INNOVATIVE CATALYST DESIGN FOR CO_2 REDUCTION REACTION: A COMBINED MULTISCALE MODELLING AND MACHINE LEARNING APPROACH

ARISA KAEWPRATUM

A Thesis Submitted in Partial Fulfillment of the Requirements for the
Master Degree of Science in Integrated Science and Innovation
Suranaree University of Technology
Academic Year 2025

การออกแบบเชิงนวัตกรรมของตัวเร่งปฏิกิริยาสำหรับปฏิกิริยารีดักชันคาร์บอนไดออกไซด์ โดยการผสมผสานระหว่างการสร้างแบบจำลองหลายระดับความละเอียดและการเรียนรู้ของเครื่อง

นางสาวอาริษา แก้วประทุม

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาบูรณาการด้านวิทยาศาสตร์และนวัตกรรม
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2568

INNOVATIVE CATALYST DESIGN FOR CO_2 REDUCTION REACTION: A COMBINED MULTISCALE MODELLING AND MACHINE LEARNING APPROACH

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for a Master's degree.

Thesis Examining Committee

(Asst. Prof. Dr. Ittipon Fongkaew)
Chairperson

(Assoc. Prof. Dr. Suwit Suthirakun)
Member (Thesis Advisor)

(Assoc. Prof. Dr. Anyanee Kamkaew)
Member

(Assoc. Prof. Dr. Rapee Utke)
Member

(Dr. Wiwat Nuansing)
Member

(Dr. Tongjai Chookajorn)
Member

(Assoc. Prof. Dr. Yupaporn Ruksakulpiwat)
Acting Vice Rector for Academic Affairs and Quality Assurance

(Prof. Dr. Santi Maensiri)
Acting Dean of the Institute of Science

อาริษา แก้วประทุม : การออกแบบเชิงนวัตกรรมของตัวเร่งปฏิกิริยาสำหรับปฏิกิริยารีดักชันคาร์บอนไดออกไซด์ โดยการการผสมผสานระหว่างการสร้างแบบจำลองหลายระดับความละเอียดและการเรียนรู้ของเครื่อง (INNOVATIVE CATALYST DESIGN FOR CO_2 REDUCTION REACTION: A COMBINED MULTISCALE MODELLING AND MACHINE LEARNING APPROACH)
อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร.สุวิทย์ สุธีรากุล, 78 หน้า

คำสำคัญ: ปฏิกิริยารีดักชันคาร์บอนไดออกไซด์ การสร้างแบบจำลองจลนพลศาสตร์ระดับจุลภาคแบบจำลองการเรียนรู้ของเครื่อง

ระดับก๊าซคาร์บอนไดออกไซด์ในชั้นบรรยากาศเป็นประเด็นที่สำคัญ จำเป็นต้องหานวัตกรรมใหม่ ๆ เพื่อลดผลกระทบ หนึ่งในแนวทางที่มีแนวโน้มที่ดีคือปฏิกิริยารีดักชันคาร์บอนไดออกไซด์ ให้กลายเป็นผลิตภัณฑ์ที่มีมูลค่า วิทยานิพนธ์นี้นำเสนอ กระบวนการทำงานแบบจำลองหลายระดับความละเอียดและการเรียนรู้ของเครื่อง (ML) ที่ออกแบบมาเพื่อค้นหาตัวเร่งปฏิกิริยา เริ่มต้นด้วยการสร้างฐานข้อมูลโครงสร้างตัวเร่งปฏิกิริยากว่า 1,632 โครงสร้าง ซึ่งประกอบด้วยโลหะบริสุทธิ์และโลหะผสม การศึกษาใช้แบบจำลอง ML ที่เรียกว่า Fairchem เพื่อคำนวณพลังงานการดูดซับ ซึ่งจะเป็นข้อมูลป้อนเข้าสำหรับ แบบจำลองไมโครไคเนติก เพื่อทำนายอัตราการเกิดปฏิกิริยาและค่าการเลือกเกิดผลิตภัณฑ์ นอกจากนี้ ยังพัฒนาแบบจำลอง ML เพื่อทำนายประสิทธิภาพของตัวเร่งปฏิกิริยา โดยโครงข่ายประสาทเทียม ให้ประสิทธิภาพดีที่สุดในการทำนายค่าการเลือกเกิดผลิตภัณฑ์ ขณะที่ตัวจัดประเภทแบบซัพพอร์ตเวกเตอร์ ทำหน้าที่ได้ดีที่สุดในการทำนายอัตราการเกิดปฏิกิริยารวม เพื่อให้เข้าถึงได้ง่าย เราสร้างเว็บไซต์ที่สามารถทำนายผลลัพธ์ได้ทันทีโดยใช้สูตรเคมีเป็นข้อมูลป้อนเข้า ข้อจำกัดสำคัญคือความคลาดเคลื่อนระหว่างการทำนายและผลการทดลองจริง ซึ่งมาจากลักษณะทั่วไปของ Fairchem และการไม่ได้คำนึงถึงปัจจัยการทดลอง เช่น ผลกระทบจากตัวทำละลายและสมบัติพลวัตอื่นๆ แนวทางแก้ไขคือการ ปรับจูนแบบจำลองด้วยข้อมูลเฉพาะ หรือใช้แนวทางการเรียนรู้แบบ delta learning โดยสรุปแล้ว งานวิจัยนี้ช่วยให้การออกแบบตัวเร่งปฏิกิริยาเป็นเรื่องง่ายขึ้นและมีส่วนอย่างมากในการค้นหาวัสดุพลังงานที่ยั่งยืน

สถานสหวิทยาการและนานาชาติ ลายมือชื่อนักศึกษา อาริษา แก้วประทุม
ปีการศึกษา 2568 ลายมือชื่ออาจารย์ที่ปรึกษา

ARISA KAEWPRATUM : INNOVATIVE CATALYST DESIGN FOR CO_2 REDUCTION REACTION: A COMBINED ULTISCALE MODELLING AND MACHINE LEARNING APPROACH.
THESIS ADVISOR : ASSOC. PROF. SUWIT SUTHIRAKUN, Ph.D. 78 PP.

Keyword: CO_2 reduction reaction, Microkinetic modelling, Machine learning models.

The levels of atmospheric CO_2 are a crucial issue, making it necessary to find innovative solutions to mitigate its impact. One promising approach is the CO_2 reduction reaction (CO_2RR) to valuable products. This thesis introduces a multiscale modeling and machine learning (ML) framework designed to accelerate the discovery of CO_2RR catalysts. The methodology began with the creation of a 1,632 structures catalyst database, including pure metals and bimetallic alloys. We used a ML potential, Fairchem, to calculate adsorption energies, which served as input for a microkinetic model to predict rates and selectivity. We also developed ML models for rapid performance prediction. An Artificial Neural Network was optimal for predicting selectivity, while a Support Vector Classifier performed best for predicting the overall rate. To enhance accessibility, a web application was created for instant predictions based on a chemical formula input. A key limitation is the discrepancy between our predictions and experimental results. This is attributed to the general nature of the Fairchem potential and the exclusion of experimental factors like solvent effects and dynamic properties. The way to improve accuracy is by fine-tuning models with specific DFT data or implementing a delta learning approach. This research provides a practical workflow for rational catalyst design, offering an alternative to traditional high-cost computational screening and contributing significantly to the field of sustainable energy material discovery.

Department of Interdisciplinary Science and Internationalization
Academic Year 2025

Student's Signature _____
Advisor's Signature _____

ACKNOWLEDGEMENTS

I am deeply grateful to my thesis advisor, Assoc. Prof. Dr. Suwit Suthirakun, and other people who give best advise me, Asst. Prof. Dr. Ittipon Fongkaew, Dr. Lappawat Ngamwongwan, and Dr. Tongjai Chookajorn for their constant support, valuable guidance, thoughtful discussions, and ongoing encouragement throughout my thesis journey. Their knowledge and dedication have been essential to the success of this research. I sincerely thank Assoc. Prof. Dr. Anyanee Kamkaew, Assoc. Prof. Dr. Rapee Utke, and Dr. Wiwat Nuansing for their time and expertise as members of my thesis committee. Their valuable feedback and guidance have significantly improved this work. I am truly thankful to the OROG for granting me a scholarship that made it possible to pursue my master's degree. I also appreciate the financial support from Suranaree University of Technology (SUT), which was crucial for completing this research. I am grateful to the NSTDA Supercomputer Center (ThaiSC) and the Institute of Science, Suranaree University of Technology, for providing the computational resources essential for this research. I also thank the faculty and staff of the Integrated Science and Innovation, Institute of Science, Suranaree University of Technology, for their support, guidance, and encouragement throughout my academic journey. I sincerely thank Dr. Lappawat Ngamwongwan and Panupol Untarabut for their guidance, patience, and practical advice on computational techniques, which greatly improved the quality of my research. I am also truly grateful to my friends and colleagues in the Computational Materials Science and Catalysis (COMSCAT) Group for their support, teamwork, and encouragement, making my academic journey both productive and enjoyable. Finally, I am truly grateful to my family and my boyfriend, Mr. Jirathip Kanphirom, for their unwavering love, support, and belief in me. I sincerely thank everyone who has supported and contributed to this work, whether directly or indirectly. Your guidance, encouragement, and assistance have been invaluable and deeply appreciated.

Arisa Kaewpratum

CONTENTS

CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF ABBREVIATIONS

CO_2RR	= CO_2 reduction reaction
CCUS	= Carbon capture, utilization, and storage
FCC	= Face center cubic
TMs	= Transition metals
XPS	= X-ray photoelectron spectroscopy
DFT	= Density functional theory
MKM	= Microkinetic modelling
ML	= Machine learning
CBFV	= Composition based feature vector
SVMs	= Support vector machines
LightGBM	= Light gradient boosting machine
ANNs	= Artificial neural networks
MAE	= Mean Absolute Error
RMSE	= Root Mean Square Error
R^2	= Coefficient of determination
COGS	= Cost of goods sold
EBTIDA	= Earnings before interest, taxes, depreciation, and amortization
EBT	= Earnings Before Taxes
BEP	= Break–even point

CHAPTER I

INTRODUCTION

The growth of human industrial activities has gradually disturbed the earth's environmental balance, leading to higher CO_2 emission (Zoubir et al., 2022) ns. The increasing of CO_2 levels in the air have caused serious problems like the greenhouse effect, which results in global warming, melting ice, and more climate change. Therefore, it is very important to reduce the amount of CO_2 in the atmosphere. Carbon capture, utilization, and storage (CCUS) is an innovative technology with the potential to address the challenges of climate change (Bagger et al., 2017). CO_2 can be captured from emission sources like power plants, energy consumption, and industrial facilities, as well as directly from the atmosphere. Once captured, the CO_2 can be used as a feedstock for chemical and fuel production or securely stored deep underground for long-term containment. To achieve net zero emissions, CO_2 utilization is considered a promising approach. So, this is the challenge for the researcher to find the solution to address this urgent mission.

For the CO_2 utilization, the electrochemical CO_2 reduction reaction (CO_2RR) is considered as the most promising technology due to the practical operating conditions, and scalability. The electrocatalytic cell consists of three main parts and two electrodes including: an anode, a separator, and a cathode. At the anode, H_2O will oxidize to O_2 and H^+ (oxidation reaction), then H^+ will diffuse through the separator into the cathode that CO_2 will be reduced to various products (reduction reaction), driven by the applied electricity (Nitopi et al., 2019). CO_2 can be reduced to valuable products, for example carbon monoxide (CO), methane (CH_4), formic acid (HCOOH), ethylene (C_2H_4), ethanol (C_2H_5OH), and others, which are highly dependent on the intrinsic properties of catalysts (Bagger et al., 2017).

From literature, the metal catalyst of CO_2RR can be obtained from major products including H_2, CO, HCOOH, and beyond CO (Schouten et al., 2011; Tang et al., 2020; Y. Wang et al., 2022). For example, the metal catalysts with major products of

H_2 are Ti, Fe, Ni, Ga, Pd, and Pt. For CO, the group of metal includes Zn, Ag, and Au. For HCOOH it involves Cd, In, Sn, Hg, Tl, and Pb. For products beyond CO, the active catalyst occurs on Cu (Y. Guan et al., 2021; Hirunsit et al., 2015; Hori et al., 1997; Peterson et al., 2010). According to the previous study, the choice of catalyst can lead to the production of different target products. However, H_2 is the result of hydrogen evolution reaction (HER) which is the competing reaction with CO_2RR (Alasali et al., 2024). Therefore, to explore the desirable product of CO_2RR the crucial process is finding the active and selective catalyst.

To improve the active and selective catalyst for CO_2RR, bimetallic catalysts are good candidates. These materials are made by combining two different metals, which work together to improve the efficiency of CO_2RR. Each metal in the catalyst plays a specific role, such as activating the CO_2 molecule or stabilizing reaction intermediates (An et al., 2017), which helps increase the activity and selectivity toward desired products (Shetty et al., 2022). By tuning the combination and composition of the two metals, researchers can enhance the performance of bimetallic catalysts, making them more effective for converting CO_2 into valuable chemicals or fuels.

As mentioned above, converting CO_2 into valuable products involves multiple steps, as illustrated in **Figure 1.1** The process typically begins with the adsorption of or activation of CO_2 on catalyst surface. This is followed by series of proton couple electron transfer steps (Bagger et al., 2017) that produces intermediate reactions, which eventually lead to the formation of the desired products (Tang et al., 2020; Y. Wang et al., 2022). The efficiency and selectivity of these pathways depend on the catalyst's ability to lower the energy barriers of the intermediate steps and prevent side reactions (Dong et al., 2019; Quan et al., 2021; Zijlstra et al., 2020). Generally, for C1 products, CH_4 is the main product at high overpotentials, while CO and HCOOH are the dominant products at lower overpotentials, along with a large amount of hydrogen, for most transition metal electrodes (Hori et al., 1985, 1986, 2002). CH_3OH is often reported as a side product at higher overpotentials on transition metal (Kuhl et al., 2014; Schouten et al., 2011). So, the optimal design of catalysts must go through a long-term trial and error process, which requires a lot of time and resources.

To better understand the mechanism of CO_2RR, microkinetic modeling is a tool that can help the researchers to enhance the process of catalyst design (X. Wang et al., 2025; Zijlstra et al., 2020). Microkinetic modeling is used to identify critical reaction intermediates, rate-determining elementary reactions, and selectivity of target product under various operating conditions (Filot, 2023). For example, Zijlstra et al. explored reaction energetics for CO_2 reduction on Cu(111) and Cu(211), obtained from DFT calculations, into microkinetic simulations to evaluate the effects of surface topology, the presence of water, and potential diffusion limitations. Their study predicted the impact of these factors on current density-potential relationships and Faradaic efficiencies (Zijlstra et al., 2020). They found that the presence of a catalytic H_2O molecule increases the overall rate and selectivity to products CO and CH_4 derived from dissociated CO_2.

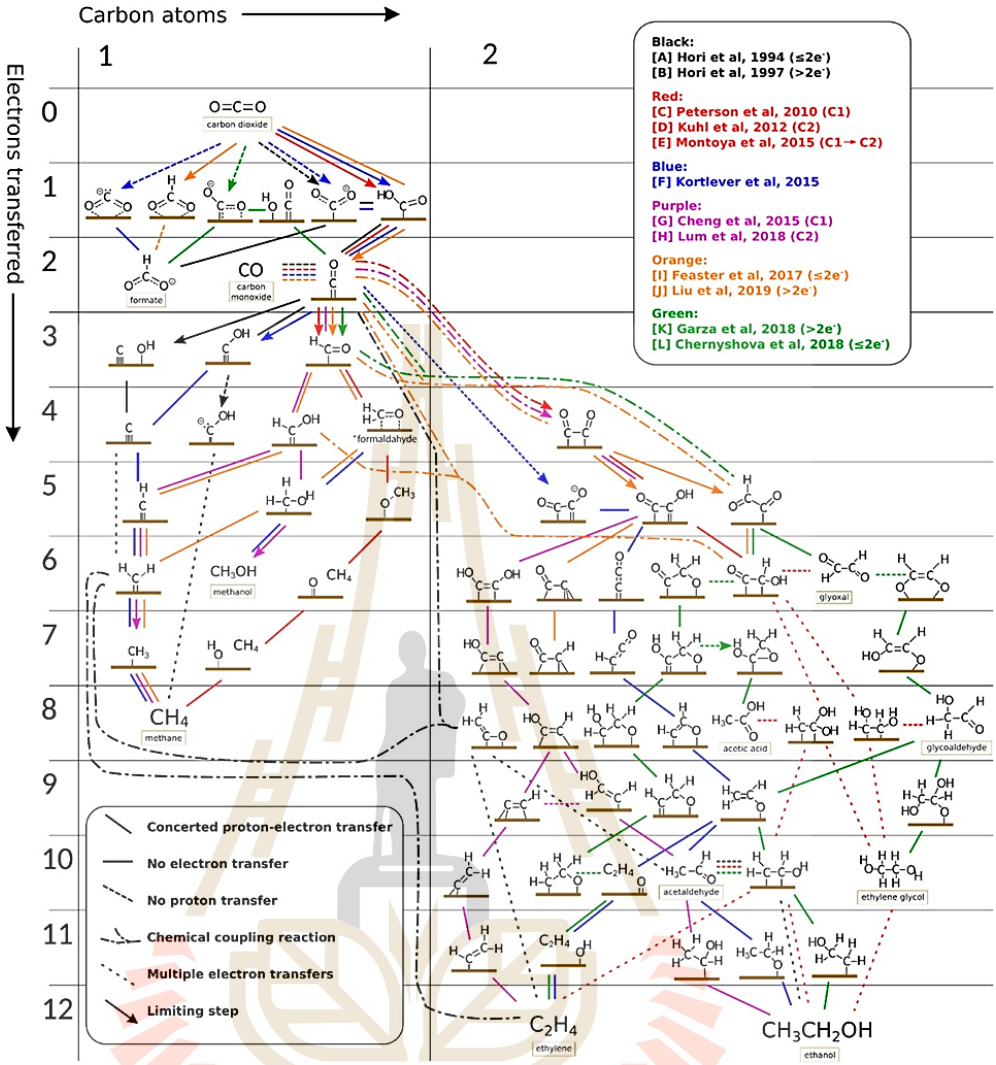


Figure 1.1 Possible reaction pathways of CO_2 reduction to C1 and C2 products on polycrystalline copper (Nitopi et al., 2019).

With rapid progress in computational science, artificial intelligence (AI) technology has become a global trend in research and development (Liu et al., 2022; Xu et al., 2021). Compared to traditional trial-and-error methods, high-throughput material calculations offer advantages by significantly reducing both the time and cost required for research (S. Guan et al., 2021; Kim et al., 2021; Zhu et al., 2019). Among AI techniques, machine learning (ML) has attracted particular interest from researchers. ML can establish complex relationships between inputs and outputs, which are often linked to the properties of potential catalysts (X. Wang et al., 2020). This powerful tool is widely used in four key areas of materials science: enabling the discovery and

synthesis of new materials (Frey et al., 2019), predicting material properties (Zhang et al., 2024), analyzing material microstructures (Y. C. Li et al., 2019), and speeding up computational simulations (Morgan and Jacobs, 2020).

ML is an effective tool for predicting catalytic activity, optimization of catalysts, identification of active sites, and design of reaction pathways in the CO_2RR. With the continuous development of ML, various studies in this field use ML to predict the properties of catalysts. Li et al. present a holistic machine-learning framework for rapid screening of bimetallic catalysts with the aid of the descriptor-based kinetic analysis (Z. Li et al., 2017). The adsorption energies of *CO and *OH on (111)-terminated metal surfaces, along with key features of active sites, use as output and input respectively, based on DFT calculations. For approximately 1,000 alloy surfaces, the artificial neural network (ANN) model demonstrated strong predictive accuracy, enabling efficient exploration of the vast chemical space of bimetallic catalysts. Ma et al. combine DFT calculation and ML prediction to predict the yield of copper-catalysed P–H insertion reactions (Ma et al., 2023). ML algorithms were used to analyze 16 descriptors obtained from DFT, transition state calculations to predict product yields. Among the algorithms tested, the Support Vector Machine (SVM) showed the highest accuracy in predictions. The results of the SVM model aligned well with experimental yield data, demonstrating its reliability.

Recently, combining ML with computational methods has led to significant advancements in electrocatalysis research. In this work, we combine multi-scale modelling and machine learning approaches to develop the framework of bimetallic catalyst design for CO_2RR. Moreover, ML model is used to predict the reaction outcome as product selectivity and reaction rate of CO_2RR.

1.1 Research objectives

This work involves three steps. First, we plan to create a clear framework for catalyst design specifically for CO_2RR. This guide will provide a solid way to do all the necessary calculations and modeling later. Second, the thesis focuses on building a Machine Learning (ML) model designed to estimate the reaction's overall rate and how

selectivity based on bimetallic composition. Finally, to make this prediction tool easy to use in the real world, we will build a simple, easy-to-use website. This website is an important result of the project, as it lets researchers and industry users quickly get the rate and selectivity predictions from the ML model, making the results useful for making decisions.

1.2 Scope and limitations

This thesis focuses on CO_2 reduction to C1 products, including HCOOH, CO, CH_3OH, and CH_4, with the goal of establishing a framework for catalyst design. We began by constructed the surface model from 34 pure metals, comprising 3d (Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn), 4d (Y, Zr, Nb, Mo, Tc, Ru, Rh, Pd, Ag, Cd), 5d (Hf, Ta, W, Re, Os, Ir, Pt, Au) transition metals, and p-block elements (Al, Ga, In, Sn, Tl, Pb). The investigation is extended to bimetallic compositions 1:1, and 1:3.

Computational data (bimetallic composition, reaction energies, and energy profile) are generated using Fairchem, a centralized repository developed by FAIR Chemistry (Chanussot et al., 2021) and serve as inputs for the microkinetic model (MKM). The MKM is employed to evaluate the reaction behavior of each catalyst, providing insights into reaction rates and product yield and selectivity. To further enhance efficiency, ML models, including various regression and classification models. The goal is to predict reaction outcomes—such as rates and selectivity—for the remaining bimetallic based on their composition. Therefore, the features are the bimetallic compositions and penitential (U), and the target are the reaction rate and product selectivity.

This approach combines multi-scale modelling and ML techniques to accelerate the identification of active and selective catalysts for CO_2 reduction, offering valuable insights into the design of materials for sustainable chemical processes.

