MSc Engineering – Energy and Power Systems
MSc dissertation

Optimisation and prediction of a catalytic process for low carbon fuel synthesis using machine learning.

Sammya Majumdar

September 2021

Supervisor: Prof. Xin Tu
Assessor: Dr. Naser Sedghi

Dissertation submitted to the University of Liverpool in partial fulfilment of the requirements for the degree of MSc – Energy and Power Systems

Abstract

Dry Reformation of Methane (DRM) has been the topic of significant research because of its property of converting two of the most significant greenhouse gases present in earth's atmosphere, carbon dioxide (CO) and methane (CH_4) to produce syngas, a mixture of carbon monoxide (CO) and hydrogen (H_2). Syngas is useful to manufacture straight chain hydrocarbons by the Fischer Tropsch process and allows the extraction of hydrogen for hydrogen power cells. However, this highly endothermic reaction has led to difficulties in the choice of catalysts as high reaction temperatures cause catalyst deactivation due to coke formation on the catalyst surface and thermal sintering. Nickel (Ni) and Cobalt (Co) over Al_2O_3 support are the two most researched catalysts in this reaction because of their high availability and low cost. This work aims to develop an Artificial Neural Network (ANN) approach to advance the understanding of the effect various reaction and operational parameters have on the methane conversion rate. Because of the large number of experimental descriptors, data mining tools such as ANNs can spot statistical patterns and correlations within the data which is not possible by manual inspection. The data is collected from the relevant literature on DRM reactions and the model developed in this paper will be tested against unseen data and ANN models already existing in the literature.

Contents

Nomenclature

Al_2O_3:	Aluminium Oxide
ANN:	Artificial Neural Network
Ar:	Argon
Ba:	Barium
BBD:	Box Behnken Design
Ce:	Cerium
CeO_2:	Cerium dioxide
CH_4:	Methane
CO:	Carbon monoxide
CO_2:	Carbon dioxide
CPU:	Central Processing Unit
Cu:	Copper
DBD:	Dielectric Barrier Discharge
DRM:	Dry Reformation of Methane
e:	Exponential
EDA:	Exploratory Data Analysis
GHSV:	Gas hourly space velocity
GPU:	Graphics Processing Unit
H_2:	Hydrogen
He:	Helium
Ir:	Iridium
K:	Potassium
La:	Lanthanum
La_2O_3:	Lanthanum Oxide
MCM-41:	Mobil composition of Matter no 41
Mg:	Magnesium
$MgAl_2O_4$:	Magnesium Aluminate
Mn:	Manganese
MAE:	Mean Absolute Error
MSE:	Mean Squared Error
N_2:	Nitrogen
Ni:	Nickel
Pd:	Palladium
Pr_2O_3:	Praseodymium (III) Oxide
PrO_2:	Praseodymium dioxide
Pt:	Platinum
RBFNN:	Radial Basis Function Neural Network
ReLU:	Rectified Linear Unit
Rh:	Rhodium
RMSE:	Root Mean Squared Error
Ru:	Ruthenium
Si_3N_4:	Silicon Nitride
SiO_2:	Silicon dioxide
Sr:	Strontium
TiO_2:	Titanium dioxide
V_2O_5:	Vanadium Pentoxide
Zr:	Zirconium
ZrO_2:	Zirconium dioxide
ZSM-5:	Zeolite Socony Mobil–5

Acknowledgement

I would like to express my sincere gratitude to my MSc project supervisor, Prof. Xin Tu, University of Liverpool for giving me the opportunity to learn and work under his supervision and guiding me throughout the course of this project.

I am grateful to my parents for their love and the sacrifices they made so that I can afford the privilege to study and pursue research at one of the premier universities in the UK. Also, I am thankful to my dear friend and classmate, Bagrat Utegenov for his motivation while working on this project and challenging me to develop new ideas and keep working on them.

List of Figures

List of Tables

1. INTRODUCTION

The measure of a civilisation's progress is often characterised by the amount of energy it consumes in a set period. Since the industrial revolution, that amount has been steadily increasing with no signs of slowing down. However, with this boom in energy consumption came the dangers of global warming. In the 21st century, the threat of climate change is more real than ever - brought on by relentless mining of fossil fuels and burning them to drive steam turbines, a dominant paradigm in energy production. Methane and carbon dioxide are known to contribute 91% greenhouse emissions in the world thereby accelerating the onslaught of climate change due to global warming. This threatens the very existence of humanity. Naturally, converting these gases into useful alternatives has piqued the interest of the scientific community. One of these alternatives is the dry reformation of methane (DRM). This chemical reaction uses CO_2 and CH_4 as reactants (usually from natural gas) to produce syngas, a mixture of carbon monoxide and hydrogen, essentially killing two birds with one stone. Equation 1 shows the DRM reaction:

$$CH_4 + CO_2 \leftrightarrow 2CO + 2H_2$$

$$\Delta H° = 247.3 \text{ kJ/mol} \qquad (1)$$

Hydrogen (H_2) from the syngas can be extracted and used in applications ranging from hydrogen fuel cells, petroleum refining to glass purification. Furthermore, syngas is used in the Fischer Tropsch process (Franz Fischer and Hans Tropsch developed this technique back in 1920 in Germany) to synthesise straight chain hydrocarbons, from methane to paraffin and sometimes alcohols. Often the products extracted by the Fischer Tropsch process are used to manufacture diesel and Jet Fuel. [1] Despite the demonstrated advantages, the choice of catalyst has remained an obstacle to achieving successful commercialisation of this process. Rapid catalyst deactivation due to carbon formation, high endothermicity, and equilibrium limitations are some of the difficulties that makes this process harder to be adapted by the industry. Coke formation on the surface of the catalyst or along the reactor walls is given by the reactions 2 and 3:

$$CH_4 \rightleftharpoons C + 2H_2 \qquad (2)$$

$$2CO \rightleftharpoons CO_2 + C \qquad (3)$$

Equation 2 is the endothermic breakdown of Methane and equation 3 is known as the Boudouard reaction. Methane breakdown occurs at higher temperatures (>700ºC) owing to its endothermicity. Boudouard reactions occurs at lower temperatures. However, it can also occur at higher temperatures if the temperature inside a section of the reactor is low due to poor heat transfer. Noble metals like Ruthenium (Ru), Rhodium (Rh), Platinum (Pt) have shown good performance as catalysts, the cost and rarity of these elements make them infeasible for use in the industry. Therefore, their cheaper alternatives like the transition metals (Nickel (Ni), Cobalt (Co)) are the focus of present research [2]. There are a few other alternatives to the DRM reaction. 1) Steam reforming – methane is reacted with steam to produce syngas with a H_2/CO ratio of 2. Often the ratio is falls below 2 due to the reverse water gas shift reaction. This is where carbon dioxide reacts with the hydrogen in syngas to produce carbon monoxide and water. 2) To combat the RWGS reaction in steam reforming, additional steam is added to the reactor. This is known as autothermal reforming of methane. Use of these various techniques is dictated by the requirement of carbon monoxide to hydrogen ratio in the syngas. For DRM, a ratio of 1 is preferred.

This work will present the application of deep learning algorithms to develop a neural network model to analyse and predict the CH_4 conversion rates based on given conditions. A dataset has been synthesised from the relevant literature in this field containing approximately 6000 data points which will act as an input to the model on which it will train on. Upon completion of training, the fidelity of the model will be tested by comparison with similar models found in the literature and testing it on unseen data. Neural networks are used to detect and learn statistical patterns in large datasets, which are then used to make predictions on data it has never seen before. Taking the inspiration from biological neurons in human brains, Warren McCulloch and Walter Pitts proposed the idea of a neural network in 1943. [3] Since then neural networks improvised to a great extent and drifted away from its resemblance to biological neurons. A neural network essentially creates a weighted sum of its inputs, applies an activation function and outputs a value. The weights are the connections between the neurons or “units” which the model learns from the data. Using the backpropagation algorithm, the network adjusts the weights every time it makes an incorrect prediction to reduce an optimisation function. Neural networks form the core of Deep Learning and are behind many of the most popular products prevalent in the industry today, image classification, speech recognition (Cortana, Siri),

recommendation systems (Netflix, Amazon Prime), and self - driving vehicles (Tesla, Comma.ai). Availability of massive datasets and superior computing power (largely owing to the GPUs from the gaming industry) has made developing large neural networks possible in reasonable time. The aim of this work is to utilise the power of these neural networks and data processing algorithms to create a model that can predict the methane conversion percentage from reaction and experimental parameters. This will also help in understanding the correlation between individual reaction features and the overall efficacy of the reaction. The followings sections are outlined as follows: Literature review of the current research being carried out in the field, a broad overview of the data pre-processing techniques to be used for the dataset, an introduction to neural networks including forward propagation, backward propagation, gradient descent and activation functions followed by a section of output analysis and possible difficulties that may occur during the course of the project. Everything will be summed up in a brief conclusion alongside some ideas for future work.

