

Contents

Abstract	vii
Acknowledgments	ix
1 Introduction	1
1.1 Motivations	1
1.2 Machine Learning Fundamentals	1
1.3 Previous Work in Machine Learning & Catalysis	5
1.3.1 Experimental Catalysis	5
1.3.2 Scaling Relations	7
1.3.3 Explored in this Work	7
2 Methods	9
2.1 Original Dataset	9
2.2 Data Disqualification	9
2.2.1 Cleaning By Beta's	10
2.2.2 Limiting to LT-WGS	12
2.2.3 Normalizing the Data	12
2.2.4 Removing Au/CeO ₂ and Pt/CeO ₂ Instances	13
2.3 Model Development	14
2.4 Attribute Determination	15
2.4.1 Primary Metal Descriptors	16
2.4.2 Support Attributes	17
2.4.3 Promoter Attributes	18
2.4.4 Resulting Domain Space	18
2.5 Determining Architecture	19

3	Results	23
3.1	Final Neural Network	23
3.2	Scaling Relations	25
3.2.1	Following Scaling Relations	25
3.2.2	Breaking Scaling Relations:	
	Au/CeO ₂ & Pt/CeO ₂ Example	25
3.3	Analysis of Fitting Untrained Data	27
3.3.1	Primary Metals	27
3.3.2	Supports	29
3.3.3	Promoters	30
4	Discussion	31
4.1	Conclusions	31
4.2	Limitations	32
4.3	Future Work	33

List of Figures

1.1	Example of an artificial neuron	2
1.2	Simple feed-forward artificial neural network	3
2.1	Distribution of the β values after removal of instances which did not have thermodynamic consistency or were near equilibrium	11
2.2	Distribution of temperatures after data removal by β values and lim- iting to LT-WGS	12
2.3	Histograms of forward rate distribution prior to normalizing	13
2.4	Normalized forward rate histogram	13
3.1	Parity plot for training performance of the ANN	23
3.2	Parity plot for testing performance of the ANN	24
3.3	ANN performance with Au/CeO ₂ and Pt/CeO ₂ instances	27
3.4	Primary Metal Interpolation Results	28
3.5	Support Interpolation Results	29
3.6	Promoter Interpolation Results	30

List of Tables

2.1	Number of data points lost by each cleaning criteria	10
2.2	ANN Performance when trained with and without Au/CeO ₂ and Pt/CeO ₂	14
2.3	ANN Model Training Parameters	15
2.4	Binding energy (eV) of relevant species on noble metal surfaces . . .	17
2.5	Comparison of binding species for primary metal attribute	17
2.6	Comparison of chemical properties for support attributes	18
2.7	Comparison of chemical properties for promoter attributes	18
2.8	Domain of the ANN after cleaning and attribute selection	19
2.9	Catalyst components in the domain	20
2.10	Optimization of ANN Architecture	21
2.11	Rigorous optimization of ANN architectures	21
3.1	Error Analysis from ANN Training	24
3.2	Sensitivity analysis results	25
3.3	Optimum catalyst search results	26

Abstract

Statistical data mining techniques, such as artificial neural networks (ANN), serve to address limitations of the current catalyst development and optimization approaches. Data mining techniques in conjunction with large catalytic data sets, should be able to assist in identifying meaningful materials descriptors, trends in catalytic reactivity and mechanistic differences between catalysts. This work explores the application of ANNs as sophisticated tools for identifying scaling relations using experimental low-temperature water-gas shift reaction data. We found that the ANN developed was capable of predicting catalytic activity for catalysts within the identified domain. Additionally, the ANN developed assisted in identifying catalysts which were not consistent with the ANN-optimized scaling relations.

Acknowledgments

I would like to take this opportunity to acknowledge those who have supported me through this scientific journey.

Prof. Dumesic and Prof. Huber, for providing the opportunity to work on this project, as well as intellectual and intuitive insight along the way.

WARF 2020 for financially supporting this project.

The educators and mentors who played a fundamental role in preparing me for a STEM career. Mrs. Jill Locke, who taught graphing using baby bunnies and never batted an eye at the many snakes and lizards I brought to her classroom. Dr. Susie Knetter for demonstrating how to wear the many hats of a successful scientist-friend-wife-mother. Prof. Detamore for introducing me to academic research and guiding me through the graduate school admission process.

The fellow scientists and friends who have been by my side for the past 18 months. My parents, Paula & Jeff Livingston, for being excellent role models and continually supporting my ambitious attitude and academic growth. From letting me be a turtle-catching, squirrel-raising little girl to driving me across town to participate in the IB Program to demonstrating curiosity and perseverance each day, they have contributed immeasurably to my successes.

My husband, Gehrig Keane, for always believing in me.

Chapter 1

Introduction

1.1 Motivations

Statistical data mining techniques, such as artificial neural networks (ANN), serve to address limitations of the current catalyst development and optimization approaches (e.g., careful catalyst synthesis, characterization, and reactivity studies, combined with detailed density functional calculations). Data mining techniques in conjunction with large catalytic data sets, should be able to assist in identifying meaningful materials descriptors, trends in catalytic reactivity (scaling relations) and mechanistic differences between catalysts. This document details the exploration of these ideas for the water-gas shift (WGS) reaction (i.e., $\text{CO} + \text{H}_2\text{O} \rightleftharpoons \text{CO}_2 + \text{H}_2$).

1.2 Machine Learning Fundamentals

Data mining, the process of identifying patterns in large datasets, is a subfield of computer science at the intersection of statistics and machine learning (ML). Data mining algorithms enable computers to make predictions for complex, highly dimensional problems based on previously-encountered data. This project employs the use of artificial neural networks (ANN). ANNs are similar to traditional curve-fitting tools, which define an input-output relationship. However, ANNs differ from traditional curve-fitting models in their ability to fit highly dimensional datasets, such as those encountered in the field of catalysis.

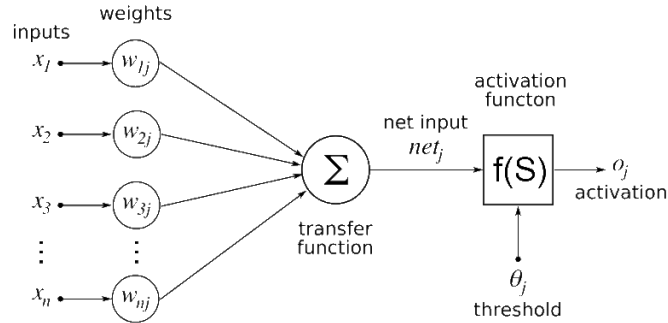
When developing an ANN, the model must first be ‘trained’ using a set of examples referred to as ‘training data’. The training algorithm optimizes the model using training data to tune the model’s mathematical parameters. While ANNs are

frequently considered ‘black boxes’, they actually consist of complex, statistically optimized mathematical functions. Trained with an appropriate data set, ANNs can be used to predict an optimum solution to a mathematically described problem.

To place this work in the larger context of ML, the ANNs discussed in this document are a form of ‘supervised’ ML. Each example or ‘instance’ in the training data set consists of input information (a vector) and corresponding output information. This training data is considered ‘labeled’ because each input instance is ‘labeled’ with a corresponding output.

Architecturally, ANNs are similar to biological neural networks, storing and transferring information with numerical values rather than electrical impulses. Similar to a biological neural network, an ANN consists of many inter-connected ‘neurons’ or ‘perceptrons’, such as the one in figure 1.1. For ‘feed-forward’ networks, each neuron receives input signals from the preceding layer and sends an output signal to each neuron in the following layer [16]. Each input signal has an indepen-

Figure 1.1: Example of an artificial neuron



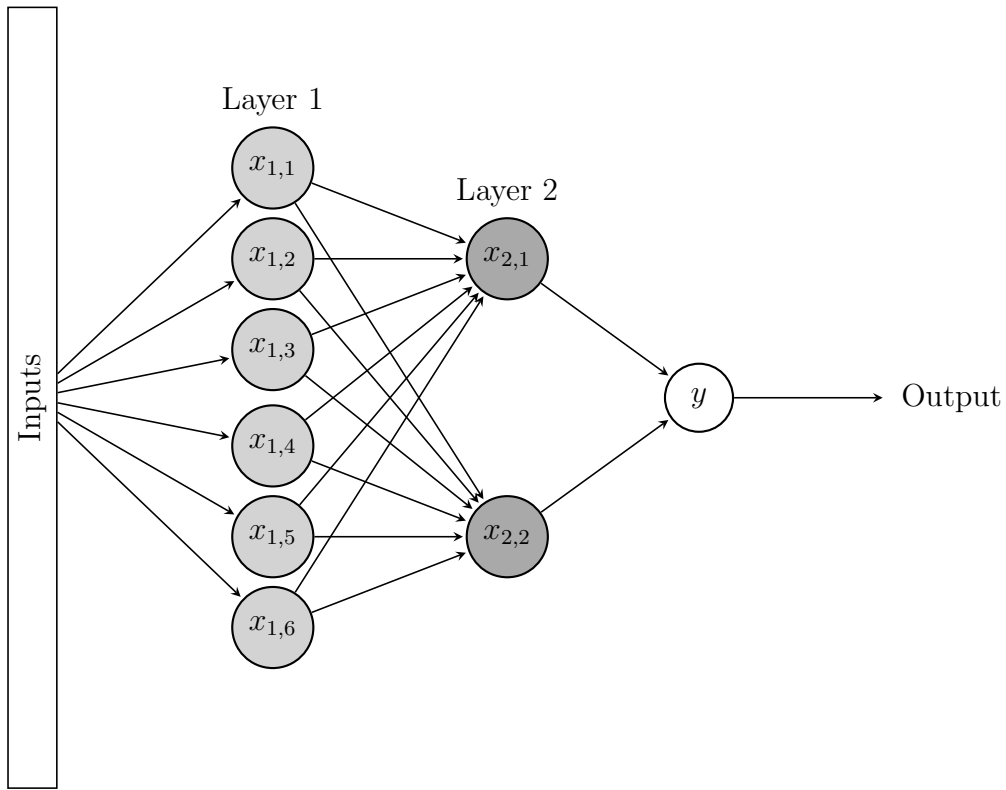
dent value. These input values are summed and then transformed by an ‘activation function’ to produce a single output signal. Activation functions are typically step, linear or sigmoid functions. Mathematically, a neuron’s output is determined by the activation function applied to the summation of the weighted inputs, shown in equations 1.1 and 1.2.

$$input = w_1x_1 + w_2x_2 + \dots + w_nx_n = \sum w_ix_i \quad (1.1)$$

$$output = f(input) = f\left(\sum w_ix_i\right) \quad (1.2)$$

When these neurons are linked together into a network, shown in figure 1.2, the mathematical model becomes a complex function of weighted summations and activation functions. While many types of neural networks exist (i.e., deep, convolutional, relational, etc.) this report focuses on feed-forward neural networks with one or two layers of hidden neurons. Furthermore, information only travels in a forward-direction, there is no recursion or feedback within the network. The networks explored in this project only have a single output neuron.

