdisciplinary field requiring consultation with data analysis experts before data collection to reduce rework costs and ensure operators collect data in a suitable format.
- The available plant requires more sensors and instruments to give insight into what is happening inside the reactor and product tank in terms of the primary and possible simultaneous

reactions. For example, sensors measuring water formation give insight into the occurrence of side reactions generating water, monitoring pressure changes gives insight into reactor blockage due to carbon deposition, and real-time temperature changes reflect the probability of possible endothermic or exothermic reactions within the reactor.

- Higher uniformity and variety of samples in terms of various attributes such as gas composition, gas flow and space velocity are required to improve the generalization capabilities of the model.
- The measurement uncertainties were not considered in the measured data and, as such, not reflected in the developed machine learning models. Accounting for the measurement uncertainties to determine its effect on the convergence of machine learning models is a topic for future work.

Bibliography

[1] McKinsey Global Institute. AI: The Next Frontier of Performance in Industrial Processing Plants. Accessed on 2023-11-07. [Online]. Available: https://www.mckinsey.com/industries/metals-and-mining/our-insights/ai-the-next-frontier-of-performance-in-industrial-processing-plants#/

[2] "Justin Trudeau's climate plans are stuck in Alberta's tar sands. The Economist," *The Economist*. [Online]. Available: https://www.economist.com/the-americas/2018/12/15/justin-trudeaus-climate-plans-are-stuck-in-albertas-tar-sands

[3] "Justin Trudeau's climate plans are stuck in Alberta's tar sands. The Economist," *The Economist*. [Online]. Available: https://www.economist.com/the-americas/2018/12/15/justin-trudeaus-climate-plans-are-stuck-in-albertas-tar-sands

[4] H. J. Guzman, "Email communication," November 24 2023, sender position: "Product Development Lead in Carbonova Corp.".

[5] M. A. Hossain, B. V. Ayodele, C. K. Cheng, and M. R. Khan, "Artificial neural network modeling of hydrogen-rich syngas production from methane dry reforming over novel Ni/CaFe2O4 catalysts," *International Journal of Hydrogen Energy*, vol. 41, no. 26, pp. 11 119–11 130, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.ijhydene.2016.04.034

[6] J. R. Ross, "Chapter 3 - how does a catalyst work?" in *Contemporary Catalysis*, J. R. Ross, Ed. Amsterdam: Elsevier, 2019, pp. 69–89. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780444634740000035

[7] J. R. Ross, "Chapter 7 - the kinetics and mechanisms of catalytic reactions," in *Contemporary Catalysis*, J. R. Ross, Ed. Amsterdam: Elsevier, 2019, pp. 161–186. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780444634740000072

[8] J. R. Ross, "Chapter 8 - mass and heat transfer limitations and other aspects of the use of large-scale catalytic reactors," in *Contemporary Catalysis*, J. R. Ross, Ed. Amsterdam: Elsevier, 2019, pp. 187–213. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780444634740000084

[9] J. J. Siirola, G. J. Powers, and D. F. Rudd, "Synthesis of system designs: III. Toward a process concept generator," *AIChE Journal*, vol. 17, no. 3, pp. 677–682, 1971. [Online]. Available: https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.690170334

[10] J. J. Siirola and D. F. Rudd, "Computer-Aided Synthesis of Chemical Process Designs From Reaction Path Data to the Process Task Network," *Industrial and Engineering Chemistry Fundamentals*, vol. 10, no. 3, pp. 353–362, 1971.

[11] G. Aurlien, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow.* O'Reilly Media, 2022, vol. Third edit. [Online]. Available: https://ezproxy.lib.ucalgary.ca/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=3406174&site=ehost-live

[12] V. Venkatasubramanian, "The promise of artificial intelligence in chemical engineering: Is it here, finally?" *AIChE Journal*, vol. 65, no. 2, pp. 466–478, 2019.

[13] F. A. A. Souza, R. Araújo, and J. Mendes, "Review of soft sensor methods for regression applications," *Chemometrics and Intelligent Laboratory Systems*, vol. 152, pp. 69–79, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169743915003263

[14] J. Mohd Ali, M. A. Hussain, M. O. Tade, and J. Zhang, "Artificial Intelligence techniques applied as estimator in chemical process systems – A literature survey," *Expert Systems with Applications*, vol. 42, no. 14, pp. 5915–5931, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417415002171

[15] D. Fooshee, A. Mood, E. Gutman, M. Tavakoli, G. Urban, F. Liu, N. Huynh, D. Van Vranken, and P. Baldi, "Deep learning for chemical reaction prediction," *Molecular Systems Design and Engineering*, vol. 3, no. 3, pp. 442–452, 2018.