1.1 On the application of Artificial Intelligence in Dry Reformation of Methane

The concept of Artificial Intelligence (AI) can be traced back to 1950 when the British computer scientist Alan Turing in his landmark paper [4], “Computing Machinery and Intelligence” asked the question: “Can machines think?”. In an endeavour to answer this question, modern AI evolved into what it is today. In the year 1956, Marvin Minsky coined the term AI in his conference “Darthmouth Summer Research Project on Artificial Intelligence”. He wanted to pique the interests of the scientific community in the technology that he believed could demonstrate human like intelligence in the following 8 to 10 years. However, the hype surrounding AI far exceeded what was practically achievable by contemporary computers. As a result of this, public interest, research funding started to dry up for a long period of time, colloquially known as the “AI winter”. Since the 1990s interest in AI began to revive and it led to the discovery of many ground-breaking algorithms in natural language processing, reinforcement learning etc. In 1997, IBM’s Deep Blue software beat the Russian chess grandmaster Gary Kasparov. Google’s Alpha Go defeated the Chinese go champion Ke Jie. Since these events, AI saw remarkable interest among researchers and engineers. Fast forward to present day, with the advent of superior computing power, AI algorithms have been integrated deep into the society to perform highly specialised tasks. From

natural language processing algorithms to power smartphone voice assistants like Apple's Siri, Microsoft's Cortana to Recommender systems used by Netflix, Amazon, Uber to improvise overall customer experience. The competition in the field of AI has never been fiercer and this led to opportunities which could potentially revolutionise every industry in today's world.

Superior computing power combined with the advent of "Big Data" means machine learning and deep learning algorithms will have large amounts of data to train on. Because quality data forms the backbone of learning algorithms, this translates to better predictive power of the ML/DL models. In the context of this work, better prediction abilities mean researchers could have a clear idea of which parameters would produce the maximum yield of methane conversion. Without the use of machine learning, collecting data would mean investigating empirical evidence using valuable resources like time, money, and personnel only to receive subpar results. Using an optimised set of parameters predicted by a ML algorithm would also mean one would have a very clear idea of what to expect even before the reformation reactors are fired up. This would make the whole procedure of DRM experiments way more efficient. This translates into researchers to using those data to create new catalyst formulations. As shown by [10], many catalyst formulations remain undiscovered, and hence focusing on development of novel catalysts and utilising intelligent learning algorithms to predict outcomes of those catalysts would accelerate the process of making DRM reaction an industrial success.

While ANNs make a prediction, different weights are assigned to the connections between the layers of neurons and these weights depend on the importance of a particular variable in making the final output. Upon investigating these connections, better understanding can be gained on how certain parameters effect the overall methane conversion yield. Consequently, tweaking those parameters during empirical studies would result in better outputs. [2] investigated the feature importance of the parameters by removing one column at a time and retraining the model every time. While this is a robust approach in understanding feature importance, it takes a lot of time and computational power to execute. A faster approach is undertaken in this work – using Permutation Importance function from the Scikit learn library in Python. A detailed discussion of this technique and the obtained results is presented in chapter 5, section 5.4.

2. BACKGROUND AND LITERATURE REVIEW

Industrially, the DRM reaction is yet to be popular. However, one paper by Peter Mølgaard Mortensen and Ib Dybkjær [5] investigated the SPARG (Sulfur PAssivated ReforminG) process in an industrial scale and demonstrated how adding H_2S to the feed can deter the carbon formation in the reformer tubes. The foundation for this procedure steams from a study in 1986 by J. R. Rostrup and Nielsen [6] where they showed that Sulfur can block the sites where nucleation of carbon takes place. As a result, the duration of carbon free DRM reaction increases by a great extent. The authors also reported a 70% H_2S coverage can eliminate whisker carbon formation but cannot stop building up of amorphous carbon on the active sites. To mitigate the formation of Nickel sulfides, hydrogen is also added to the feed with a H_2S/H_2 ratio below 0.9. It was concluded that higher hydrocarbons present in natural gas were responsible for severe coking and therefore the process was optimised by removing the sulfur, and pre-reforming the higher hydrocarbons over Ni based catalyst. Later in the stage, H_2S was added to the feed. The adding of water in the pre reformation stage of the reaction made conducting the pure "Dry Reformation" challenging. In spite of the challenges, the feasibility was SPARG technique was demonstrated by a large-scale plant is Houston, Texas in the United States which conducted carbon free reformation process for 4 years.

Ayodele et al [7] applied Leven–Marquardt algorithm, Bayesian regularization and conjugated gradient algorithm to develop three neural network models to predict CO and H2 production rates from DRM reaction over Co/Pr_2O_3 catalyst. They used CH_4 and CO_2 partial pressure, and reaction temperature as input parameters to the neural network. The authors deployed 13 and 15 hidden neural units and reported the most accurate model was with the 13 neuron Level Marquardt algorithm model which most closely agreed with the experimental data, thereby producing a minimum error of $2.0526 \times 10E{-}17$. However, the authors did not report using catalyst preparation methods, W/F flow rate, On Stream Time as parameters which could potentially limit the understanding of the effect of these features on the final product yield.

Alsafar et al [8] studied the effect of nitrogen (N2) flow rate, reaction temperature and CH_4/CO_2 ratio on carbon deposition on alumina supported Cobalt catalyst. They used a

Radial Basis Function neural network (RBFNN) on a dataset of 85 data points for each model. This study illustrated that reaction temperature is by far the most important factor leading to the most important drawback of the DRM reaction: catalyst deactivation due to carbon deposition. However, the small dataset and the limited number of input parameters may have affected the generalisation capacity of the models developed and therefore limited its use in practical situations despite agreeing closely with experimental results.

In another study by Ayodele et al [9], the authors studied the effect of reaction temperature, CH_4 and CO_2 partial pressures and their respective flow rates on the final CO and H2 yield. They deployed Artificial Neural Network (ANN) and Box Behnken Design (BBD) models and reported accuracies of 74.84% for CH_4 and 76.49% for CO respectively at an optimal temperature of 728 ºC. Performance of the neural network was determined with three parameters: Mean Squared Error (MSE), training error and coefficient of determination (R^2). This further demonstrated the applicability of ANN models in catalysis research.

Sener et al [2] applied decision trees and ANN to find the optimal conditions to extract methane from the DRM reaction from a dataset of 5521 data points collected from 101 publications. According to their study, operational conditions have relatively more effect on the final output (CH_4 conversion) than catalyst preparation variables. Their finding was that the reaction temperature is the most important parameter in the DRM reaction that affects the conversion rate. This agreed with the findings of the previous literature as well as this work. This is followed by active metal and catalyst support type had the highest effect regarding catalyst preparation variables. The neural network architecture they deployed had one hidden layer of 20 neurons with the sigmoid activation function. Fidelity of the model was validated by 4-fold cross validation and testing on unseen data. They achieved error rates of less than 20% for most of the data the model was tested on. The dataset compiled in this paper is the backbone of this dissertation.

As evident from Chapter 4, the data that this work is based on is highly concentrated in a very narrow range of experimental conditions. For example: Nickel is the most used catalyst in the dataset, wet – impregnation method of catalyst preparation constitutes about 55.82% of the data. Similar characteristics can be observed from all other parameters as well. This highly affects the predictability power of a machine learning

model. An ML model is only as good as the data it trains on and this limits the ability of the model in a large section of the lesser studied experimental space. To overcome this problem, Smith et. al. [10] used principal component analysis (PCA) and constrained principal component analysis to map the high dimensional experimental data into a low dimensional information space. They reported that since the reported literature on DRM were clustered in a small region in the information space, the predictability power of the ANNs were highly affected. Furthermore, the traditional descriptors which are reported alongside DRM reaction papers were insufficient for the ML models to find formulations for novel catalysts. It further revealed that reaction condition variance did not provide as much information as the variance of catalyst formulations. They also attributed this to the fact that 90% of the reported literature had variances in the reaction conditions and not on the formulations. There were only 187 unique variations of catalysts in the literature. This result implies that there remains a lot of undiscovered catalyst formulations which may form the motivation for future work.

3. ARTIFICIAL NEURAL NETWORKS

3.1 Introduction

Neural networks are an interconnected layer of neurons (or units) designed to map inputs from one vector space into another vector space, the output. A typical ANN has three types of layers, the input layer, where the data is fed, the output layer and the hidden layer between them. Neurons of one layer are connected to every neuron in its preceding and successive layers. The connections are defined by a real number known as the weight coefficient. The weight coefficient is determined based on the importance of that connection. Optimising the weight coefficients is termed as the learning of the ANN.