Figure 1.2: Simple feed-forward artificial neural network



Consider the case with two hidden layers and a single output where the first hidden layer contains six neurons and the second layer contains two neurons, demonstrated in figure 1.2. Note that each of the neurons has the properties shown in figure 1.1. To illustrate the mathematics contained in the ‘black box’, begin by considering the output, y as a function of the neurons $x_{2,1}$ and $x_{2,2}$ and their associated weights in the preceding layer, equation 1.3. Similarly, the output from the $x_{2,1}$ and $x_{2,2}$ neurons is determined by the neurons in first hidden layer ($x_{1,1}$, $x_{1,2}$, $x_{1,3}$, $x_{1,4}$, $x_{1,5}$, $x_{1,6}$). The mathematical relationship between the first and second hidden

layers is demonstrated in equations 1.4 and 1.5. Finally, the output for each neuron in the first layer is given as a function of the network inputs (in_1 to in_n), equation 1.6.

$$y = f(x_{2,1}, x_{2,2}) \quad (1.3)$$

$$x_{2,1} = f_{2,1}(x_{1,1}, x_{1,2}, x_{1,3}, x_{1,4}, x_{1,5}, x_{1,6}) \quad (1.4)$$

$$x_{2,2} = f_{2,2}(x_{1,1}, x_{1,2}, x_{1,3}, x_{1,4}, x_{1,5}, x_{1,6}) \quad (1.5)$$

$$x_{1,i} = f_{1,i}(in_1, in_2, \dots, in_n) \quad (1.6)$$

To determine the overall function representing the ANN shown, begin by nesting the input-output functions for each layer. First, replace the $x_{2,1}$ and $x_{2,2}$ inputs in equation 1.3 with equations 1.4 and 1.5 to yield equation 1.7. Then replacing the $x_{1,i}$ inputs in equation 1.7 with their corresponding equation 1.6 form yields the overall ANN representation in equation 1.8.

$$y = f[f_{2,1}(x_{1,1}, x_{1,2}, \dots, x_{1,6}), \quad (1.7)$$

$$f_{2,2}(x_{1,1}, x_{1,2}, \dots, x_{1,6})]$$

$$y = f[f_{2,1}(f_{1,1}(in_1, in_2, \dots, in_n) \quad (1.8)$$

$$f_{1,2}(in_1, in_2, \dots, in_n),$$

$$\dots,$$

$$f_{1,6}(in_1, in_2, \dots, in_n)],$$

$$f_{2,2}[f_{1,1}(in_1, in_2, \dots, in_n)$$

$$f_{1,2}(in_1, in_2, \dots, in_n),$$

$$\dots,$$

$$f_{1,6}(in_1, in_2, \dots, in_n))]$$

In general, increasing the size of the training data set improves predictions. A

common heuristic for determining the minimum size of the training data set is that there should be twice as many instances in the training data set as weights in the ANN [16]. For the most general feedforward ANN, the number of weights is determined by equation 1.9. For ANNs discussed in this work with two hidden layers and a single output, the number of weights is given by 1.10.

$$weights = n_{inputs} * n_{Layer1} + n_{Layer1} * n_{Layer2} + \dots + n_{LastLayer} * n_{outputs} \quad (1.9)$$

$$weights = n_{inputs} * n_{Layer1} + n_{Layer1} * n_{Layer2} + n_{Layer2} \quad (1.10)$$

1.3 Previous Work in Machine Learning & Catalysis

ML was initially proposed in the 1950’s with backpropagation methods incorporated in 1986 [34, 6]. Three years later, in 1989, Shigeharu Kito & Tadashi Hattori pioneered ML as a tool for catalyst optimization [19]. ANNs have since been explored for optimizing materials/catalyst properties and reaction conditions as well as approximating computationally intensive calculations [31, 4, 20, 35]. As experimental techniques, ML methods and computational power advance, new and exciting opportunities to merge these fields develop.

1.3.1 Experimental Catalysis

The earliest and most-explored application of ML in catalysis has been in using ANNs to approximate the reactivity or selectivity for differing catalyst compositions. This movement was led by Kito & Hattori who, in 1989, proposed INCAP (INtegration of Catalyst Activity Patterns), a system they claimed was capable of predicting the reaction mechanism, important catalytic functions, plausible side reactions, unfavorable products resulting from said side reactions and proposing catalyst candidates [19]. Kito & Hattori continued publishing on ML in catalysis

through 2006 [13]. Using ANNs they developed models to predict the acid strength of mixed oxides, product selectivity for the oxidative dehydrogenation of methane, catalyst deactivation in methanol conversion, identification of factors controlling catalytic activity [18, 17, 20, 13].

Kohji Omata & Muneyoshi Yamada also made significant contributions to the intersection of catalysis and ML. Together they published a series of papers exploring Radial Basis Function Networks (RBFN), which are similar to ANNs, for catalyst optimization. These networks optimized **CO loading** and preparation [27], promoter optimization for oxidation of CO and dry reforming of methane [26, 25, 21]. They also developed RBFNs for optimizing the preparation and composition of a Cu-Zn-Al-Sc-B-Zr catalyst for methanol synthesis, ternary-oxide Cu catalysts for steam reforming of methanol and promoters for Ni/activated carbon catalysts for carbonylation of methanol [31, 28, 36]. Finally, Omata & Yamada used RBFNs to optimize a Cu-Zn catalyst for dimethyl ether synthesis and a ZSM-5 catalyst for dimethyl ether to olefin conversion [30, 29]. For each application, their training data was largely generated ‘in house’ using high throughput techniques. As attributes, Omata & Yamada repeatedly chose to use physical properties (molecular weight, boiling point, etc) and little work was reported on attribute optimization.

Finally, Avelino Corma has published on several diverse catalytic applications for ANNs and proposed a Design of Experiments framework. Corma began using ANNs to predict catalyst activity and selectivity for oxidative dehydrogenation of ethane [4]. The developed ANN used molar catalyst composition to predict reactant conversion, product yield and overall selectivity. More recently, Corma et al. used ANNs with high throughput experimental techniques to optimize a Ti catalyst for olefin epoxidation [5]. Additionally, he looked at ANNs and decision trees to successfully predict material properties (e.g., stability) based on composition for the high temperature synthesis of zeolites [3]. Finally, Corma et al. describe a ‘soft-computing framework’ which combines a genetic algorithm, an ANN and high-throughput experimentation to rapidly optimize a catalyst within the defined

domain [37]. Additionally, Corma has demonstrated the use of ANNs to optimize not only catalyst composition, but also reaction conditions and scale-up approaches, focusing on n-paraffin isomerization [33, 32].

1.3.2 Scaling Relations

In catalysis, scaling relations have been successfully used to drive development of novel catalytic materials and predict catalytic activity [11]. Defining a scaling relation as any relationship in which a change to one variable causes a predictable and defined change in another variable and revisiting equations 1.2 through 1.8, it is clear that ANNs can be framed as a sophisticated scaling relationship tool. While each neuron alone performs a predictable transformation on its inputs, the network as a whole also fits within the definition of a scaling relationship.

Currently, scaling relations for heterogeneous catalysis relate elementary step activation energies to adsorption energies and thus reactivity of key intermediates [9]. However, it is well understood that catalysts cannot demonstrate activities above the peak of the volcano plot while obeying scaling relations [1]. Therefore, identifying materials that promote scaling relation violations may have significant implications for future catalyst development.

1.3.3 Explored in this Work

This work expands on the foundational work presented by increasing the catalyst domain space, incorporating diverse training data, optimizing attribute selection and incorporating DFT-derived descriptors. The laborious effort of Odabasi, Gunay and Yildirim resulted in a database of 4,360 data points for the water-gas shift (WGS) reaction derived from publications spanning 10 years (2002 to 2012) [24]. The data used for this work follows from the database constructed and published by Odabasi, Gunay and Yildirim.

Unlike the previous ANN studies cited, which largely rely on high-throughput data from individual research groups, the data in this study was collected in nu-

merous laboratories with diverse experimental equipment and procedures. While many of the works presented consider a narrow project scope (i.e., only varying a single parameter such as promoter or active metal), this work considers many different supports, active metals, promoters, synthesis techniques and reactor conditions. To my knowledge, this is the first study using ANNs to combine DFT attributes (binding energies) with chemical properties to predict the fundamentally-explained forward rate. Furthermore, this work begins discussion on applying ML to derive scaling relations and evaluate meaningful catalyst attributes.

Chapter 2

Methods

2.1 Original Dataset

A database of WGS reaction data was assembled and provided as Supplementary Materials by Odabasi et al. [24]. This data set is composed of 4,360 data points from 97 publications between 2002 and 2012. The Odabasi et al. explored trends in this data using data mining techniques including decision trees, ANNs and support vector machines. For their applications, the data was organized and described categorically by support, primary metal and promoter metal attributes. The optimized NN, trained with MATLAB’s neural network toolbox, predicted CO conversion. They report a training performance of $R^2=0.972$, $MSE = 5.42$ and a testing performance of $R^2=0.905$, $MSE = 10.05$ [24].

2.2 Data Disqualification

This study expands on the results of Odabasi et al. by developing an ANN with new attributes and outputs. Prior to ANN development, the original data was ‘cleaned’ to remove data that did not align with the objectives of this project or the constraints defined. This process reduced the training data set from 4,360 instances to 1,499 instances. The removal of data is summarized in table 2.1. Note that some instances violated two or more of the constraints discussed. Therefore, table 2.1 reports both the total number of data points that violate each constraint as well as the number of data points removed after each sequential constraint is applied.

Because the water-gas shift reaction is often studied in the context of steam-reforming of methane, 223 instances used co-feeding of methane. To eliminate con-

ditions with competing reaction pathways, these data points were removed from the training data set. Similarly, the 98 instances with an O₂ co-feed were removed.