1.3 References

Alasali, F., Abuashour, M. I., Hammad, W., Almomani, D., Obeidat, A. M., and Holderbaum, W. (2024). A review of hydrogen production and storage materials for efficient integrated hydrogen energy systems. In *Energy Science and Engineering* (Vol. 12, Issue 5, pp. 1934–1968). John Wiley and Sons Ltd. https://doi.org/10.1002/ese3.1723

An, H., Ha, H., Yoo, M., and Kim, H. Y. (2017). Understanding the atomic-level process of CO-adsorption-driven surface segregation of Pd in (AuPd)147 bimetallic nanoparticles. *Nanoscale*, *9*(33), 12077–12086. https://doi.org/10.1039/c7nr04435f

Bagger, A., Ju, W., Varela, A. S., Strasser, P., and Rossmeisl, J. (2017). Electrochemical CO2 Reduction: A Classification Problem. *ChemPhysChem*, *18*(22), 3266–3273. https://doi.org/10.1002/cphc.201700736

Chanussot, L., Das, A., Goyal, S., Lavril, T., Shuaibi, M., Riviere, M., Tran, K., Heras-Domingo, J., Ho, C., Hu, W., Palizhati, A., Sriram, A., Wood, B., Yoon, J., Parikh, D., Zitnick, C. L., and Ulissi, Z. (2021). Open Catalyst 2020 (OC20) Dataset and Community Challenges. *ACS Catalysis*, *11*(10), 6059–6072. https://doi.org/10.1021/acscatal.0c04525

Dong, H., Liu, C., Li, Y., and Jiang, D. E. (2019). Computational screening of M/Cu core/shell nanoparticles and their applications for the electro-chemical reduction of CO2 and CO. *Nanoscale*, *11*(23), 11351–11359. https://doi.org/10.1039/c9nr01936g

Filot, Ivo. (2023). *Introduction to microkinetic modeling*. Technische Universiteit Eindhpoven.

Frey, N. C., Wang, J., Vega Bellido, G. I., Anasori, B., Gogotsi, Y., and Shenoy, V. B. (2019). Prediction of synthesis of 2D metal carbides and nitrides (MXenes) and their precursors with positive and unlabeled machine learning. *ACS Nano*, *13*(3), 3031–3041.

Guan, S., Shang, C., and Liu, Z. (2021). Structure and dynamics of energy materials from machine learning simulations: A topical review. *Chinese Journal of Chemistry*, *39*(11), 3144–3154.

Guan, Y., Suo, W., Zhang, Z., Wang, Y., Sun, S., and Liu, G. (2021). Insights on the Catalytic Active Site for CO2 Reduction on Copper-based Catalyst: A DFT study. *Molecular Catalysis*, *511*. https://doi.org/10.1016/j.mcat.2021.111725

Hirunsit, P., Soodsawang, W., and Limtrakul, J. (2015). CO2 electrochemical reduction to methane and methanol on copper-based alloys: Theoretical insight. *Journal of Physical Chemistry C*, *119*(15), 8238–8249. https://doi.org/10.1021/acs.jpcc.5b01574

Hori, Y., Kikuchi, K., Murata, A., and Suzuki, S. (1986). PRODUCTION OF METHANE AND ETHYLENE IN ELECTROCHEMICAL REDUCTION OF CARBON DIOXIDE AT COPPER ELECTRODE IN AQUEOUS HYDROGENCARBONATE SOLUTION. *Chemistry Letters*, *15*(6), 897–898. https://doi.org/10.1246/cl.1986.897

Hori, Y., Kikuchi, K., and Suzuki, S. (1985). PRODUCTION OF CO AND CH4 IN ELECTROCHEMICAL REDUCTION OF CO2 AT METAL ELECTRODES IN AQUEOUS HYDROGENCARBONATE SOLUTION. *Chemistry Letters*, *14*(11), 1695–1698. https://doi.org/10.1246/cl.1985.1695

Hori, Y., Takahashi, I., Koga, O., and Hoshi, N. (2002). Selective Formation of C2 Compounds from Electrochemical Reduction of CO2 at a Series of Copper Single Crystal Electrodes. *The Journal of Physical Chemistry B*, *106*(1), 15–17. https://doi.org/10.1021/jp013478d

Hori, Y., Takahashi, R., Yoshinami, Y., and Murata, A. (1997). *Electrochemical Reduction of CO at a Copper Electrode*.

Kim, J., Kang, D., Kim, S., and Jang, H. W. (2021). Catalyze materials science with machine learning. *ACS Materials Letters*, *3*(8), 1151–1171.

Kuhl, K. P., Hatsukade, T., Cave, E. R., Abram, D. N., Kibsgaard, J., and Jaramillo, T. F. (2014). Electrocatalytic Conversion of Carbon Dioxide to Methane and Methanol on Transition Metal Surfaces. *Journal of the American Chemical Society*, *136*(40), 14107–14113. https://doi.org/10.1021/ja505791r

Li, Y. C., Wang, Z., Yuan, T., Nam, D. H., Luo, M., Wicks, J., Chen, B., Li, J., Li, F., De Arquer, F. P. G., Wang, Y., Dinh, C. T., Voznyy, O., Sinton, D., and Sargent, E. H. (2019). Binding Site Diversity Promotes CO2 Electroreduction to Ethanol. *Journal*

of the American Chemical Society, *141*(21), 8584–8591. https://doi.org/10.1021/jacs.9b02945

Li, Z., Wang, S., Chin, W. S., Achenie, L. E., and Xin, H. (2017). High-throughput screening of bimetallic catalysts enabled by machine learning. *Journal of Materials Chemistry A*, *5*(46), 24131–24138.

Liu, F., Wang, Z., Wang, Z., Zhong, J., Zhao, L., Jiang, L., Zhou, R., Liu, Y., Huang, L., Tan, L., Tian, Y., Zheng, H., Fang, Q., Zhang, L., Zhang, L., Wu, H., Bai, L., and Zhou, K. (2022). High-Throughput Method–Accelerated Design of Ni-Based Superalloys. *Advanced Functional Materials*, *32*(28), 2109367. https://doi.org/https://doi.org/10.1002/adfm.202109367

Ma, Y., Zhang, X., Zhu, L., Feng, X., Kowah, J. A. H., Jiang, J., Wang, L., Jiang, L., and Liu, X. (2023). Machine Learning and Quantum Calculation for Predicting Yield in Cu-Catalyzed P–H Reactions. *Molecules*, *28*(16). https://doi.org/10.3390/molecules28165995

Morgan, D., and Jacobs, R. (2020). Opportunities and challenges for machine learning in materials science. *Annual Review of Materials Research*, *50*(1), 71–103.

Nitopi, S., Bertheussen, E., Scott, S. B., Liu, X., Engstfeld, A. K., Horch, S., Seger, B., Stephens, I. E. L., Chan, K., Hahn, C., Nørskov, J. K., Jaramillo, T. F., and Chorkendorff, I. (2019). Progress and Perspectives of Electrochemical CO2 Reduction on Copper in Aqueous Electrolyte. In *Chemical Reviews* (Vol. 119, Issue 12, pp. 7610–7672). American Chemical Society. https://doi.org/10.1021/acs.chemrev.8b00705

Peterson, A. A., Abild-Pedersen, F., Studt, F., Rossmeisl, J., and Nørskov, J. K. (2010). How copper catalyzes the electroreduction of carbon dioxide into hydrocarbon fuels. *Energy and Environmental Science*, *3*(9), 1311–1315. https://doi.org/10.1039/c0ee00071j

Quan, Y., Zhu, J., and Zheng, G. (2021). Electrocatalytic Reactions for Converting CO2 to Value-Added Products. In *Small Science* (Vol. 1, Issue 10). John Wiley and Sons Inc. https://doi.org/10.1002/smsc.202100043

Schouten, K. J. P., Kwon, Y., van der Ham, C. J. M., Qin, Z., and Koper, M. T. M. (2011). A new mechanism for the selectivity to C1 and C2 species in the electrochemical

reduction of carbon dioxide on copper electrodes. *Chemical Science*, *2*(10), 1902–1909. https://doi.org/10.1039/C1SC00277E

Shetty, S., Gayen, M., Agarwal, S., Chatterjee, D., Singh, A., and Ravishankar, N. (2022). Tuning Catalytic Activity in Ultrathin Bimetallic Nanowires via Surface Segregation: Some Insights. *Journal of Physical Chemistry Letters*, *13*(3), 770–776. https://doi.org/10.1021/acs.jpclett.1c03852

Tang, M. T., Peng, H., Lamoureux, P. S., Bajdich, M., and Abild-Pedersen, F. (2020). From electricity to fuels: Descriptors for C1 selectivity in electrochemical CO2 reduction. *Applied Catalysis B: Environmental*, *279*. https://doi.org/10.1016/j.apcatb.2020.119384

Wang, X., Ma, F., Wei, H., Pan, J., Li, W., Zhao, J., and Yang, X. (2025). A DFT and microkinetic modeling study of pressure effects on electroreduction reduction of CO2 to ethanol. *Applied Surface Science*, *681*. https://doi.org/10.1016/j.apsusc.2024.161421

Wang, X., Ye, S., Hu, W., Sharman, E., Liu, R., Liu, Y., Luo, Y., and Jiang, J. (2020). Electric dipole descriptor for machine learning prediction of catalyst surface–molecular adsorbate interactions. *Journal of the American Chemical Society*, *142*(17), 7737–7743.

Wang, Y., Zheng, M., Wang, X., and Zhou, X. (2022a). Electrocatalytic Reduction of CO2 to C1 Compounds by Zn-Based Monatomic Alloys: A DFT Calculation. *Catalysts*, *12*(12). https://doi.org/10.3390/catal12121617

Wang, Y., Zheng, M., Wang, X., and Zhou, X. (2022b). Electrocatalytic Reduction of CO2 to C1 Compounds by Zn-Based Monatomic Alloys: A DFT Calculation. *Catalysts*, *12*(12). https://doi.org/10.3390/catal12121617

Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., and Qiu, C. (2021). *Artificial intelligence: A powerful paradigm for scientific research. The Innovation, 2 (4), 100179*.

Zhang, Q.-M., Wang, Z.-Y., Zhang, H., Liu, X.-H., Zhang, W., and Zhao, L.-B. (2024). *Supplementary information Micro-kinetic Modelling of CO Reduction Reaction on Single Atom Catalysts Accelerated by Machine Learning*.

Zhu, X., Yan, J., Gu, M., Liu, T., Dai, Y., Gu, Y., and Li, Y. (2019). Activity origin and design principles for oxygen reduction on dual-metal-site catalysts: a combined density functional theory and machine learning study. *The Journal of Physical Chemistry Letters*, *10*(24), 7760–7766.

Zijlstra, B., Zhang, X., Liu, J. X., Filot, I. A. W., Zhou, Z., Sun, S., and Hensen, E. J. M. (2020). First-principles microkinetics simulations of electrochemical reduction of CO2 over Cu catalysts. *Electrochimica Acta*, *335*. https://doi.org/10.1016/j.electacta.2020.135665

Zoubir, O., Atourki, L., Ait Ahsaine, H., and BaQais, A. (2022). Current state of copper-based bimetallic materials for electrochemical CO2 reduction: a review. In *RSC Advances* (Vol. 12, Issue 46, pp. 30056–30075). Royal Society of Chemistry. https://doi.org/10.1039/d2ra05385c

CHAPTER II

LITERATURE REVIEW

2.1 Overview of carbon dioxide reduction reaction (CO_2RR)

Reducing atmospheric CO_2 is essential to mitigate the impact of greenhouse gases on global warming. Electrocatalytic CO_2 reduction reactions offer a promising approach for CO_2 utilization by converting it into valuable hydrocarbon products (Jouny et al., 2018). The electrocatalytic cell, as illustrated in **Figure 2.1**, consists of three main components: two electrodes (an anode and a cathode) and a separator. At the anode (blue), water undergoes oxidation to produce O_2 and H+ ions (oxidation reaction). The H+ ions then migrate through the separator (dotted line) to the cathode (orange), where CO_2 is reduced to various products (reduction reaction) under the influence of applied electricity. The general form of cathodic reaction is following:

$$xCO_2 + nH+ + e- \longrightarrow \text{product} + yH2O$$

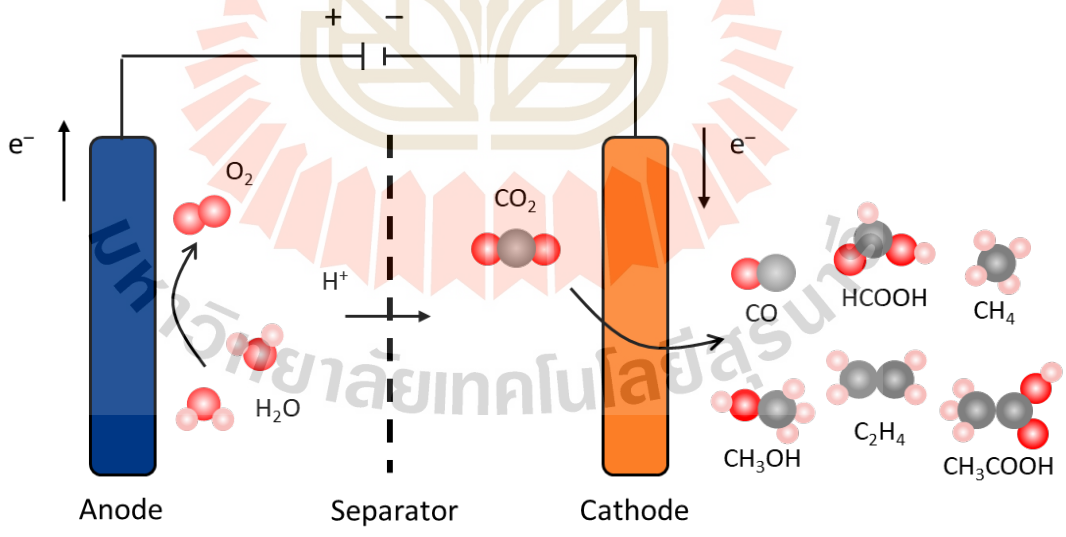


Figure 2.1 Schematic diagram of the component of electrolytic cell (Lei et al., 2023).

From literature, the metal catalyst of CO_2RR can be obtained from major products including, H_2, CO, HCOOH, and beyond CO (Schouten et al., 2011; Tang et al., 2020; Y. Wang et al., 2022). For example, the metal group with major products of H2

are Ti, Fe, Ni, Ga, Pd, and Pt. CO, the group of metal including Zn, Ag, and Au. HCOOH including Cd, In, Sn, Hg, Tl, and Pb. For beyond CO, the active catalyst occurs on Cu (Guan et al., 2021; Hirunsit et al., 2015; Hori et al., 1997; Peterson et al., 2010), which are illustrated in **Figure 2.2**. So, according to the previous study indicate that, the choice of catalyst can lead to the production of different target products. However, H2 is the result of hydrogen evolution reaction (HER) which is the competing reaction with CO_2RR (Alasali et al., 2024). Therefore, to explore the desirable product of CO_2RR the crucial process is finding the active and selective catalyst.

2.2 Classification of catalyst for CO_2RR

As I mentioned, Metal catalysts have gained significant attention for CO_2RR at the cathode, as different metals exhibit varying activities and selectivity toward specific value-added products (Hori, 2008). This section we will explore the efficient of catalyst and their composition for CO_2RR.

2.2.1 Monometallic catalyst

Metal catalysts can produce various products. These catalysts can also generate a diverse range of products through the CO_2RR. These catalysts can be roughly classified into four categories, as show in **Figure 2.2**: (1) H_2: metals including Ni, Fe, Pt, Ti, and other related metals, which have a higher hydrogen evolution potential and are prone to promoting HER side reactions; (2) Metals can strongly adsorb *COOH intermediates such as Au, (Y. Wang et al., 2022) Ag, (Han et al., 2021) Zn, (Rosen et al., 2015) Pd, (Chang et al., 2020) and Ni, (X. Li et al., 2017) but exhibit weak adsorption of CO intermediates, resulting in CO being the main product of electrocatalytic CO_2 reduction; (3) Metal catalysts which can effectively stabilize CO_2*– intermediate to prevent further reduction and promote surface electron/proton reactions such as Pb, (P. Wang et al., 2021) Zn, (Xiang et al., 2021) Sn, (Chen et al., 2020) and Cd (Huang et al., 2019) with HCOOH as the main product of CO_2 electrocatalytic reduction; (4) Cu-based catalysts which can produce various C1, C2, and C2+ products (Gao et al., 2019) and so on. Furthermore, the product selectivity of various metal catalysts can be further adjusted by changing their chemical compositions, surface structures, or electrolytes (Feng et al., 2020).

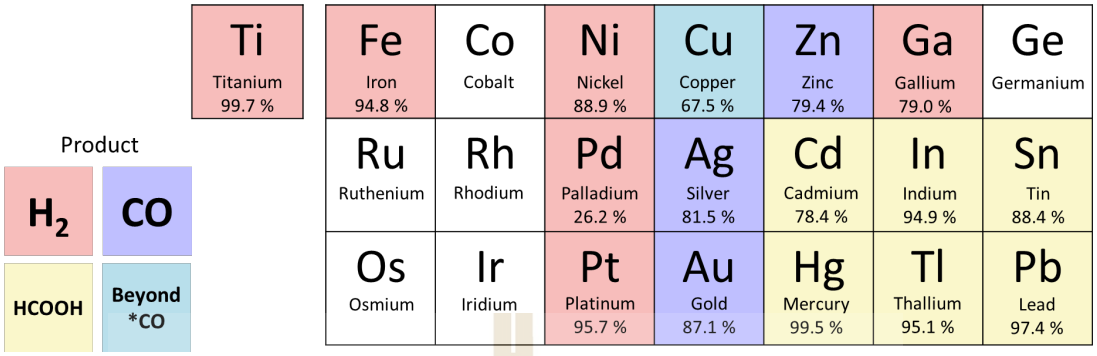


Figure 2.2 A classification of metal catalyst by target product of CO_2RR (Bagger et al., 2017).

2.2.2 Bimetallic catalyst

To improve the activity and product selectivity of CO_2RR, bimetallic exhibit the good performance (Farsi & Deskins, n.d.; Shetty et al., 2022). In electrocatalysts, bimetallic materials can provide a synergistic interface effect different from each single metal (Aljabour et al., 2016; Dhifallah et al., 2016; Saqlain et al., 2016; Yu et al., 2022). An interface effect refers to the interaction between two or more different elements forming alloys or heterojunctions with similar structures in a kind of catalyst. Interface effects play a crucial role in the application of electrocatalysts. Specifically, the catalytic activities, stabilities, selectivity, and controlled cost of catalysts can also provide the potential for fast development of bimetallic (Zhu et al., 2020).

In principle, the strength of adsorbate like CO on metallic surface could be used to predict the catalyst selectivity in CO_2 reduction. Kim et al. reported the CO_2 reduction results for AuCu alloy nanomaterials with different ratio of Au and Cu (Kim et al., 2014). They found that Au_3Cu nanoparticles (NPs) showed a high CO with FE of 66% and a mass activity of 230 A/g at −0.73 V vs. reversible hydrogen electrode (RHE) in **Figure 2.3**.

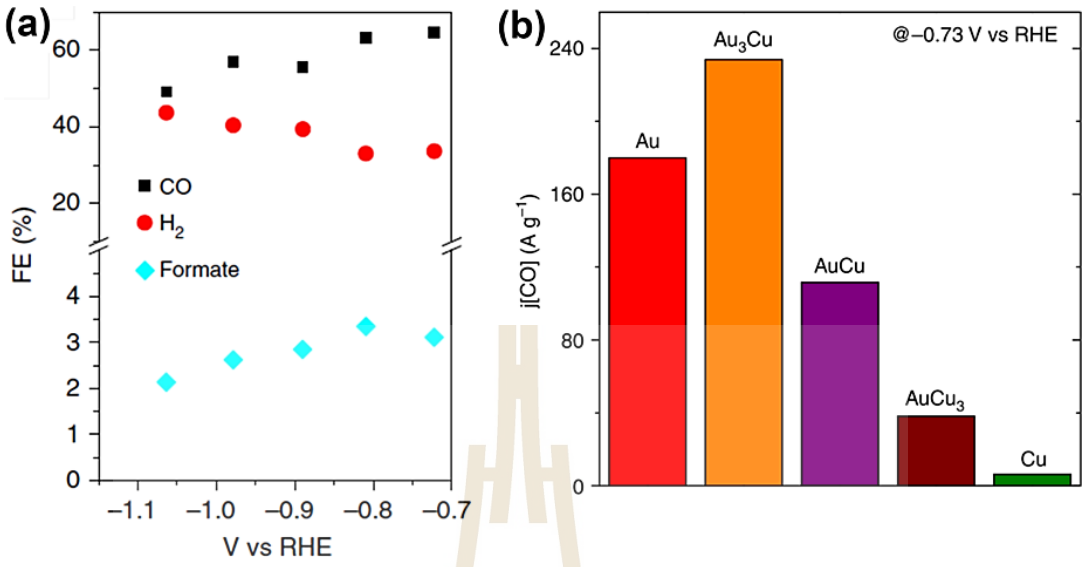


Figure 2.3 a) FE of CO, H_2 and HCOO produced with Au_3Cu, b) mass activity of CO using AuCu bimetallic NPs with different composition at −0.73 V vs RHE.

For C1 hydrocarbon production, the most accepted pathway was sequential hydrogenation of adsorbed CO* until CH_4. The potential determining step is the formation of CHO* (Nie et al., 2013) or COH*, and the subsequent reaction steps were exergonic (Dong et al., 2018). CO* required COOH* as an intermediate from initial CO_2 reduction (Hussain et al., 2018). If HCOO* formed, the final product would be HCOOH (Q. Li et al., 2017). For another C1 product, methanol, was also observed during CO_2RR. However, in this case, the intermediate was COH*, which required that hydrogen should bond to oxygen instead of carbon (Nie et al., 2014) as in **Figure 2.4**. The formation of COH* usually showed a higher kinetic barrier than that of CHO* (Hirunsit et al., 2015).