3.2 Forward Propagation

During training each neuron calculates the weighted sum of its input and a bias term with the following equation:

$$y = wx + b \quad (4)$$

where y is the output of the neuron, w is the weight of the connections between subsequent layers of the network, and b is the bias term. This output is then passed through an activation function to calculate the output for the neuron in the next layer.

$$z = \text{sigmoid}(y) \quad (5)$$

Activation functions are used to introduce non linearities in the network and to fit the values of the neuron outputs between a reasonable range. There are different kinds of activation functions for e.g. - Linear, ReLU, leaky ReLU, sigmoid, logistic sigmoid, the application of which depends on the complexity of the output, as some activation functions make training faster and some more accurate. For the purposes of the concerned problem, the sigmoid and ReLU functions will be used. The final output of this process is then applied to an optimisation function or error function, commonly RMSE (Root Mean Square Error) shown in equation 6. This calculates the difference between the ANN predicted output and the actual output. This is known as the forward propagation step in training.[11]

$$\text{RMSE} = \sum \frac{1}{2}(y - y_0)^2 \quad (6)$$

3.3 Backward Propagation

The weights are adjusted by the back propagation algorithm. [12] In this algorithm the gradient descent method optimisation technique is used to minimise the error between the calculated and the actual output. The partial derivative of the RMSE error with respect to the weight and the bias term is calculated to determine the degree of effect each synaptic connection has on the input and it is deducted from the previous weight proportionally (determined by the learning rate). Equation 7 and 8 illustrates the back propagation algorithm.

$$w_{ij}^{(k+1)} = w_{ij}^{k} - \alpha \left({}^{\delta E}/_{\delta w_{ij}} \right)^{k} \quad (7)$$

$$b_{i}^{(k+1)} = b_{i}^{k} - \alpha \left({}^{\delta E}/_{\delta b_{i}} \right)^{k} \quad (8)$$

where w_{ij} is the synaptic connection between weights of neurons i and j, α is the learning rate (α>0), b_{ij} is the bias term of the i^{th} neuron, k is the iteration number and E is the calculated error. One complete pass of forward and backward propagation

through the entire dataset is termed as one epoch. After iterating through several epochs, the ANN minimises the optimisation function by finding the global optima.

The following diagram (figure 3.3.1) illustrates the deep ANN architecture.

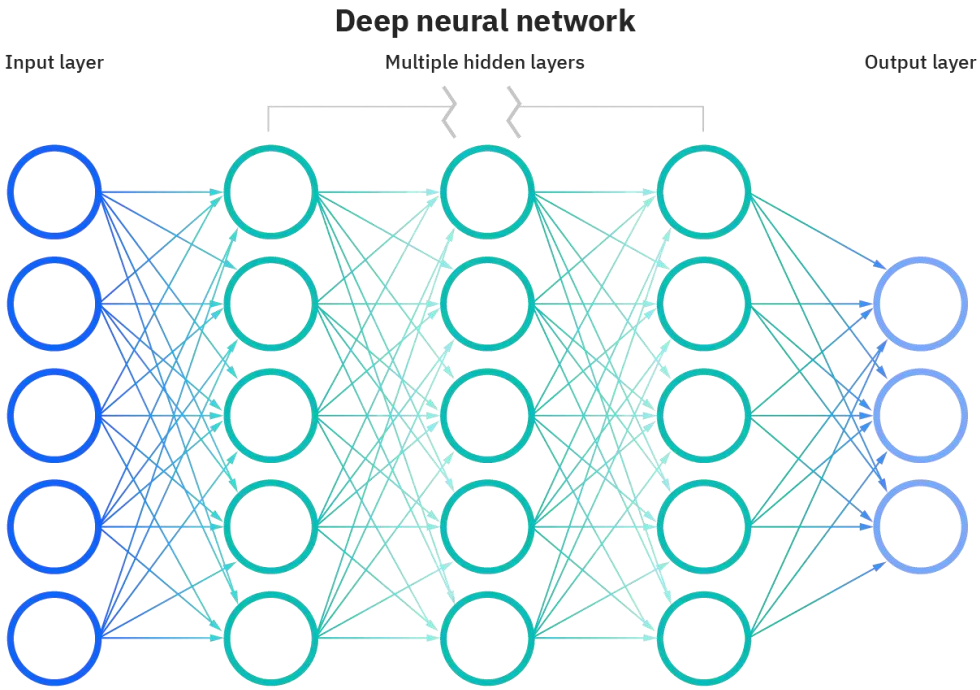


Figure 3.3.1: Artificial Neural Network architecture. [13]

3.4 Activation Functions

The output of a single neuron of a single layer is a weighted sum of its inputs and the bias term added to it. For the neuron to fire based on the relative importance of its input, a nonlinear function is required. This serves two purposes, first it allows the stacking of several layers of neurons together and second it makes the firing of the neurons computationally efficient by setting the boundary of its output to a relatively narrower range. Some examples of popular activation functions are sigmoid, tanH, ReLU, leaky ReLU etc. In this work, sigmoid and ReLU activation functions are mainly used due to the nature of the problem. These are described by the following equations:

$$sigmoid(y) = \frac{1}{1+e^{-y}} \qquad (9)$$

$$ReLU(y) = \max(0, y) \quad (10)$$

Figures 3.1 and 3.2 illustrate the ReLU and the sigmoid activation functions, respectively.* [14]

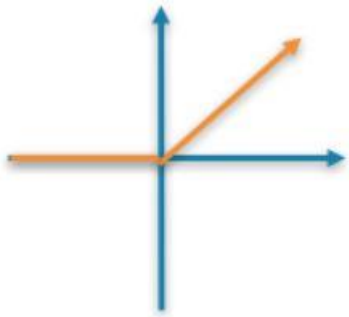


*Figure 3.1: ReLU activation function**

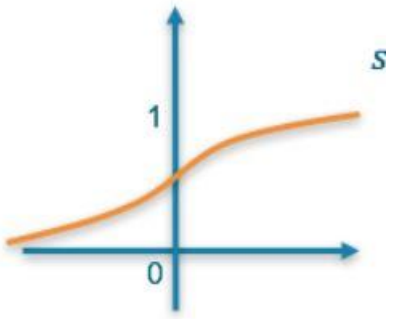


*Figure 3.2: Sigmoid activation function**

4. EXPLORATORY DATA ANALYSIS (EDA)

4.1 Dataset for the ML models

As evident from the literature, the generalisation ability of learning algorithms is directly dependant on the amount and quality of data it trains on. Hence, in any experiment involving deep learning techniques, constructing a meaningful dataset for training, validation and the testing of models is of paramount importance. Domain knowledge is required in making sure that the features and the outputs are correlated in some way or the other. Otherwise, the ANN will find patterns in the data that have no physical relevance. Overfitting and underfitting of the dataset is also taken into consideration when compiling data. Overfitting is where the ANNs essentially "memorises" the training data and shows great training performance and performs very poorly on data it didn't see before i.e., the test/validation set. Underfitting is where the ANN fails to detect the non-linearities in the data and under performs on all aspects. Care must be taken to make sure that the testing/validation datasets are completely isolated from the training set as this would lead to inaccurately high performance in experimental data and low performance when deployed in real experiments.

The dataset used in this work was compiled by Sener et. al. [2] by combing through 403 publications from 2005 – 2014. The final dataset contained 5521 datapoints extracted from 101 publications out of the 403 initially reviewed. The authors reported majority of papers used features and variables that weren't repeated enough times or

wasn't in a format which could allow data extraction, they had to be excluded from the dataset. The output variable, CH_4 conversion was chosen since it was reported in all the publications that were reviewed. The final dataset contained 63 independent features influencing the final output, CH_4 conversion expressed as a percentage. A future version of this work will aim to augment this dataset from research conducted between 2015 – 2021 alongside some other improvements.

4.2 Dataset Analysis

The following table (table 4.1) illustrates the input features for the ANN and their corresponding ranges and instances:

Parameters	**Range/Instances**
Base Metal	Co, Ce, Mn, La, Ni, Pt, Mg, Ir, Au, K, Li, Ru, Rh, Pd, Zr
Promoter	Ba, Ca, Cu, Sr
Catalyst preparation	Incipient to wetness impregnation, wet impregnation, co-impregnation, co-precipitation, sol-gel, sequential impregnation
Calcination time (h)	0-12
Calcination temperature (ºC)	25-900
Reduction time (h)	0-10
Reduction temperature (ºC)	0-850
Reduction H_2%	0-100
Catalyst support	γ-Al_2O_3(Gamma), α-Al_2O_3(Alpha), Mesoporous Al_2O_3 (MAl), La_2O_3, SBA-15, BaO, $BaTiO_3$, CaO, MgO, Nanocrystalline MgO, $MgAl_2O_4$(nano cry), CeO_2, SiO_2, Si_3N_4, TiO_2, PrO_2, V_2O_5, ZrO_2, Mesoporous nanocrystalline ZrO_2, MnO, MCM-41, Y_2O_3, ZSM-5, H-ZSM-5

Reaction temperature (ºC)	350-900
Feed CH_4 % volume	10-83.33
Feed CO_2 % volume	10-83.33
Feed Ar % volume	0-80
Feed N_2 % volume	0-75
Feed He % volume	0-80
GHSV (mg-min/mL)	0.15-88.235
Output variable (CH_4 conversion %)	0-100

Table 4.1: Dataset Analysis

As seen from table 1, the instances of different parameters are not in the same datatype (e.g., Catalyst preparation the values are in string while in reaction temperature its integers) and the numerical data are not always in the same range of values (e.g. - reaction temperature varies from 350 to 900 while calcination time varies from 0-12). To get around this, each string parameter is assigned one column and the values are one-hot encoded - the wet impregnation column will have a value of 1 every time it is used in a reaction and 0 when it's not. This is the same for base metal types, promoters, and catalyst support. To make the generalisation ability of the ANN model even better, the data from different ranges will be normalised using the following formula (equation 13):

$$X_{normalised} = \frac{X - X_{minimum}}{X_{maximum} - X_{minimum}} \quad (11)$$

As shown by previous studies that reaction temperature is the most important factor that determines the conversion rate, and the fact that DRM reaction is an endothermic one, high temperatures are employed for optimum performance and can be illustrated in figure 4.2. It shows that highest number of publications (2211) used temperatures around 700 ºC while fewest use temperatures around 400 ºC to 450 ºC.