To maintain consistency among the catalyst supports (single metal oxide), any data points collected using zeolite (ZEO), hydroxyapatite (HAP) or activated carbon (ACC) based supports were removed (149 data points). Additionally, the 953 instances with mixed metal oxide supports were removed. For similar reasons, all promoters in the final training data set were in their elemental form. To achieve this, 21 data points with yttrium-stabilized zirconia (YSZ) promotion were removed from the dataset.

Table 2.1: Number of data points lost by each cleaning criteria

Removal Criteria	Order Independent	Sequential Removal	Remaining Data Points
Original Data	-	-	4360
ZEO Support	66	66	4292
HAP Support	58	58	4236
ACC Support	25	25	4211
YSZ Promoter	21	21	4190
CH ₄ Co-Feed	223	223	3967
O ₂ Co-Feed	98	32	3935
$\beta > 0.8$	355	265	3670
$\beta < 0$	124	66	3604
β Not Calculable	177	176	3428
T <150°C	173	118	3310
T >350°C	730	440	2870
Mixed Support	953	642	2228

2.2.1 Cleaning By Beta’s

To compare and predict intrinsic catalyst activity, the data was cleaned to eliminate any instances which may have been affected by equilibrium or transport limitations. Furthermore, some instances in the dataset indicated thermodynamically impossible behavior (i.e., conversion that is impossible due to limiting reactants). The thermodynamic consistency of each data point was evaluated by determining the β value, where β is derived by equations 2.1 to 2.6. For each instance, β was determined using the primary metal loading, reaction temperature, total flowrate, CO conversion

and NIST thermodynamic parameters [23].

For an irreversible reaction the reverse rate constant, k_{rev} , must be zero and therefore, $\beta = 0$. When a reversible reaction is at thermodynamic equilibrium, $\beta = 1$. Because relevant kinetic data must be determined in a region without thermodynamic limitations, only instances with $0 < \beta < 0.8$ were retained. The resulting data distribution is given in figure 2.1.

$$rate = rate_{for} - rate_{rev} \quad (2.1)$$

$$rate = k_{for}[CO][H_2O] - k_{rev}[CO_2][H_2] \quad (2.2)$$

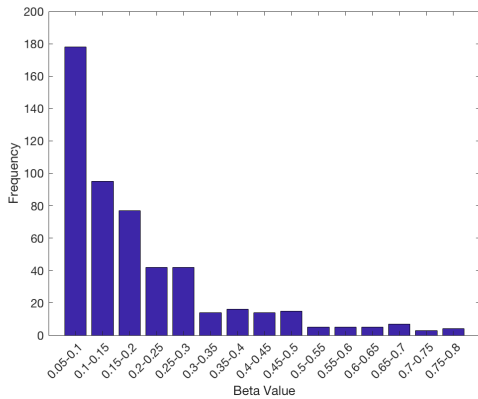
$$rate = k_{for}[CO][H_2O] \left(1 - \frac{k_{for}[CO][H_2O]}{k_{rev}[CO_2][H_2]} \right) \quad (2.3)$$

$$\beta = \frac{k_{for}[CO][H_2O]}{k_{rev}[CO_2][H_2]} = \frac{[CO][H_2O]}{K_{eq}[CO_2][H_2]} \quad (2.4)$$

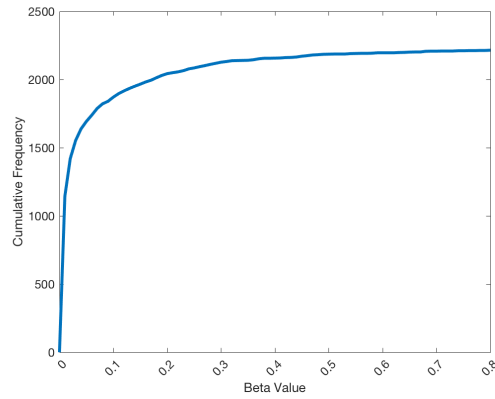
$$rate = rate_{for}(1 - \beta) \quad (2.5)$$

$$rate_{for} = \frac{rate}{1 - \beta} \quad (2.6)$$

Figure 2.1: Distribution of the β values after removal of instances which did not have thermodynamic consistency or were near equilibrium



(a) β -value distribution histogram

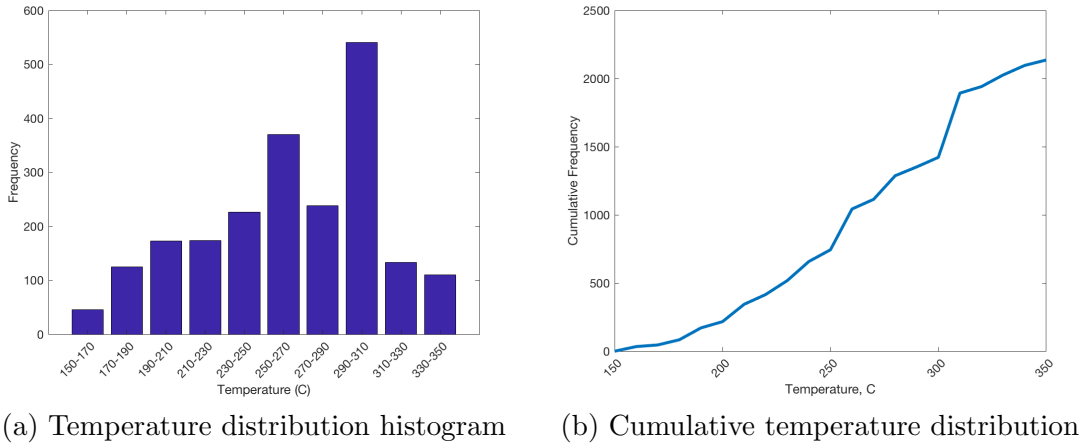


(b) Cumulative β -value distribution

2.2.2 Limiting to LT-WGS

WGS reaction research has been reported over a large range of temperatures, but can be segregated into low-temperature (LT) WGS for 150°C to 350°C and high temperature (HT) WGS for reactions above 350°C. These two categories have distinctly different catalysts and proposed mechanisms. This work constrained the data set to instances within the low-temperature (LT) WGS reaction region, defined by reaction temperatures between 150°C and 350°C.

Figure 2.2: Distribution of temperatures after data removal by β values and limiting to LT-WGS



2.2.3 Normalizing the Data

Training ANNs with vastly different attribute distributions delays the learning process and complicates initialization of the network [15]. Therefore, data normalization, or ‘feature scaling’, is frequently used to achieve similar performance with a reduced number of iterations. All non-categorical features were scaled using the standardization equation, 2.7, which is frequently used in machine learning [12].

$$x' = \frac{x - \bar{x}}{\sigma} \quad (2.7)$$

The ANN output, forward rate, originally demonstrated a log-normal distribution, shown in figure 2.3. This data distribution significantly impeded training and

testing performance. To correct for this distribution, the natural logarithm of the forward rates was taken, normalizing the probability distribution to a zero mean, demonstrated in figure 2.4.

Figure 2.3: Histograms of forward rate distribution prior to normalizing

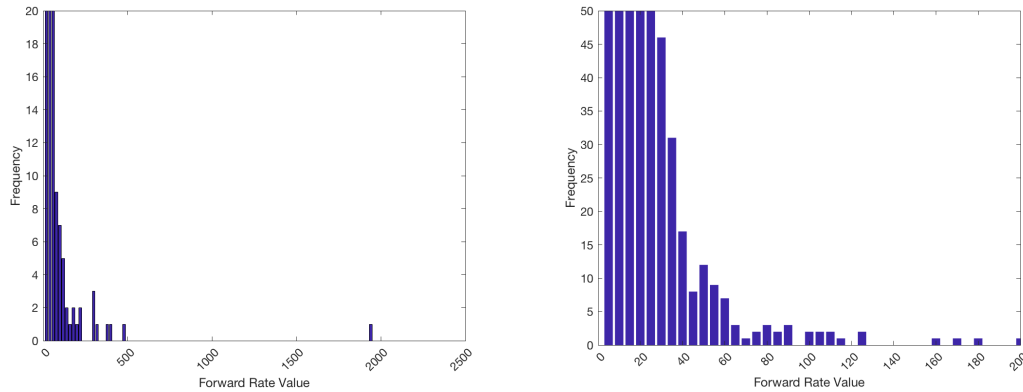
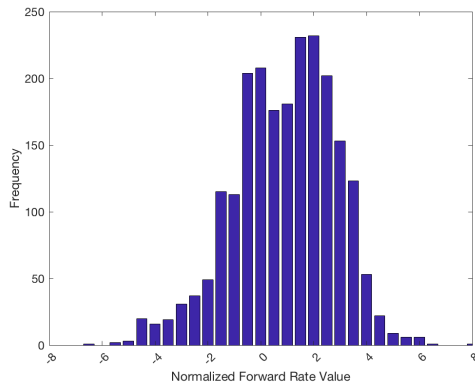


Figure 2.4: Normalized forward rate histogram



2.2.4 Removing Au/CeO₂ and Pt/CeO₂ Instances

Early in the model development process, it was identified that outlying instances were consistently associated with Au/CeO₂ and Pt/CeO₂ catalysts. This finding had a significant implication on the training performance, demonstrated in table 2.2. As a results, 749 instances with Au/CeO₂ or Pt/CeO₂ catalysts were removed from the training data set, reducing the final size from 2,228 data points to 1,499 data points. This removal is further analyzed in the results section.

Table 2.2: ANN Performance when trained with and without Au/CeO₂ and Pt/CeO₂

Data Set	R^2	MSE
With Au/CeO ₂ & Pt/CeO ₂	0.91	0.32
Without Au/CeO ₂ & Pt/CeO ₂	0.95	0.24
Difference	0.03	0.08

2.3 Model Development

The MATLAB Neural Network Toolbox was used to create and train ANNs. Specifically, the `fitnet()` function was used to define network architecture, backpropagation method and training performance parameters.

Overfitting is a significant consideration for ANN development. Overfitting occurs when a neural network receives too much training with the training data and thus, loses its ability to generalize for new instances or inputs. As a result, the error with respect to the training data is small but the error with respect to non-training data is much larger. Overfitting can be avoided by monitoring and controlling the training procedure through a process called ‘cross-validation’ [22]. With cross-validation, the entire training set is iteratively, randomly divided into a training set and a validating set. The training set is used to train the ANN while the validation set is used to evaluate the network performance on a ‘blind’ set of data. While the network is training, the error on both the training and the validation sets decreases. When overfitting begins, the error on the training set continues to decrease while the error on the validation set begins to increase. The MATLAB `fitnet()` tool uses k-fold cross validation to reduce overfitting of the ANN. k-fold cross validation is used to reduce variability, which averages the results over many rounds of dividing the data into different training and validating sets.