[16] M. R. Dobbelaere, P. P. Plehiers, R. Van de Vijver, C. V. Stevens, and K. M. Van Geem, "Machine Learning in Chemical Engineering: Strengths, Weaknesses, Opportunities, and Threats," *Engineering*, vol. 7, no. 9, pp. 1201–1211, sep 2021.

[17] D. Pakhare and J. Spivey, "A review of dry (CO2) reforming of methane over noble metal catalysts," *Chemical Society Reviews*, vol. 43, no. 22, pp. 7813–7837, 2014. [Online]. Available: http://dx.doi.org/10.1039/C3CS60395D

[18] J. R. Ross, "Chapter 12 - catalytic reactions involving syngas, hydrogen, or carbon monoxide for the production of intermediates and chemicals," in *Contemporary Catalysis*, J. R. Ross, Ed. Amsterdam: Elsevier, 2019, pp. 273–290. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780444634740000126

[19] M. M. Barroso Quiroga and A. E. Castro Luna, "Kinetic Analysis of Rate Data for Dry Reforming of Methane," *Industrial & Engineering Chemistry Research*, vol. 46, no. 16, pp. 5265–5270, 2007. [Online]. Available: https://doi.org/10.1021/ie061645w

[20] B. Nematollahi, M. Rezaei, E. N. Lay, and M. Khajenoori, "Thermodynamic analysis of combined reforming process using Gibbs energy minimization method: In view of solid carbon formation," *Journal of Natural Gas Chemistry*, vol. 21, no. 6, pp. 694–702, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1003995311604210

[21] R. Y. Chein, Y. C. Chen, C. T. Yu, and J. N. Chung, "Thermodynamic analysis of dry reforming of CH4 with CO2 at high pressures," *Journal of Natural Gas Science and Engineering*, vol. 26, pp. 617–629, 2015.

[22] D. Pakhare, C. Shaw, D. Haynes, D. Shekhawat, and J. Spivey, "Effect of reaction temperature on activity of pt- and ru-substituted lanthanum zirconate pyrochlores (la2zr2o7) for dry (co2) reforming of methane (drm)," *Journal of CO2 Utilization*, vol. 1, pp. 37–42, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212982013000127

[23] J. Sehested, "Four challenges for nickel steam-reforming catalysts," *Catalysis Today*, vol. 111, no. 1, pp. 103–110, 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092058610500708X

[24] W.-J. Jang, J.-O. Shim, H.-M. Kim, S.-Y. Yoo, and H.-S. Roh, "A review on dry reforming of methane in aspect of catalytic properties," *Catalysis Today*, vol. 324, pp. 15–26, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092058611830083X

[25] E. Yoo, D.-S. Choi, J. Kim, Y.-H. Kim, N.-Y. Kim, and J. B. Joo, "Effects of Operating Parameters and Feed Gas Compositions on the Dry Reforming of Methane over the Ni/Al2O3 Catalyst," *Catalysts*, vol. 13, no. 3, p. 602, 2023.

[26] C. H. Bartholomew and R. J. Farrauto, *Fundamentals of industrial catalytic processes*. John Wiley & Sons, 2011.

[27] C. Shi, S. Wang, X. Ge, S. Deng, B. Chen, and J. Shen, "A review of different catalytic systems for dry reforming of methane: Conventional catalysis-alone and plasma-catalytic system," *Journal of CO2 Utilization*, vol. 46, p. 101462, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212982021000299

[28] M. S. Ferrandon, C. Byron, G. Celik, Y. Zhang, C. Ni, J. Sloppy, R. A. McCormick, K. Booksh, A. V. Teplyakov, and M. Delferro, "Grafted nickel-promoter catalysts for dry reforming of methane identified through high-throughput experimentation," *Applied Catalysis A: General*, vol. 629, p. 118379, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926860X21003938

[29] X. Chen, K. Honda, and Z. G. Zhang, "A comprehensive comparison of CH4-CO2 reforming activities of NiO/Al2O3 catalysts under fixed- and fluidized-bed operations," *Applied Catalysis A: General*, vol. 288, no. 1-2, pp. 86–97, 2005.

[30] Y. Gorishniy, I. Rubachev, V. Khrulkov, and A. Babenko, "Revisiting Deep Learning Models for Tabular Data," *Advances in Neural Information Processing Systems*, vol. 23, no. NeurIPS, pp. 18 932–18 943, 2021.

[31] V. Borisov, T. Leemann, K. Seßler, J. Haug, M. Pawelczyk, and G. Kasneci, "Deep Neural Networks and Tabular Data: A Survey," no. June, pp. 1–22, 2021. [Online]. Available: http://arxiv.org/abs/2110.01889

[32] Q. Sun and Z. Ge, "A Survey on Deep Learning for Data-Driven Soft Sensors," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 9, pp. 5853–5866, 2021.