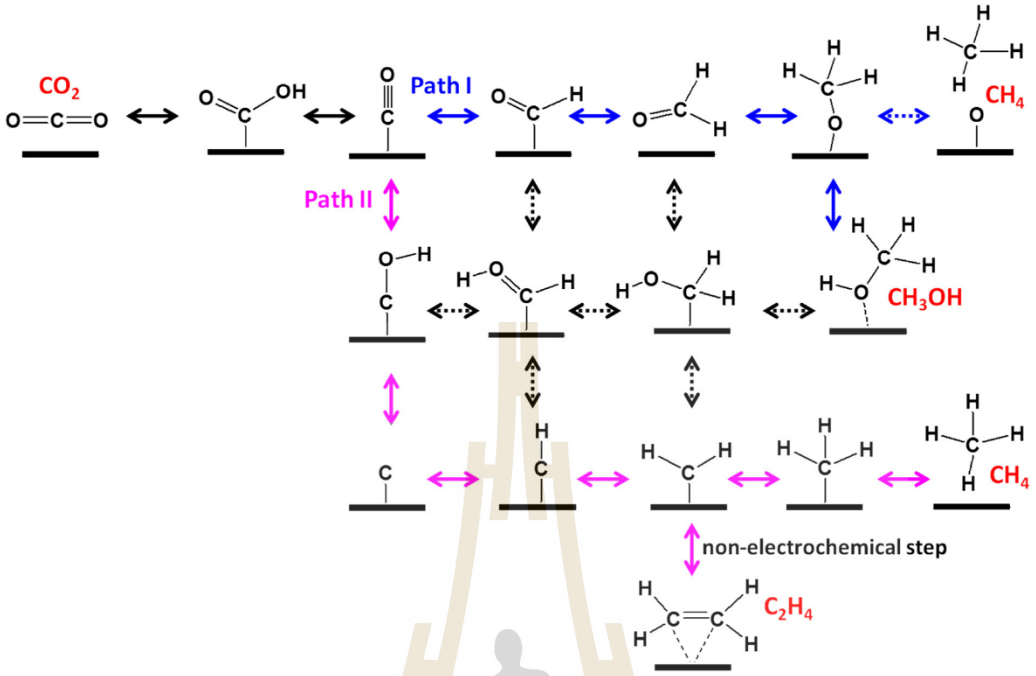


Figure 2.4 Proposed mechanism for CO_2RR electrochemical reduction on Cu(1 1 1), producing methane, methanol, and ethylene (Nie et al., 2014).

Many researchers have studied bimetallic to improve selectivity of C1 products. For example, Hahn et al. investigated the properties of thin film AuPd alloys (Hahn et al., 2015). Electrocatalytic results showed that selectivity and activity of the as-synthesized AuPd alloy films were higher for $HCOO^-$ than either pure Au or Pd metals, indicating that both metals can act synergistically. Zhang et al. synthesized 2D hierarchical Pd/SnO_2 nanosheets for CO_2RR (Zhang et al., 2018). They showed high FE (54.8%) for CO_2 to methanol at -0.24 V with a total current density of −2.5 mA/cm2. Another study of Cu-Pd NPs synthesized by underpotential showed a higher formate FE and a good stability (Yin et al., 2016). The results were attributed to decreasing the adsorption strength of CO via modification of the Pd, leading to an improved CO* adsorption and increasing FE %.

As previously mentioned, many bimetallic catalysts have been designed to enhance CO adsorption, and hydrocarbon product selectivity (Zhu et al., 2018). However, the important consideration for designing improved bimetallic CO2RR catalysts is to tune the intermediate binding energy, which can be achieved by selecting the appropriate bimetallic composition with guidance from computational study (Dong et al., 2019; Z. Li et al., 2017).

2.2.3 Bimetallic composition

Various strategies have been employed to enhance the activity of bimetal catalysts. The catalyst interface can be engineered to contain multiple sites that contribute to the binding of key intermediates in the pathway (He et al., 2017). Alloying of bimetals is one of the common methods to enhance catalytic performance (Ferrando et al., 2008; Hirunsit et al., 2015; Y. Wang et al., 2022; Zhao et al., 2020). **Figure 2.5**. show the possible mixing pattern in bimetallic alloys.

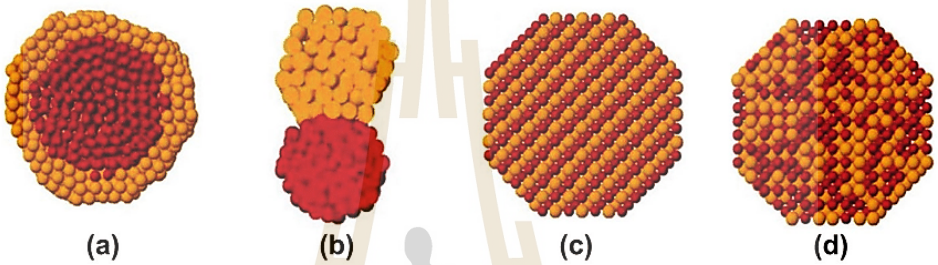


Figure 2.5 possible mixing pattern of bimetallic alloy (a) the core–shell model and (b) a phase-separated alloy. Solid-solution alloys include (c) an ordered with long-range order and (d) random alloys with short-range order (Wu et al., 2016).

Core/shell nanostructured materials can lead to both electronic modification and surface strain effects. For example, Li et al. reported that Cu/SnO_2 NPs have a uniform SnO_2 shell on a spherical Cu core (Q. Li et al., 2017). They found that SnO_2 thickness-dependent effect on CO_2RR. NPs with thicker SnO_2 shell (1.8 nm) able to produce formate with high FE, as shown in **Figure 2.6**. They also explored the free energy from DFT calculation and found that, COOH* is the key intermediate for CO formation while the HCOO* intermediate can only produce formate.

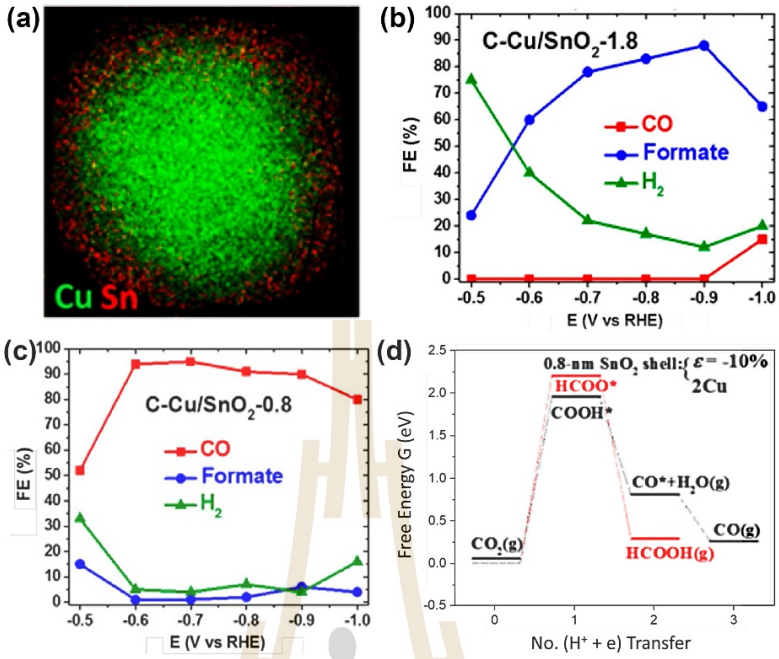


Figure 2.6 (a) TEM image of Cu/SnO_2 core/shell NPs. FE for CO, H_2 and formate during CO_2RR by using (b) C–Cu/SnO_2-0.8 catalyst and (c) C–Cu/SnO_2-1.8 catalyst. (d) Free energy diagrams for CO and formate formation. (Q. Li et al., 2017)

Ma et al. showed that phase-separated Cu–Pd catalysts can reach high 50% FE for C_2H_4 (S. Ma et al., 2017). Ordered, disordered, and phase separated Cu–Pd catalysts were synthesized, as shown in **Figure 2.7**. The major products are CO, CH_4, C_2H_4, and C_2H_5OH. Cu NPs and phase separated CuPd showed strongest and weakest binding with CO, both showed similar activity. Therefore, the arrangement is important, with the orientation of the intermediate on the surface playing a key role for enhanced products formation.

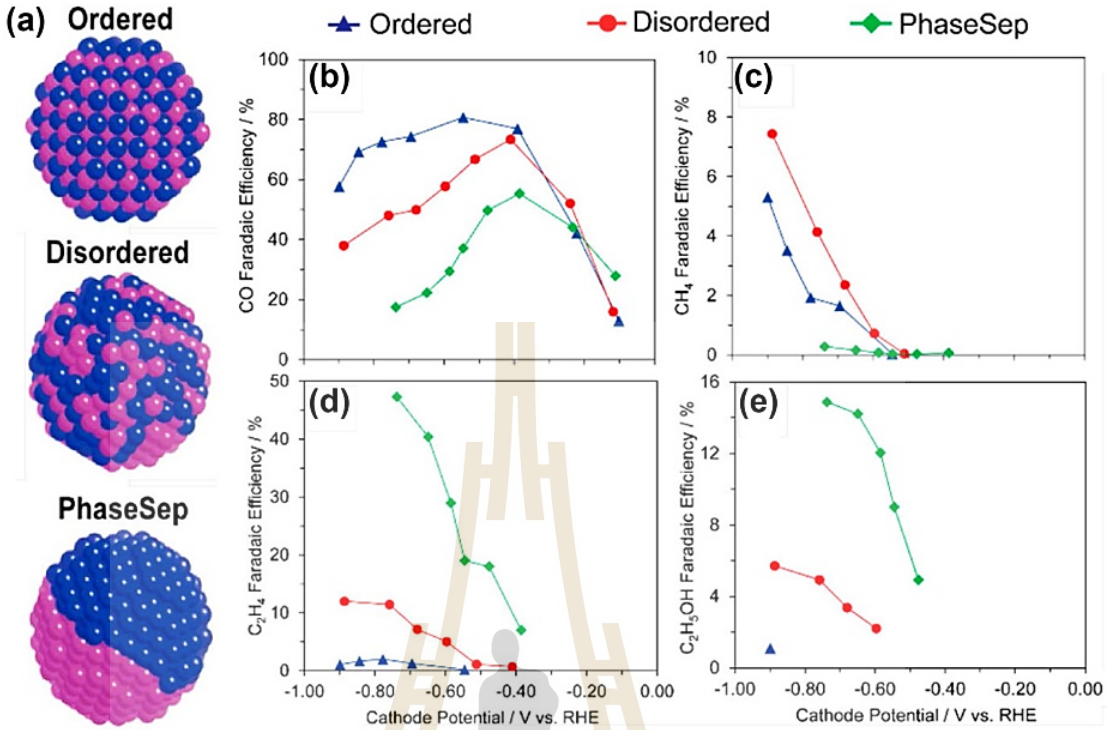


Figure 2.7 (a) The model of ordered, disordered and phase separated CuPd nanoalloys. FEs for (b) CO; (c) CH_4; (d) C_2H_4; (e) C_2H_5OH for CuPd catalyst (S. Ma et al., 2017).

For another alloy model, Liu et al. investigated Cu_3M (M = Ag, Au, and Zn) catalyst to provide a theoretical basis for exploring the CO_2RR catalyst. They found that CO adsorption behavior and the surface electronic structure affect catalyst activity (F. Liu et al., 2022). Lu et al. study the activity of different surface alloy orientations of CoCu on CO_2RR activity. They found that free energy of CO_2 reduction CH_3OH on alloy surface is more thermodynamically than original one (S. Liu et al., 2019).

However, to successfully evaluate the mechanisms of reactivity, the surface and bulk compositions, and structure before and after the CO_2RR should be considered. The identification of intermediates formed upon activation and reduction of CO_2 is important to understanding the reaction mechanism.

2.3 Microkinetic modelling (MKM)

The microkinetic model is a powerful tool for identifying the behavior of CO_2RR catalysts under operating conditions (Filot, 2023). A general approach to developing a microkinetic model involves first creating a reaction scheme that includes all elementary reaction steps for the conversion of reactants into products, as shown in **Figure 2.8**. This means accounting for all reactants, products, and reaction intermediates in the mechanism. However, while it is important to include all relevant intermediates on the catalyst surface, this should be balanced with the ability to describe intermediates that can be experimentally measured or theoretically predicted. Next, the thermodynamic properties of all adsorbate species and the kinetic parameters for all elementary reactions are calculated or estimated. Since the reaction rate is typically influenced by a few key factors, such as the activation energies and the adsorption energies of the intermediates.

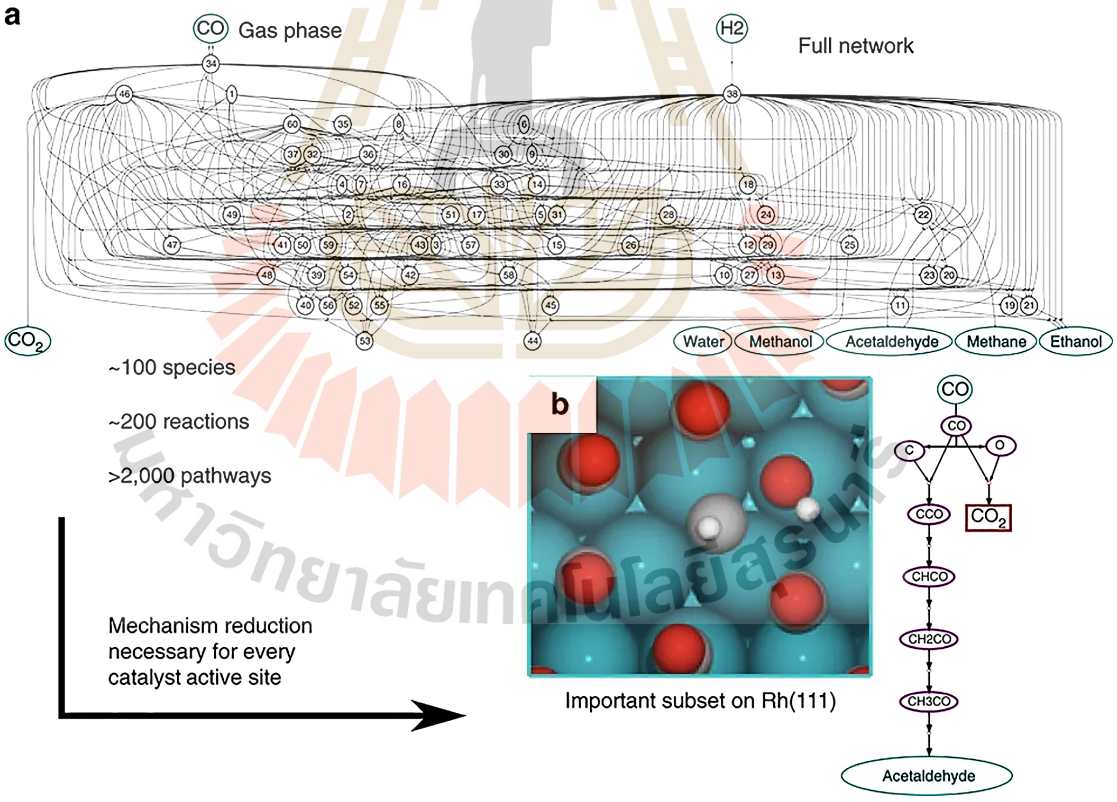


Figure 2.8 (a) Reaction network for the conversion of syngas to carbon dioxide, water, methane, methanol, acetaldehyde, and ethanol. (b) Reduced network for acetaldehyde production from syngas on Rh(111) (Ulissi, Medford, et al., 2017).

Many studies have used microkinetic models to explore the CO_2RR mechanism. For example, Zijlstra et al. performed microkinetic to compute CO_2 conversion and product formation rates as a function of the applied electrochemical potential over Cu catalyst (Zijlstra et al., 2020). The results show that hydrogen evolution and CO_2 reduction compete on these surfaces. The stepped Cu(211) surface exhibits higher activity compared to the Cu(111). While Cu(211) primarily produces HCOO as the main CO_2 reduction product, the presence of a catalytic H_2O molecule enhances the overall rate and improves selectivity toward more reduced products, such as CO and CH_4. Wang et al. combining between DFT calculation and microkinetic to study pressure effect for CO_2 reduction to ethanol (X. Wang et al., 2025). The proposed ethanol formation pathway is $CO_2(g) \longrightarrow *CO_2 \longrightarrow *COOH \longrightarrow *CO + H_2O\ (g) \longrightarrow *CHO \longrightarrow *CHOCO \longrightarrow *CHOCHO \longrightarrow *CH_2OCHO \longrightarrow *CH_2OHCHO \longrightarrow *CH_2CHO + H_2O\ (g) \longrightarrow *CH_3CHO \longrightarrow *CH_3CH_2O \longrightarrow *CH_3CH_2OH \longrightarrow CH_3CH_2OH(g)$, as shown in **Figure 2.9**. The high coverage of *CO has significant effects on the *CHO–*CO addition step. The increased coverage of *CO first facilitates ethanol production.

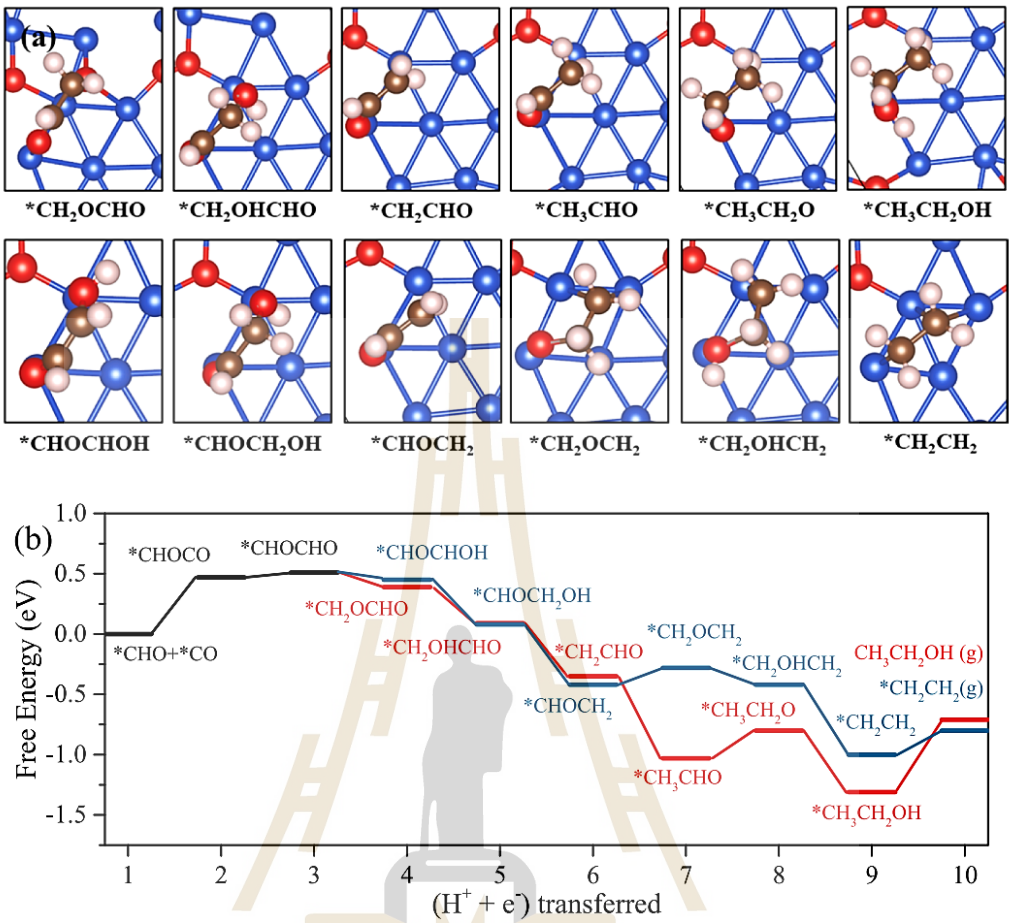


Figure 2.9 (a) Adsorption configurations of the species and (b) free energy diagram from *CHOCO to ethanol and C_2H_4 (X. Wang et al., 2025).

2.4 Machine learning prediction for catalyst screening

Recently, machine learning (ML) has been applied to high-throughput screening of catalysts and has shown to be effective and reliable across a wide range of catalysts. ML algorithms, such as Decision tree, Light Gradient-Boosting Machine (LightGBM), Support Vector Machine (SVM), and Artificial Neural Networks (ANNs), are used to build predictive models for important catalyst properties, including stability, activity, and selectivity (Goldsmith et al., 2018).

Wexler et al. investigated the catalytic activity of the hydrogen evolution reaction on $Ni_2P(0001)$ nanoclusters, focusing on the effects of different adsorption site structures (Wexler et al., 2018). The influence of surface doping on the surface structure, charge states, and HER activity, as shown in **Figure 2.10**, was calculated using DFT. A regularized random forest machine learning algorithm was applied to identify the relative importance of structural and charge descriptors, revealing that the Ni–Ni bond length is a key factor in determining the HER activity of $Ni_2P(0001)$. Li et al. proposed a comprehensive machine learning framework for the rapid screening of bimetallic catalysts for methanol electro-oxidation (Z. Li et al., 2017). The adsorption energies of *CO and *OH on (111)-terminated metal surfaces, along with active sites, were used as input and output variables, derived from DFT calculations. For around 1,000 idealized alloy surfaces, the trained ML models demonstrated strong predictive accuracy, enabling efficient exploration of the vast chemical space of bimetallic catalysts.

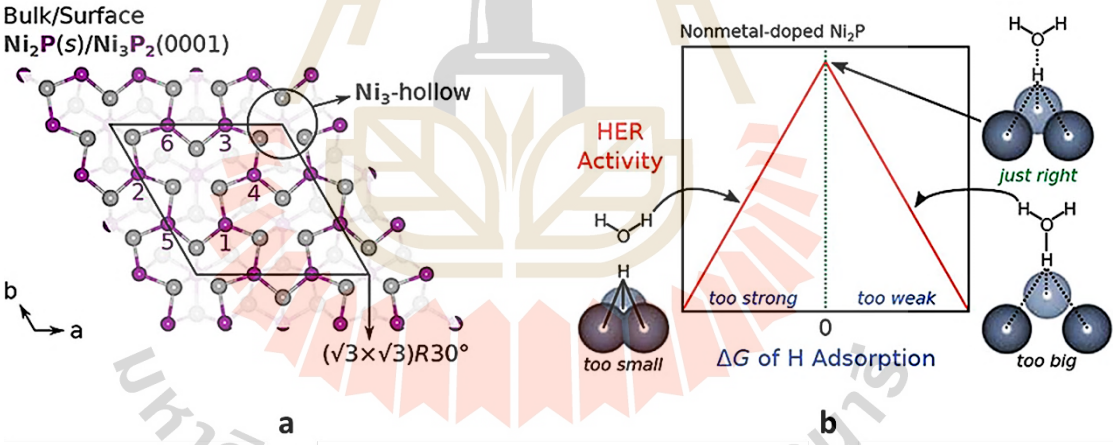


Figure 2.10 (a) Structure of the $Ni_3P_2(0001)$ surface of Ni_2P showing the (√3 × √3) R30° supercell. (b) The discovered Ni–Ni bond length is the most important descriptor of HER activity (Wexler et al., 2018).