For each ANN presented here, the original data set was first divided into model training data (80%) and independent testing data (20%). The model training data was then further divided by MATLAB’s `fitnet()` function into a training data set (70%), a validation data set (30%) and used to develop an ANN. Therefore, the data divisions compound to 56% for training, 24% for validation and 20% for testing.

The `fitnet()` function was modified by increasing the maximum number of iterations to 2,500 and adjusting the backpropagation algorithm from the default Levenberg-Marquardt regularization to Bayesian regularization. Both of these adjustments improved model performance, as evaluated by the mean-squared error and the coefficient of determination. These training adjustments did not result in overfitting. Adjusting other parameters (maximum number of epochs, performance goal, learning rate and performance gradient) was explored and did not significantly improve model performance. The summary of model training parameters is provided in table 2.3.

Table 2.3: ANN Model Training Parameters

Parameter	Specification
Error Function	Mean-squared error
Backpropagation	Bayesian Regularization
Maximum Iterations	2,500
Training Data	56% of total data
Cross Validation Data	24% of total data
Testing Data	20% of total data

The performance of the ANN was then evaluated with the independent testing data using both R^2 and the MSE. Note that due to the k-means cross validation procedure used by MATLAB’s `fitnet()` function, the testing performance provided by the `fitnet()` function is consistently larger than the independently determined error. All results presented henceforth are determined with the independent testing data unless otherwise noted.

2.4 Attribute Determination

Predictions made by ANNs are limited to interpolation within the domain space. Therefore, representing the data to maximize the domain space also maximizes the types of catalysts that can be evaluated using the ANN. A critical modification was migrating the data from the originally categorical representation to a continuous representation that accurately captured complex relationships between attributes.

This adjustment allows predictions to be made for catalysts that were not explicitly included in the training data set, but whose descriptors lie within the defined domain.

To determine the most relevant features in describing the training data, many descriptors were compared for each component of the catalyst (support, primary metal, promoter). The final ANN incorporated the attributes which improved ANN testing performance. This was quantified using the following algorithm:

```

Randomize the data  $m$  times
For each randomization, train  $n$  ANNs
For each ANN evaluate performance (testing  $R^2$  and MSE)
Record the best performance from the  $n$  trained ANNs
Report the average error of the best performance over  $m$  randomizations

```

2.4.1 Primary Metal Descriptors

To build the most complete domain, a descriptor is needed to relate the separate active noble metals to one another. A well-understood descriptor is the energy of a reactive species binding to the surface of the metal. It has been demonstrated that activity is highly correlated to optimizing the binding energy of a surface [11]. For this case study, the binding energies of C, O, H, CO and OH were considered, as each is expected to be present on the catalyst surface and play a prevalent role in the WGS mechanism [10].

For this project, the adsorption energy of carbon on the specified surface of each noble metal was used as the active metal descriptor, given in table 2.4. By using a descriptor that exists on a continuous scale to relate the different surfaces, it allows predictions to be made for unstudied surfaces with intermediate descriptor values. To ensure consistency in calculations, all adsorption energies were taken from a single publication by Herron et al. [14].

Binding energies are typically highly correlated, and therefore, the primary metal should be sufficiently described with a single binding energy [11]. To determine the adsorbate that best describes the primary metal, ANNs were trained using one of the

Table 2.4: Binding energy (eV) of relevant species on noble metal surfaces

Surface	CO	O	C	H	OH
Au(111)	0.3	-1.94	-3.06	-1.84	-1.06
Cu(111)	-0.33	-3.58	-3.75	-2.2	-2.13
Pt(111)	-1.34	-3.18	-5.81	-2.57	-1.65
Pd(111)	-1.49	-3.14	-5.62	-2.62	-1.65
Ir(111)	-1.52	-4.14	-6.46	-2.61	-2.12
Rh(111)	-1.51	-4.13	-5.67	-2.64	-2.14
Ru(0001)	-3.47	-4.73	-6.16	-2.75	-2.64

five possible attributes listed in table 2.4 and the algorithm presented with $m = 5$ and $n = 25$. Both the MSE and the R^2 error, reported in table 2.5 were determined with independent testing data and used to identify the optimum adsorbate for use as a primary metal descriptor. Based on the results given in table 2.4, using carbon as the adsorbate produced the least error in the model.

Table 2.5: Comparison of binding species for primary metal attribute

Binding Species	MSE	R^2
C	0.49	0.87
CO	0.55	0.86
H	0.56	0.85
O	0.58	0.85
OH	0.59	0.85

2.4.2 Support Attributes

A similar procedure was used to determine the best support attributes. Only single metal oxide supports were considered. Each support was described by the chemical properties of the metal or metalloid. Non-metal oxide supports, such as zeolite, hydroxyapatite and activated carbon were removed from the data set as described. For all remaining supports, the following atomic descriptors were explored: ionization energy, reduction potential, electronegativity, most reduced state and most oxidized state. The value of each attribute for each support is given in appendix 1. The analysis results are listed in table 2.6. The first ionization energy and electronegativity were chosen as the most meaningful support attributes.

Table 2.6: Comparison of chemical properties for support attributes

Support Descriptor	MSE	R^2
First Ionization Energy	0.26	0.93
Electronegativity	0.29	0.93
MW of Metal Oxide	0.32	0.92
MW of Metal	0.33	0.91
Highest Oxidation State	0.36	0.91
Reduction Potential	0.38	0.90
Lowest Reduction State	0.39	0.90

2.4.3 Promoter Attributes

Chemical properties were also chosen for promoter attributes. Ionization energies, electronegativity, atomic radius, covalent radius, most oxidized state, most reduced state, ionic radius, reduction potential and charge/ionic radius were explored as descriptors. The value of each attribute for each promoter is given in appendix 2. Using the algorithm described with $m = 5$ and $n = 25$, the best promoter attributes were determined. For these trials, the primary metal descriptor was BE_C and the support descriptors were first ionization energy and electronegativity. From the results in table 2.7, electronegativity and charge/ionic radius were chosen as promoter attributes.

Table 2.7: Comparison of chemical properties for promoter attributes

Promoters	R^2	MSE
Charge/Ionic Radius	0.93	0.275
Electronegativity	0.92	0.296
First Ionization Energy	0.92	0.300
Redox	0.92	0.302
Molecular Weight	0.92	0.306
Most Reduced State	0.92	0.312
Most Oxidized State	0.92	0.315
Covalent Radius	0.88	0.438

2.4.4 Resulting Domain Space

The domain space after cleaning the data set and determining the optimum attributes is given in 2.8. This corresponds to instances with the supports, pri-

mary metals and promoters listed in table 2.9. The remaining categorical synthesis methods are incipient wetness impregnation (IWI), wet impregnation (WI), co-impregnation (CI), sequential impregnation (SI), homogenous deposition precipitation (HDP), flame spray pyrolysis (FSP), deposition precipitation (DP).

Table 2.8: Domain of the ANN after cleaning and attribute selection

Attribute		Minimum Value	Maximum Value
Binding Energy Carbon	(eV)	-6.46	-3.06
Primary Metal Loading	(wt%)	0.1	39.8
Promoter, Charge/Ionic Radius	(charge/pm)	0.06 (0)	1.52
Promoter, Electronegativity	(Pauling Scale)	0.79 (0)	1.91
Promoter Loading	(wt%)	0	20
Support Metal, First Ionization Energy	(kJ/mol)	534.4	786.5
Support Metal, Electronegativity	(Pauling Scale)	1.1	1.9
Synthesis Method		IWI, WI, CI, SI, HDP, FSP, DP	
Calcination Temperature	(Celsius)	25	650
Calcination Time	(hours)	0	10
Reaction Temperature	(Kelvin)	423	623
Reactant Feed, H ₂ vol%		0	60
Reactant Feed, CO vol%		0.2	12
Reactant Feed, H ₂ O vol%		2	60
Reactant Feed, CO ₂ vol%		0	15
Time on Stream	(min)	0.8	4314
Total Inlet Flowrate/Catalyst Mass	(mL/min/g)	0.028	17.6

2.5 Determining Architecture

A critical step in ANN development is determining the ‘optimum’ architecture for the ANN. While no ‘rules’ or heuristics exist for specifying how to determine this, it does depend on the amount of training data, the number of inputs and the complexity of the problem. For this project, all architectures explored had either one or two hidden layers with the second hidden layer being equal to or smaller than the first hidden layer. Furthermore, a given network should have twice as many training data points as it does network weights, to decrease the risk of overfitting. Due to the large size of the training data set, this was not a concern for the architectures explored. For each ANN, the optimum architecture was determined with a ‘brute

Table 2.9: Catalyst components in the domain

Supports	Primary Metals	Promoters
CaO	Cu	Li
Y ₂ O ₃	Ru	Na
TiO ₂	Rh	K
ZrO ₂	Pd	Rb
HfO ₂	Ir	Cs
MnO	Pt	Ca
Fe ₂ O ₃	Au	Sr
Co ₃ O ₄		Y
La ₂ O ₃		Ti
CeO ₂		Zr
ThO ₂		Cr
Al ₂ O ₃		Mn
SiO ₂		Fe
		Co
		Ni
		Re
		Zn
		La
		Ce
		Nd
		Ho
		Er
		Tm

force’ search approach. In other words, a large number of ANNs were trained and evaluated, using the algorithm described with $m = 5$ and $n = 5$ for a total of 25 trained ANNs per architecture. For each architecture, the overall ‘best’ and the average MSE was reported. After completing the broad architecture search given in table 2.10, a more rigorous comparison was performed with $m = 5$ and $n = 20$ for the top three architectures identified with the broad search. These results are presented in 2.11. The results in both searches agree that the optimum ANN architecture has six neurons in the first hidden layer and two neurons in the second hidden layer. Thus, the final ANN has 23 inputs, one for each attribute and synthesis method in 2.8, six hidden neurons in the first layer, two neurons in the second hidden layer and a single output for the $\ln(rate_{forward})$.