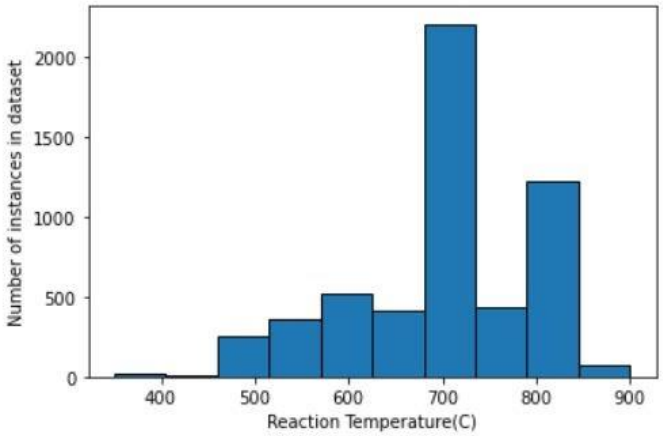


Figure 4.2: Reaction temperature distribution in the dataset in Celsius

Reduction temperature varies from 0 – 850 ºC, while 700 ºC being the most applied as shown by the figure (figure 4.3) below:

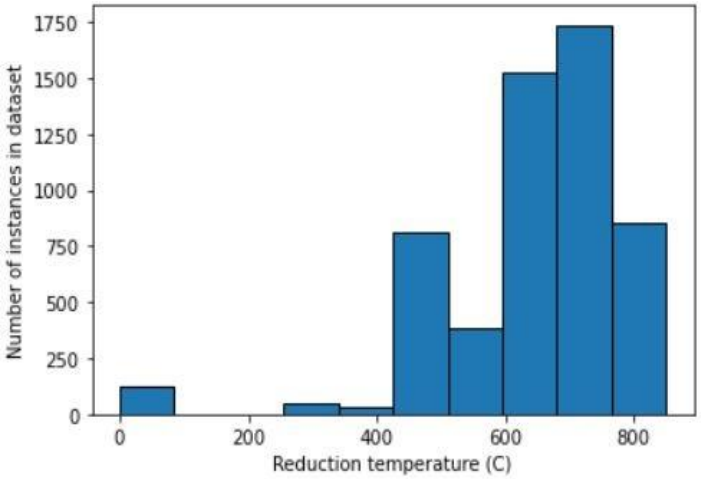


Figure 4.3: Reduction temperature distribution in the dataset in C.

CH_4 and CO_2 is fed into the reactor, since many authors added N_2, Ar and He to the mix. Hence a ratio of CH_4/CO_2 is taken and is illustrated in Figure 4.4, where it is evident that most authors preferred to keep an equal proportion of methane and carbon dioxide without adding any external gases. The maximum value noticed in the dataset is 5 and the minimum is 0.2. As seen from [4] in industrial plants H_2S is added in the feed tubes as well which is not yet studied by majority of the papers in the dataset.

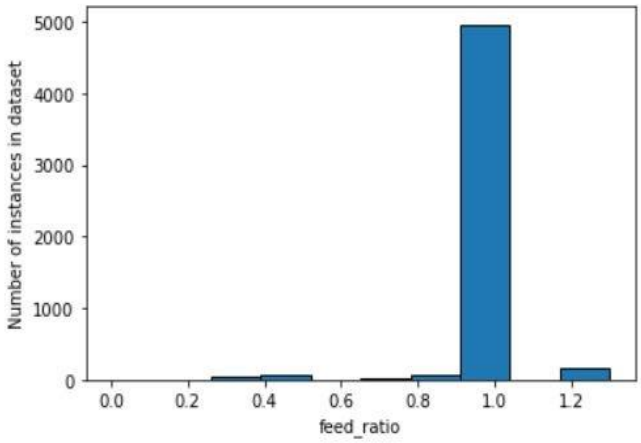


Figure 4.4: CH_4/CO_2 ratio distribution in the data set.

The prepared catalyst is calcined at high temperatures to remove impurities and is usually done for a maximum of 12 hours. Although the most optimum time reported in the literature is around 4 to 5 hours. Figure 4.5 depicts the distribution of the calcination time found in the dataset.

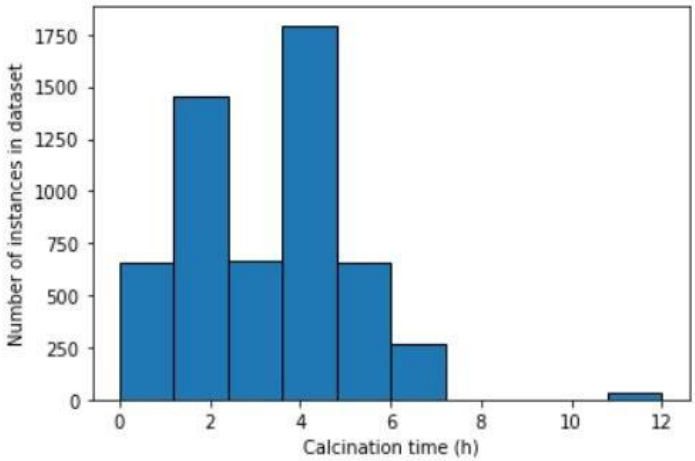


Figure 4.5: Calcination time in hours

Because of the rarity and high cost of noble metals, transition metals are mostly used as base metals to prepare the catalyst for the DRM reaction. Nickel is the most popular choice of researchers for the catalyst because of its availability and potential to be accepted for industrial production. The following pie chart (figure 4.6) shows the proportion of different base metals found in the dataset. Nickel accounts for 65.18% of all the base metals found in the literature, followed by 11.4% Platinum and the rest comprises other metals.

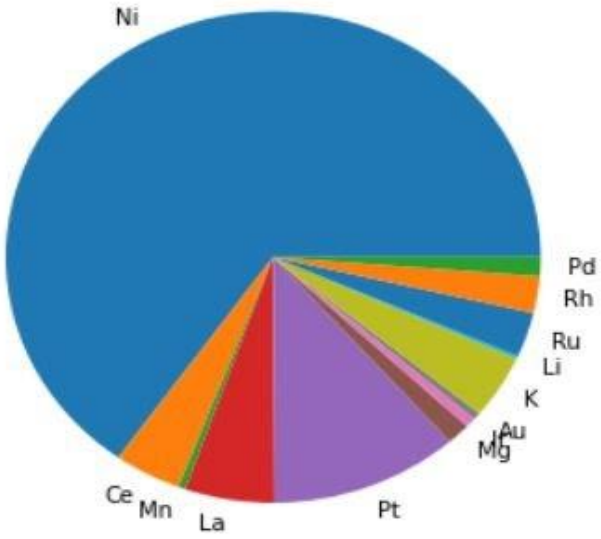


Figure 4.6: Pie-chart depicting the proportion of base metals in the dataset.

Figure 4.7(below) shows the relative proportion of catalyst preparation methods.

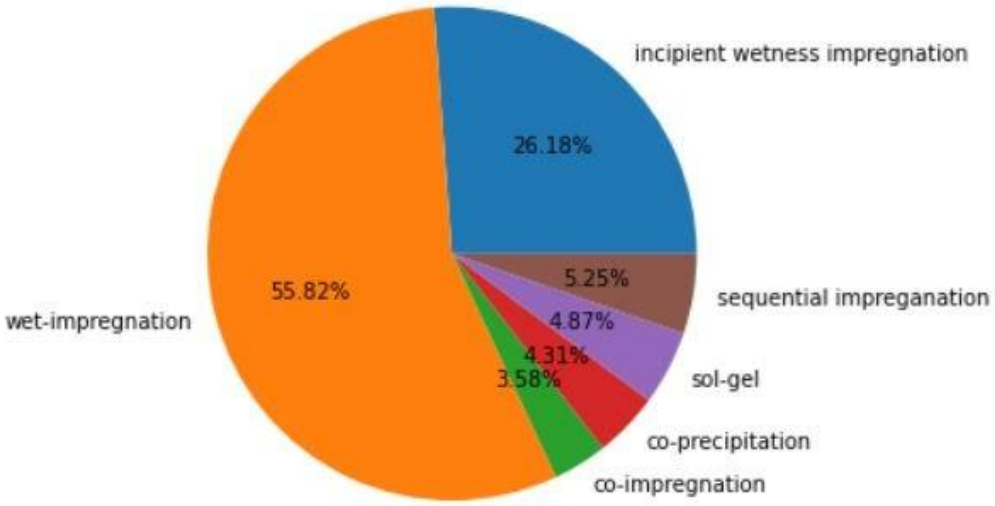


Figure 4.7: Pie-Chart illustrating the relative proportion of catalyst preparation techniques in the dataset.

