Parametric Optimization and Intelligent CAD Automation for Bicycle Frame Design via Multi Agent

Chon Kit Chan[1], Wen-Yi Yang[2], Jin H. Huang[3]
[1] Department of Aerospace and Systems Engineering
[2] Master's Program of Electro acoustics
[3] Department of Mechanical and Computer Aided Engineering
Feng Chia University
Taichung, Taiwan (R.O.C.)
e-mails: p1130484@o365.fcu.edu.tw

***Abstract*—This research presents an intelligent design platform integrating Large Language Models (LLMs), Retrieval Augmented Generation (RAG), a multi-agent system, and the Model Context Protocol (MCP). It addresses the inefficiencies, high costs, and heavy reliance on manual expertise in traditional bicycle design workflows, which constrain both innovation and responsiveness to market demands. This challenge is especially pressing as the bicycle industry faces increasing demands for customization, small-batch production, and rapid product development. While prior approaches such as Machine Learning (ML) and data-driven optimization have shown promise, they remain confined to isolated tasks and lack a unified, end-to-end framework. The proposed system leverages these technologies to generate context-aware design recommendations tailored to user intent and automates Computer Aided Design (CAD) model generation, thereby substantially reducing development time and manual workload. As a proof of concept, the platform is first applied to rim design the component with the richest dataset before scaling to other components and full frame development. Experimental results demonstrate that the system can reduce design cycles from weeks to hours, providing higher efficiency, lower costs, and improved accuracy of design recommendations. The goal is to accelerate intelligent, flexible, and automated design workflows for the next generation of bicycle products.**

***Keywords*—Bicycle rim design; large language model; retrieval-augmented generation; multi-agent system; intelligent CAD automation.**

I. INTRODUCTION

Taiwan assumes a critical position in the global bicycle industry, renowned for its leading manufacturing technologies. However, the evolving market demand for highly customized, small batch, and diversified products presents persistent difficulties to Taiwan's bicycle industry, particularly in structural innovation and aesthetic design. Traditional design workflows are predominantly dependent on the experience and manual efforts of engineers. This includes extensive data collection (e.g., rim profiles, aerodynamic characteristics, weight, stiffness) and iterative design modifications. Such processes are not only time-consuming, often requiring over a month to complete an initial design, but also incur high Research and Development (R&D) costs and involve a large volume of repetitive tasks.

The current bicycle design workflow faces several critical challenges. Firstly, low efficiency prolongs design cycles and hampers responsiveness to market shifts. Secondly, a strong dependence on manual labor and iterative testing inflates R&D costs. Thirdly, significant time spent by designers on tedious data collection and model adjustments limits their capacity for creative innovation. Lastly, excessive reliance on senior engineers' experience complicates systematic knowledge management, application, and transfer.

However, existing approaches remain insufficient to overcome these challenges. Traditional CAD and engineering workflows are heavily manual, making them slow, costly, and difficult to scale. While data-driven and ML methods have recently been introduced, most prior efforts are restricted to narrow tasks such as parameter prediction or dataset construction. These approaches lack a unified framework that connects data analysis, knowledge retrieval, and CAD automation, limiting their applicability to complex, multi-component bicycle designs.

To overcome these limitations, this study introduces an innovative intelligent design platform integrating LLMs, RAG, a multi-agent system, and the MCP. The platform's core objective is twofold: to intelligently generate design parameters and recommendations aligned with user defined intent, and to automate the generation of CAD models, thereby alleviating the workload of manual drafting. Implementation of these features is anticipated to significantly shorten product development cycles, reduce R&D costs, and enhance both design efficiency and quality.

The remainder of this paper is organized as follows: Section 2 provides a review of relevant literature and technologies; Section 3 provides an in-depth explanation of the proposed AI system architecture and research methodology; Section 4 showcases the experimental results and offers analysis in Results and Discussion; and Section 5 summarizes the key contributions of this study and discusses directions for future research.