Table 2.10: Optimization of ANN Architecture

1st Hidden Layer	2nd Hidden Layer	Weights in Network	Average MSE	Best MSE
5	0	120	0.47	0.35
5	1	121	0.39	0.27
5	2	127	0.52	0.36
5	3	133	0.44	0.25
5	4	139	0.49	0.27
6	0	144	0.43	0.33
6	1	145	0.47	0.38
6	2	152	0.35	0.23
6	3	159	0.43	0.25
6	4	166	0.39	0.28
7	0	168	0.43	0.25
7	1	169	0.60	0.34
7	2	177	0.41	0.26
7	3	185	0.58	0.34
7	4	193	0.37	0.30
8	0	192	0.36	0.23
8	1	193	0.43	0.27
8	2	202	0.47	0.37
8	3	211	0.53	0.28
8	4	220	0.56	0.37

Table 2.11: Rigorous optimization of ANN architectures

1st Hidden Layer	2nd Hidden Layer	Weights in Network	Average MSE	Best MSE
5	1	121	0.41	0.31
6	2	152	0.40	0.19
8	0	192	0.38	0.29

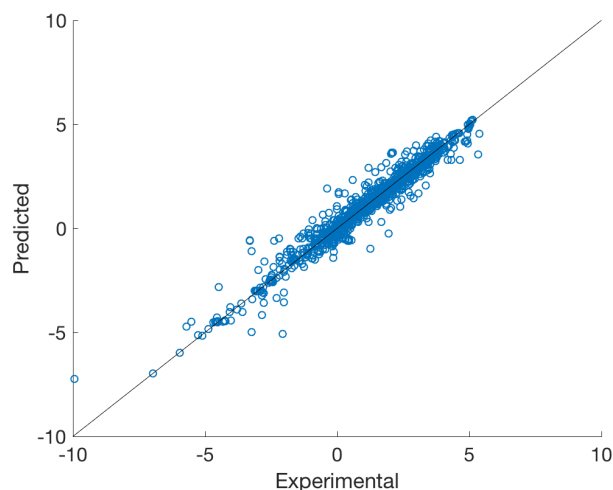
Chapter 3

Results

3.1 Final Neural Network

The 1,499 instances in the cleaned training data set and their corresponding attributes were used to develop an ANN with the optimized architecture given in 2.11. The resulting network has twenty three input attributes, six neurons in the first hidden layer, two neurons in the second hidden layer and predicts the $\ln(\text{forward rate})$. The parity plots demonstrating the resulting fit of the optimized ANN are given in figures 3.1 for the training set and 3.2 for the testing set. The corre-

Figure 3.1: Parity plot for training performance of the ANN



sponding errors are given in table 3.1. Based on the error analysis, the model is not believed to be overtrained. This is consistent with the ratio of training data points to weights in the ANN. These results demonstrate that using the modified continuous attributes to describe each catalyst component, rather than categorically representing the data, results in comparable ANN performance. A sensitivity analysis was performed to evaluate the relative contribution of each attribute to

Figure 3.2: Parity plot for testing performance of the ANN

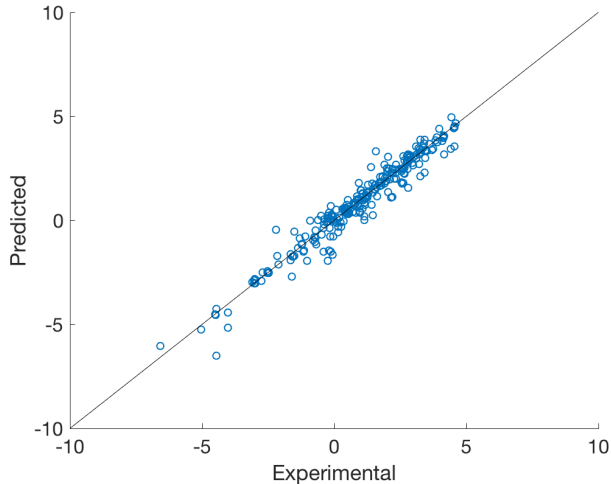


Table 3.1: Error Analysis from ANN Training

Data Set	R^2	MSE
Training	0.945	0.232
Testing	0.946	0.240

the ANN. To determine this, each attribute was removed from the input data set. Then following the algorithm presented in the Methods section ($m = 5$, $n = 5$), the resulting ANN performance was evaluated. The results are given in table 3.2. An increase in the MSE or decrease in the R^2 value from the base case indicates that the attribute contributes significantly to the ANN prediction. The attributes with the largest sensitivity were binding energy of carbon and reaction temperature. Removing the binding energy of carbon, which was used to relate the primary catalyst metals, led to an increase in the MSE from 0.42 to 0.58. The reaction temperature also contributed significantly to the predicted rate, increasing the MSE from 0.42 to 1.24. The MSE and R^2 reported for all other attributes are considered within reasonable error.

Table 3.2: Sensitivity analysis results

Attribute	MSE	R^2
No Attributes Removed, Base Case	0.420	0.90
Binding Energy Carbon	0.577	0.87
Primary Metal Loading, wt%	0.454	0.89
Promoter Charge/Ionic Radius	0.385	0.91
Promoter Electronegativity	0.366	0.92
Promoter Loading, wt%	0.348	0.91
Support, First Ionization Energy	0.403	0.91
Support, Electronegativity	0.414	0.90
Synthesis	0.419	0.90
Calcination Temperature.	0.316	0.92
Calcination Time	0.415	0.90
Reaction Temperature	1.244	0.70
Reactant Feed, H ₂ vol%	0.449	0.90
Reactant Feed, CO vol%	0.430	0.90
Reactant Feed, H ₂ O vol%	0.446	0.90
Reactant Feed, CO ₂ vol%	0.371	0.91
Time on Stream.	0.357	0.92
Total Inlet Flowrate/Catalyst Mass	0.463	0.90

3.2 Scaling Relations

3.2.1 Following Scaling Relations

As with any scaling relation, the ANN can be used to predict the optimum catalyst that falls within the scaling relation. After training the model with the 1499 instances (neglecting Au/CeO₂ and Pt/CeO₂), the optimum catalyst was found to have the following properties listed in table 3.3. This corresponds to a 20wt%Pt/Al₂O₃ catalyst promoted with 15wt%Ni, synthesized using incipient wetness impregnation.

3.2.2 Breaking Scaling Relations:

Au/CeO₂ & Pt/CeO₂ Example

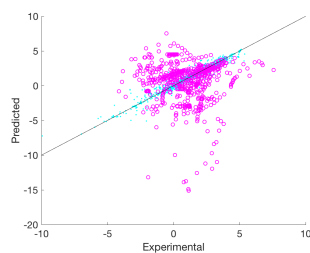
Because ANNs can be thought of as scaling relationship tools for highly dimensional datasets, identifying outliers, or finding points for which the prediction does not match the experimental value indicates that the reaction, either due to catalyst

Table 3.3: Optimum catalyst search results

Attribute	Optimum
Binding Energy Carbon (eV)	-5.81
Primary Metal Loading (wt%)	0.2
Promoter Charge/Ionic Radius (charge/pm)	0.3571
Promoter Electronegativity (Pauling Scale)	1.91
Promoter Loading (wt%)	15
Support Metal, First Ionization Energy (kJ/mol)	577.50
Support Metal, Electronegativity (Pauling Scale)	1.61
Synthesis Method	IWI
Calcination Temperature (Celsius)	500
Calcination Time (hours)	4
Reaction Temperature (Celsius)	350
Reactant Feed, H ₂ vol%	0
Reactant Feed, CO vol%	3
Reactant Feed, H ₂ O vol%	6
Reactant Feed, CO ₂ vol%	0
Time on Stream (min)	60
Total Inlet Flowrate/Catalyst Mass (mL/min/g)	2.000

properties or experimental error, does not follow the expected scaling relationship. It can also be identified whether there is a favorable breaking of scaling relations or an unfavorable breaking of scaling relations. A predicted value lower than the experimental value indicates a favorable breaking of scaling relations. A predicted value higher than the experimental value indicates an unfavorable breaking of scaling relations.

As discussed in the methods, it was found that instances with Pt/CeO₂ and Au/CeO₂ catalysts were consistently outliers. Therefore, these instances were removed from the training data set. After training the model, presented in figure 3.1, this hypothesis was confirmed by examining the model's ability to predict the forward rate for these instances. The prediction results are given in figure 3.3. Based on these results, the ANN developed without Au/CeO₂ and Pt/CeO₂ instances is unable to make acceptable predictions. This result is justified by the experimentally drawn conclusion that Au/CeO₂ and Pt/CeO₂ catalysts operate through a unique mechanism [8, 2, 7].

Figure 3.3: ANN performance with Au/CeO₂ and Pt/CeO₂ instances

● - Training Data, ○ - Au/CeO₂ and Pt/CeO₂ Data

3.3 Analysis of Fitting Untrained Data

Using continuous attributes to relate naturally categorical data should theoretically allow for predictions on materials that were not present in the original training data set. To test this, networks were trained with the original training data set (excluding Au/CeO₂ and Pt/CeO₂), neglecting any instances with a given catalyst property (i.e., leaving out all attributes with Cu as the primary metal). The ability of the ANN to interpolate within the domain for unique catalysts was then evaluated by using the excluded data to compare the experimental forward rate to the predicted forward rate.

For all catalyst components explored below, it appears that an increase in the number of data points withheld increases prediction error. For cases such as Al₂O₃, where nearly 30% of the training data set is withheld, it may be expected that the resulting ANN is not nearly as robust.

3.3.1 Primary Metals

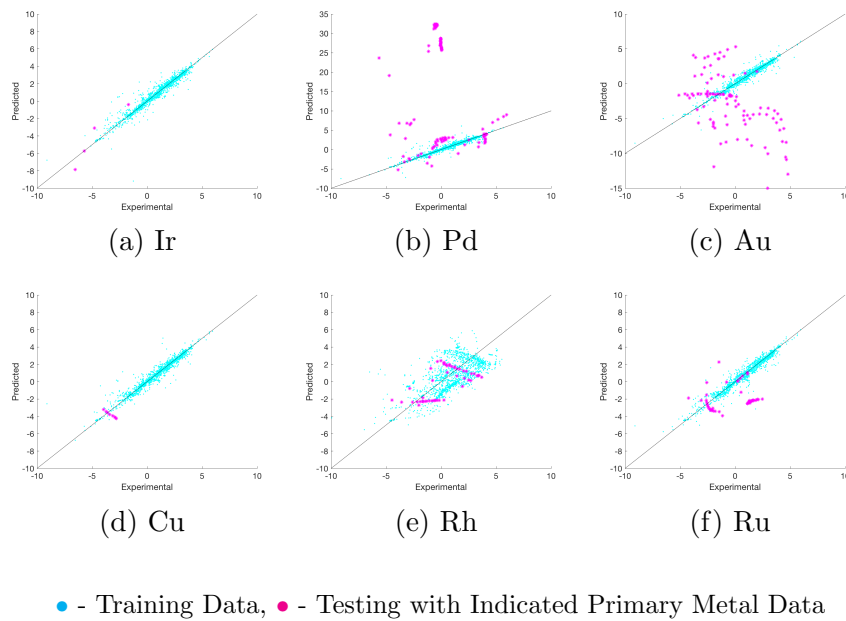
This was performed for all of the primary metals except Pt, due to the large number of instances with Pt-catalysts (1193 instances), given in figure 3.4. Acceptable predictions were observed for the primary metals Ir, Cu, Rh and Ru. However, the model did not successfully interpolate within the domain for Au and Pd. This could be attributed to **either** the number of data points, 99 for Au and 105 for Pd.