[33] S. Feng, H. Zhou, and H. Dong, "Using deep neural network with small dataset to predict material defects," *Materials Design*, vol. 162, pp. 300–310, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0264127518308682

[34] H. N. Mhaskar and T. Poggio, "Deep vs. shallow networks: An approximation theory perspective," *Analysis and Applications*, vol. 14, no. 6, pp. 829–848, 2016.

[35] P. Nkulikiyinka, Y. Yan, F. Güleç, V. Manovic, and P. T. Clough, "Prediction of sorption enhanced steam methane reforming products from machine learning based soft-sensor models," *Energy and AI*, vol. 2, 2020.

[36] C. J. Taylor, A. Pomberger, K. C. Felton, R. Grainger, M. Barecka, T. W. Chamberlain, R. A. Bourne, C. N. Johnson, and A. A. Lapkin, "A Brief Introduction to Chemical Reaction Optimization," *Chemical Reviews*, 2022.

[37] B. V. Ayodele and C. K. Cheng, "Modelling and optimization of syngas production from methane dry reforming over ceria-supported cobalt catalyst using artificial neural networks and Box–Behnken design," *Journal of Industrial and Engineering Chemistry*, vol. 32, pp. 246–258, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1226086X15004025

[38] B. V. Ayodele, S. I. Mustapa, M. A. Alsaffar, and C. K. Cheng, "Artificial intelligence modelling approach for the prediction of CO-rich hydrogen production rate from methane dry reforming," *Catalysts*, vol. 9, no. 9, 2019.

[39] M. A. Alsaffar, B. V. Ayodele, and S. I. Mustapa, "Scavenging carbon deposition on alumina supported cobalt catalyst during renewable hydrogen-rich syngas production by methane dry reforming using artificial intelligence modeling technique," *Journal of Cleaner Production*, vol. 247, p. 119168, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652619340387

[40] F. Elmaz, Ö. Yücel, and A. Y. Mutlu, "İstatistiksel ve Makine Öğrenmeye Dayalı Yaklaşımlarla Kobalt Katalizör Üzerinden Metan Kuru Reformundan Elde Edilen Sentez Gazının Tahmini Modellemesi," *International Journal of Advances in Engineering and Pure Sciences*, mar 2020.

[41] B. V. Ayodele, M. A. Alsaffar, S. I. Mustapa, R. Kanthasamy, S. Wongsakulphasatch, and C. K. Cheng, "Carbon dioxide reforming of methane over Ni-based catalysts: Modeling the effect of process parameters on greenhouse gasses conversion using supervised machine learning algorithms," *Chemical Engineering and Processing - Process Intensification*, vol. 166, p. 108484, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0255270121001847

[42] Z. Hao, C. Kefa, and M. Jianbo, "Combining neural network and genetic algorithms to optimize low NOx pulverized coal combustion," *Fuel*, vol. 80, no. 15, pp. 2163–2169, 2001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0016236101001041

[43] S. Nandi, Y. Badhe, J. Lonari, U. Sridevi, B. S. Rao, S. S. Tambe, and B. D. Kulkarni, "Hybrid process modeling and optimization strategies integrating neural networks/support vector regression and genetic algorithms: study of benzene isopropylation on Hbeta catalyst," *Chemical Engineering Journal*, vol. 97, no. 2, pp. 115–129, 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1385894703001505

[44] E. B. Gueguim Kana, J. K. Oloke, A. Lateef, and M. O. Adesiyan, "Modeling and optimization of biogas production on saw dust and other co-substrates using Artificial Neural network

and Genetic Algorithm," *Renewable Energy*, vol. 46, pp. 276–281, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960148112002145

[45] R. Soleimani, N. A. Shoushtari, B. Mirza, and A. Salahi, "Experimental investigation, modeling and optimization of membrane separation using artificial neural network and multi-objective optimization using genetic algorithm," *Chemical Engineering Research and Design*, vol. 91, no. 5, pp. 883–903, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0263876212003012

[46] A. Velásco-Mejía, V. Vallejo-Becerra, A. U. Chávez-Ramírez, J. Torres-González, Y. Reyes-Vidal, and F. Castañeda-Zaldivar, "Modeling and optimization of a pharmaceutical crystallization process by using neural networks and genetic algorithms," *Powder Technology*, vol. 292, pp. 122–128, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0032591016300274

[47] Y. Shin, Z. Kim, J. Yu, G. Kim, and S. Hwang, "Development of NOx reduction system utilizing artificial neural network (ANN) and genetic algorithm (GA)," *Journal of Cleaner Production*, vol. 232, pp. 1418–1429, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652619318128

[48] D. Q. Gbadago, J. Moon, M. Kim, and S. Hwang, "A unified framework for the mathematical modelling, predictive analysis, and optimization of reaction systems using computational fluid dynamics, deep neural network and genetic algorithm: A case of butadiene synthesis," *Chemical Engineering Journal*, vol. 409, p. 128163, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1385894720342790