Nørskov applied ML to model the nickel-gallium bimetallic surface for electrochemical CO_2 reduction (Ulissi, Tang, et al., 2017), as illustrated in **Figure 2.11**. Each atom's local geometry creates a fingerprint, processed by a neural network to predict its impact on adsorption energy. This method significantly reduced the number of DFT calculations required, achieving estimates of activity with an order of magnitude

fewer computations. The results indicated that the most promising active site motifs are isolated nickel atoms surrounded by gallium atoms.

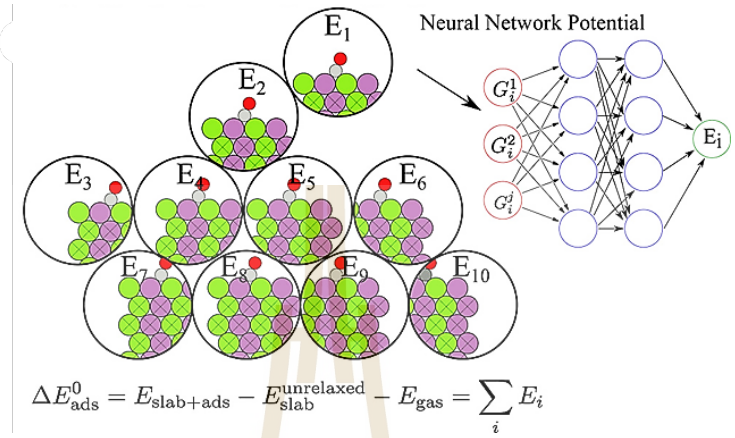


Figure 2.11 The neural network potential used to directly relax and predict adsorption energies for small molecules. The local region around each atom is used to generate a geometric fingerprint as a feature.

Additionally, product selectivity can also be predicted using machine learning approaches. Ma et al. combined computational calculations and ML to predict product yields for P-H insertion reactions (Y. Ma et al., 2023). ML algorithms were used to analyze 16 descriptors derived from DFT-calculated transition states to predict product yields. Among the algorithms tested, the Support Vector Machine (SVM) achieved the highest prediction accuracy of 97%, with a correlation exceeding 80% in Leave-One-Out Cross-Validation (LOOCV). The model demonstrated a strong correlation and agreement between predicted and experimental yields, as illustrated in **Figure 2.12**.

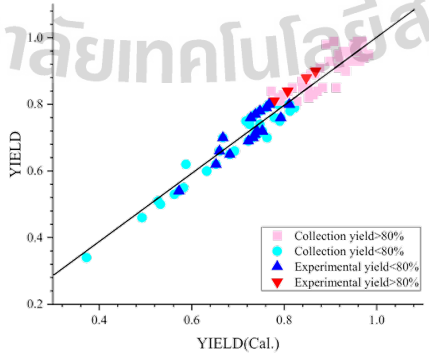


Figure 2.12 The correlation between experimental results and predicted curve (Y. Ma et al., 2023).

2.5 References

Alasali, F., Abuashour, M. I., Hammad, W., Almomani, D., Obeidat, A. M., and Holderbaum, W. (2024). A review of hydrogen production and storage materials for efficient integrated hydrogen energy systems. In *Energy Science and Engineering* (Vol. 12, Issue 5, pp. 1934–1968). John Wiley and Sons Ltd. https://doi.org/10.1002/ese3.1723

Aljabour, A., Apaydin, D. H., Coskun, H., Ozel, F., Ersoz, M., Stadler, P., Sariciftci, N. S., and Kus, M. (2016). Improvement of Catalytic Activity by Nanofibrous CuInS2 for Electrochemical CO2 Reduction. *ACS Applied Materials & Interfaces*, *8*(46), 31695–31701. https://doi.org/10.1021/acsami.6b11151

Bagger, A., Ju, W., Varela, A. S., Strasser, P., and Rossmeisl, J. (2017). Electrochemical CO2 Reduction: A Classification Problem. *ChemPhysChem*, *18*(22), 3266–3273. https://doi.org/10.1002/cphc.201700736

Chang, Q., Kim, J., Lee, J. H., Kattel, S., Chen, J. G., Choi, S.-I., and Chen, Z. (2020). Boosting Activity and Selectivity of CO2 Electroreduction by Pre-Hydridizing Pd Nanocubes. *Small*, *16*(49), 2005305. https://doi.org/https://doi.org/10.1002/smll.202005305

Chen, Z., Gao, M.-R., Duan, N., Zhang, J., Zhang, Y.-Q., Fan, T., Zhang, J., Dong, Y., Li, J., Liu, Q., Yi, X., and Luo, J.-L. (2020). Tuning adsorption strength of CO2 and its intermediates on tin oxide-based electrocatalyst for efficient CO2 reduction towards carbonaceous products. *Applied Catalysis B: Environmental*, *277*, 119252. https://doi.org/https://doi.org/10.1016/j.apcatb.2020.119252

Dhifallah, M., Dhouib, A., Aldulaijan, S., Renzo, F., and Guesmi, H. (2016). First-principles study of Au-Cu alloy surface changes induced by gas adsorption of CO, NO, or O2. *Journal of Chemical Physics*, *145*(2). https://doi.org/10.1063/1.4955104

Dong, H., Li, Y., and Jiang, D. E. (2018). First-Principles insight into electrocatalytic reduction of CO2 to CH4 on a copper nanoparticle. *Journal of Physical Chemistry C*, *122*(21), 11392–11398. https://doi.org/10.1021/acs.jpcc.8b01928

Dong, H., Liu, C., Li, Y., and Jiang, D. E. (2019). Computational screening of M/Cu core/shell nanoparticles and their applications for the electro-chemical reduction

of CO2 and CO. *Nanoscale*, *11*(23), 11351–11359. https://doi.org/10.1039/c9nr01936g

Farsi, L., and Deskins, N. A. (n.d.). *First Principles Analysis of Surface Dependent Segregation in Bimetallic Alloys*.

Feng, Y., Yang, H., Zhang, Y., Huang, X., Li, L., Cheng, T., and Shao, Q. (2020). Te-Doped Pd Nanocrystal for Electrochemical Urea Production by Efficiently Coupling Carbon Dioxide Reduction with Nitrite Reduction. *Nano Letters*, *20*(11), 8282–8289. https://doi.org/10.1021/acs.nanolett.0c03400

Ferrando, R., Jellinek, J., and Johnston, R. L. (2008). Nanoalloys: From theory to applications of alloy clusters and nanoparticles. In *Chemical Reviews* (Vol. 108, Issue 3, pp. 845–910). American Chemical Society. https://doi.org/10.1021/cr040090g

Filot, Ivo. (2023). *Introduction to microkinetic modeling*. Technische Universiteit Eindhpoven.

Gao, D., Arán-Ais, R. M., Jeon, H. S., and Roldan Cuenya, B. (2019). Rational catalyst and electrolyte design for CO2 electroreduction towards multicarbon products. *Nature Catalysis*, *2*(3), 198–210. https://doi.org/10.1038/s41929-019-0235-5

Goldsmith, B. R., Esterhuizen, J., Liu, J., Bartel, C. J., and Sutton, C. (2018). *Machine learning for heterogeneous catalyst design and discovery*.

Guan, Y., Suo, W., Zhang, Z., Wang, Y., Sun, S., and Liu, G. (2021). Insights on the Catalytic Active Site for CO2 Reduction on Copper-based Catalyst: A DFT study. *Molecular Catalysis*, *511*. https://doi.org/10.1016/j.mcat.2021.111725

Hahn, C., Abram, D. N., Hansen, H. A., Hatsukade, T., Jackson, A., Johnson, N. C., Hellstern, T. R., Kuhl, K. P., Cave, E. R., and Feaster, J. T. (2015). Synthesis of thin film AuPd alloys and their investigation for electrocatalytic CO2 reduction. *Journal of Materials Chemistry A*, *3*(40), 20185–20194.

Han, X., Liu, L., Yuan, J., Zhang, X., and Niu, D. (2021). Polyacrylamide-Mediated Silver Nanoparticles for Selectively Enhancing Electroreduction of CO2 towards CO in Water. *ChemSusChem*, *14*(2), 721–729. https://doi.org/https://doi.org/10.1002/cssc.202002481

He, J., Dettelbach, K. E., Salvatore, D. A., Li, T., and Berlinguette, C. P. (2017). High-throughput synthesis of mixed-metal electrocatalysts for CO2 reduction. *Angewandte Chemie*, *129*(22), 6164–6168.

Hirunsit, P., Soodsawang, W., and Limtrakul, J. (2015). CO2 electrochemical reduction to methane and methanol on copper-based alloys: Theoretical insight. *Journal of Physical Chemistry C*, *119*(15), 8238–8249. https://doi.org/10.1021/acs.jpcc.5b01574

Hori, Y. i. (2008). Electrochemical CO2 reduction on metal electrodes. *Modern Aspects of Electrochemistry*, 89–189.

Hori, Y., Takahashi, R., Yoshinami, Y., and Murata, A. (1997). *Electrochemical Reduction of CO at a Copper Electrode*.

Huang, J., Guo, X., Huang, X., and Wang, L. (2019). Metal (Sn, Bi, Pb, Cd) in-situ anchored on mesoporous hollow kapok-tubes for outstanding electrocatalytic CO2 reduction to formate. *Electrochimica Acta*, *325*, 134923. https://doi.org/https://doi.org/10.1016/j.electacta.2019.134923

Hussain, J., Jónsson, H., and Skúlason, E. (2018). Calculations of Product Selectivity in Electrochemical CO2 Reduction. *ACS Catalysis*, *8*(6), 5240–5249. https://doi.org/10.1021/acscatal.7b03308

Jouny, M., Luc, W., and Jiao, F. (2018). General techno-economic analysis of CO2 electrolysis systems. *Industrial & Engineering Chemistry Research*, *57*(6), 2165–2177.

Kim, D., Resasco, J., Yu, Y., Asiri, A. M., and Yang, P. (2014). Synergistic geometric and electronic effects for electrochemical reduction of carbon dioxide using gold–copper bimetallic nanoparticles. *Nature Communications*, *5*(1), 4948.

Lei, Y., Wang, Z., Bao, A., Tang, X., Huang, X., Yi, H., Zhao, S., Sun, T., Wang, J., and Gao, F. (2023). Recent advances on electrocatalytic CO2 reduction to resources: Target products, reaction pathways and typical catalysts. *Chemical Engineering Journal*, *453*, 139663. https://doi.org/https://doi.org/10.1016/j.cej.2022.139663

Li, Q., Fu, J., Zhu, W., Chen, Z., Shen, B., Wu, L., Xi, Z., Wang, T., Lu, G., and Zhu, J. (2017). Tuning Sn-catalysis for electrochemical reduction of CO2 to CO via the

core/shell Cu/SnO2 structure. *Journal of the American Chemical Society*, *139*(12), 4290–4293.

Li, X., Bi, W., Chen, M., Sun, Y., Ju, H., Yan, W., Zhu, J., Wu, X., Chu, W., Wu, C., and Xie, Y. (2017). Exclusive Ni–N4 Sites Realize Near-Unity CO Selectivity for Electrochemical CO2 Reduction. *Journal of the American Chemical Society*, *139*(42), 14889–14892. https://doi.org/10.1021/jacs.7b09074

Li, Z., Wang, S., Chin, W. S., Achenie, L. E., and Xin, H. (2017). High-throughput screening of bimetallic catalysts enabled by machine learning. *Journal of Materials Chemistry A*, *5*(46), 24131–24138.

Liu, F., Wang, Z., Wang, Z., Zhong, J., Zhao, L., Jiang, L., Zhou, R., Liu, Y., Huang, L., Tan, L., Tian, Y., Zheng, H., Fang, Q., Zhang, L., Zhang, L., Wu, H., Bai, L., and Zhou, K. (2022). High-Throughput Method–Accelerated Design of Ni-Based Superalloys. *Advanced Functional Materials*, *32*(28), 2109367. https://doi.org/https://doi.org/10.1002/adfm.202109367

Liu, S., Zhao, Z. J., Yang, C., Zha, S., Neyman, K. M., Studt, F., and Gong, J. (2019). Adsorption Preference Determines Segregation Direction: A Shortcut to More Realistic Surface Models of Alloy Catalysts. *ACS Catalysis*, *9*(6), 5011–5018. https://doi.org/10.1021/acscatal.9b00499

Ma, S., Sadakiyo, M., Heima, M., Luo, R., Haasch, R. T., Gold, J. I., Yamauchi, M., and Kenis, P. J. A. (2017). Electroreduction of carbon dioxide to hydrocarbons using bimetallic Cu–Pd catalysts with different mixing patterns. *Journal of the American Chemical Society*, *139*(1), 47–50.

Ma, Y., Zhang, X., Zhu, L., Feng, X., Kowah, J. A. H., Jiang, J., Wang, L., Jiang, L., and Liu, X. (2023). Machine Learning and Quantum Calculation for Predicting Yield in Cu-Catalyzed P–H Reactions. *Molecules*, *28*(16). https://doi.org/10.3390/molecules28165995

Nie, X., Esopi, M. R., Janik, M. J., and Asthagiri, A. (2013). Selectivity of CO2 reduction on copper electrodes: The role of the kinetics of elementary steps. *Angewandte Chemie - International Edition*, *52*(9), 2459–2462. https://doi.org/10.1002/anie.201208320

Nie, X., Luo, W., Janik, M. J., and Asthagiri, A. (2014). Reaction mechanisms of CO2 electrochemical reduction on Cu(1 1 1) determined with density functional theory. *Journal of Catalysis*, *312*, 108–122. https://doi.org/10.1016/j.jcat.2014.01.013

Peterson, A. A., Abild-Pedersen, F., Studt, F., Rossmeisl, J., and Nørskov, J. K. (2010). How copper catalyzes the electroreduction of carbon dioxide into hydrocarbon fuels. *Energy and Environmental Science*, *3*(9), 1311–1315. https://doi.org/10.1039/c0ee00071j

Rosen, J., Hutchings, G. S., Lu, Q., Forest, R. V, Moore, A., and Jiao, F. (2015). Electrodeposited Zn Dendrites with Enhanced CO Selectivity for Electrocatalytic CO2 Reduction. *ACS Catalysis*, *5*(8), 4586–4591. https://doi.org/10.1021/acscatal.5b00922

Saqlain, M. A., Hussain, A., Siddiq, D. M., Leenaerts, O., and Leitão, A. A. (2016). DFT Study of Synergistic Catalysis of the Water-Gas-Shift Reaction on Cu-Au Bimetallic Surfaces. *ChemCatChem*, *8*(6), 1208–1217. https://doi.org/10.1002/cctc.201501312

Schouten, K. J. P., Kwon, Y., van der Ham, C. J. M., Qin, Z., and Koper, M. T. M. (2011). A new mechanism for the selectivity to C1 and C2 species in the electrochemical reduction of carbon dioxide on copper electrodes. *Chemical Science*, *2*(10), 1902–1909. https://doi.org/10.1039/C1SC00277E

Shetty, S., Gayen, M., Agarwal, S., Chatterjee, D., Singh, A., and Ravishankar, N. (2022). Tuning Catalytic Activity in Ultrathin Bimetallic Nanowires via Surface Segregation: Some Insights. *Journal of Physical Chemistry Letters*, *13*(3), 770–776. https://doi.org/10.1021/acs.jpclett.1c03852

Tang, M. T., Peng, H., Lamoureux, P. S., Bajdich, M., and Abild-Pedersen, F. (2020). From electricity to fuels: Descriptors for C1 selectivity in electrochemical CO2 reduction. *Applied Catalysis B: Environmental*, *279*. https://doi.org/10.1016/j.apcatb.2020.119384

Ulissi, Z. W., Medford, A. J., Bligaard, T., and Nørskov, J. K. (2017). To address surface reaction network complexity using scaling relations machine learning and DFT calculations. *Nature Communications*, *8*(1), 14621.

Ulissi, Z. W., Tang, M. T., Xiao, J., Liu, X., Torelli, D. A., Karamad, M., Cummins, K., Hahn, C., Lewis, N. S., and Jaramillo, T. F. (2017). Machine-learning methods enable exhaustive searches for active bimetallic facets and reveal active site motifs for CO2 reduction. *Acs Catalysis*, *7*(10), 6600–6608.

Wang, P., Yang, H., Xu, Y., Huang, X., Wang, J., Zhong, M., Cheng, T., and Shao, Q. (2021). Synergized Cu/Pb Core/Shell Electrocatalyst for High-Efficiency CO2 Reduction to C2+ Liquids. *ACS Nano*, *15*(1), 1039–1047. https://doi.org/10.1021/acsnano.0c07869

Wang, X., Ma, F., Wei, H., Pan, J., Li, W., Zhao, J., and Yang, X. (2025). A DFT and microkinetic modeling study of pressure effects on electroreduction reduction of CO2 to ethanol. *Applied Surface Science*, *681*. https://doi.org/10.1016/j.apsusc.2024.161421

Wang, Y., Zheng, M., Wang, X., and Zhou, X. (2022). Electrocatalytic Reduction of CO2 to C1 Compounds by Zn-Based Monatomic Alloys: A DFT Calculation. *Catalysts*, *12*(12). https://doi.org/10.3390/catal12121617

Wexler, R. B., Martirez, J. M. P., and Rappe, A. M. (2018). Chemical pressure-driven enhancement of the hydrogen evolving activity of Ni2P from nonmetal surface doping interpreted via machine learning. *Journal of the American Chemical Society*, *140*(13), 4678–4683.

Wu, D., Kusada, K., and Kitagawa, H. (2016). Recent progress in the structure control of Pd–Ru bimetallic nanomaterials. *Science and Technology of Advanced Materials*, *17*, 583–596. https://doi.org/10.1080/14686996.2016.1221727

Xiang, Q., Li, F., Wang, J., Chen, W., Miao, Q., Zhang, Q., Tao, P., Song, C., Shang, W., Zhu, H., Deng, T., and Wu, J. (2021). Heterostructure of ZnO Nanosheets/Zn with a Highly Enhanced Edge Surface for Efficient CO2 Electrochemical Reduction to CO. *ACS Applied Materials & Interfaces*, *13*(9), 10837–10844. https://doi.org/10.1021/acsami.0c20302

Yin, Z., Gao, D., Yao, S., Zhao, B., Cai, F., Lin, L., Tang, P., Zhai, P., Wang, G., and Ma, D. (2016). Highly selective palladium-copper bimetallic electrocatalysts for the electrochemical reduction of CO2 to CO. *Nano Energy*, *27*, 35–43.

Yu, Y., Huang, W., Liu, Z., Hu, Z., and Wang, L. (2022). First-principles study of surface segregation in bimetallic Cu3M(1 1 1) (M = Au, Ag, and Zn) alloys in presence of adsorbed CO. *Computational Materials Science*, *212*. https://doi.org/10.1016/j.commatsci.2022.111550

Zhang, W., Qin, Q., Dai, L., Qin, R., Zhao, X., Chen, X., Ou, D., Chen, J., Chuong, T. T., and Wu, B. (2018). Electrochemical reduction of carbon dioxide to methanol on hierarchical Pd/SnO2 nanosheets with abundant Pd–O–Sn interfaces. *Angewandte Chemie International Edition*, *57*(30), 9475–9479.

Zhao, M., Brouwer, J. C., Sloof, W. G., and Böttger, A. J. (2020). Surface segregation of Pd–Cu alloy in various gas atmospheres. *International Journal of Hydrogen Energy*, *45*(41), 21567–21572. https://doi.org/10.1016/j.ijhydene.2020.05.268

Zhu, W., Tackett, B. M., Chen, J. G., and Jiao, F. (2018). Bimetallic Electrocatalysts for CO2 Reduction. In *Topics in Current Chemistry* (Vol. 376, Issue 6). Springer International Publishing. https://doi.org/10.1007/s41061-018-0220-5

Zhu, W., Tackett, B. M., Chen, J. G., and Jiao, F. (2020). Bimetallic Electrocatalysts for CO2 Reduction. In M. Shao (Ed.), *Electrocatalysis* (pp. 105–125). Springer International Publishing. https://doi.org/10.1007/978-3-030-43294-2_4

Zijlstra, B., Zhang, X., Liu, J. X., Filot, I. A. W., Zhou, Z., Sun, S., and Hensen, E. J. M. (2020). First-principles microkinetics simulations of electrochemical reduction of CO2 over Cu catalysts. *Electrochimica Acta*, *335*. https://doi.org/10.1016/j.electacta.2020.135665

CHAPTER III

METHODOLOGY

In this thesis, we provide a framework for catalyst design by combining multi-scale modelling and machine learning as shown in **Figure 3.1**. In this framework, the process begins with generating raw data on bimetallic compositions and energy profiles total of 1,717 data using Fairchem. These data are then used as inputs for the microkinetic model to provide the reaction rate and selectivity of each catalyst under experimental conditions. Finally, a machine learning model is employed to predict the reaction outcomes (reaction rate and selectivity) for the remaining catalysts based on their composition. This integrated approach enables efficient exploration and design of catalysts for CO_2 reduction.