According to previous studies, catalyst preparation has the highest effect among preparational conditions on the final output. In the dataset used in this paper, wet impregnation has the highest occurrence at 55.82%, followed by incipient to wetness impregnation at 26.18%. Only a few papers reported using co-impregnation technique for preparation of the catalyst at only 3.51%.

5. METHODOLOGY

5.1 Baseline Model

In order to gauge performance of various models developed, a simple linear regression approach has been taken to build a baseline model. Two performance metrics are chosen for this approach: Mean Absolute Error (MAE) and R^2 score provided by the Scikit Learn library. This model will also provide a fair idea of what can be expected in terms of prediction performance on unseen data. A multiple linear regression is a machine learning algorithm that learns a linear relationship between two or more feature variables from one target variable (methane conversion percentage in the present context). A multiple linear regression model is of the following form:

$$y = Wx + c \quad (12)$$

Where,

$$x = \begin{matrix} X_1^T \\ X_2^T \\ X_n^T \end{matrix} \quad (13) \qquad W = \begin{matrix} W_1 \\ W_2 \\ W_n \end{matrix} \quad (14) \qquad c = \begin{matrix} c_1 \\ c_2 \\ c_n \end{matrix} \quad (15) \qquad \text{and } y = \begin{matrix} y_1 \\ y_2 \\ y_n \end{matrix} \quad (16)$$

And T denotes transpose of the matrix, and Wx is the vector dot product between W and x.

The dataset is first split into a training set and a testing set in the ratio of 9:1. Upon testing the model on unseen data, it is found that the R^2 score is 0.65 and the MAE is 11.84. The MAE signifies that this model is on average 11% over or under in its predictions compared to the true values and the R^2 score signifies that the model explains 65% variation around the mean. Manual inspection of the dataset did not indicate any bias and hence these metrics can be employed to gauge performance.

Further inspection of the dataset revealed the presence of outliers in the data where methane conversion was 0%. Removing the outliers showed slight improvement in the model performance as can be seen in Table 5.1.1.

Model	R^2 score	MAE
With outliers	0.65	11.84
Without outliers	0.69	11.21

Table 5.1.1: Baseline model performance with and without outliers.

5.2 Artificial Neural Network Approach

ANNs are powerful tools to learn nonlinear relationships hidden deep inside large datasets which are not visible to manual inspection. Hence to create a model with higher performance metrics, a neural network approach is now undertaken. Different ANN architectures are studied in this work with varying number of hidden layers with varying number of nodes in every hidden layer. Adam optimiser is used for all the models and is trained for 300. Stochastic Gradient Descent algorithm did not perform well on this dataset when used on architectures with more than hidden one layer. The "exploding gradient" problem was noticed. This happens during the backpropagation step when error gradients consistently over the value of 1 are summed together resulting in huge weight values and eventually results in NaN (Not a Number). This makes the updating the weights of the network impossible rendering the model unstable. [15] To get around this problem, Adam optimiser was deployed. Introduced in 2015 by Deiderik P. Kingma and Jimmy Lei Ba [16], Adam is a combination of the SGD with momentum and the RMSprop algorithm. The weights and biases are updated using exponential weighted averages of the weight and bias gradients and making use of three hyperparameters, viz. $\beta_1, \beta_2 \; and \; \varepsilon$. In the Keras implementation of this optimizer the default values of these hyperparameters are used: $\beta_1 = 0.9, \beta_2 = 0.999 \; and \; \varepsilon = 10^{-8}$. Table 5.2 represents the performance on the training set of 4 different architectures of ANNs tested. Several iterations of training revealed no significant differences in performance before and after removing outliers and hence only the performance metrics in the dataset with outliers is reported as it can help improvise the generalisation ability of the models. The input layer of the networks contains 62 nodes as there are 62 feature descriptors in the dataset.

Model	Hidden layers	Node architecture	R^2 score	MAE	Epochs
Model A	1	32,1	0.79	7.53	300
Model B	1	64,1	0.84	6.52	300
Model C	2	32,64,1	0.88	5.38	300
Model D	3	32,64,128, 1	0.96	3.35	300

Table 5.2.1: Architecture and performance of different ANNs

The following figures (5.2.1, 5.2.2, 5.2.3, 5.2.4) illustrate the learning curves of the four models (A, B, C, D) respectively. MAE is plotted against the number of epochs. A learning curve is a plot that illustrates the performance of the model as training progresses. At the beginning of the training since the weights are randomly assigned, the loss function value is very high. Through back propagation as the weights and biases are adjusted to reduce the loss function, the curves are seen to fall. This is a clear indication that the models are improving with each iteration. It is evident however that around epoch 270 the curve has started to flatten with no further improvement implying the models have found the global minima of the loss function. A learning curve with error very close to zero would indicate that the model is overfitting and to mitigate this number of epochs have been restricted to 300.

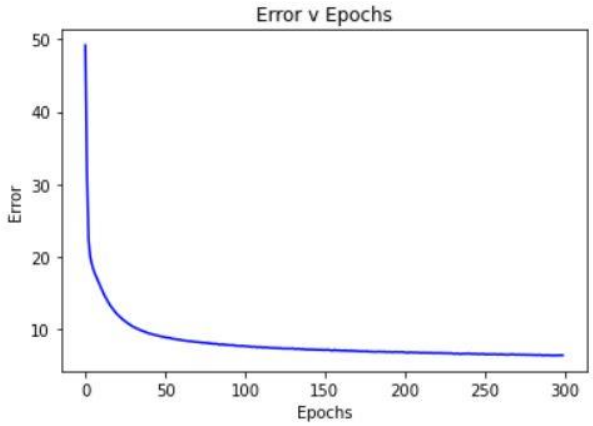


Figure 5.2.1: Learning curve of model A

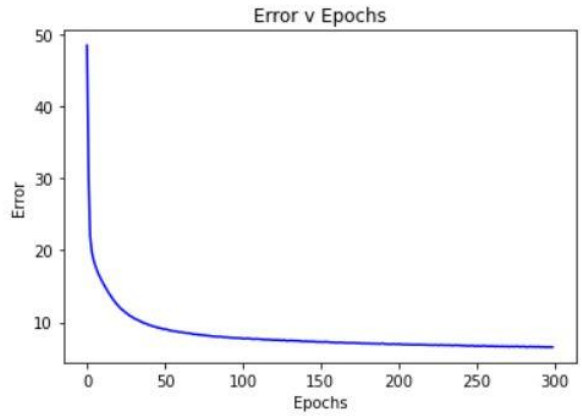


Figure 5.2.2: Learning curve of model B.

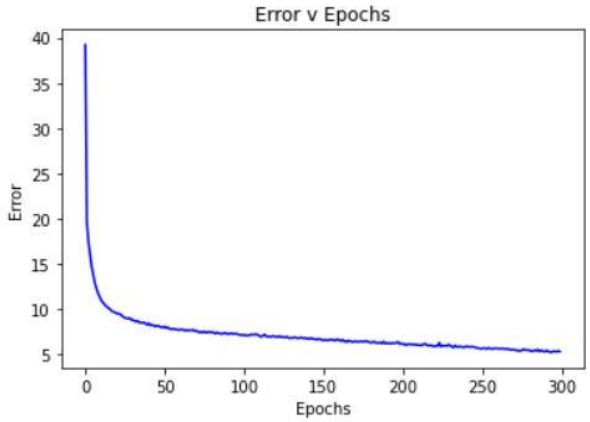


Figure 5.2.3: Learning curve of model C.

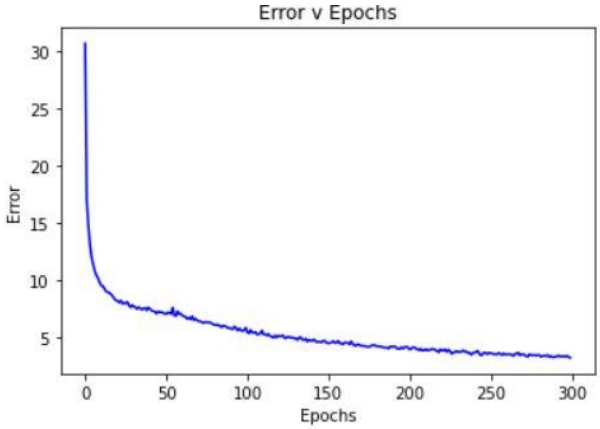


Figure 5.2.4: Learning curve of model D.

The results clearly indicate the Neural Networks perform significantly better than the baseline linear regression model. Between the models tested, there appears to be a correlation in lower training error and increased model complexity. Dropouts and early stopping methods are not investigated since the model's performance was quite satisfactory. Only the model with the best performance (model D) is tested on unseen

data and is explored further. It showed a test set MAE value of 3.74 very closely agreeing with the training MAE and a R^2 score of 0.94.