It is suggested that the lack of fit for Au is due to an alteration to the domain,

resulting from the segregation of the Au instances. Au has the weakest of the binding energy (-3.06eV) of all the primary metals considered. By removing Au from the training data set, we reduce the domain from $-6.46 < BE_C < -3.06\text{eV}$ to $-6.46 < BE_C < -3.75\text{eV}$. Thus, the model's predictions are extrapolations outside of the training domain, which is well understood to lead to a decrease **is** predictive accuracy [32].

For Pd it appears the prediction segregates into two dominant groups, one which follows the scaling relation and another that deviates. The predicted rate for the deviating data was much larger than the expected rate. Unlike the Au justification, the binding energy attribute for Pd is in the middle of the data set (-5.62eV). Multiple factors may contribute to this result. It is possible that the data cluster not following the scaling relationship has an unusual feature that was not within the domain space upon removal of all instances with Pd as the primary metal. It is also possible that the catalyst behaves more like the Au/CeO₂ and Pt/CeO₂ catalysts described above, due to **some** methodological difference in the catalyst preparation. Regardless, additional work would be needed to confirm the source of this discrepancy.

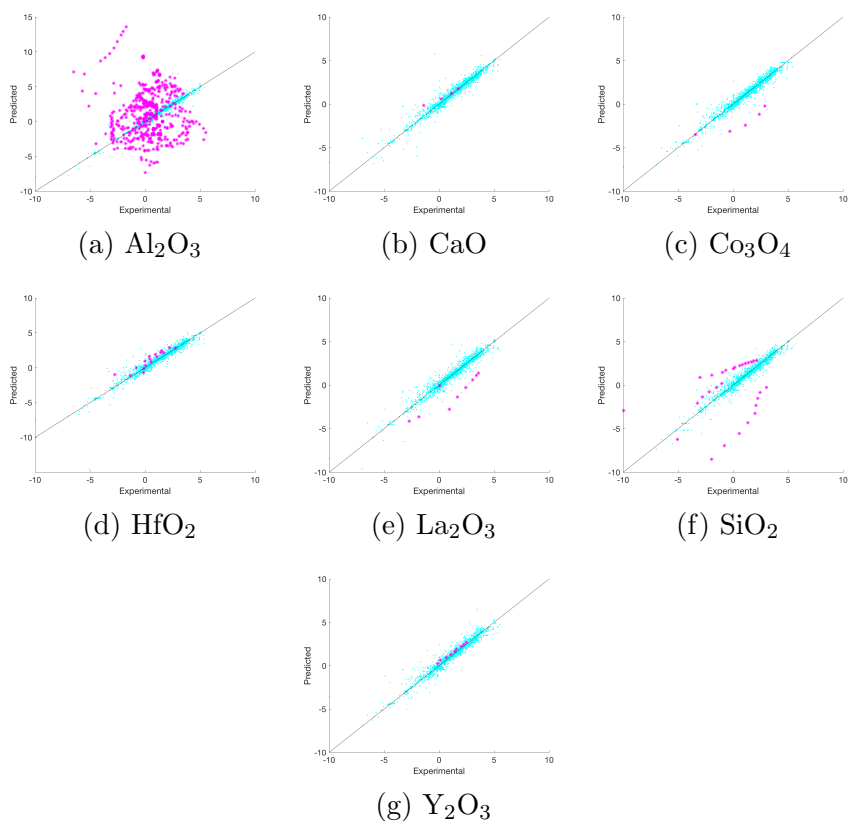
Figure 3.4: Primary Metal Interpolation Results



3.3.2 Supports

The procedure was repeated for instances of varying supports (Al_2O_3 , CaO , Co_3O_4 , HfO_2 , La_2O_3 , SiO_2 , Y_2O_3), selectively choosing supports that span the domain. 3.5. Of the supports explored, only Al_2O_3 and SiO_2 yielded poor predictions. It is suggested that in the case of Al_2O_3 supports removing the 556 data points negatively impacted ANN training, as previously discussed. For SiO_2 29 data points were removed. Thus, it is unlikely that the quantity of data removed impacted the training. However, SiO_2 had the maximum support electronegativity (1.9, Pauling Scale) in the domain. Similar to the argument for Au, it is likely that removal of these data points decreased the domain space and the model had to extrapolate to make predictions. As a result, the error of the predictions increased.

Figure 3.5: Support Interpolation Results

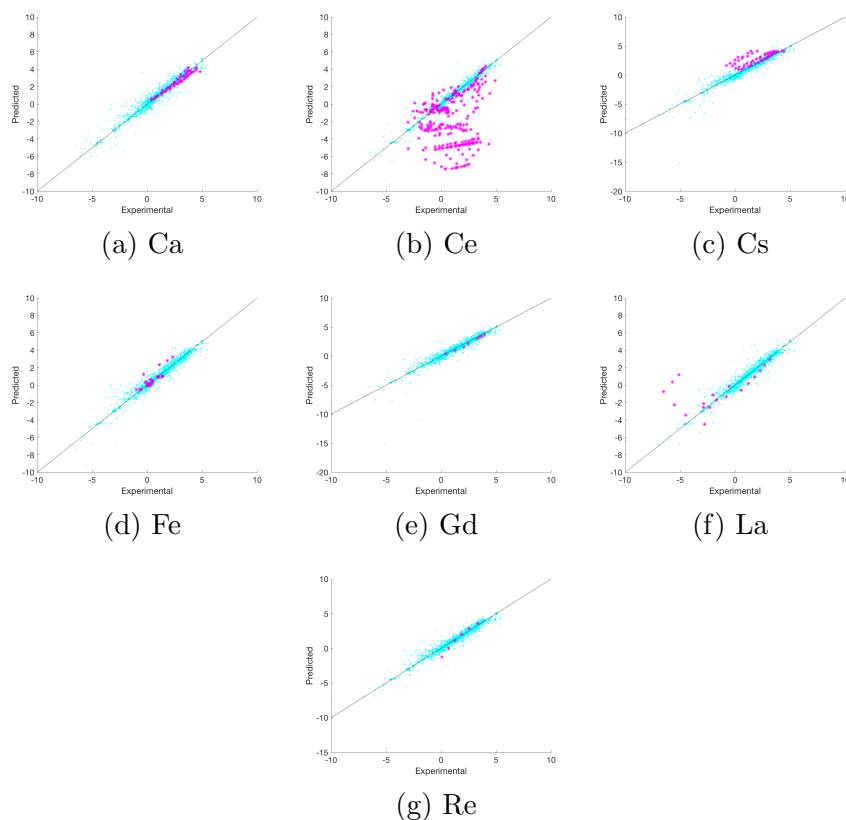


● - Training Data, ● - Testing with Indicated Support Data

3.3.3 Promoters

The ability to predict promoter contributions was also evaluated by considering the promoters Ca, Ce, Cs, Fe, Gd, La, Re, shown in figure 3.6. Again, the group with the largest number of instances, Ce (297 instances) demonstrated the largest prediction error, although the associated attributes (electronegativity = 1.12, Charge/Ionic Radius = 0.46) were well within the domain. Therefore, it is suggested that either the number of data points removed affected training performance or using a Ce promoter alters the reaction mechanism, resulting in a catalyst that breaks the scaling relation.

Figure 3.6: Promoter Interpolation Results



● - Training Data, ● - Testing with Indicated Promoter Data

Chapter 4

Discussion

4.1 Conclusions

ANNs are a promising tool for future catalyst development, particularly when considering well-studied reactions with ample published data, such as the water-gas shift reaction. As demonstrated, ANNs are capable of merging diverse data from many sources into a single mathematical model capable of providing catalytic insight. The model developed in this work is capable of predicting catalysts which follow scaling relations and identifying catalysts that do not. Furthermore, the process of rigorously optimizing an ANN is a valuable approach to analyzing readily available catalytic data. Specifically, identifying appropriate attributes and evaluating their relative importance via a sensitivity analysis leads to improved understanding of underlying catalytic properties.

Because ANNs rely on an initial, well-characterized training data set, they cannot serve as an independent platform for catalyst development. However, used in conjunction with modern computational techniques and cutting-edge experimental methods, ANNs have the potential to bridge these areas of research and guide intelligent catalyst design. ANNs can be applied in catalytic research to (1) discern the underlying chemical properties that drive catalytic activity, (2) to predicting optimum catalyst compositions according to the developed scaling relation and (3) given experimental data, to identify catalytic materials capable of breaking the scaling relation.

4.2 Limitations

A developed ANN is limited by the quality of training data and validity of attribute selection. The model’s predictions will reflect any biases in the training data. Therefore, rigorous data evaluation and cleaning, as done here with the β approach, is critical to ensuring model validity. Additionally, the ANN is trained by minimizing an error function that compares an expected value to the predicted value. As a result, the trained model will incorporate some bias based on the distribution of the training data. If domain coverage is sparse in one region and dense in another, the predictions for inputs lying in the dense area will have less error than those in the sparse area. Therefore, while this work uses the overall MSE and the coefficient of determination to evaluate model performance, sparse areas of the domain will result in poorer predictions than indicated by these statistical analysis.

Furthermore, rigorous model development requires fundamental catalytic knowledge to hypothesize valuable attribute selection. Once attributes are hypothesized, a trial-and-error process remains for evaluating the most meaningful descriptors. Thus, the scientific process is not something that can be abandoned with ML techniques. Additionally, identifying attributes that result in adequate performance does not necessarily indicate a causal relationship. Rather, an attribute may correlate with the output without being causally related. Therefore, blindly accepting attributes does not necessarily yield physical insight about the catalytic system.