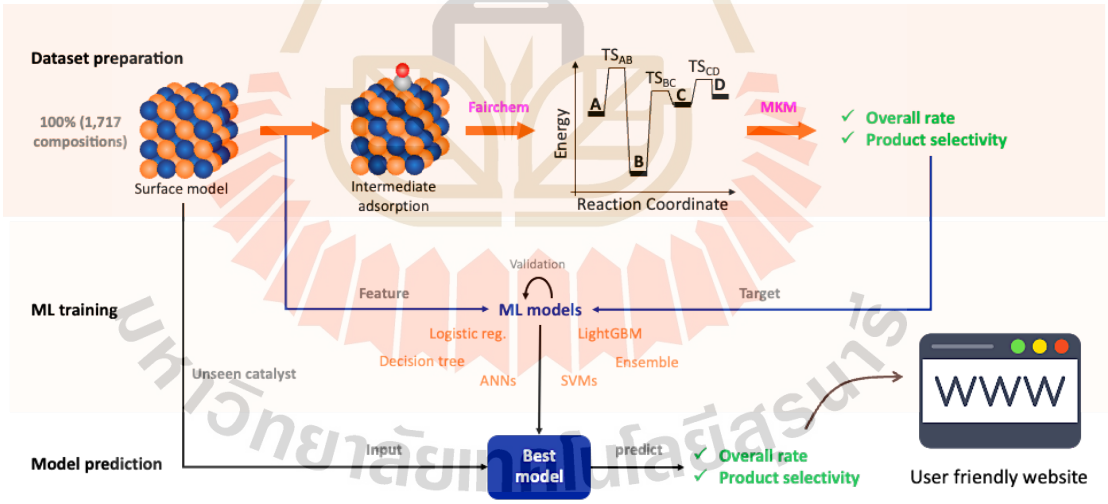


Figure 3.1 workflow of catalyst design for CO_2RR.

3.1 Fairchem

To accelerate calculations for developing a framework for catalyst design in CO_2RR, all computations were performed using Fairchem, a centralized repository developed by FAIR Chemistry. This platform serves as a comprehensive resource for

materials science and quantum chemistry, housing data, models, demos, and application tools. Collaborative projects that contribute or use the models and approaches in Open Catalyst 2020 (OC20) (Chanussot et al., 2021).

The OC20 dataset, consisting of 1,281,040 DFT relaxations (~264,890,000 single-point evaluations) across a wide swath of materials, surfaces, and adsorbates (N, C, and O), as shown in **Figure 3.2**. The dataset consists of three key tasks representative of routine catalyst modeling and includes predefined train, validation, and test splits. This setup enables straightforward comparisons with future model development efforts.

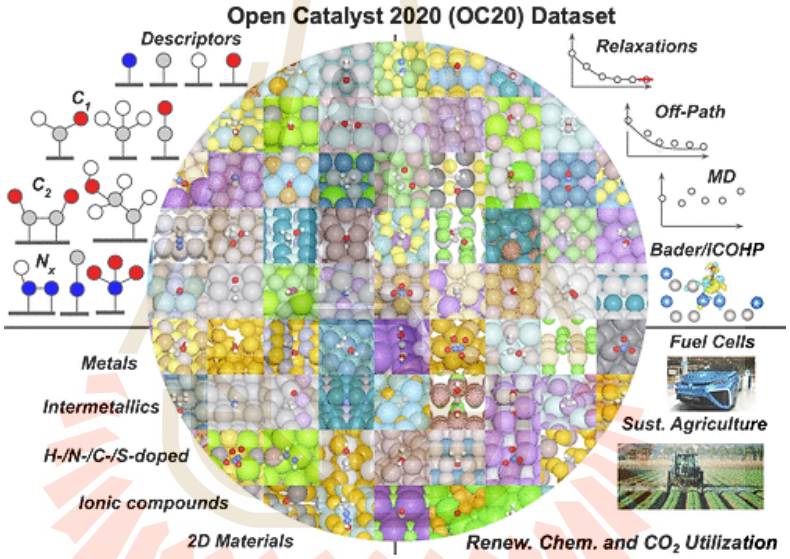


Figure 3.2 Adsorbates materials, calculations, and impact areas of the OC20 dataset. (Chanussot et al., 2021).

This thesis, we use Predicting Energy and Forces from a Structure (S2EF), EquiformerV2-153M-S2EF-OC20-All+MD which is the baseline pre-trained model. All approaches predict structure energies in their forward pass and per-atom forces by the negative gradient of the predicted energy with respect to atomic positions. Adsorption energy, and reaction energy of each catalyst is calculated by this model.

3.2 Computational hydrogen electrode (CHE) model

To calculate the reaction energy of CO_2 reduction to C1 product on each metal catalyst surface we explore the methods call computational hydrogen electrode (CHE). In this model the reference potential is set at the standard hydrogen electrode (0 V vs. SHE). At this potential the chemical potential of the reaction of $H^+ + e^- \rightarrow \frac{1}{2}H_2$ is equilibrate at pH = 0 and 1 bar of hydrogen in gas phase at 298 K.

Based on the reference energy of CO_2RR in theoretical calculation, the reaction energy can be calculated from the energy of current state subtract by the previous one. But, because of this thesis we use Fairchem as the calculation tools, so the calculation of reference energy is a bit change. The gas phase references, E_{gas}, for each adsorbate was computed as a linear combination of N_2, H_2O, CO, and H_2 resulting in the atomic energies from **Table 3.1**.

Table 3.1 The chemical potential of adsorbate per atom used to calculate the gas phase reference energy for an adsorbate molecule (Chanussot et al., n.d.).

Adsorbate species per atom	Chemical potential (eV)
H	−3.447
C	−7.204
O	−7.282
N	−8.083

Therefore, based on CHE model and this gas reference energy the reaction energy of each elementary reaction of CO_2 reduction to C1 product can be calculate follow equation,

The electrochemical step; A* ⟶ B*

$$\Delta E = \left(E_{B^*}^{FC} + \mu_B\right) - \left(E_{A^*}^{FC} + \mu_A\right) - \frac{1}{2}E_{H_2}^{DFT}$$

Where $E_{B^*/A*}^{FC}$ represent the energy of intermediate adsorption from Fairchem. $\mu_{B/A}$ represent the chemical potential of gas, and $E_{H_2}^{DFT}$ represent the energy of hydrogen gas on DFT.

The non electrochemical step; B* ⟶ B + *

$$\Delta E = E_B^{DFT} - \left(E_{B^*}^{FC} + \mu_B\right)$$

Where $E_{B^*}^{FC}$ represent the energy of intermediate adsorption from Fairchem. μ_B represent the chemical potential of gas, and E_B^{DFT} represent the energy of gas on DFT.

3.3 Microkinetic modelling

To better understand the reaction behaviour of each catalyst, we then explore the microkinetic modelling. The general concept of microkinetic model is that it takes energy profile and predict the reaction rate and product selectivity. The steps to construct a microkinetic model are as follows:

1. Construct the set of elementary reaction steps.
2. Derive rate expressions for each of the elementary reaction steps.
3. Convert the set of rate expressions to a set of ordinary differential equations.
4. Define boundary conditions for the system (e.g. partial pressures), the initial values (e.g. initial surface concentrations) and any model parameters (e.g. temperature).
5. Solve the system of ordinary differential equations.
6. Interpret the results (using our chemical intuition).

For example, A unimolecular reaction on a catalytic surface is given by the kinetic network below. In this network, A adsorbs on the catalytic surface, is then converted to B and finally B desorbs from the surface.

$$A \longrightarrow A^* \longrightarrow B^* \longrightarrow B$$

Which is composed of three elementary reaction steps,

$$A + * \rightleftharpoons A^* \quad (r1)$$

$$B + * \rightleftharpoons B^* \quad (r2)$$

$$A^* \rightleftharpoons B^* \quad (r3)$$

The overall reaction: $A \longrightarrow B$

Each elementary step is characterized by a rate equation based on its rate constant, which follows the Arrhenius equation:

$$\mathbf{k_i = A_i \exp\left(\frac{-E_a(U)}{k_B T}\right)}$$

where $\mathbf{A_i}$ is an effective pre-factor. As an approximation, the pre-factor was set to equal $\frac{\mathbf{k_B T}}{\mathbf{h}}$ s-1 for all the surface elementary reactions. $\mathbf{E_a}$ is the energy barrier in eV.

However, if the elementary reaction has the adsorption and desorption process, the rate constant can be explored from the following:

$$\mathbf{k_{ads} = \frac{A}{\sqrt{2\pi m k_B T}}}$$

$$\mathbf{k_{des} = \frac{k_{ads}}{\exp\left(\frac{-E_a(U)}{k_B T}\right)}}$$

Where A is an active site of catalyst surface, and m represents the molecular mass (kg).

Then, the rate expression of each elementary step can be written based on the rate law,

$$r1 = k_1 P_A[\theta_*] - k_{-1}[\theta_A]$$

$$r2 = k_2 P_B[\theta_*] - k_{-2}[\theta_B]$$

$$r3 = k_3[\theta_A] - k^{-2}[\theta_B]$$

where $k_{+/-}$ represent the rate constants of the forward and backward reactions, respectively. $\theta_{A,B}$ represents the coverage of intermediate. $P_{A,B}$ represents the pressure of A and B in gas phase.

The set of rate expressions to a set of differential equations,

$$\frac{d\theta_A}{d_t} = r_1 - r_3$$

$$\frac{d\theta_B}{d_t} = r_3 - r_2$$

The next step is defining the condition (e.g. pressure, temperature, and surface concentration). Then, solve the system of ordinary differential equations and analyse the result from the simulation.

3.4 Machine learning models

This thesis is anchored in the application of machine learning (ML), a powerful subset of artificial intelligence that enables computational systems to learn patterns from data and make predictions without being explicitly programmed. Unlike traditional computational chemistry methods that rely on first-principles calculations or predefined mechanistic models, ML models can be trained on large datasets of experimental or simulated results to discover hidden, non-linear relationships between

a reaction's input parameters and its observed outcomes. This approach is particularly well-suited for predicting complex phenomena like reaction rates and selectivity, where the interplay of multiple variables (e.g., catalyst structure, temperature, pressure, and solvent effects) can be challenging to model through conventional means.

In this work, the machine learning pipeline will serve as a core component for predicting the rate and selectivity of the targeted chemical reactions. The process involves a series of critical steps: a curated dataset is first used to define the input features, which are then fed into an algorithm that learns a function, $y=f(X)$, where y represents the predicted rate or selectivity and X is the set of relevant input features. Following a rigorous training phase, the model's performance will be validated using a held-out test set to ensure its predictive accuracy and generalization capability. This methodology allows for the rapid exploration of vast chemical spaces, guiding the identification of optimal reaction conditions and enabling the design of novel catalytic systems with high efficiency and specificity.

3.4.1 Regression models

The regression models, a powerful class of supervised machine learning algorithms designed to predict continuous numerical outputs. In this context, the task of predicting reaction rates and selectivity translates directly to a regression problem, as these outcomes are continuous variables. The goal of the regression model is to learn the underlying function that maps a set of input features (e.g., catalyst composition, reaction conditions) to the corresponding continuous target values which is the selectivity in this work.

For the selectivity prediction, we employ 4 difference regression model including, Light Gradient Boosting Machine (LightGBM), Support Vector Regression (SVR), Ensemble model, and Artificial Neural Network (ANN).

Light Gradient Boosting Machine (Ke et al., 2017) is a highly efficient and fast-performing gradient boosting framework that utilizes tree-based learning algorithms. A key distinction of LightGBM is its leaf-wise tree growth strategy, which differs from the level-wise approach used by many other boosting algorithms. This method prioritizes splitting the leaf that promises the greatest reduction in loss, leading to faster convergence and often higher accuracy. To further enhance its efficiency, particularly with large datasets, LightGBM incorporates two innovative techniques: Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB). GOSS focuses on instances with larger gradients (i.e., those that are not well-trained), while EFB bundles mutually exclusive features to reduce the dimensionality of the data. As a result, LightGBM is known for its ability to handle large-scale data with lower memory usage and significantly faster training speeds.

Support Vector Regression (Basak et al., 2007) is a powerful regression method based on the principles of Support Vector Machines (SVMs), which are primarily used for classification. Unlike traditional regression models that aim to minimize the sum of errors, SVR seeks to find a hyperplane that best fits the data points within a defined margin or epsilon-insensitive tube. Data points that fall within this tube are not penalized, making the model more robust to outliers. The critical data points that lie on or outside the tube's boundary are called "support vectors," and they are the only ones that influence the model's final form. A key advantage of SVR is its ability to handle non-linear relationships by using various kernel functions, such as the Radial Basis Function (RBF) kernel, which effectively maps the data into a higher-dimensional space where a linear relationship can be found.

Ensemble model (Mishra et al., 2025) is a sophisticated machine learning approach that combines the predictions of multiple individual models, or "base learners," to produce a single, more accurate, and robust prediction. The underlying principle is that a collective of diverse models can overcome the weaknesses of any

single model, leading to improved overall performance. Common ensemble methods for regression include bagging, boosting, and stacking. Bagging (like in Random Forest) trains parallel models on different random subsets of the data and averages their predictions. Boosting (like in LightGBM) builds models sequentially, with each new model focusing on correcting the errors of its predecessors. Stacking is a more complex technique where a "meta-model" learns to combine the predictions of the base models. By leveraging the strengths of different algorithms, ensemble methods effectively reduce issues like high variance or bias, resulting in a more generalized and reliable model.

Artificial Neural Networks (Grossi & Buscema, 2007) are a class of machine learning models inspired by the human brain, composed of interconnected "neurons" organized in layers. The term "deep learning" specifically refers to ANNs with multiple hidden layers, which enables them to learn complex, hierarchical representations of data. In a regression task, data flows from an input layer through these hidden layers, where each neuron performs a weighted sum of its inputs and applies an activation function. The network's final prediction is generated by the output layer. The learning process involves comparing the network's prediction to the true value using a loss function (e.g., Mean Squared Error) and then using a process called backpropagation and an optimization algorithm (e.g., Gradient Descent) to iteratively adjust the weights and biases to minimize the error. A primary strength of deep ANNs is their ability to automatically extract and learn meaningful features from raw data, making them particularly effective for problems with intricate and non-linear patterns without the need for manual feature engineering.

3.4.2 Classification models

A classification model is a type of machine learning model designed to categorize data points into predefined, discrete groups or classes. Unlike regression, which predicts a continuous numerical value, classification predicts a categorical label. These models are typically trained on a labelled dataset, where each data point is paired with its correct class label. During training, the model learns the unique characteristics and patterns that define each class. Once trained, the model can then be used to predict the class of new, unseen data. Common classification algorithms include Logistic Regression, Support Vector Machines, Decision Trees, and LightGBM, each with its own approach to creating decision boundaries that separate the different classes.

Logistic Regression (Peng et al., 2002) is a fundamental and widely used supervised machine learning algorithm for binary classification problems. Unlike its name, it is not a regression model that predicts continuous values, but rather a classification model that estimates the probability of a data point belonging to a certain class. It achieves this by applying a sigmoid (S-shaped) function to a linear equation, which transforms the output into a value between 0 and 1, representing a probability. The model then uses a predefined threshold (commonly 0.5) to classify the data point. For example, in a medical context, it can predict the probability of a patient having a disease based on test results. This simplicity and the interpretability of its coefficients make it a popular choice for problems where a clear understanding of feature importance is required.

Decision Trees (Packwood et al., 2022) is a versatile supervised learning algorithm that can be used for both classification and regression tasks. It works by creating a tree-like, flowchart structure where each internal node represents a test on a feature (e.g., "Is the temperature greater than 25°C?"), each branch represents the outcome of that test, and each leaf node represents the final class prediction. The

model works by recursively partitioning the data into smaller, more homogeneous subsets based on the features that provide the most information gain or purity at each split. The process continues until all, or most, of the data is classified. Decision Trees are intuitive and easy to visualize, making them highly interpretable. However, they are also prone to overfitting, which means they can become too complex and memorize the training data, leading to poor performance on new data.

3.5 References

Basak, D., Pal, S., and Patranabis, D. C. (2007). Support Vector Regression. In *Neural Information Processing-Letters and Reviews* (Vol. 11, Issue 10).

Chanussot, L., Das, A., Goyal, S., Lavril, T., Shuaibi, M., Riviere, M., Tran, K., Heras-Domingo, J., Ho, C., Hu, W., Palizhati, A., Sriram, A., Wood, B., Yoon, J., Parikh, D., Zitnick, ⊥ C Lawrence, Ulissi, Z., †co-First Author, ¶, and Research, A. I. (n.d.). *Supporting Information: The Open Catalyst 2020 (OC20) Dataset and Community Challenges*.

Chanussot, L., Das, A., Goyal, S., Lavril, T., Shuaibi, M., Riviere, M., Tran, K., Heras-Domingo, J., Ho, C., Hu, W., Palizhati, A., Sriram, A., Wood, B., Yoon, J., Parikh, D., Zitnick, C. L., and Ulissi, Z. (2021). Open Catalyst 2020 (OC20) Dataset and Community Challenges. *ACS Catalysis*, *11*(10), 6059–6072. https://doi.org/10.1021/acscatal.0c04525

Grossi, E., and Buscema, M. (2007). Introduction to artificial neural networks. In *European Journal of Gastroenterology and Hepatology* (Vol. 19, Issue 12, pp. 1046–1054). https://doi.org/10.1097/MEG.0b013e3282f198a0

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. (2017). *LightGBM: A Highly Efficient Gradient Boosting Decision Tree*. https://www.researchgate.net/publication/378480234

Mishra, D., Tripathi, S. M., Chaurasia, A., and Chaurasia, P. K. (2025). A Review on Ensemble Learning Methods: Machine Learning Approach. *International Journal of Research Publication and Reviews*, *6*(2), 3795–3803. https://doi.org/10.55248/gengpi.6.0225.0971

Packwood, D., Nguyen, L. T. H., Cesana, P., Zhang, G., Staykov, A., Fukumoto, Y., and Nguyen, D. H. (2022). Machine Learning in Materials Chemistry: An Invitation. *Machine Learning with Applications*, *8*, 100265. https://doi.org/10.1016/j.mlwa.2022.100265

Peng, C. Y. J., Lee, K. L., and Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *Journal of Educational Research*, *96*(1), 3–14. https://doi.org/10.1080/00220670209598786

CHAPTER IV
RESULTS AND DISCUSSION

This chapter present the result and discossion of workflow for catalyst design for CO_2RR. The study begin with surface model generation from 34 monometallic then, the adsorption of associated intermediates are calculated to use for generate the energy profile which is the input for microkinetic. After we get the rate and selectivity from that particular microkinetic, machine learning is start to predict rate and selectivity for the other catalyst based on their composition. Fainally, the data after prediction will collect in the database and represent as the demo website for the utilization.

4.1 Surface model generation

We began by constructed the surface model from 34 pure metals, comprising 3d (Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn), 4d (Y, Zr, Nb, Mo, Tc, Ru, Rh, Pd, Ag, Cd), 5d (Hf, Ta, W, Re, Os, Ir, Pt, Au) transition metals, and p-block elements (Al, Ga, In, Sn, Tl, Pb). The investigation is extended to bimetallic compositions 1:1, and 1:3. The model was constructed using a supercell size of 4 × 4 with four atomic layers total 64 atoms. The two bottom layers were fixed during optimization. The vacuum gap was added by 20 Å in the z-direction. The crystal structure of all surfaces is focus on simple FCC(111) facets. This $L1_2$ crystal is widely use in thermodynamic study on CO_2 conversion to C1 product (Hirunsit, 2013; Hirunsit et al., 2015). The constructed structure is shown in **Figure 4.1**.

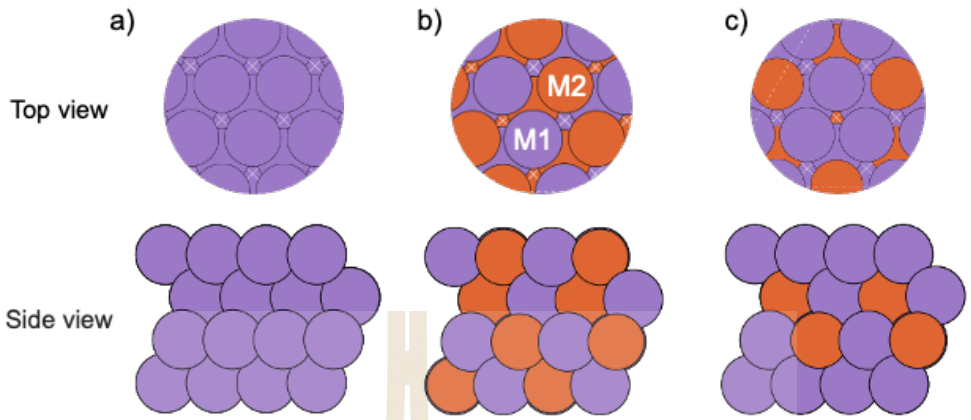


Figure 4.1 The geometric structure of metal catalyst a) monometallic, b) bimetallic composition 1:1, and c) bimetallic composition 1:3. The purple atom represent atom of metal1, and the orange atom represent atom of metal2.

The lattice constant of all monometallic we focus to use the value in only FCC phase for easy to create bimetallic composition. For composition 1:1 and 1:3, the lattice constant is calculated form the average of lattice constant of two metal elements.

$$\text{Lattice constant of composition 1:1} = \frac{\text{metal1}+\text{metal2}}{2}$$

$$\text{Lattice constant of composition 1:3} = \frac{\text{metal1}+3(\text{metal2})}{4}$$

After generating all total 1,717 surface model (34 monometallic, 561 bimetallic ratio 1:1, and 1,122 bimetallic ratio 1:3) we have optimized the structure using machine learning potential name Fairchem with EquiformerV2 model. The number of final structures is change because some of bimetallic model is distort during optimization. Therefore, we will consider only 1,632 structures (34 monometallic, 537 bimetallic 1:1, and 1,061 bimetallic 1:3) for further study.

4.2 CO_2RR intermediate adsorption on catalyst surface

This thesis carried out all possible pathways of CO_2 reduction to C1 products, so the associated intermediates are considered. The intermediate is adsorbed on the catalyst surface with all possible configurations. For monometallic catalysts, the adsorption sites consist of the top site (T), bridge site (B), hollow-fcc (H_{fcc}), and hollow-hcp (H_{hcp}). For bimetallic catalysts with compositions 1:1 and 1:3, they have many configurations because they exhibit different surface orientations. The adsorption sites are shown in **Figure 4.2**.