The following figure (figure 5.2.5) shows the training MAE and validation MAE curve of model D with respect to number of epochs. It appears that there has been slight overfitting near the end of the training cycle as seen from deviation of the validation MAE from the training MAE. In an ideal scenario, the validation/test curve would overlap with the training curve indication 100% perfect prediction in every datapoint. In a more practical sense, learning algorithms perform slightly worse on data it had not seen before. While a large deviation from the training curve would indicate severe overfitting, a slight deviation as shown in this case is considered normal. The predicted methane conversion values and the ground truth values closely agree with each other.

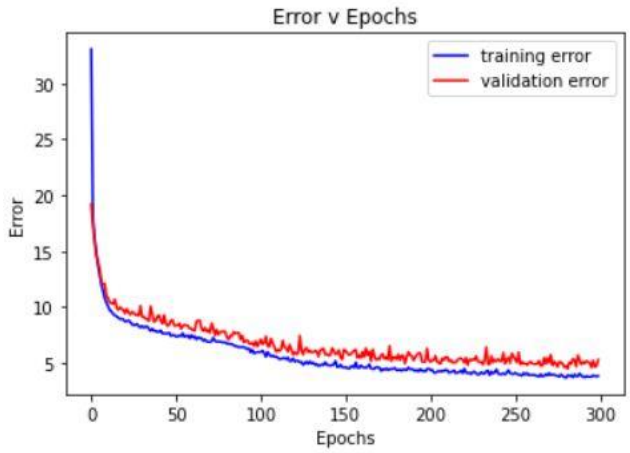


Figure 5.2.5: Training and Validation MAE curve of model D.

Upon training model D for another 600 epochs, it was observed that the deviation between the training error and the validation error increased consistently indicating clear case of overfitting. Hence, it can be concluded that 300 is the optimal number of epochs the model should train in a dataset of this size. However, in future if more data points are added to the training set, more epochs might be useful to train the model better.

Table 5.2.2 lists the performance metrics of model D across training, validation, and test sets. Validation set was 20% of the training set and test set was 10% of the original dataset.

Dataset	R^2 score	MAE	Loss
Train set	0.96	3.35	35.35
Validation set	0.88	5.28	67.87
Test set	0.94	3.74	N/A

Table 5.2.2: Final performance of model D across seen and unseen data

5.3 Performance comparison and scope of further work

As evident from tables 5.1.1 and table 5.2.2, the ANNs perform significantly better both in terms of training and testing on unseen data. However, there remains the issue of overfitting near the end of the training cycle which could potentially limit the generalisation ability of the model. This will form the basis for the next stage of this work. Sener et. al. [2] reported a RMSE and R^2 metrics of 4.23 and 0.97 respectively on the train set where model D achieved RMSE of 5.83 and R^2 value of 0.96. However, when testing on unseen data, the model presented in this work performed better with a RMSE of 6.82 as compared to 8.66 and R^2 of 0.94 as compared to Sener's 0.89. This indicates that the model presented in this work achieved better generalisation performance.

Next part of this work will focus on two components of one task: 1) Adding more data into the dataset from 2015 onwards. 2) Making sure all catalyst formulations, supports, promoters etc. have equal instances of examples in the dataset. Owing to the lack of sufficient examples in some instances (evident from the EDA section, Chapter 4), the model will demonstrate subpar performance when faced with those respective features.

A more comprehensive model can be created by retraining model D on data collected from reformation of methane using steam reforming, autothermal reforming and SPARG techniques. Dataset will have to be augmented with features relevant to these techniques and examples from the empirical data presented by researchers. Interest has been growing in investigating the synergy between plasma and catalysis in the scientific community. [17] has investigated using a combination of non-thermal plasma in cylindrical DBD reactor alongside conventional catalysts to optimise the DRM

reaction. Contrary to popular belief that plasma and catalyst work independently of each other, the authors reported that while the plasma creates an electric field and is responsible for the formation of reactive species (ions, electrons, neutral particles), the catalysts lower the activation barriers for many intermediate reactions in the DRM process. Most notably, plasmas can increase the surface area of the catalyst and form hotspots and enhance charge deposition. These are known to better activate the reactants in DRM. [18] studied the performance of the DRM reaction in a coaxial DBD reactor immersed in oil bath and suggested using a high GHSV and/or low applied voltage to increase the efficiency based on predictive modelling. Xin et. al. [19] studied the performance of DRM reactions with Ni catalyst over 4 supports (MgO, γ-Al_2O_3, SiO_2, TiO_2) inside a coaxial DBD reactor. They reported a CH_4 conversion of 44% with Ni/γ-Al_2O_3 with the highest energy and fuel production efficiency. From these papers it is evident that, adding the relevant features can enable a broader learning opportunity of the model by expanding to plasma catalytic techniques: Voltage in DBD reactor, power dissipation, reactor configuration, discharge volume, specific input energy, sinusoidal AC voltage frequency and diluent gases.

5.4 Understanding relative importance of experimental parameters.

While the final methane conversion yield can give a clear indication of how well the experiment is performing, understanding the parameters governing the output of the experiment is equally important. This can form a set of guidelines interpreted from the machine learning model, which can be used by researchers to efficiently tweak the parameters and produce higher volume of syngas. In this work, permutation importance function is used from the Scikit learn library. This function randomly shuffles every column of data in the dataset and uses a trained model (model D in this work) to observe how much the predictions change from the ground truth values. A smaller variation means that feature has little impact on how the predictions turn out to be and a higher variation suggests a greater importance of that feature on the output. The following diagram (Figure 5.4.1) illustrates the top 5 features which are used while making a prediction and their corresponding weights. The weights signify the degree of

impact individual features had on the final prediction.

Weight	Feature
1.1707 ± 0.1021	react Temperature (C)
0.4202 ± 0.0495	GHSV (mgmin/mL)
0.3869 ± 0.0509	γ-Al2O3(Gamma)
0.3349 ± 0.0385	Wet impregnation
0.2959 ± 0.0323	Ni
0.2851 ± 0.0777	Pt

Figure 5.4.1: Relative importance of feature variables in DRM.

As shown in figure 5.4.1, the reaction temperature parameter is the most important factor that the ANN model considers when making a prediction. This is for several reasons. As Jean Michel Lavoie [20] reports, CO_2 is very stable and oxidised and making it interact as an oxidant with other molecules requires very high temperatures. The temperature requirement of dry reformation process is in fact more than steam reforming or autothermal reforming. Using a proper catalyst can reduce the temperature at which the reactants are activated, while keeping the syngas yield high. The lower energy requirement is justified by the fact that external heat may contribute to further GHG emissions defying the main objective of catalysis research. Empirical evidence from Shamsi et. al. [21] shows how difference in reaction temperature can influence methane conversion over the same catalytic arrangement (Ca/Ni/K). At 750 ºC, CH_4 conversion was 44% and at 850 ºC, the conversion reached a maximum of 99.5%. Hence, choosing an optimum reaction temperature is of paramount importance when testing new catalytic formulations.

The second most important parameter used in predictions of the ANN is Gas Hourly Flow Rate or GHSV. It has been reported by Lotina et. al. [22] that methane and carbon dioxide conversions reach equilibrium as the GHSV is increased. Furthermore, a low value of GHSV means that the period of contact between the reactants and the catalyst is low resulting in low conversion amounts. They also reported that conversion amount of CO_2 and CH_4 should be equal and ratio of CO and H_2 in the syngas should ideally be 1 according to reaction stoichiometry. Any deviation from this value indicates the presence of secondary reactions (RWGS and/or carbon gasification by water). Therefore, GHSV should be controlled in a very strict manner to reduce unwanted reactions and achieve unity ratio of the products.

The importance of support in chemical reactions is well documented in the literature. However, in several instances a support is expected to be inert and simply provide sufficient surface area for dispersion of the catalyst. However, in the case of DRM, the support plays a crucial part in determining the overall conversion amounts and running time of the reaction. Aparicio et. al. in [23] showed the role of supports in the reaction by using Ru catalysts over Silica and Alumina supports (Ru/SiO_2 and Ru/Al_2O_3). Their results indicate that Ru/SiO_2 catalysts showed severe deactivation in first 3 hours of the reaction but Ru/Al_2O_3 showed no loss of activity in the same time frame. Unsurprisingly, Gamma – Alumina (γ-Al_2O_3) is the third most important feature that the ANN model uses to make predictions. In an industrial context, alumina's hardness, high melting point and low electrical conductivity can ensure stability in large number of chemical reactions including the DRM. It is reported in the literature that CH_4 conversion is highest when catalyst nanoparticles are dispersed on support [24] and because the γ-Al_2O_3 support can be made with a high surface area, this is widely preferred in the scientific community. [25] reported using ordered mesoporous alumina (<2 nm pores) with Ni catalysts and achieved near equilibrium conversions for a period of 100 hours with no catalyst deactivation. This occurred because the pores on the support helped keep the catalyst pellets in plate and thereby preventing sintering. Hence it can be concluded using alumina-based supports have a positive impact on the stability and performance of the DRM reaction.