[49] V. Khezri, E. Yasari, M. Panahi, and A. Khosravi, "Hybrid Artificial Neural Network–Genetic Algorithm-Based Technique to Optimize a Steady-State Gas-to-Liquids Plant," *Industrial & Engineering Chemistry Research*, vol. 59, no. 18, pp. 8674–8687, 2020. [Online]. Available: https://doi.org/10.1021/acs.iecr.9b06477

[50] H. Fang, J. Zhou, Z. Wang, Z. Qiu, Y. Sun, Y. Lin, K. Chen, X. Zhou, and M. Pan, "Hybrid method integrating machine learning and particle swarm optimization for smart chemical process operations," *Frontiers of Chemical Science and Engineering*, vol. 16, no. 2, pp. 274–287, 2022. [Online]. Available: https://doi.org/10.1007/s11705-021-2043-0

[51] F. N. Alotaibi, A. S. Berrouk, and M. Saeed, "Optimization of yield and conversion rates in methane dry reforming using artificial neural networks and the multiobjective genetic algorithm," *Industrial & Engineering Chemistry Research*, vol. 62, no. 42, pp. 17 084–17 099, 2023. [Online]. Available: https://doi.org/10.1021/acs.iecr.3c01813

[52] "Carbonova dataset," 2022, proprietary.

[53] Tin Kam Ho, "Random Decision Forests Tin Kam Ho Perceptron training," *Proceedings of 3rd International Conference on Document Analysis and Recognition*, pp. 278–282, 1995. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/598994/

[54] I. Goodfellow, Y. Bengio, and ·. A. Courville, "Deep Learning," *Foreign Affairs*, vol. 91, no. 5, pp. 1689–1699, 2016.

[55] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," *Advances in Neural Information Processing Systems*, no. 1, pp. 153–160, 2007.

[56] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Computer Science*, vol. 7, p. e623, 2021.

[57] S. Yasyerli, S. Filizgok, H. Arbag, N. Yasyerli, and G. Dogu, "Ru incorporated Ni-MCM-41 mesoporous catalysts for dry reforming of methane: Effects of Mg addition, feed composition and temperature," *International Journal of Hydrogen Energy*, vol. 36, no. 8, pp. 4863–4874, 2011.

[58] A. Serrano-Lotina and L. Daza, "Influence of the operating parameters over dry reforming of methane to syngas," *International Journal of Hydrogen Energy*, vol. 39, no. 8, pp. 4089–4094, 2014.

[59] S. Mirjalili, *Evolutionary Algorithms and Neural Networks: Theory and Applications*, ser. Studies in Computational Intelligence. Springer International Publishing, 2018. [Online]. Available: https://books.google.ca/books?id=eS9iDwAAQBAJ

[60] D. Simon, *Evolutionary Optimization Algorithms*. Wiley, 2013. [Online]. Available: https://books.google.ca/books?id=gwUwIEPqk30C

[61] I. Rahimi, A. H. Gandomi, F. Chen, and E. Mezura-Montes, "A Review on Constraint Handling Techniques for Population-based Algorithms: from single-objective to multi-objective optimization," *Archives of Computational Methods in Engineering*, vol. 30, no. 3, pp. 2181–2209, 2022. [Online]. Available: https://doi.org/10.1007/s11831-022-09859-9

[62] P. Hajela and C. Y. Lin, "Genetic search strategies in multicriterion optimal design," *Structural optimization*, vol. 4, no. 2, pp. 99–107, 1992. [Online]. Available: https://doi.org/10.1007/BF01759923

[63] H. Wang, M. Olhofer, and Y. Jin, "A mini-review on preference modeling and articulation in multi-objective optimization: current status and challenges," *Complex Intelligent Systems*, vol. 3, no. 4, pp. 233–245, 2017.

[64] A. F. Gad, "PyGAD: An Intuitive Genetic Algorithm Python Library," no. April 2020, 2021. [Online]. Available: http://arxiv.org/abs/2106.06158

[65] S. Marsili Libelli and P. Alba, "Adaptive mutation in genetic algorithms," *Soft Computing*, vol. 4, no. 2, pp. 76–80, 2000.

[66] B. Miao, S. S. K. Ma, X. Wang, H. Su, and S. H. Chan, "Catalysis mechanisms of CO2 and CO methanation," *Catal. Sci. Technol.*, vol. 6, no. 12, pp. 4048–4058, 2016. [Online]. Available: http://dx.doi.org/10.1039/C6CY00478D

[67] J. Sehested, S. Dahl, J. Jacobsen, and J. R. Rostrup-Nielsen, "Methanation of CO over Nickel: Mechanism and Kinetics at High H2/CO Ratios," *The Journal of Physical Chemistry B*, vol. 109, no. 6, pp. 2432–2438, 2005. [Online]. Available: https://doi.org/10.1021/jp040239s