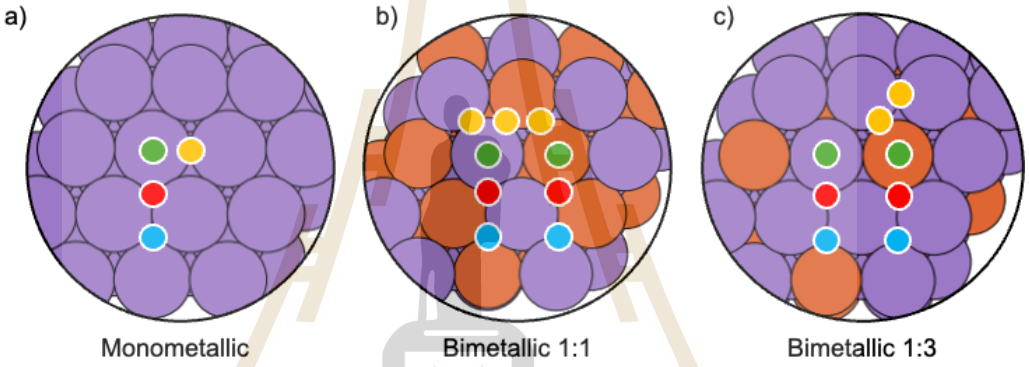
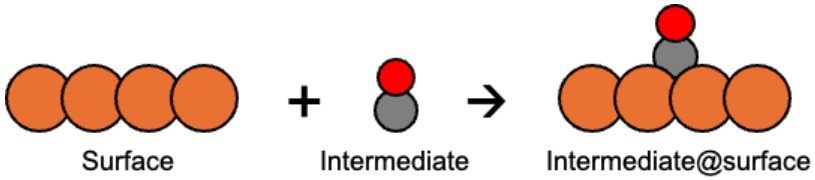


Figure 4.2 The adsorption site of different catalyst surface a) monometallic, b) bimetallic composition 1:1, and c) bimetallic composition 1:3. The green color, represent top site, yellow represent bridge site, red represent H_{fcc} site, and blue represent H_{hcp} site.

The adsorption energy of intermediate can determine trend of catalytic properties for the reduction to target products (Bagger et al., 2017). The adsorption energy can be calculated from this following equation, as shown in **Figure 4.3**.



$$E_{ads} = E_{int@surf} - (E_{surf} + E_{int})$$

Figure 4.3 The equation to calculate adsorption energy of intermediate.

where $E_{int@surf}$ is the total energy of the adsorbate-catalyst system, E_{surf} is the energy of the clean catalyst surface, and E_{int} is the energy of the isolated intermediate molecule in the gas phase. A more negative adsorption energy indicates a stronger interaction between the intermediate and the catalyst surface. This value is a crucial descriptor for understanding the stability of reaction intermediates and predicting the most favorable reaction pathways.

The further reduction of CO_2 is primarily constrained by the catalyst's ability to adsorb CO intermediate without any effect of apply potential (Bagger et al., 2019). The binding strength of CO is particularly important because it is a key intermediate in the CO_2 reduction to C1 products. A catalyst that binds CO too strongly can lead to poisoning of the active sites, while one that binds it too weakly may not facilitate its further reduction. Therefore, a balance in the CO adsorption energy is necessary for efficient catalysis. However, this thesis does not use the directly DFT calculation to calculate adsorption energy of intermediate. So, the value from the calculation maybe out from realistic. As shown in **Figure 4.4**, a direct comparison of the CO adsorption energy on a copper catalyst reveals a discrepancy between the value calculated by the Fairchem model and a typical literature value from DFT.

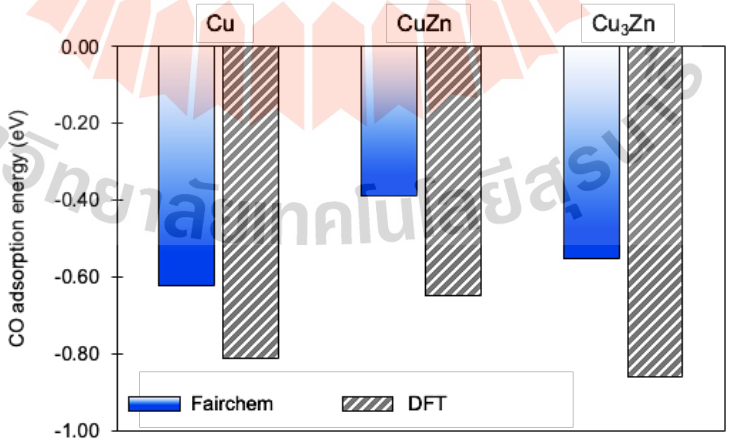


Figure 4.4 The comparison of adsorption energy between Fairchem vs. DFT (K. Li et al., 2024; Zoubir et al., 2022) of CO intermediate on different catalyst surface.

As previously mentioned, the Fairchem model is designed for rapid calculation and reproduces DFT-level accuracy at a fraction of the computational cost. While the absolute CO adsorption energy values from Fairchem may not be identical to those obtained from a specific DFT calculation found in the literature, the results shown in **Figure 4.4** demonstrate that they are not drastically different. This small deviation is expected and can arise from differences in the DFT functional used for training the model, the structural configurations in the training set, or other minor variations.

The calculations of adsorption energies for all relevant intermediates, such as *COOH, *CO, and *CHO, on all possible sites for each catalyst surface (monometallic and bimetallic) will provide a comprehensive understanding of the reaction landscape. This systematic investigation allows for the identification of the most stable configurations for each intermediate and, consequently, the most likely reaction pathway. This data will be used to construct the energy profiles for the entire CO_2RR pathway in the next section, providing a quantitative basis for comparing the performance of the different catalysts.

4.3 Energy profile of CO_2RR pathway on each catalyst surface

The reaction mechanism of CO_2 reduction to C1 product is very complex system. The CO_2RR pathway can be affect by many factors such as, catalyst surface (Bagger et al., 2017), interaction of intermediate on surface (Liu et al., 2019), and apply bias potential (Tomboc et al., 2020). This thesis, we provide all possible pathway for CO_2 reduction to C1 product with relevant of apply bias potentials. The schematic of reaction pathway as shown in **Figure 4.5**.

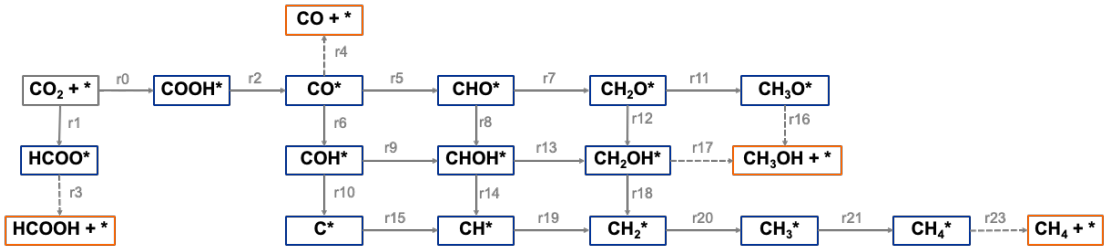


Figure 4.5 A comprehensive reaction network for the electrochemical reduction of CO_2 on a catalytic surface, with active sites denoted by the asterisk (*). This mechanistic scheme illustrates the sequential proton-coupled electron transfer (PCET) steps that lead to the formation of four distinct C1 products: formic acid, carbon monoxide, methanol, and methane.

The reaction is initiated by the one-electron and one-proton reduction of CO_2 to form two primary adsorbed intermediates. The initial pathway proceeds via step r0, leading to the formation of the carboxyl intermediate, *COOH. An alternative initial pathway, step r1, the formate intermediate, *HCOO. These two intermediates serve as crucial branching points for the subsequent reaction pathways. The first product, formic acid (HCOOH), is formed directly from the *HCOO intermediate via step r3, which involves a single PCET step. Simultaneously, the *COOH intermediate undergoes step r2 to produce the key intermediate, adsorbed carbon monoxide (*CO). Desorption of this species via step r4 to form the second product, gaseous carbon monoxide (CO).

Further reduction of the *CO intermediate initiates the pathways toward more reduced products. Two main routes are depicted: one through the *CHO intermediate, (step r5), and the other through the *COH intermediate, (step r6). The *CHO pathway proceeds through successive PCET steps (r7, r11) to form *CH_3O. This intermediate is then reduced in step r16 which subsequently desorbs to produce the third product, gaseous methanol (CH_3OH).

The *COH pathway provides a separate route to the methane product. Subsequent reduction of *COH (step r8) leads to the *CHOH intermediate, which is further reduced to *CH_2OH in step r13. From this point, the mechanism shows a convergence: the *CH_2OH intermediate can be reduced to the methanol and key methyl intermediate, *CH_2, via step r18. Alternative routes to the *CH intermediate also exist, proceeding through steps r9/r14 (via *CH) and r10/r15 (via *C). The *CH_3 intermediate is then a direct precursor to methane, undergoing a final PCET step (r20) to form the desorption (r21) to produce the final product, gaseous methane (CH_4).

The presence of multiple converging and diverging pathways underscores the complex nature of CO_2 electroreduction. The observed product selectivity on a given catalyst is a direct result of the relative kinetics of these competing reaction steps, influenced by factors such as surface binding energies of the intermediates and the applied electrode potential. The example of energy profile of CO_2RR over Cu catalyst as shown in **Figure 4.6**. This mechanism provides a fundamental framework for understanding and optimizing catalysts for specific product outcomes.

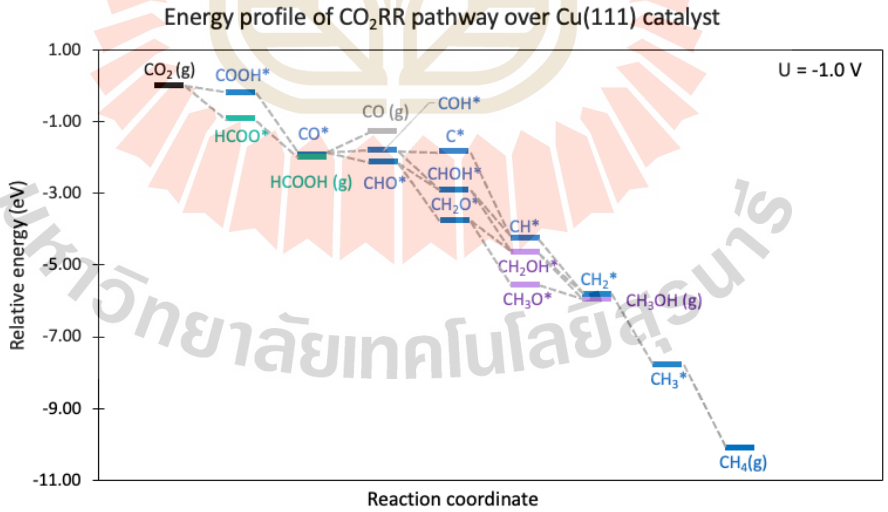


Figure 4.6 The energy profile of CO_2 reduction to C1 product over Cu catalyst with apply potential -1.0 V.

The energy profile shows that all key intermediates are stabilized relative to gaseous CO_2, as they all lie below the zero-energy reference line with apply potential. For example, the formation of the initial intermediates, *COOH and *HCOO, is an exergonic (stabilizing) step, with *COOH being slightly more stabilized. The reduction of these intermediates to the key *CO intermediate is also a stabilizing process, as *CO is at a significantly lower energy than its precursors. This strong stabilization of *CO on the Cu surface is a defining characteristic of Cu catalysts (Bagger et al., 2017, 2019; Hirunsit, 2013), enabling the subsequent deep reduction to hydrocarbons.

The overall trend shows a continuous and strong stabilization as the reaction proceeds, with each successive intermediate having a lower energy than the one before it. The most thermodynamically favorable intermediates and products are those with the lowest energy. In this profile, the final product, gaseous methane (CH4), is the most stabilized species, residing at the lowest energy level. This indicates that the reaction pathway leading to methane is the most thermodynamically favorable, consistent with the fact that Cu is a primary catalyst for CH_4 production. Conversely, an intermediate that is too strongly stabilized could potentially "poison" the catalyst by blocking active sites and making its conversion to the next intermediate a difficult, uphill process. However, this profile does not show any such steps, suggesting a favorable cascade toward the final products. As shown in **Table 4.1**, the reaction energy for each step can be calculated as the energy difference between the product and the reactant of that specific step, with a negative value signifying an exergonic reaction.

Table 4.1 The list of elementary reaction and reaction energy equation of CO_2 reduction to C1 product.

Code	Reaction	Reaction energy equations
R0	$CO_2(g) + * + 1\times\frac{1}{2}H_2(g) \rightarrow *COOH$	$\Delta E_0 = (\Delta E^{FC}_{COOH^*} + \mu_{COOH}) - \frac{1}{2}E^{DFT}_{H_2} - E^{DFT}_{CO_2}$
R1	$CO_2(g) + * + 1\times\frac{1}{2}H_2(g) \rightarrow *HCOO$	$\Delta E_1 = (\Delta E^{FC}_{HCOO^*} + \mu_{HCOO}) - \frac{1}{2}E^{DFT}_{H_2} - E^{DFT}_{CO_2}$
R2	$*COOH + 1\times\frac{1}{2}H_2(g) \rightarrow *CO + 1\times H_2O(l)$	$\Delta E_2 = (\Delta E^{FC}_{CO^*} + \mu_{CO}) - (E^{FC}_{COOH^*} + \mu_{COOH}) + E^{DFT}_{H_2O} - \frac{1}{2}E^{DFT}_{H_2}$
R3	$*HCOO + 1\times\frac{1}{2}H_2(g) \rightarrow HCOOH + *$	$\Delta E_3 = E^{DFT}_{HCOOH} - (E^{FC}_{HCOO^*} + \mu_{HCOO}) - \frac{1}{2}E^{DFT}_{H_2}$
R4	$*CO \rightarrow CO(g) + *$	$\Delta E_4 = E^{DFT}_{CO} - (E^{FC}_{CO^*} + \mu_{CO})$
R5	$*CO + 1\times\frac{1}{2}H_2(g) \rightarrow *CHO$	$\Delta E_5 = (E^{FC}_{CHO^*} + \mu_{CHO}) - (E^{FC}_{CO^*} + \mu_{CO}) - \frac{1}{2}E^{DFT}_{H_2}$
R6	$*CO + 1\times\frac{1}{2}H_2(g) \rightarrow *COH$	$\Delta E_6 = (E^{FC}_{COH^*} + \mu_{COH}) - (E^{FC}_{CO^*} + \mu_{CO}) - \frac{1}{2}E^{DFT}_{H_2}$
R7	$*CHO + 1\times\frac{1}{2}H_2(g) \rightarrow *CH_2O$	$\Delta E_7 = (E^{FC}_{CH_2O^*} + \mu_{CH_2O}) - (E^{FC}_{CHO^*} + \mu_{CHO}) - \frac{1}{2}E^{DFT}_{H_2}$
R8	$*CHO + 1\times\frac{1}{2}H_2(g) \rightarrow *CHOH$	$\Delta E_8 = (E^{FC}_{CHOH^*} + \mu_{CHOH}) - (E^{FC}_{CHO^*} + \mu_{CHO}) - \frac{1}{2}E^{DFT}_{H_2}$
R9	$*COH + 1\times\frac{1}{2}H_2(g) \rightarrow *CHOH$	$\Delta E_9 = (E^{FC}_{CHOH^*} + \mu_{CHOH}) - (E^{FC}_{COH^*} + \mu_{COH}) - \frac{1}{2}E^{DFT}_{H_2}$
R10	$*COH + 1\times\frac{1}{2}H_2(g) \rightarrow *C + 1\times H_2O(l)$	$\Delta E_{10} = (E^{FC}_{C^*} + \mu_C) - (E^{FC}_{COH^*} + \mu_{COH}) + E^{DFT}_{H_2O} - \frac{1}{2}E^{DFT}_{H_2}$
R11	$*CH_2O + 1\times\frac{1}{2}H_2(g) \rightarrow *CH_3O$	$\Delta E_{11} = (E^{FC}_{CH_3O^*} + \mu_{CH_3O}) - (E^{FC}_{CH_2O^*} + \mu_{CH_2O}) - \frac{1}{2}E^{DFT}_{H_2}$
R12	$*CH_2O + 1\times\frac{1}{2}H_2(g) \rightarrow *CH_2OH$	$\Delta E_{12} = (E^{FC}_{CH_2OH^*} + \mu_{CH_2OH}) - (E^{FC}_{CH_2O^*} + \mu_{CH_2O}) - \frac{1}{2}E^{DFT}_{H_2}$
R13	$*CHOH + 1\times\frac{1}{2}H_2(g) \rightarrow *CH_2OH$	$\Delta E_{13} = (E^{FC}_{CH_2OH^*} + \mu_{CH_2OH}) - (E^{FC}_{CHOH^*} + \mu_{CHOH}) - \frac{1}{2}E^{DFT}_{H_2}$
R14	$*CHOH + 1\times\frac{1}{2}H_2(g) \rightarrow *CH + 1\times H_2O(l)$	$\Delta E_{14} = (E^{FC}_{CH^*} + \mu_{CH}) - (E^{FC}_{CHOH^*} + \mu_{CHOH}) + E^{DFT}_{H_2O} - \frac{1}{2}E^{DFT}_{H_2}$
R15	$*C + 1\times\frac{1}{2}H_2(g) \rightarrow *CH$	$\Delta E_{15} = (E^{FC}_{CH^*} + \mu_{CH}) - (E^{FC}_{C^*} + \mu_C) - \frac{1}{2}E^{DFT}_{H_2}$
R16	$*CH_3O + 1\times\frac{1}{2}H_2(g) \rightarrow CH_3OH(g) + *$	$\Delta E_{16} = E^{DFT}_{CH_3OH} - (E^{FC}_{CH_3O^*} + \mu_{CH_3O}) - \frac{1}{2}E^{DFT}_{H_2}$
R17	$*CH_2OH + 1\times\frac{1}{2}H_2(g) \rightarrow CH_3OH(g) + *$	$\Delta E_{17} = E^{DFT}_{CH_3OH} - (E^{FC}_{CH_2OH^*} + \mu_{CH_2OH}) - \frac{1}{2}E^{DFT}_{H_2}$
R18	$*CH_2OH + 1\times\frac{1}{2}H_2(g) \rightarrow *CH + 1\times H_2O(l)$	$\Delta E_{18} = (\Delta E^{FC}_{CH_2^*} + \mu_{CH_2}) - (E^{FC}_{CH_2OH^*} + \mu_{CH_2OH}) + E^{DFT}_{H_2O} - \frac{1}{2}E^{DFT}_{H_2}$
R19	$*CH + 1\times\frac{1}{2}H_2(g) \rightarrow *CH_2$	$\Delta E_{19} = (E^{FC}_{CH_2^*} + \mu_{CH_2}) - (E^{FC}_{CH^*} + \mu_{CH}) - \frac{1}{2}E^{DFT}_{H_2}$
R20	$*CH_2 + 1\times\frac{1}{2}H_2(g) \rightarrow *CH_3$	$\Delta E_{20} = (E^{FC}_{CH_3^*} + \mu_{CH_3}) - (E^{FC}_{CH_2^*} + \mu_{CH_2}) - \frac{1}{2}E^{DFT}_{H_2}$
R21	$*CH_3 + 1\times\frac{1}{2}H_2(g) \rightarrow CH_4(g) + *$	$\Delta E_{21} = E^{DFT}_{CH_4} - (E^{FC}_{CH_3^*} + \mu_{CH_3}) - \frac{1}{2}E^{DFT}_{H_2}$

4.4 Microkinetic model

To better understand the behavior of catalysts for CO_2RR under experimental conditions, we explored microkinetic modeling based on our proposed mechanisms. Our microkinetic simulations were generated from a custom Python code (Python version 3.9.19). The core concept of this approach is to use the calculated energy profile to predict the reaction rate and selectivity of our target products (Filot, 2023).

In the previous section (section 4.3), we already calculated the reaction energies and generated the energy profiles for all catalysts which is the input for microkinetic. For the initial condition of microkinetic model, the partial pressure of CO_2 is equal to 1 bar and other species consist of HCOOH, CO, CH_3OH, and CH_4 are equal to 1.00×10^{-5} bar. The temperature is set as 300 K. The transition state theory (TST) was used to calculate the rate constants for each elementary step i. The rate constant of surface reaction is calculated by the following the Arrhenius equation:

$$\mathrm{k_i = A_i exp\left(\frac{-\Delta G^{\ddagger})}{k_B T}\right)}$$

where A_i is an effective pre-factor. As an approximation, the pre-factor was set to equal $\frac{\mathrm{k_B T}}{\mathrm{h}}$ s^{-1} for all the surface elementary reactions. The energy barrier of electrochemical step was set as 0.26 eV for all the $H^+ + e^-$ steps demonstrating the fast kinetics (Tripković et al., 2010).

The rate constant of adsorption reactions (k_{ads}) and desorption reaction (k_{des}) are calculated using collision theory,

$$\mathrm{k_{ads} = \frac{A}{\sqrt{2\pi m k_B T}}}$$

$$\mathrm{k_{des} = \frac{k_{ads}}{exp\left(\frac{-\Delta G^{\ddagger}}{k_B T}\right)}}$$

Where A is an active site of catalyst surface, which is equal to 4.00×10^{-20} m^2, and m represents the molecular weight (kg) of adsorbed species. Note that, this thesis we did not consider any adsorption step.

The net rate of each elementary step mM $\leftrightarrows$ nN was determined as,

$$r = k_{+}\theta_{M}^{m} - k_{-}\theta_{N}^{n}$$

where $k_{+/-}$ represent the rate constants of the forward and backward reactions, respectively. θ_{int} represents the coverage of intermediate and this modeled using the steady-state approximation. The sum of coverages of all the intermediates and free sites θ_{*} is 1 ML, as follows:

$$\Sigma\theta_{int} + \theta_{*} = 1$$

$$\frac{d\theta_{int}}{d_t} = 0$$

The coverage species on all of catalyst surface are considered as one active site. The steady-state approximation was used to solve the ordinary differential equations (ODE). The rate equation and the ODE of each elementary step are listed in **Table 4.2** and **Table 4.3**.

Additionally, the percentage of relative rate of each product can be determined from the rates of desorption species. The selectivity for 4 products: HCOOH, CO, CH_3OH, and CH_4 are calculated using the following equation:

$$S_i = \frac{R_i}{R_{tot}} \times 100\%$$

Where R_i represents the rate of specific desorbed species, and R_{tot} is the summation rate of desorbed species.

Table 4.2 The elementary steps and the rate equations of CO_2RR on the catalyst surface used in microkinetic modelling.