The next parameter in the list is "wet impregnation" i.e., catalyst preparation. Wet impregnation is done by introducing the precursor metal to an aqueous solution. Then it is calcined at high temperatures to remove impurities and volatile compounds. Capillary action makes sure the solution is absorbed in the support through the pores. The relationship between catalyst preparation techniques and catalytic performance was examined by Huang et. al. [26]. The authors demonstrated how catalyst preparation variables has a distinct effect on the weight loading of the catalyst on the support and the interaction between the metal and support. Because of this, choice of catalyst preparation technique remains high in the list of important parameters in the DRM reaction. Some of the highest methane conversion percentages (>80%) has been reported by authors who used wet impregnation as their catalyst preparation method. [27]

From figure 4.7 it is seen that Nickel and Platinum are the two most widely researched catalysts in the field of DRM and hence they are found on the top 5 most important features. Hou et. al. [28] studied the effect of noble metals (Pt, Rh, Ru, Pd, Ir) and non-noble metals (Ni, Co) and reported higher catalytic activities on Ni and Co but higher coking resistance ability on the noble metals. However, this created an imbalance in the dataset with regards to some of the lesser researched elements (Mg, La, Au, Ce, K, etc) which can potentially limit the generalisation ability of the model presented in this work. Keeping the database updated with newer active metals, novel catalysts, then retraining the model will help mitigate this issue.

6. CONCLUSION

DRM reaction has been, for many years, the centre of interest for researchers looking to clean up the environment by putting methane and carbon dioxide to good use. The problem of optimising the conditions of the reaction, choice preparation of the catalysts has made significant breakthroughs in the recent years but still has many miles to go before it can be adopted by the industry. This work made use of powerful artificial intelligence algorithms to the DRM reaction field to guide future experimentation.

The proposed ANN model achieved higher test results on the data when compared with the model developed in [2]. At the same time, some degree of overfitting has been noticed during the training phase, which can be expected when making predictions on unseen data. The data was processed using normal regularisation techniques and one hot encoding before feeding it to the model. Upon training 4 models and selecting the model with the highest R^2 score on the test data, permutation importance library was used to understand relative importance of individual parameters. Based on the results, 5 parameters were presented which had the most impact on the prediction of final CH_4 yield. These results closely agree with the existing empirical evidence in the literature.

Since the quality and variability of the input data is paramount in learning algorithms, some motivations on future work have also been suggested. These include, but not limited to augmenting dataset with up-to-date research and including data from alternative DRM procedures like plasma catalysis, autothermal reforming, steam reforming. Technological advancements in the deep learning community means newer models of neural networks are constantly developed which when applied to the field of DRM can greatly aid in making this reaction a commercial success.

7. REFERENCES

[1] Jan van de Loosedrecht et. al., 2013, “Fisher-Tropsch synthesis: Catalysts and chemistry”, Elsevier, Vol. 2, Chapter 7.20.

[2] Ayşe Neslihan Şener et. al., 2017, “Statistical review of dry reforming of methane literature using decision tree and artificial neural network analysis.”, Catalysis Today, Vol. 299, pp. 289-302.

[3] McCulloch, W.S., Pitts, 1943, “A logical calculus of the ideas immanent in nervous activity”, Bulletin of mathematical biophysics, Vol. 5, pp. 115-133.

[4] A. M. Turing, 1950, “I. – Computing machinery and intelligence.”, Mind, Vol. LIX 236, pp.433 -460.

[5] Peter Mølgaard Mortensen, Ib Dybkjær, 2015, “Industrial scale experience on steam reforming of CO2-rich gas.”, Applied Catalysis A: General, Vol. 495, pp. 141 – 151.

[6] J. R. Rostrup, Nielsen, 1984, “Sulfur passivated Nickel Catalysts for Carbon free Steam Reforming of Methane”, Journal of Catalysis, Vol. 85, pp. 31 – 43.

[7] Bamidele Victor Ayodele et. al., 2019, “Artificial Intelligence Modelling Approach for the Prediction of CO-Rich Hydrogen Production Rate from Methane Dry Reforming”, Catalysts, Vol. 9, pp. 738.

[8] May Ali Alsaffar, Bamidele Victor Ayodele, Siti Indati Mustapa, 2019, “Scavenging carbon deposition on alumina supported cobalt catalyst during renewable hydrogen-rich syngas production by methane dry reforming using artificial intelligence modelling technique”, Journal of Cleaner Production, Vol. 247, pp. 119-168.

[9] Bamidele Victor Ayodele, Chin Kui Cheng, 2015, “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.

[10] Alexander Smith, Andrea Keane, 2020, “A machine learning framework for the analysis and prediction of catalytic activity from experimental data”, Applied Catalysis B: Environmental, Vol. 269, pp. 118-257.

[11] Daniel Svosil et. al., 1997, “Introduction to multi-layer feed forward neural networks”, Chemometrics and intelligent laboratory systems, Vol. 39, Issue 1, pp. 43-62.

[12] David E. Rumelhart et. al., 1986, “Learning representations by back-propagating errors”, Nature, Vol. 323, pp. 533-536.

[13] IBM Cloud Education, 2020, “AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What’s the Difference?”, IBM Cloud Education, Available at: https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks [accessed September 2, 2021]

[14] Stacey Ronaghan, 2018, “Deep Learning: Which loss and activation functions should I use?” [online], Towards Data Science, Available at: https://towardsdatascience.com/deep-learning-which-loss-and-activation-functions-should-i-use-ac02f1c56aa8 [accessed April 2, 2021].

[15] Pascanu R, Youshua Bengio, 2013, “On the difficulty of training recurrent neural networks”, Proceedings of the 30th International Conference on Machine Learning, Vol. 3, pp.1310-1318.

[16] Kingma, Diederik & Ba, Jimmy, 2014, “Adam: A Method for Stochastic Optimization.”, International Conference on Learning Representations, Journal paper.

[17] Asif Hussain Khoja, Muhammad Tahir, Nor Aishah Saidina Amin, 2019, “Recent developments non-thermal catalytic DBD plasma reactor for dry reforming of methane.”, Energy Conversion and Management, Vol. 183, pp. 529 – 560.

[18] Duc Ba Nguyen, Won Gyu Lee, 2015, “Analytical approach to estimate and predict the CO_2 reforming of CH_4 in a Dielectric Barrier Discharge.”, Chemical Engineering communications, Vol. 203, pp. 917 – 923.

[19] Xin Tu, Ya-Ling He, Bryony Ashford, Danhua Mei, 2016, “Plasma-catalytic reforming of biogas over Ni catalysts in a dielectric barrier discharge reactor: Effect of catalyst supports”, Plasma processes and polymers, Vol. 14.

[20] Jean-Michel Lavoie, 2014, “Review on dry reforming of methane, a potentially more environmentally-friendly approach to the increasing natural gas exploitation”, Frontiers in Chemistry, Vol. 2, pp. 2296 – 2646.

[21] Abolghasem Shamsi, 2006, “Partial Oxidation and Dry Reforming of Methane over Ca/Ni/K(Na) Catalysts”, Catalysis Letters, Vol. 109, pp. 189 – 193.

[22] A. Serrano Lotina, L. Daza, 2014, “Influence of operating parameters over dry reforming of methane to syngas”, International Journal of hydrogen energy, Vol. 39, pp. 4089 – 4094.

[23] P. Ferreira-Aparicioa, I. Rodr´ıguez-Ramos, J.A. Anderson, A. Guerrero-Ruiz, 2000, “Mechanistic aspects of dry reforming of methane over ruthenium catalysts”, Applied Catalysis A: General, Vol. 202, pp. 183 - 196

[24] Jin-Hong Kim, Dong Jin Suh, Tae-Jin Park, Kyung-Lim Kim, 2000, “Effect on metal particle size on coking during CO_2 reforming of CH_4 over Ni-alumina aerogel catalyst”, Applied Catalysis A: General, Vol. 197, pp. 191-200.

[25] Leilei Xu, Huahua Zhao, Huanling Song, Lingjun Chou, 2012, “Ordered mesoporous alumina supported nickel-based catalysts for carbon dioxide reforming of methane”, International Journal of hydrogen energy, Vol. 37, pp. 7497 – 7511.

[26] Y. J. Huang, J. A. Schwarz, 1986, “The effect of catalyst preparation on catalytic activity: The catalytic activity of Ni/Al2O3 catalysts prepared by wet impregnation.”, Applied Catalysis, Vol. 30, pp. 239 – 253.

[27] Ahmed S. Al-Fatesh, Muhammad A. Naeem, Anis H. Fakeeha and Ahmed E. Abasaeed, 2014, “Role of La2O3 as Promoter and Support in Ni/γ-Al2O3 Catalysts for Dry Reforming of Methane”, Chinese Journal of Chemical Engineering, Vol. 22, pp. 28 – 37.

[28] Zhaoyin Houa, Ping Chena, Heliang Fanga, Xiaoming Zhenga, TatsuakiYashimab, 2006, “Production of synthesis gas via methane reforming with CO_2 on noble metals and small amount of noble-(Rh-) promoted Ni catalysts”, International Journal of Hydrogen Energy, Vol. 31, pp. 555 – 561.