Finally, previously published work merging experimental results and ANNs has been in optimizing a catalyst or reaction conditions to promote a desired outcome. However, viewing the ANN as a complex scaling relation, these optimizations and predictions are limited to catalysts that fall on the scaling relation. Realistically, the most significant catalytic breakthroughs often occur when scaling relations are broken. Therefore, solely exploring catalysts which obey the ANN-predicted scaling relation may have limited value.

4.3 Future Work

The approaches and results contained in this report could be further fortified by pursuing the future work proposed here. The attribute identification and analysis was relatively crude and not mathematically substantiated. Alternatively, clustering techniques could be used to identify the attributes prior to and independent of ANN development. In addition, identifying a relationship to allow inclusion of mixed supports in this model would add value by expanding the domain. Similarly, identifying attributes that allowed successful integration of the Au/CeO₂ and Pt/CeO₂ data would also be useful. This would reintroduce a significant number of data points and may also have implications for explaining the unique mechanism observed from these catalysts. Along with improving attribute identification, a more rigorous sensitivity analysis could be performed. This may lead to additional insights on the relative importance of given attributes.

It is worth acknowledging that the current model is only capable of identifying whether an instance breaks or follows the scaling relation. In cases where the scaling relation is not followed, the ANN is unable to identify how or why the prediction was wrong. Therefore, a particularly valuable future direction may be in determining a means of identifying where the divergence from the scaling relation occurred. Successfully identifying the ‘breaking point’ of a scaling relation may provide insight regarding mechanistic differences between catalysts.

Support Chemical Properties

Support	First Ionization Energy (kJ/mol)	Reduction Reaction	Reduction Potential	Electronegativity	Most	Most
					Oxidized State	Reduced State
Al ₂ O ₃	578	$\text{Al}^{3+} \xrightarrow{+3e} \text{Al}$	-1.662	1.61	3	-2
MgO	738	$\text{Mg}^{+} \xrightarrow{+e} \text{Mg}$	-2.7	1.31	2	1
CeO ₂	534	$\text{Ce}^{3+} \xrightarrow{+3e} \text{Ce}$	-2.336	1.12	4	1
TiO ₂	659	$\text{Ti}^{2+} \xrightarrow{+2e} \text{Ti}$	-1.63	1.54	4	-2
MnO	717	$\text{Mn}^{2+} \xrightarrow{+2e} \text{Mn}$	-1.185	1.55	7	-3
Y ₂ O ₃	600	$\text{Y}^{3+} \xrightarrow{+3e} \text{Y}$	-2.372	1.22	3	1
Tb ₄ O ₇	566	$\text{Tb}^{3+} \xrightarrow{+3e} \text{Tb}$	-2.28	1.2	4	1
HfO ₂	659	$\text{Hf}^{4+} \xrightarrow{+4e} \text{Hf}$	-1.55	1.3	4	-2
La ₂ O ₃	538	$\text{La}^{3+} \xrightarrow{+3e} \text{La}$	-2.379	1.1	3	1
Co ₃ O ₄	760	$\text{Co}^{2+} \xrightarrow{+2e} \text{Co}$	-0.28	1.88	5	-3
ThO ₂	587	$\text{Th}^{4+} \xrightarrow{+4e} \text{Th}$	-1.899	1.3	4	1
SiO ₂	787	$\text{SiO}_2 + 4\text{H} \xrightarrow{+4e} \text{Si} + 2\text{H}_2\text{O}$	0.857	1.9	4	-4
Fe ₂ O ₃	763	$\text{Fe}^{3+} \xrightarrow{+3e} \text{Fe}$	-0.037	1.83	7	-4
Sm ₂ O ₃	545	$\text{Sm}^{3+} \xrightarrow{+3e} \text{Sm}$	-2.304	1.17	4	1
Gd ₂ O ₃	593	$\text{Gd}^{3+} \xrightarrow{+3e} \text{Gd}$	-2.297	1.2	3	1
Yb ₂ O ₃	603	$\text{Yb}^{3+} \xrightarrow{+3e} \text{Yb}$	-2.19	1.1	3	1
CaO	590	$\text{Ca}^{+} \xrightarrow{+e} \text{Ca}$	-3.8	1	2	1
ZrO ₂	640	$\text{Zr}^{4+} \xrightarrow{+4e} \text{Zr}$	-1.45	1.33	4	-2

Promoter Chemical Properties

Promoter	First	Covalent	Most	Most	Ionic	Reduction		
	MW	Ionization	Radius	Reduced	Oxidized	Radius	Potential	
	Energy	(ang)	State	State	(pm)	$\frac{Charge}{IonicRadius}$		
Li	6.94	5.39	1.28	0.00	1.00	76.00	0.1316	-3.040
Na	22.99	5.14	1.66	-1.00	1.00	102.00	0.0980	-2.710
K	39.10	4.34	2.03	-1.00	1.00	138.00	0.0725	-2.931
Rb	85.47	4.18	2.20	-1.00	1.00	152.00	0.0658	-2.980
Cs	132.90	3.89	2.44	-1.00	1.00	167.00	0.0599	-3.026
Mg	24.31	7.65	1.41	1.00	2.00	72.00	0.2778	-2.372
Ca	40.08	6.11	1.76	1.00	2.00	100.00	0.2000	-2.868
Sr	87.62	5.70	1.95	1.00	2.00	118.00	0.1695	-2.899
Y	88.91	6.22	1.76	1.00	3.00	90.00	0.3333	-2.372
La	138.91	5.58	2.07	1.00	3.00	103.20	0.2907	-2.379
Ce	140.12	5.54	2.04	1.00	4.00	87.00	0.4598	-2.336
Nd	144.24	5.53	2.01	2.00	4.00	98.30	0.4069	-2.323
Sm	150.36	5.64	1.98	1.00	4.00	95.80	0.4175	-2.304
Gd	157.25	6.15	1.96	1.00	3.00	93.50	0.3209	-2.279
Ho	164.93	6.02	1.92	1.00	3.00	90.10	0.3330	-2.330
Er	167.26	6.11	1.89	1.00	3.00	89.00	0.3371	-2.331
Tm	168.93	6.18	1.90	2.00	3.00	88.00	0.3409	-2.319
Yb	174.97	5.43	1.87	1.00	3.00	86.80	0.3456	-2.190
Ti	47.87	13.58	1.48	-2.00	4.00	60.50	0.6612	-1.370
Zr	91.22	6.63	1.64	-2.00	4.00	72.00	0.5556	-1.450
V	50.94	6.75	1.44	-3.00	5.00	54.00	0.9259	-1.175
Cr	52.00	6.77	1.30	-4.00	6.00	44.00	1.3636	-0.744
Mn	54.94	7.43	1.29	-3.00	7.00	46.00	1.5217	-1.185
Re	186.21	7.83	1.41	-3.00	7.00	53.00	1.3208	0.300
Fe	55.85	7.90	1.24	-4.00	7.00	64.50	1.0853	-0.037
Co	58.93	7.88	1.18	-3.00	5.00	54.50	0.9174	-0.280
Ni	58.69	7.64	1.17	-4.00	2.00	56.00	0.3571	-0.257
Zn	65.38	9.39	1.20	-2.00	2.00	74.00	0.2703	-0.762

Bibliography

- [1] Mie Andersen et al. “Scaling-Relation-Based Analysis of Bifunctional Catalysis: The Case for Homogeneous Bimetallic Alloys”. In: *ACS Catalysis* 7.6 (2017), pp. 3960–3967. DOI: [10.1021/acscatal.7b00482](https://doi.org/10.1021/acscatal.7b00482). URL: <http://dx.doi.org/10.1021/acscatal.7b00482>.
- [2] T. Bunluesin, R.J. Gorte, and G.W. Graham. “Studies of the water-gas-shift reaction on ceria-supported Pt, Pd, and Rh: Implications for oxygen-storage properties”. In: *Applied Catalysis B: Environmental* 15.1 (1998), pp. 107–114. ISSN: 0926-3373. DOI: [https://doi.org/10.1016/S0926-3373\(97\)00040-4](https://doi.org/10.1016/S0926-3373(97)00040-4). URL: <http://www.sciencedirect.com/science/article/pii/S0926337397000404>.
- [3] Avelino Corma et al. “A New Mapping/Exploration Approach for HT Synthesis of Zeolites”. In: *Chemistry of Materials* 18.14 (2006), pp. 3287–3296. DOI: [10.1021/cm060620k](https://doi.org/10.1021/cm060620k). URL: <http://dx.doi.org/10.1021/cm060620k>.
- [4] Avelino Corma et al. “Application of Artificial Neural Networks to Combinatorial Catalysis: Modeling and Predicting ODHE Catalysts”. In: *ChemPhysChem* 3.11 (2002), pp. 939–945. ISSN: 1439-7641.
- [5] A. Corma et al. “Optimisation of olefin epoxidation catalysts with the application of high-throughput and genetic algorithms assisted by artificial neural networks (softcomputing techniques)”. In: *Journal of Catalysis* 229.2 (2005), pp. 513–524. ISSN: 0021-9517. DOI: <https://doi.org/10.1016/j.jcat.2004.11.024>. URL: <http://www.sciencedirect.com/science/article/pii/S002195170400555X>.
- [6] Ronald Williams David Rumelhart Geoffrey Hinton. “Learning representations by back-propagating errors”. In: *Nature* 323 (1986), pp. 533–536. URL: <http://elderlab.yorku.ca/~elder/teaching/cosc6390psyc6225/readings/hinton%201986.pdf>.