Code	Reaction	Rate equation
R0	$CO_2(g) + * + 1\times\frac{1}{2}H_2(g) \rightarrow *COOH$	$R0 = k_0 P_{CO_2}[\theta_*] - k_{-0}[\theta_{COOH}]$
R1	$CO_2(g) + * + 1\times\frac{1}{2}H_2(g) \rightarrow *HCOO$	$R1 = k_1 P_{CO_2}[\theta_*] - k_{-1}[\theta_{HCOO}]$
R2	$*COOH + 1\times\frac{1}{2}H_2(g) \rightarrow *CO + 1\times H_2O(l)$	$R2 = k_2[\theta_{COOH}] - k_{-2}[\theta_{CO}]$
R3	$*HCOO + 1\times\frac{1}{2}H_2(g) \rightarrow HCOOH + *$	$R3 = k_3[\theta_{HCOO}] - k_{-3}P_{HCOOH}[\theta_*]$
R4	$*CO \rightarrow CO(g) + *$	$R4 = k_4[\theta_{CO}] - k_{-4}P_{CO}[\theta_*]$
R5	$*CO + 1\times\frac{1}{2}H_2(g) \rightarrow *CHO$	$R5 = k_5[\theta_{CO}] - k_{-5}[\theta_{CHO}]$
R6	$*CO + 1\times\frac{1}{2}H_2(g) \rightarrow *COH$	$R6 = k_6[\theta_{CO}] - k_{-6}[\theta_{COH}]$
R7	$*CHO + 1\times\frac{1}{2}H_2(g) \rightarrow *CH_2O$	$R7 = k_7[\theta_{CHO}] - k_{-7}[\theta_{CH_2O}]$
R8	$*CHO + 1\times\frac{1}{2}H_2(g) \rightarrow *CHOH$	$R8 = k_8[\theta_{CHO}] - k_{-8}[\theta_{CHOH}]$
R9	$*COH + 1\times\frac{1}{2}H_2(g) \rightarrow *CHOH$	$R9 = k_9[\theta_{COH}] - k_{-9}[\theta_{CHOH}]$
R10	$*COH + 1\times\frac{1}{2}H_2(g) \rightarrow *C + 1\times H_2O(l)$	$R10 = k_{10}[\theta_{COH}] - k_{-10}[\theta_C]$
R11	$*CH_2O + 1\times\frac{1}{2}H_2(g) \rightarrow *CH_3O$	$R11 = k_{11}[\theta_{CH_2O}] - k_{-11}[\theta_{CH_3O}]$
R12	$*CH_2O + 1\times\frac{1}{2}H_2(g) \rightarrow *CH_2OH$	$R12 = k_{12}[\theta_{CH_2O}] - k_{-12}[\theta_{CH_2OH}]$
R13	$*CHOH + 1\times\frac{1}{2}H_2(g) \rightarrow *CH_2OH$	$R13 = k_{13}[\theta_{CHOH}] - k_{-13}[\theta_{CH_2OH}]$
R14	$*CHOH + 1\times\frac{1}{2}H_2(g) \rightarrow *CH + 1\times H_2O(l)$	$R14 = k_{14}[\theta_{CHOH}] - k_{-14}[\theta_{CH}]$
R15	$*C + 1\times\frac{1}{2}H_2(g) \rightarrow *CH$	$R15 = k_{15}[\theta_C] - k_{-15}[\theta_{CH}]$
R16	$*CH_3O + 1\times\frac{1}{2}H_2(g) \rightarrow CH_3OH(g) + *$	$R16 = k_{16}[\theta_{CH_3O}] - k_{-16}P_{CH_3OH}[\theta_*]$
R17	$*CH_2OH + 1\times\frac{1}{2}H_2(g) \rightarrow CH_3OH(g) + *$	$R17 = k_{17}[\theta_{CH_2OH}] - k_{-17}P_{CH_3OH}[\theta_*]$
R18	$*CH_2OH + 1\times\frac{1}{2}H_2(g) \rightarrow *CH + 1\times H_2O(l)$	$R18 = k_{18}[\theta_{CH_2OH}] - k_{-18}[\theta_{CH}]$
R19	$*CH + 1\times\frac{1}{2}H_2(g) \rightarrow *CH_2$	$R19 = k_{19}[\theta_{CH}] - k_{-19}[\theta_{CH_2}]$
R20	$*CH_2 + 1\times\frac{1}{2}H_2(g) \rightarrow *CH_3$	$R20 = k_{20}[\theta_{CH_2}] - k_{-20}[\theta_{CH_3}]$
R21	$*CH_3 + 1\times\frac{1}{2}H_2(g) \rightarrow CH_4(g) + *$	$R21 = k_{21}[\theta_{CH_3}] - k_{-21}P_{CH_4}[\theta_*]$

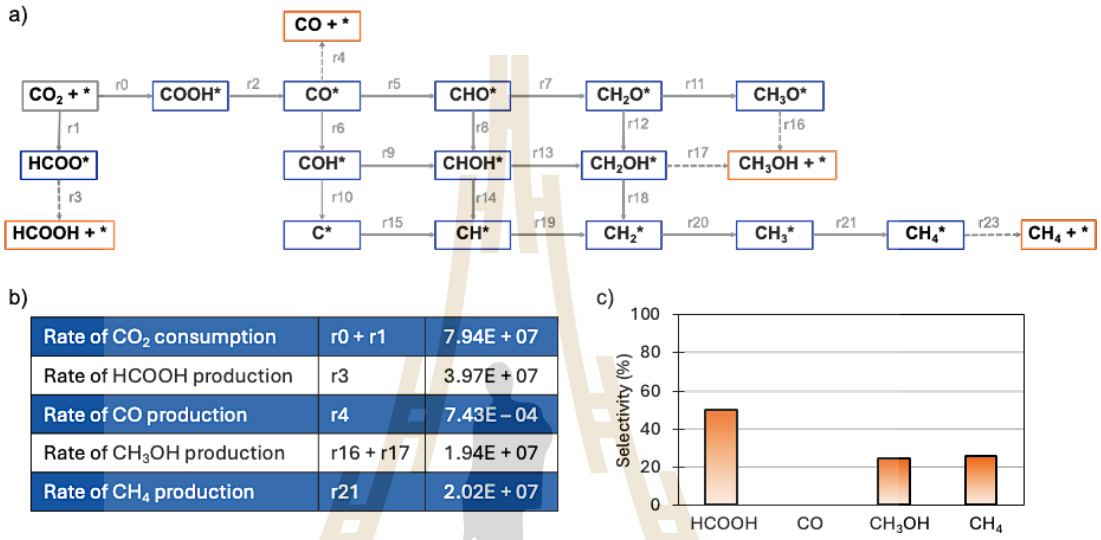
Table 4.3 The ODE equation for each elementary step.

Coverage species	ODE
*COOH	$\frac{d\theta_{COOH}}{dt} = R0 + R2$
*HCOO	$\frac{d\theta_{HCOO}}{dt} = R1 - R3$
*CO	$\frac{d\theta_{CO}}{dt} = R2 - R4 - R5 - R6$
*CHO	$\frac{d\theta_{CHO}}{dt} = R5 - R7 - R8$
*COH	$\frac{d\theta_{COH}}{dt} = R6 - R9 - R10$
*CH_2O	$\frac{d\theta_{CH_2O}}{dt} = R7 - R11 - R12$
*CHOH	$\frac{d\theta_{CHOH}}{dt} = R8 + R9 - R13 - R14$
*C	$\frac{d\theta_{C}}{dt} = R10 - R15$
*CH_3O	$\frac{d\theta_{CH_3O}}{dt} = R11 - R16$
*CH_2OH	$\frac{d\theta_{CH_2OH}}{dt} = R12 + R13 - R17 - R18$
*CH	$\frac{d\theta_{CH}}{dt} = R14 + R15 - R19$
*CH_2	$\frac{d\theta_{CH_2}}{dt} = R18 + R19 - R20$
*CH_3	$\frac{d\theta_{CH_3}}{dt} = R20 - R21$
*	$\frac{d\theta_{*}}{dt} = R3 + R4 + R16 + R17 + R21 - R0 - R1$

As an example, we can consider a microkinetic model for CO_2RR on a Cu catalyst, which is widely recognized as the most promising candidate for the reduction of CO_2 beyond the CO intermediate to form valuable C1 and C2 products. The rate of each elementary reaction step in this model is fundamentally governed by the corresponding energy profile. The example of result from microkinetic as shown in **Figure 4.7**.

Specifically, the forward and reverse reaction rates are control by the activation barriers and reaction energies of each step, which are derived directly from the computed energy profile. A large reaction energy in any step, especially a highly uphill process, will significantly affect the overall reaction dynamics. Such an uphill step will create a high-energy bottleneck, where the production of the subsequent intermediate is thermodynamically suppressed. This leads to a substantial decrease in the rate of

that specific step and, consequently, a slowdown in the overall catalytic turnover. For instance, if the reduction of a key intermediate, such as *COOH to *CO, has a large positive reaction energy, the coverage of *COOH will remain low, and the overall reaction will be kinetically limited at this point.



Rate of CO_2 consumption	r0 + r1	7.94E + 07
Rate of HCOOH production	r3	3.97E + 07
Rate of CO production	r4	7.43E – 04
Rate of CH_3OH production	r16 + r17	1.94E + 07
Rate of CH_4 production	r21	2.02E + 07

Figure 4.7 The output from microkinetic model over Cu catalyst a) energy profile of CO_2RR at U = –1.0 V, a) reaction network, b) the summation of consumption and production rate, and c) percent selectivity of each C1 products.

Moreover, to consider for the influence of the electrode potential on rate and selectivity of product on each catalyst, we have applied potential with relevant value from 0.0 V to –1.5 V. Finally, the output from microkinetic model is consist of the overall rate and selectivity of 4 products at 10 potentials value for all catalyst. The total number of datasets is equal 16,320 which is used to investigate in next machine learning part.

4.5 Machine learning prediction for rate and selectivity

The conventional design of catalysts often relies on a trial-and-error approach, which is time-consuming and resource-intensive. In computational electrocatalyst screening, the problem becomes even more difficult. Because of the huge variety of

potential materials and surface adsorbates, using density functional theory (DFT) for calculations is extremely time-consuming (Cheng et al., 2013; H. Li et al., 2020). A compelling example is the screening of CO_2 reduction activity over Ni-Ga alloys, which involved four distinct compositions and 40 different facet terminations. This matrix of parameters led to 583 unique adsorption sites, necessitating over 70,000 single-point DFT calculations and underscoring the significant computational bottleneck of traditional screening methods (Ulissi et al., 2017).

The machine learning models were trained on a comprehensive dataset derived from the output of a microkinetic model. This model simulated the electrocatalytic reduction of CO_2 over various bimetallic catalysts under relevant applied bias potentials. The flowchart of machine learning process as shown in **Figure 4.8**.

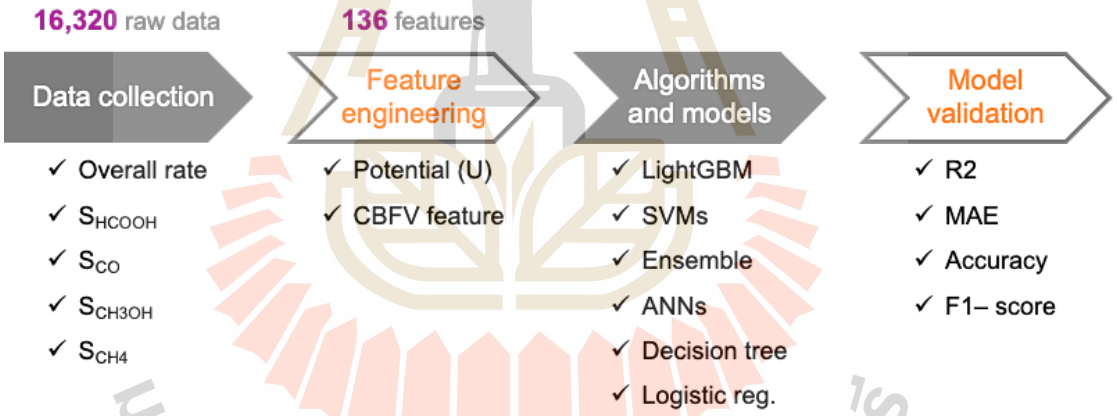


Figure 4.8 The flowchart of machine learning process for predicting the rate and selectivity of catalyst for CO_2RR.

The dataset comprises a total of 16,320 data points, with each point representing a unique catalytic system and its corresponding performance metrics. The key outputs from the microkinetic simulations that were used as target variables for our models include the overall reaction rate, defined as the consumption rate of CO_2, and the selectivity values for four distinct products. For feature engineering, we utilized the Composition-Based Featurization Vector (CBFV) method, which systematically

converts the elemental composition of each bimetallic catalyst into a robust feature vector. This approach ensures that the machine learning models are trained on chemically relevant descriptors. Prior to model training, a critical data cleaning step was performed to remove outliers, resulting in a final refined dataset of 15,956 data points, which was then used for all subsequent analysis and model development. For the selectivity prediction, we used the regression model, whereas the classification model was used to predict the overall rate of bimetallic catalyst. The data distribution as shown in **Figure 4.9**.

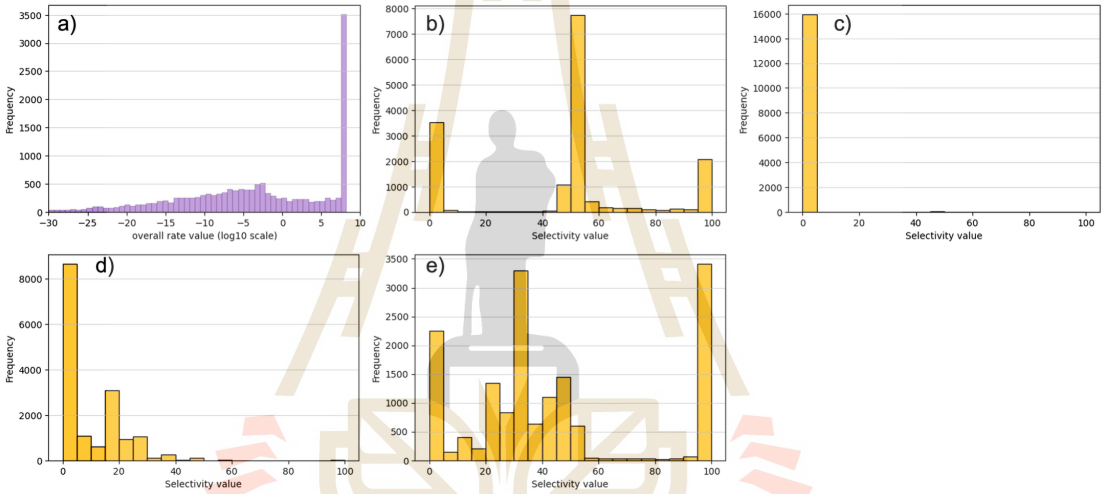


Figure 4.9 The data distribution of target a) overall rate, b) S_{HCOOH}, c) S_{CO}, d) S_{CH3OH}, and e) S_{CH4}.

4.5.1 Featurization

In machine learning, transforming raw data into meaningful inputs is a crucial step for building accurate predictive models. This process, known as feature engineering, involves selecting or creating relevant quantitative parameters—or features—that best represent the characteristics of the properties that we want to predict.

This thesis explores a composition-based feature vector approach using the CBFV method. To generate these vectors, we utilized the Oliynyk feature set, which represents a specific collection of elemental property descriptors such as atomic

number, atomic weight, and covalent radius. This set provides a rich, chemically intuitive representation of the bimetallic catalyst composition. Furthermore, to account for the crucial influence of reaction conditions, we included the applied potential (U) as an additional feature, allowing our models to capture the electrochemical driving force of the reaction. In total, this process resulted in a comprehensive feature set consisting of 264 unique descriptors for each data point.

To streamline the model and avoid overfitting from a high-dimensional feature space, a feature selection process was employed. We used Pearson correlation to identify and remove highly correlated features, ensuring that the model is trained on a set of independent and informative descriptors. This method quantifies the linear relationship between each pair of features, allowing us to systematically eliminate redundant information. The detail of feature selection as shown in **Figure 4.10**.

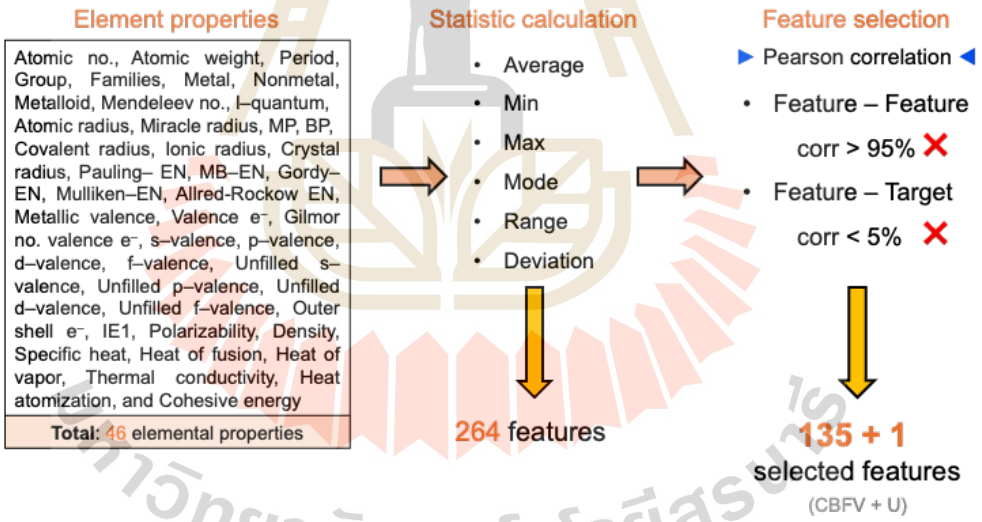


Figure 4.10 The detail of feature selection by Pearson correlation matrix. Beginning with list of elemental properties which is containing in CBFV features. Then, talking elemental descriptor and compute statistic across element and weighted by stoichiometry (Durdy et al., 2022).

After this selection step, the initial 264 features were reduced to a more concise and optimized set of 136 features. This refined feature set was subsequently used to train all machine learning models. The selected feature was shown in **Figure 4.11**.

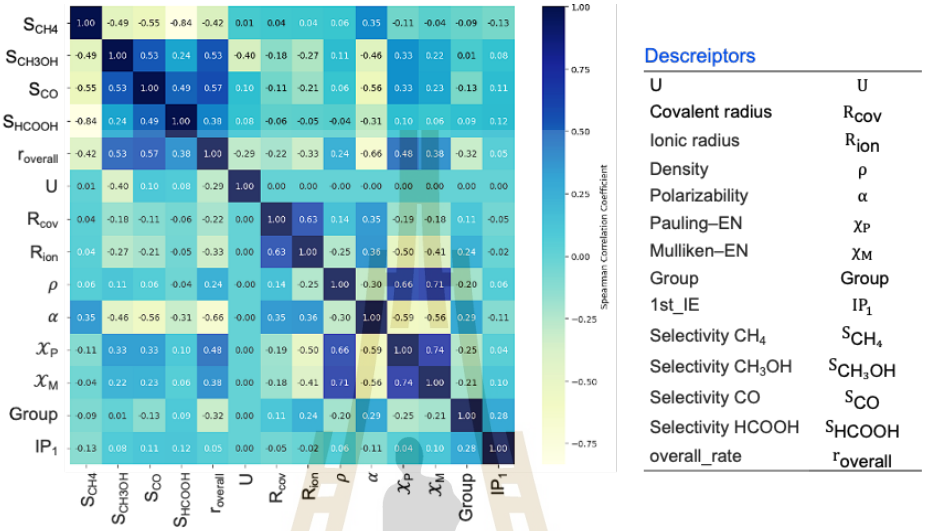


Figure 4.11 The correlation matrix of feature selection.

4.5.2 Model prediction of rate and selectivity

Selecting the optimal machine learning model is a critical step in obtaining reliable predictions for catalyst performance in CO_2RR. To capture the complex, non-linear relationships between catalyst composition and catalytic activity, we tested and compared several widely used models, seeking to identify the most suitable algorithms for this specific problem.

To ensure robust validation of our models, the prepared dataset was divided into a training set and a testing set with an 80/20 ratio. We investigated the predictive power of four different regression algorithms for selectivity prediction, which are described in the subsequent sections. After training and optimizing the hyperparameters for each model, their performance was evaluated. This evaluation process allowed for a comprehensive assessment of each algorithm's accuracy and enabled a direct comparison of their predictive capabilities on unseen data.

Table 4.4 The evaluation of test set of 4 regression models to predict the selectivity of C1 products.

	S_{HCOOH}		S_{CO}		S_{CH3OH}		S_{CH4}	
	R^2	MAE	R^2	MAE	R^2	MAE	R^2	MAE
LightGBM	0.79	6.88	0.20	0.22	0.78	2.72	0.81	7.61
SVR	0.59	9.68	0.12	0.26	0.51	4.47	0.65	10.37
Ensemble	0.75	7.87	0.25	0.21	0.74	2.99	0.79	8.08
ANN	0.79	7.19	0.16	0.14	0.76	3.13	0.81	7.61

After evaluating multiple machine learning models, the Artificial Neural Network (ANN) was selected as the most suitable algorithm for predicting both the overall rate and the selectivity of CO_2RR to various products. As shown in **Table 4.4**, while models like LightGBM, ensemble method, and SVM achieved slightly better performance for certain single-target predictions, they are inherently designed to predict only one target property at a time. This would require training and deploying separate models for each target variables a process that is less efficient.

In contrast, the multi-output capability of the ANN model allows it to learn the complex relationships between catalyst composition and all target properties simultaneously. This approach provides a more holistic and integrated solution, making it the most straightforward and effective method for predicting the multi-faceted performance of bimetallic catalysts. Therefore, despite similar predictive accuracy, the ANN model's ability to provide a complete, single-pass prediction of all relevant outputs makes it the superior choice for this thesis.

The result from the prediction gives significant different from the experiment study. As shown in **Figure 4.12**, the experimental data shows a significant increase in FE% for a particular product as the percentage of Zn increases from 0% to 100%. In contrast, the Fairchem prediction shows a much flatter trend, with little change in selectivity across the same range of Zn content. This divergence highlights a key

limitation of the Fairchem model: it does not fully capture the complex, synergistic effects that likely occur in the experiment of CuZn alloy system. The model's inability to accurately predict the strong correlation between zinc content and product selectivity suggests that it may not be adequately accounting for factors such as the experimental condition, the electronic structure changes, surface morphology, or the role of specific catalytic sites that are crucial for the reaction pathway. The experimental results indicate that these factors are critically important for tuning the selectivity of the electrochemical CO_2 reduction reaction.