Appendix 1: Project Specification

1.1 Introduction

Carbon Dioxide and Methane contributes to 91% of the total greenhouse gas content in the Earth's atmosphere, most of which originates from the heavily industrialised society that we live in. To tackle the ongoing crisis of greenhouse gas emission, significant research is being carried out on technologies aimed at utilizing the abundance of these gases for production of less harmful alternatives.

$$CH_4 + CO_2 = 2CO + 2H_2$$

$$\Delta H° = 247.3 \text{ kJ/mol}$$

Dry Reformation of Methane (DRM) is a reaction which takes CO_2 and CH_4 as reactants and produces synthesis gas (syngas) as a product. Synthesis gas is a mixture of carbon monoxide, hydrogen (usually in the ratio of 1:1) and traces of carbon dioxide. Syngas is important in the production of long chain hydrocarbons and oxygenated chemicals through Fischer-Tropsch process.[1] The choice of catalyst in this reaction is of paramount importance and several catalytic materials have been studied in the scientific community. Although noble metals like Ruthenium (Ru), Rhodium (Rh), Platinum (Pt) have shown excellent performance, their high cost and rarity pose a major challenge to the commercialisation of the process. Supported nickel (Ni) catalysts, because of their high availability and low cost are more generally used albeit they come with their own set of challenges - deactivation at high temperatures, thermal sintering and coke formation on the catalyst surface.[2] Overcoming this barrier will allow the DRM reaction to be widely commercialised and would be a huge step in fighting climate change.

The aim of this work is to use machine learning algorithms to investigate the optimal combination of reaction conditions to produce the highest CH_4 conversion rate. These descriptors include - reaction temperature, calcination temperature, reaction time, calcination time, technique of catalyst preparation (co-impregnation, co-precipitation, sol-gel, sequential impregnation), catalyst support, reduction time and temperature. This project will focus on augmenting a dataset containing the reaction conditions data from the scientific publications in the field and train a machine learning model with it.

The results will help us investigate the relative correlation between the amount of methane conversion and the reaction conditions and hence better understand how to optimise those conditions to maximise conversion amount.

A dataset from a previous research will be augmented with new data points collected from the relevant literature. The data will then be normalised, and one hot encoded (for non-numeric data types e.g., catalyst preparation method). This processed data will act as an input to an Artificial Neural Network, which will then minimise the loss function (mean squared error) to calculate the total CH_4 conversion expressed as a percentage. The model accuracy can be tested with a validation/test dataset which will be done on a unseen fraction of the original dataset (20%) not used for training. The code for the neural network will be written in Python 3 on the Jupyter notebook environment and trained on a Nvidia RTX 2070 GPU.

1.2 Work Outline

Neural Networks work best when there is a large amount of data to train on. Therefore, in the first part of the project, a dataset from a past research will be taken and augmented with new data points from recent publications on Dry Reformation of Methane. The data points will include catalyst preparation variables (co-impregnation, co-precipitation etc) and reaction variables (reduction time, reduction temperature, reaction time etc), promotor type, sacrificial agent - which will act as input to the Neural Network model. Application of Machine Learning techniques will allow us to understand the correlation between the input variables and CH_4 conversion amount which is impossible to figure out by the naked eye from a large database. Understanding the effect of the variables on the conversion rate from the model can be used to aid future experimental work in the field.

Since the data will consist of different data types, and widely different ranges (catalyst preparation - string datatype, reaction time - float datatype, reaction time: 3 to 10 hours, reaction temperature - 350 to 900 C), it will require some pre-processing to make finding patterns in the data easier for the NN model. The numerical data will first be normalised by deducting the minimum value of the variable and dividing by the range of the variable. Nonnumeric variables will be expressed in One Hot Encoded vectors.

At this stage, the Python code for the NN model will be written and the model will be trained on 80% of the dataset. In terms of hardware, a Nvidia RTX 2070 GPU because of its dedicated CUDA cores which will train the model faster than conventional CPUs. The accuracy of the model will be tested on unseen data, the remaining 20% of the original dataset. Based on the results further improvements to the model will be made. The final model will be tested against the models developed by other researchers in this field.

1.3 Gantt Chart

TASKS	WEEK 1	WEEK 2	WEEK 3	WEEK 4	WEEK 5	WEEK 6	WEEK 7	WEEK 8	WEEK 9	WEEK 10	WEEK 11	WEEK 12	WEEK 13	WEEK 14
Literature Review														
Dataset development														
Data pre-processing														
Python Code														
Model training and testing														
Result Analysis														
Conclusion and report preparation														

Milestones:

1. ◆: Literature Review.
2. ◆: Dataset development complete.
3. ◆: Data pre-processing complete.
4. ◆: Python code development complete.
5. ◆: Model training and testing ready.

6. ◆: Result analysis complete.
7. ◇: Project report ready.

1.4 Risk Management

Risk	Likelihood	Severity	Risk rating	Mitigation
Supervisor quitting	1	5	5	Look for new supervisor
COVID19 infection	3	3	9	Seek medical assistance
GPU crash	2	3	6	Train model on Google Colab env.
Bad quality dataset	3	3	9	Rebuild dataset from scratch
Computer Crash	2	4	8	Use library computer
Inaccurate results	3	5	15	Make changes to code and data

0 – 5 = Low Risk 6 – 10 = Moderate Risk 11 – 15 = High Risk 16 – 25 = extremely high unacceptable risk		Severity of the potential injury/damage				
		Insignificant damage to Property, Equipment or Minor Injury 1	Non-Reportable Injury, minor loss of Process or slight damage to Property 2	Reportable Injury moderate loss of Process or limited damage to Property 3	Major Injury, Single Fatality critical loss of Process/damage to Property 4	Multiple Fatalities Catastrophic Loss of Business 5
Likelihood of the hazard happening	Almost Certain 5	5	10	15	20	25
	Will probably occur 4	4	8	12	16	20
	Possible occur 3	3	6	9	12	15
	Remote possibility 2	2	4	6	8	10
	Extremely Unlikely 1	1	2	3	4	5

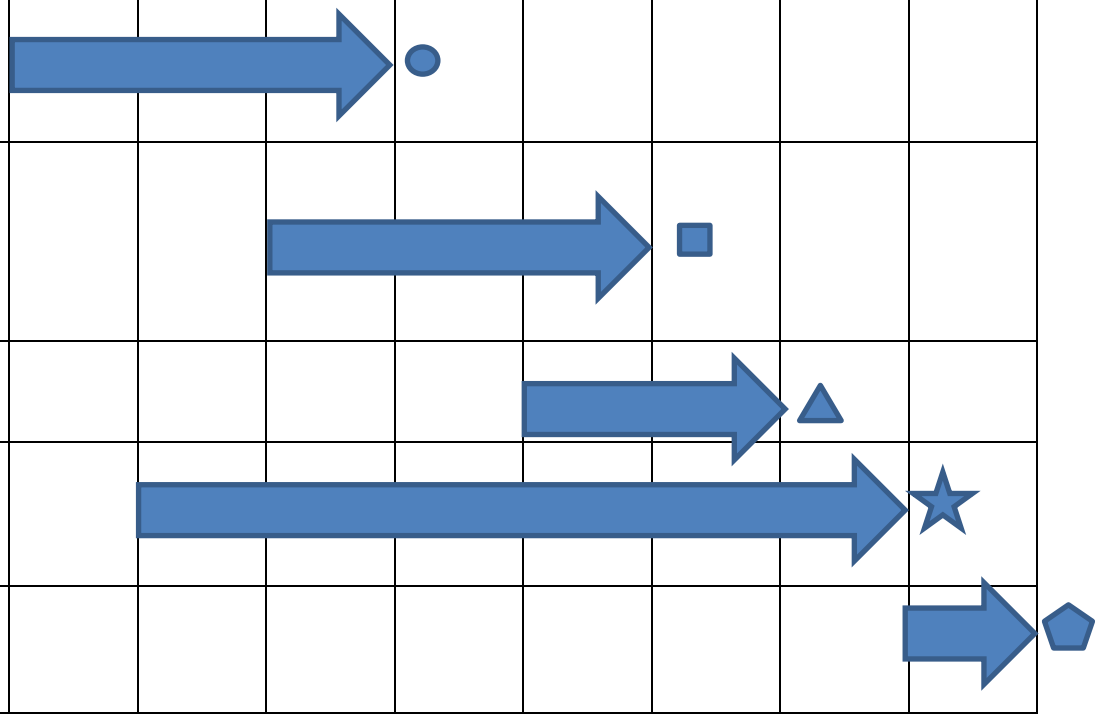
References:

[1] Z. Alipour, M. Rezaei, 2014, F. Meshkani, J. Ind. Eng. Chem. 20, 2858–2863

[2] Ayşe Neslihan Şener et. al, 2017, “Statistical review of dry reforming of methane literature using decision tree and artificial neural network analysis”, Catalysis today, Vol 299, 920-5861

Appendix 2: Updated Gnatt Chart

TASKS	WEEK 5	WEEK 6	WEEK 7	WEEK 8	WEEK 9	WEEK 10	WEEK 11	WEEK 12
Model improvisation								
Further testing and optimisation								
Result analysis								
Report preparation								
Submission								



Milestones:

1. : Final model completion.
2. : Testing against similar models in the literature and optimisation completion.
3. : Result Analysis completion.
4. : Report preparation and conclusion completion.
5. : Final dissertation submission.