II. RELATED WORK

In recent years, Artificial Intelligence (AI) technologies, particularly LLMs, have made significant advancements across diverse domains and are increasingly demonstrating potential in complex engineering design and automation workflows. LLMs have shown practical value in specialized fields, such as medicine [1], and their capacity for natural language comprehension and generation has opened new avenues for enhancing human machine interaction, for instance, in human robot communication [2]. Academic research further indicates that LLMs can continuously refine their performance via self-improvement mechanisms [3] and have been successfully applied in areas like personalized recommendation systems [4].

Within intelligent design automation, especially for product specific applications like bicycle design, data-driven approaches have garnered considerable attention. For example, the BIKED dataset, developed by Regenwetter et al., provides valuable data resources for bicycle design and has facilitated the establishment of ML benchmarks and applications in this domain [5] [6]. These studies underscore the feasibility of employing data and ML for analyzing and optimizing design parameters.

This study utilizes an SQL Agent that interacts with databases via LLMs, aligning with current advancements in Text to SQL technologies. Numerous academic efforts have concentrated on augmenting the SQL generation capabilities of LLMs. These include integrating domain knowledge to improve accuracy (e.g., the Knowledge to SQL approach) [7], conducting systematic reviews of LLM based Text to SQL methodologies [8], and employing strategies like model fusion to enhance overall performance [9] and accuracy on real world relational databases [10]. These advancements provide a solid foundation for developing more natural and intelligent data querying and analysis interfaces.

The design platform proposed herein integrates RAG to bolster the model's domain-specific knowledge and query response accuracy, particularly concerning bicycle rim design. Furthermore, the platform adopts a Multi-Agent System architecture. This architecture is designed to decompose complex design tasks and delegate them to specialized agents with distinct functionalities, thereby improving the efficiency and modularity of the overall design workflow.

While the aforementioned technologies (LLMs, RAG, Multi-Agent Systems, MCP) have individually shown progress in specific applications, their integrated application to parametric optimization and CAD automation—particularly for complex components like bicycle rims and frames requiring multifaceted engineering considerations — remains a challenging yet promising research avenue. This study aims to offer an innovative solution to the limitations of traditional design workflows through the development of a unified intelligent design platform.

III. Methodology and AI System Architecture

This section introduces the methodology and overall system architecture of the proposed intelligent design platform. By integrating LLMs, RAG, a Multi-Agent System, and the MCP into a unified framework, the approach aims to overcome the inefficiencies and limitations of traditional bicycle design workflows. The subsections describe the system design, the construction of each agent, and the user interface.

A. *Overall AI System Architecture*

The AI-based bicycle rim design system developed in this study features a backend architecture as illustrated in Figure 1. The system is designed to handle user tasks, such as: "I would like to design a bicycle rim with low aerodynamic drag, low weight, and high stiffness. Please provide design recommendations and generate a CAD drawing."

At the core of the system is the Main Agent, responsible for receiving and interpreting user tasks, decomposing them, and distributing them to specialized sub agents within the Multi-Agent framework. It then consolidates their responses to generate a complete reply for the user. Considering that bicycle design often involves parameter analysis, textual reference consultation, and CAD drafting, this study introduces three key sub agents. The first is the SQL Agent, which automatically generates SQL queries to retrieve parameter data from the database; if the required data is not available, it leverages ML techniques to predict the parameters. The second is the RAG Agent, which enriches the system's domain knowledge in bicycle rim design to ensure the quality and accuracy of generated responses. The third is the CAD Agent, which uses the verified parameters—proposed collaboratively by the SQL and RAG Agents—to generate both 2D profiles and 3D models.

This study also employs prompt engineering to establish a horizontal communication mechanism among agents, enabling effective information flow between them. This approach enhances the overall reliability of user responses and helps mitigate the risk of potential "data silos."

In this study, Claude was adopted as the primary LLM based on response quality, computational efficiency, and practical usability. This choice ensured stable and reliable performance across different tasks without requiring additional reconciliation between multiple models.