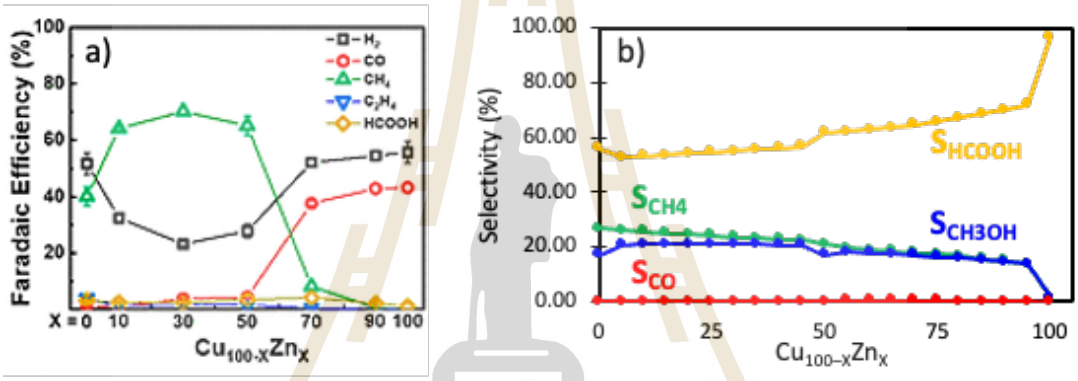


Figure 4.12 The validation result of selectivity on CuZn catalyst from a) experimental literature (Jeon et al., 2019), and b) model prediction.

For the overall rate prediction, we explore four difference classification models to classify the group of rate zone which is represent the performance of production of target product. The group are including excellent, acceptable, and bad rate corresponding to each catalyst as shown in **Figure 4.13**.

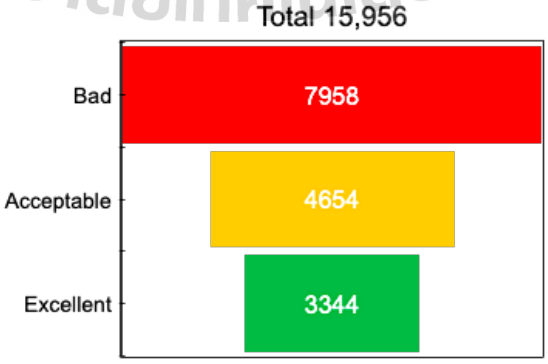


Figure 4.13 classifying the performance of overall rate

The threshold to classify each group is define from the approximate the rate constant where energy barrier was set as fast PCET step (~5.70×10^7 s^{-1}). After training and tuning hyperparameters of each model, their performance was evaluate as shown in **Table 4.5**.

Table 4.5 The evaluation of 4 classification models to predict the overall rate of each catalyst.

		Excellent		Acceptable		Bad	
	Accuracy	Precision	F1-score	Precision	F1-score	Precision	F1-score
LightGBM	0.86	0.92	0.92	0.77	0.76	0.89	0.90
Decision tree	0.81	0.87	0.87	0.69	0.68	0.86	0.87
Logistic regression	0.76	0.81	0.81	0.61	0.59	0.82	0.84
SVC	0.86	0.93	0.92	0.78	0.74	0.86	0.89

As shown in **Table 4.5**, the LightGBM and Decision Tree models achieved very high training accuracy (99% and 97% respectively), their performance dropped significantly on the test data (86% for LightGBM and 81% for Decision Tree). This large drop between the training and test scores is a clear sign of overfitting. These models likely memorized the training data and are not reliable for predicting the performance of new, unseen catalysts. Whereas, the Logistic Regression model showed a much smaller gap between its training (79%) and test (76%) accuracy, which is a good sign that it generalizes well. However, its overall performance is the lowest of all the models.

For the SVC model. It provides the best combination of performance and generalization. With a training accuracy of 92% and a test accuracy of 86%, the performance drop is much smaller than the tree-based models, indicating that it is a more robust and reliable model for new data. Furthermore, when we specifically look at the metrics for the excellent class on the test set, the SVC model has a high precision of 93% and a recall of 91%, with an F1-score of 92%. This means the model is very good at identifying truly excellent catalysts and doesn't miss many of them.

This is the most critical aspect of the model for our application. The LightGBM model also performed well on the test set, but its severe overfitting is a major risk. Therefore, the SVC model exhibit strong test performance, combined with its resistance to overfitting, makes it the most suitable model for this classification task.

Finally, a support vector classifier from our classification model and a neural network from our regression model were finally employed to predict the rate and selectivity of unseen catalyst compositions. For user convenience, we have developed a user-friendly website. Users can simply input the chemical formula and select the desired potential. The output will then display a bar chart illustrating the selectivity of each product, along with a performance metric for the rate. The example of web interface as shown in **Figure 4.14**.

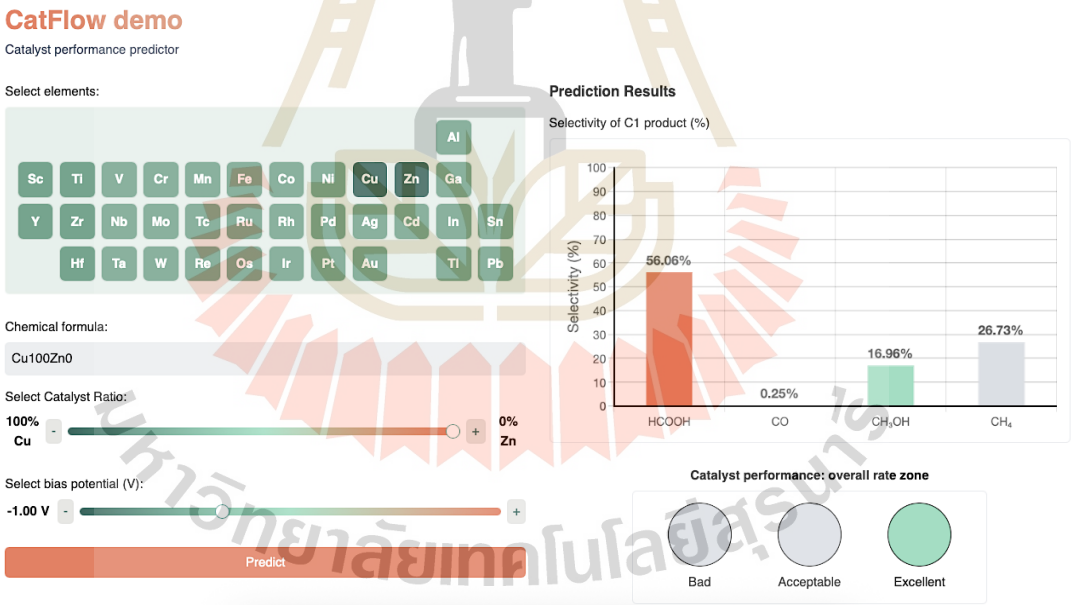


Figure 4.14 The example of web interface for the rate and selectivity predictor.

4.6 Cost structures

The cost structure for the CO_2RR catalyst prediction ML model business is designed under a One–Time–Purchase business model. This model dictates that most costs are categorized as fixed costs for development and annual operations, while the cost of goods sold (COGS) remains relatively low since the model's computation does not rely on continuous server usage.

4.6.1 Revenue and total cost assumptions

Based on the annual operating projection, the business relies on an assumed sales volume of 200 units (Licenses) per year, with a price set at $2,500 per unit, as detailed in **Table 4.6**.

Table 4.6 The detail of revenue and cost assumption.

Name	Value ($)	Unit
Sale Revenue	200	pcs.
	2,500	$
Total Sale Revenue	500,000	$
Cost of Good Sold (COGS)	227,000	$
	1,135	$/unit
Gross Profit	273,000	$
Selling & Marketing Expenses	100,000	$
General & Administrative Expenses	100,000	$
EBITDA	73,000	$
Operation Profit / EBIT	73,000	$
EBT (Earnings Before Taxes)	73,000	$
Taxes (20%)	14,600	$
Net Profit	**58,400**	**$**
	4,867	**$/month**

The COGS represents the variable cost incurred for every single license sold. The total annual COGS is $227,000, which translates to $1,135 per unit. The COGS for this business includes the necessary expenses for delivery and installation of the model onto the customer's system. This covers costs for storing the large model files on a cloud server, data transfer fees for the customer's download, and minor fees for building/preparing the final file package. The COGS of $1,135 per unit accounts for only 45.4% of the sale price ($2,500), resulting in a high gross profit.

Total fixed costs amount to $200,000 per year, categorized under others cost, representing operational expenses independent of sales volume:

Selling & Marketing Expenses: $100,000 per year (~$8,333 per month), which is the primary expenditure for attracting B2B clients through channels like Google Ads and content marketing.

General & Administrative Expenses: $100,000 per year (~$8,333 per month), covering staff salaries, rent, software licenses, and ongoing R&D expenditures.

The cost analysis indicates the business generates a net profit of $58,400 per year in the first year, which is a respectable profit margin given the high initial investment in personnel and R&D.

4.6.2 Break – even point (BEP)

The Break-Even Point calculation determines the number of licenses required to cover all fixed costs, which can calculate from the following equation,

$$\text{BEP (units)} = \frac{\textbf{Total fixed costs}}{\textbf{selling price per unit} - \textbf{COGS}}$$

$$\text{BEP (units)} = \frac{\$200{,}000}{\$2{,}500 - \$1{,}135} = \sim 147\ units$$

Consequently, the business needs to sell 147 licenses to reach its break-even point. This represents 73.5% of the targeted annual sales of 200 units. Any unit sold beyond the 147th unit contributes directly to the company's net profit.

4.7 Expansion opportunities

This section outlines potential avenues for future research and development to enhance the Fairchem model's accuracy and expand the utility of the associated platform.

4.7.1 Enhancing model accuracy

The discrepancy results between DFT and Fairchem model are shown in **Figure 4.15**. To improve the predictive accuracy of the Fairchem model, a delta learning approach can be implemented. This method involves training a separate model to predict the difference (or delta) between the Fairchem prediction and the actual, high-accuracy DFT data. By adding this correction to the Fairchem output, we can achieve DFT-level accuracy at a fraction of the computational cost. This approach is particularly effective because it allows the Fairchem model to handle the bulk of the calculation while the delta model focuses on correcting the nuanced errors.

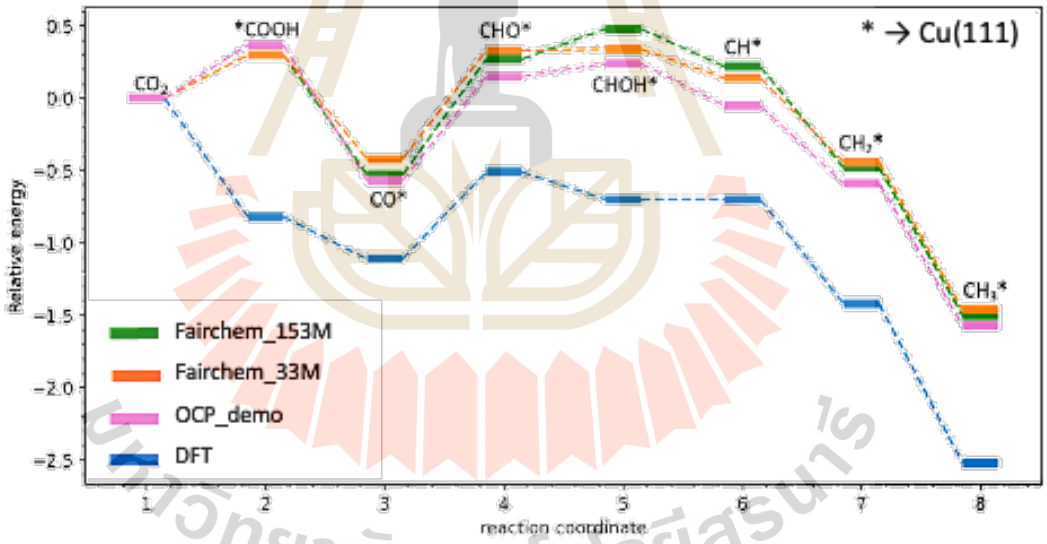


Figure 4.15 The energy profile of CO_2RR on Cu catalyst between DFT and Fairchem model.

Furthermore, the model can be fine-tuned with additional training data specifically focused on different reaction pathways of CO_2RR. This targeted training would enable the model to make more precise predictions for specific catalytic systems and conditions, moving beyond general adsorption energy calculations to more complex reaction scenarios.

4.7.2 Improving practical application

To make the platform more valuable for real-world applications, several practical considerations should be integrated. First, the user interface could be improved by incorporating constraints based on the stability of different bimetallic catalysts. Instead of allowing users to choose any arbitrary ratio, the system could guide them towards stable, synthesizable compositions, which would make the predictions more realistic and actionable for experimental chemists. Additionally, to provide more realistic utilization data, the platform could integrate economic and regional factors. For instance, including the cost and availability of precursor materials for different catalysts could help researchers and engineers make more informed decisions about which materials are not only catalytically active but also commercially viable.

Moreover, the model's utility can be significantly expanded by extending its predictive capabilities beyond CO_2RR. The current framework can be adapted to calculate the adsorption energies of other key species involved in reactions with natural gas or biomass. By training the model on the adsorption behavior of these different molecules, the platform could become a versatile tool for designing catalysts for a wide range of chemical processes, thereby broadening its impact across various fields of energy and chemical engineering. This scalability would make the platform a more comprehensive resource for catalyst design and discovery.

4.8 References

Bagger, A., Ju, W., Varela, A. S., Strasser, P., and Rossmeisl, J. (2017). Electrochemical CO_2 Reduction: A Classification Problem. *ChemPhysChem*, *18*(22), 3266–3273. https://doi.org/10.1002/cphc.201700736

Bagger, A., Ju, W., Varela, A. S., Strasser, P., and Rossmeisl, J. (2019). Electrochemical CO_2 Reduction: Classifying Cu Facets. *ACS Catalysis*, *9*(9), 7894–7899. https://doi.org/10.1021/acscatal.9b01899

Cheng, D., Negreiros, F. R., Aprà, E., and Fortunelli, A. (2013). Computational Approaches to the Chemical Conversion of Carbon Dioxide. *ChemSusChem*, *6*(6), 944–965. https://doi.org/https://doi.org/10.1002/cssc.201200872

Durdy, S., Gaultois, M. W., Gusev, V. V., Bollegala, D., and Rosseinsky, M. J. (2022). Random projections and kernelised leave one cluster out cross validation: universal baselines and evaluation tools for supervised machine learning of material properties. *Digital Discovery*, *1*(6), 763–778. https://doi.org/10.1039/d2dd00039c

Filot, Ivo. (2023). *Introduction to microkinetic modeling*. Technische Universiteit Eindhpoven.

Hirunsit, P. (2013). Electroreduction of carbon dioxide to methane on copper, copper-silver, and copper-gold catalysts: A DFT study. *Journal of Physical Chemistry C*, *117*(16), 8262–8268. https://doi.org/10.1021/jp400937e

Hirunsit, P., Soodsawang, W., and Limtrakul, J. (2015). CO_2 electrochemical reduction to methane and methanol on copper-based alloys: Theoretical insight. *Journal of Physical Chemistry C*, *119*(15), 8238–8249. https://doi.org/10.1021/acs.jpcc.5b01574

Jeon, H. S., Timosnenko, J., Scholten, F., Sinev, I., Herzog, A., Haase, F. T., and Cuenya, B. R. (2019). Operando insight into the correlation between the structure and composition of CuZn nanoparticles and their selectivity for the electrochemical CO_2 reduction. *Journal of the American Chemical Society*, *141*(50), 19879–19887. https://doi.org/10.1021/jacs.9b10709

Li, H., Xu, S., Wang, M., Chen, Z., Ji, F., Cheng, K., Gao, Z., Ding, Z., and Yang, W. (2020). Computational design of (100) alloy surfaces for the hydrogen evolution reaction. *J. Mater. Chem. A*, *8*(35), 17987–17997. https://doi.org/10.1039/D0TA04615A

Li, K., Li, L., Chang, X., Shi, X., Li, X., Pei, C., Zhao, Z. J., and Gong, J. (2024). Origin of Autocatalytic Behavior of Water over CuZn Alloy in CO_2 Hydrogenation. *Chem and Bio Engineering*, *1*(3), 274–282. https://doi.org/10.1021/cbe.3c00124

Liu, K., Wang, J., Shi, M., Yan, J., and Jiang, Q. (2019). Simultaneous Achieving of High Faradaic Efficiency and CO Partial Current Density for CO_2 Reduction via Robust, Noble-Metal-Free Zn Nanosheets with Favorable Adsorption Energy. *Advanced Energy Materials*, *9*(21). https://doi.org/10.1002/aenm.201900276

Tomboc, G. M., Choi, S., Kwon, T., Hwang, Y. J., and Lee, K. (2020). Potential Link between Cu Surface and Selective CO_2 Electroreduction: Perspective on Future Electrocatalyst Designs. *Advanced Materials*, *32*(17). https://doi.org/10.1002/adma.201908398

Tripković, V., Skúlason, E., Siahrostami, S., Nørskov, J. K., and Rossmeisl, J. (2010). The oxygen reduction reaction mechanism on Pt(1 1 1) from density functional theory calculations. *Electrochimica Acta*, *55*(27), 7975–7981. https://doi.org/10.1016/j.electacta.2010.02.056

Ulissi, Z. W., Tang, M. T., Xiao, J., Liu, X., Torelli, D. A., Karamad, M., Cummins, K., Hahn, C., Lewis, N. S., Jaramillo, T. F., Chan, K., and Nørskov, J. K. (2017). Machine-learning methods enable exhaustive searches for active Bimetallic facets and reveal active site motifs for CO_2 reduction. *ACS Catalysis*, *7*(10), 6600–6608. https://doi.org/10.1021/acscatal.7b01648

Zoubir, O., Atourki, L., Ait Ahsaine, H., and BaQais, A. (2022). Current state of copper-based bimetallic materials for electrochemical CO_2 reduction: a review. In *RSC Advances* (Vol. 12, Issue 46, pp. 30056–30075). Royal Society of Chemistry. https://doi.org/10.1039/d2ra05385c

CHAPTER V
CONCLUSION

This thesis presents a comprehensive computational and machine learning-based framework for the design of catalysts for the carbon dioxide reduction reaction. The goal was to establish an efficient workflow to predict the catalytic performance of various materials, thereby accelerating the discovery of new catalysts.

The multi-stage method involved creating a large database of 1,632 catalyst structures, including pure metal surfaces and bimetallic alloys with 1:1 and 1:3 compositions. We used Fairchem, a machine learning potential tool, to calculate the adsorption energies of key reaction intermediates to model the complex reaction network for converting CO_2 to products. This data then served as the input for a microkinetic model (MKM), which provided a robust prediction of reaction rates and selectivities for four primary products: formic acid, carbon monoxide, methanol, and methane. The output from this modeling was crucial for the next step.

Finally, we explored machine learning models to provide rapid predictions of catalyst performance. Using catalyst composition and applied potential as features, we evaluated various regression and classification models. The results demonstrated that an Artificial Neural Network (ANN) provided the best performance for selectivity prediction, while a Support Vector Classifier (SVC) was optimal for predicting the overall reaction rate. To make these models easy to use, we developed a simple website where users can input a catalyst's chemical formula and get an instant prediction.

An important limitation of this study is the discrepancy between our model-predicted results and existing experimental data. This can be attributed to several

factors. First, the Fairchem model was trained on a broad range of DFT data and is not specifically optimized for the complex reaction environment of CO_2RR. Additionally, our models do not account for critical experimental conditions, such as solvent effects, electrolyte interactions, and dynamic surface properties, which can significantly influence catalytic activity. This highlights a common challenge in bridging the gap between theoretical modeling and real-world experimental results.

To improve accuracy, a potential solution is to fine-tune the machine learning models with more specific data from detailed computational studies like DFT. Another idea is to use delta learning, where a simple model provides an initial prediction, and a second, more powerful model learns to correct the difference (or delta) between that prediction and the actual, more accurate value.

In conclusion, this thesis provides a practical framework for the rational design of CO_2RR catalysts. By integrating machine learning into the computational workflow, this research offers a valuable alternative to traditional high-cost computational screening. The developed framework, the explored ML models, and the available web application represent a significant step toward making catalyst design tools more accessible to the scientific community.

APPENDIX

APPENDIX

PUBLICATION AND PRESENTATIONS

A.1 List of publication

- Watwiangkham, A., Santatiwongchai, J., **Kaewpratum, A.**, Ngamwongwan, L., Chookajorn, T., Suthirakun, S., & Hirunsit, P. (2025). Cooperative phase-separated Cu–Zn catalyst for selective CO2 electrochemical reduction toward ethanol: theoretical insights into the role of the interface sites. Journal of Catalysis, 116393.
- Khajondetchairit, P., **Kaewpratum, A.**, Rittiruam, M., Saelee, T., Hirunsit, P., Suthirakun, S., & Praserthdam, S. (2024). Enhanced hydrogen adsorption-desorption reversibility found on NiAl alloy: A first-principles study. International Journal of Hydrogen Energy, 68, 8-16.

A.2 List of presentations

- **Kaewpratum, A.**; Ngamwongwan, L.; Fongkaew, I.; Suthirakuna S., Innovative Catalyst Design for CO2 Reduction Reaction: A Combined Multiscal Modelling and Machine learning Approach, The 28th International Annual Symposium on Computational Science and Engineering, Thammasat University, Thailand, 30th July – 1st August 2025. (Poster presentation)
- **Kaewpratum, A.**; Untarabut, P.; Chookajorn T.; Suthirakuna S., First Principles Study of CuZn segregation on CO2RR activity, The 26th International Annual Symposium on Computational Science and Engineering, A-One the Royal Cruise Hostel Pattaya, Thailand, 20th July – 22nd July 2023. (Oral presentation

CURRICULUM VITAE

Arisa Kaewpratum was born on November 5th, 1999, in Nakhonratchasima, Thailand. She received his B.Sc. in Chemistry in 2021 from the School of Chemistry, Institute of Science, Suranaree University of Technology, Thailand. During her bachelor's degree, she started research in computational materials science for developing adsorbent materials for hydrogen adsorption under the supervision of Assoc. Prof. Dr. Suwit Suthirakun. In the master's degree program at Integreated Science and Innovation, Suranaree University of Technology, Thailand, she changed to focus on electrocatalysis CO_2RR under the supervision of Assoc. Prof. Dr. Suwit Suthirakun. She has published articles during her master's program (as list in appendix).