- [7] Weiling Deng and Maria Flytzani-Stephanopoulos. “On the Issue of the Deactivation of Au–Ceria and Pt–Ceria Water–Gas Shift Catalysts in Practical Fuel-Cell Applications”. In: *Angewandte Chemie* 118.14 (2006), pp. 2343–2347. ISSN: 1521-3757. DOI: 10.1002/ange.200503220. URL: <http://dx.doi.org/10.1002/ange.200503220>.
- [8] Qi Fu, Howard Saltsburg, and Maria Flytzani-Stephanopoulos. “Active Non-metallic Au and Pt Species on Ceria-Based Water-Gas Shift Catalysts”. In: *Science* 301.5635 (2003), pp. 935–938. ISSN: 0036-8075. DOI: 10.1126/science.1085721. eprint: <http://science.sciencemag.org/content/301/5635/935.full.pdf>. URL: <http://science.sciencemag.org/content/301/5635/935>.
- [9] Terry Zhi Hao Gani and Heather Janine Kulik. “Understanding and Breaking Scaling Relations in Single-Site Catalysis: Methane-to-methanol Conversion by Fe(IV)=O”. In: (Oct. 2017). DOI: 10.26434/chemrxiv.5525572.v1.
- [10] Amit A. Gokhale, James A. Dumesic, and Manos Mavrikakis. “On the Mechanism of Low-Temperature Water Gas Shift Reaction on Copper”. In: *Journal of the American Chemical Society* 130.4 (2008), pp. 1402–1414. DOI: 10.1021/ja0768237. URL: <http://dx.doi.org/10.1021/ja0768237>.
- [11] Jeffrey Greeley. “Theoretical Heterogeneous Catalysis: Scaling Relationships and Computational Catalyst Design”. In: *Annual Review of Chemical and Biomolecular Engineering* 7.1 (2016), pp. 605–635. DOI: 10.1146/annurev-chembioeng-080615-034413.
- [12] Joel Grus. *Data Science from Scratch: First Principles with Python*. Beijing: O’Reilly, 2015. ISBN: 978-1-4919-0142-7. URL: <http://my.safaribooksonline.com/97814919-01427>.
- [13] Tadashi Hattori and Shigeharu Kito. “Analysis of factors controlling catalytic activity by neural network”. In: *Catalysis Today* 111.3 (2006), pp. 328–332. ISSN: 0920-5861. DOI: <https://doi.org/10.1016/j.cattod.2005.10>.

044. URL: <http://www.sciencedirect.com/science/article/pii/S092058610500800X>.
- [14] Jeffrey A. Herron et al. “Trends in Formic Acid Decomposition on Model Transition Metal Surfaces: A Density Functional Theory study”. In: *ACS Catalysis* 4.12 (2014), pp. 4434–4445. DOI: 10.1021/cs500737p. eprint: <http://dx.doi.org/10.1021/cs500737p>. URL: <http://dx.doi.org/10.1021/cs500737p>.
- [15] Sergey Ioffe and Christian Szegedy. “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift”. In: *Proceedings of the 32nd International Conference on Machine Learning (ICML-15)*. Ed. by David Blei and Francis Bach. JMLR Workshop and Conference Proceedings, 2015, pp. 448–456.
- [16] Lee Jacobson. *Introduction to Artificial Neural Networks - Part 1*. Dec. 2013. URL: <http://www.theprojectspot.com/tutorial-post/introduction-to-artificial-neural-networks-part-1/7>.
- [17] Shigeharu Kite, Tadashi Hattori, and Yuichi Murakami. “Estimation of catalytic performance by neural network a product distribution in oxidative dehydrogenation of ethylbenzene”. In: *Applied Catalysis A: General* 114.2 (1994), pp. L173–L178. ISSN: 0926-860X. DOI: [https://doi.org/10.1016/0926-860X\(94\)80169-X](https://doi.org/10.1016/0926-860X(94)80169-X).
- [18] Shigerharu Kito and Tadashi Hattori. “Estimation of the Acid Strength of Mixed Oxides by a Neural Network”. In: *Ind. Eng. Chem. Res.* 31 (1991), pp. 979–981. URL: <http://pubs.acs.org/doi/pdf/10.1021/ie00003a046>.
- [19] Shigerharu Kito, Tadashi Hattori, and Yuichi Murakami. “Expert systems approach to computer-aided design of catalysts”. In: *Applied Catalysis* 48.1 (1989), pp. 107–121. ISSN: 0166-9834. DOI: [https://doi.org/10.1016/S0166-9834\(00\)80269-9](https://doi.org/10.1016/S0166-9834(00)80269-9). URL: <http://www.sciencedirect.com/science/article/pii/S0166983400802699>.

- [20] S. Kito et al. “Application of neural network to estimation of catalyst deactivation in methanol conversion”. In: *Catalysis Today* 97.1 (2004). Advanced Methods in Characterization of Catalysts and Materials, pp. 41–47. ISSN: 0920-5861. DOI: <https://doi.org/10.1016/j.cattod.2004.04.052>. URL: <http://www.sciencedirect.com/science/article/pii/S0920586104003499>.
- [21] Yasukazu Kobayashi, Kohji Omata, and Muneyoshi Yamada. “Screening of Additives to a Co/SrCO₃ Catalyst by Artificial Neural Network for Preferential Oxidation of CO in Excess H₂”. In: *Industrial & Engineering Chemistry Research* 49.4 (2010), pp. 1541–1549. DOI: 10.1021/ie901435h. eprint: <http://dx.doi.org/10.1021/ie901435h>. URL: <http://dx.doi.org/10.1021/ie901435h>.
- [22] Ron Kohavi. “A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection”. In: Morgan Kaufmann, 1995, pp. 1137–1143.
- [23] P.J. Linstrom and W.G. Mallard. *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*. Gaithersburg MD, 20899: National Institute of Standards and Technology. DOI: doi:10.18434/T4D303.
- [24] Cagla Odabasi, M Gaenay, and Ramazan Yildirim. “Knowledge extraction for water gas shift reaction over noble metal catalysts from publications in the literature between 2002 and 2012”. In: 39 (Apr. 2014).
- [25] K. Omata, Y. Kobayashi, and M. Yamada. “Artificial neural network aided virtual screening of additives to a Co/SrCO₃ catalyst for preferential oxidation of CO in excess hydrogen”. In: *Catalysis Communications* 8.1 (2007), pp. 1–5. ISSN: 1566-7367. DOI: <https://doi.org/10.1016/j.catcom.2006.04.018>. URL: <http://www.sciencedirect.com/science/article/pii/S1566736706001233>.
- [26] Kohji Omata, Yasukazu Kobayashi, and Muneyoshi Yamada. “Artificial Neural Network Aided Screening and Optimization of Additives to Co/SrCO₃ Catalyst for Dry Reforming of Methane under Pressure”. In: *Energy & Fuels* 23.4

- (2009), pp. 1931–1935. DOI: 10.1021/ef8010004. eprint: <http://dx.doi.org/10.1021/ef8010004>. URL: <http://dx.doi.org/10.1021/ef8010004>.
- [27] Kohji Omata, Yasukazu Kobayashi, and Muneyoshi Yamada. “Artificial neural network-aided design of Co/SrCO₃ catalyst for preferential oxidation of CO in excess hydrogen”. In: *Catalysis Today* 117.1 (2006). The 5th World Congress on Oxidation Catalysis, pp. 311–315. ISSN: 0920-5861. DOI: <https://doi.org/10.1016/j.cattod.2006.05.032>. URL: <http://www.sciencedirect.com/science/article/pii/S0920586106002811>.
- [28] Kohji Omata and Muneyoshi Yamada. “Prediction of Effective Additives to a Ni/Active Carbon Catalyst for Vapor-Phase Carbonylation of Methanol by an Artificial Neural Network”. In: *Industrial & Engineering Chemistry Research* 43.20 (2004), pp. 6622–6625. DOI: 10.1021/ie049609p. URL: <http://dx.doi.org/10.1021/ie049609p>.
- [29] Kohji Omata et al. “Artificial Neural Network (ANN)-Aided Optimization of ZSM-5 Catalyst for the Dimethyl Ether to Olefin (DTO) Reaction from Neat Dimethyl Ether (DME)”. In: *Industrial & Engineering Chemistry Research* 48.13 (2009), pp. 6256–6261. DOI: 10.1021/ie801757p. eprint: <http://dx.doi.org/10.1021/ie801757p>. URL: <http://dx.doi.org/10.1021/ie801757p>.
- [30] Kohji Omata et al. “Design and Development of Cu-Zn Oxide Catalyst for Direct Dimethyl Ether Synthesis Using an Artificial Neural Network and Physicochemical Properties of Elements”. In: *Industrial & Engineering Chemistry Research* 45.14 (2006), pp. 4905–4910. DOI: 10.1021/ie050640g. eprint: <http://dx.doi.org/10.1021/ie050640g>. URL: <http://dx.doi.org/10.1021/ie050640g>.
- [31] Kohji Omata et al. “Simultaneous Optimization of Preparation Conditions and Composition of the Methanol Synthesis Catalyst by an All-Encompassing Calculation on an Artificial Neural Network”. In: *Industrial & Engineering*

- Chemistry Research* 43.13 (2004), pp. 3282–3288. DOI: 10.1021/ie034173j. eprint: <http://dx.doi.org/10.1021/ie034173j>. URL: <http://dx.doi.org/10.1021/ie034173j>.
- [32] Jose M. Serra et al. “Can artificial neural networks help the experimentation in catalysis?” In: *Catalysis Today* 81.3 (2003). European Workshop on Combinatorial Catalysis, pp. 393–403. ISSN: 0920-5861. DOI: [https://doi.org/10.1016/S0920-5861\(03\)00137-8](https://doi.org/10.1016/S0920-5861(03)00137-8).
- [33] Jose M. Serra et al. “Neural networks for modelling of kinetic reaction data applicable to catalyst scale up and process control and optimisation in the frame of combinatorial catalysis”. In: *Applied Catalysis A: General* 254.1 (2003). Combinatorial Chemistry, pp. 133–145. DOI: [https://doi.org/10.1016/S0926-860X\(03\)00283-7](https://doi.org/10.1016/S0926-860X(03)00283-7).
- [34] A. M. TURING. “COMPUTING MACHINERY AND INTELLIGENCE”. In: *Mind* LIX.236 (1950), pp. 433–460. DOI: 10.1093/mind/LIX.236.433. URL: <http://dx.doi.org/10.1093/mind/LIX.236.433>.
- [35] Zachary W. Ulissi et al. “To address surface reaction network complexity using scaling relations machine learning and DFT calculations”. In: *Nature Communications* 8 (2017). DOI: 10.1038/ncomms14621. URL: <https://www.nature.com/articles/ncomms14621.pdf>.
- [36] Tetsuo Umegaki et al. “Development of a high performance Cu-based ternary oxide catalyst for oxidative steam reforming of methanol using an artificial neural network”. In: *Applied Catalysis A: General* 351.2 (2008), pp. 210–216. ISSN: 0926-860X. DOI: <https://doi.org/10.1016/j.apcata.2008.09.019>. URL: <http://www.sciencedirect.com/science/article/pii/S0926860X08005851>.
- [37] S. Valero et al. “DoE framework for catalyst development based on soft computing techniques”. In: *Computers & Chemical Engineering* 33.1 (2009), pp. 225–238. ISSN: 0098-1354. DOI: <https://doi.org/10.1016/j.compchemeng>.

2008.08.012. URL: <http://www.sciencedirect.com/science/article/pii/S0098135408001658>.