B. *Training Data*

The AI training data used in this study is diverse, supporting various system functionalities. The first component is structured data, which includes the correlation between geometric dimensions and corresponding performance metrics (such as aerodynamic drag, weight, and stiffness) of specific bicycle rim models (e.g., Model A). This study utilizes a dataset generated by the Taiwan Bicycle Research Center (CHC) through Finite Element Method (FEM) simulations. The second component is textual data. To enhance the model's knowledge base, this study incorporates RAG techniques. The related training texts are primarily sourced from extensive online technical materials—such as design manuals, technical reports, and research articles—collected using the research capabilities of LLMs (e.g., OpenAI ChatGPT, XAI Grok, and Google Gemini). The third component consists of CAD file data. As one of the study's objectives is to enable automated rim CAD drawing generation, a standardized CAD dataset was constructed to train the AI model in learning the geometric features and drawing logic of rims, which also serves as a basis for verifying the generated outputs.

C. *Core AI Technologies*

The design platform developed in this study integrates several cutting-edge AI technologies to enable highly automated and intelligent design capabilities. At its core is the LLM, which is trained on deep learning architectures such as Transformers and excels in natural language understanding and generation. Within the system, the LLM plays a central role in semantic parsing, task reasoning, and user interaction for various intelligent agents. ML, a crucial subfield of AI, builds predictive models through data-driven algorithms; in this system, ML techniques are primarily applied within the SQL Agent, which uses trained regression models to predict parameters when they are unavailable in the database. Intelligent agents are autonomous computational entities capable of perceiving their environment, reasoning, and taking actions to achieve specific goals. The agents in this system combine the semantic capabilities of the LLM with operational tools (e.g., database query modules and CAD control APIs), allowing for multi-step reasoning and automation of complex tasks. The MCP is a system architecture designed for multi-module AI systems. It emphasizes modular encapsulation of various models and functional units, coordinated via a unified command protocol, which enhances the system's manageability and scalability by enabling dynamic integration and flexible execution of tools. Lastly, RAG combines information retrieval with generative modeling by first retrieving relevant documents or text fragments from external knowledge bases and then feeding them to the LLM as context for content generation and question answering. This significantly improves performance in knowledge intensive tasks in terms of accuracy, reliability, and interpretability.

The evaluation in this study was based on standard predictive performance metrics, including R^2, MSE, MAE, and RMSE, reported in the experimental section. A goal specification was also defined, e.g., "generate a rim CAD model with weight < 450 g and stiffness > 120 N/mm." This goal was decomposed into parameter retrieval, property estimation, and CAD generation, handled by the SQL, RAG, and CAD Agents. A goal was considered achieved once all sub-goals met predefined thresholds, with satisfactory levels verified against industry standards. In practice, the CHC industry partner confirmed that the generated results met their design requirements and were considered satisfactory, further validating the achievement criteria.

D. Construction Process and Details of Each Agent

The overall system architecture and workflow of this study are illustrated in Figure 2. After defining the goals and framework of the AI system, each agent was independently constructed and trained. The development of the RAG Agent began with the collection of a large corpus of domain specific texts related to bicycle rim design. These texts were preprocessed by extracting relevant content, segmenting it into appropriately sized chunks, and converting them into vector representations. Subsequently, a retrieval mechanism based on a Faiss vector database was established, and the RAG model was trained and fine-tuned to optimize both retrieval efficiency and generation quality. To guide the reasoning pathway and operational logic of the RAG Agent, carefully crafted prompts were designed. Finally, the trained RAG model and its supporting tools—such as the Faiss search engine and metadata filtering modules—were encapsulated as MCP modules and subjected to multiple rounds of testing and validation.

The construction process of the SQL Agent begins with cleaning and organizing the optimized parameters obtained through Finite Element Method (FEM) simulations and storing them in a structured database. The next step involves developing functionalities that enable the agent to automatically generate SQL query statements and execute database searches. Since the FEM dataset contains only around 500 data entries, ML is employed to predict values that fall outside this range or are missing from the database. To this end, a multi-output regression model is adopted and trained for prediction tasks. Prompts are also designed to guide the SQL Agent's query logic and the conditions under which the prediction model should be triggered. Finally, the SQL query engine, SQL code generator, and ML based prediction executor are encapsulated into an MCP module, and both the query results and prediction outputs are validated for accuracy and reasonableness.

The development of the CAD Agent begins with the preparation of standardized bicycle rim CAD files as training data. The agent is then trained to understand the drawing logic and geometric specifications of CAD models. Utilizing FreeCAD MCP Tools, which include a Python code generation module and a FreeCAD API controller — the agent sends generated Python drawing commands via API to the FreeCAD software for automated generation of 2D and 3D CAD models. During this process, prompts are also designed to guide the CAD Agent in translating input parameters into appropriate drawing instructions. All related functionalities are encapsulated into an MCP module, followed by tests to verify the accuracy and completeness of the generated models.

Once the three core agents are trained and packaged, they are integrated into the Main Agent for system level integration testing and validation. If the system passes the validation, it is ready for deployment; otherwise, the process returns to the corresponding agent's training phase for refinement and optimization.

E. Frontend User Interface (UI)

The frontend User Interface (UI) of the system is designed to provide a convenient and intuitive interaction experience and is implemented using CURSOR, as shown in Figure 3. Users can simply enter their design requirements or questions in natural language through the chat box on the right side of the interface (highlighted as Area 2 in red in Figure 3) to interact with the AI design system. The left side of the interface (Area 1 in red) displays the currently active agents and available MCP tools, enhancing the transparency of system operations.

IV. RESULTS AND DISCUSSION

To evaluate the effectiveness of the AI design system, we employed the latest large language model, Claude 4 Sonnet [11], as the core reasoning engine for each agent. A simulated question and answer scenario replicating a realistic design task was created to assess each agent's response capabilities, the efficiency of horizontal communication between agents, and the overall quality of information integration and final response by the Main Agent. An example test prompt is as follows: "I would like to design a bicycle rim with moderate aerodynamic drag, stiffness, and weight. Please recommend an optimal set of dimensional parameters, analyze the reasonableness of these parameters based on the design manual, then draw the 2D and 3D models in FreeCAD, and finally provide a summary and design recommendations." The following sections present the functionality and analysis of each agent:

A. SQL Agent Response

In response to the user's query, "I am currently looking to design a bicycle rim with moderate aerodynamic drag, moderate stiffness, and moderate weight. Please recommend an optimal set of size parameters," the SQL Agent first queried the internal database. As shown in the red box in Figure 4, the agent recommended the following optimal parameters: outer width of 34 mm, inner width of 19 mm, and total length/depth of 79 mm. An initial assessment suggests that this response is reasonable — for instance, greater aerodynamic requirements often correspond to increased rim depth. However, the actual validity of these parameters requires further analysis by the RAG Agent.

B. RAG Agent Response

Following the parameters provided by the SQL Agent, the user then asked, "Please analyze the reasonableness of this set of parameters." The response from the RAG Agent, as shown in the red box in Figure 5, indicates that it successfully communicated horizontally with the SQL Agent, retrieved the parameters, and performed an analysis based on the embedded knowledge from the bicycle rim design manual. In this analysis, the RAG Agent offered suggestions regarding the feasibility of each dimension and the types of bicycles for which the parameters would be appropriate. Due to space limitations in the paper, only a portion of the key content is presented here. The full analysis includes multiple aspects, and users can further inquire about how to refine the current parameters for improved design suitability.

C. CAD Agent Response

This study trained the CAD Agent using 2D profile drawings, such as the example training data shown in Figure 6. After the SQL and RAG Agents confirm the design parameters, the CAD Agent receives instructions to perform modeling. As shown in the red box in Figure 7, the CAD Agent confirms the modeling dimensions—outer width of 34 mm, inner width of 19 mm, and total depth of 79 mm—consistent with the previously recommended values. Figure 8 highlights the 2D profile and 3D solid model successfully generated by the CAD Agent in FreeCAD, precisely matching the input parameters. These results demonstrate that the CAD Agent can execute instructions effectively, significantly reducing the time and manual effort required for CAD drafting by the user.

D. Main Agent Response

Figure 9 presents the response from the Main Agent of the AI-based bicycle rim design system. It consolidates the processes and analyses of all the aforementioned agents and provides concrete design recommendations. This system enables users to efficiently achieve low volume, high variation design objectives. Moreover, the agents are equipped with real time online search capabilities, enhancing their domain knowledge and adaptability.

E. ML Multi-Output Regression Model Training Result

Figure 10 displays the training results of the ML multi-output regression model. In this study, a parameter prediction model was trained using the following variables: bicycle rim cross section outer width (mm), inner width (mm), total length (mm), FEM simulation rigidity (N/m), FEM simulation weight (g), and FEM simulation wind resistance (gf). Performance metrics were used as input features, while dimensional parameters served as outputs. This allows the model to predict a broader range of parameter combinations based on user requirements, rather than being limited to the 500 FEM simulated samples. Prior to training, the linear relationships between dimensions and performance indicators were verified. As shown in Figure 10, the model performed well overall. However, one parameter exhibited signs of overfitting, with an MSE of 0 and an R^2 of 1, which occurred because the parameter in question was a constant value.

F. ML Multi-Output Regression Model Training Result

This study adopted Claude 4 [12] as the primary LLM due to its stable reasoning performance, consistent response quality, and efficient integration with external tools. Compared with alternatives such as ChatGPT or Gemini, Claude 4 delivered more reliable outcomes for multi-agent collaboration and CAD automation tasks. For real-world deployment, the proposed agentic framework is technically feasible, as its modular, MCP-based design enables seamless integration with existing CAD and database systems. Nonetheless, additional factors including data security, system latency, and compatibility with large-scale production workflows must be addressed. This collaboration with CHC indicated that the system outputs aligned with industry expectations, suggesting strong potential for adoption in actual production environments.

V. Conclusion and Future Work

This study successfully developed an intelligent CAD automation design platform that integrates LLMs, RAG, a Multi-Agent System, and the MCP, specifically aimed at optimizing bicycle rim parameters and reengineering the traditional design workflow. By incorporating the SQL Agent for parameter querying and prediction, the RAG Agent for domain knowledge enhancement, and the CAD Agent for 2D and 3D modeling, the system effectively addresses the high time and labor costs associated with manual data collection and repetitive drafting in conventional design processes. Experimental results

 ISBN: 978-1-68558-293-7

demonstrate that the platform can reduce design time from several weeks to just a few hours, significantly improving both efficiency and the accuracy of design recommendations.

Looking ahead, the intelligent design platform proposed in this study holds significant potential for further development and expansion. First, its current capabilities in rim design can be gradually extended to other critical bicycle components, such as frames and forks, ultimately enabling intelligent design assistance at the full vehicle level. Second, the integration of materials databases and related analytical modules is envisioned, allowing the system to consider the impact of material selection on performance and cost during the early stages of design. Third, deeper integration with advanced Computer Aided Engineering software could automate more complex simulations involving structural mechanics, aerodynamics, and beyond. Fourth, continuous improvements to the user interface and interactive experience are necessary to develop more intuitive and intelligent human machine collaboration modes, thereby lowering the barrier to use. The ultimate goal is to explore higher level automated product development paradigms, applying AI to conceptual design, engineering design, manufacturing planning, and even product lifecycle management—further advancing smart manufacturing and digital transformation.

ACKNOWLEDGMENT

This work was supported by the National Science and Technology Council of Taiwan under Grant NSTC 113-2221-E-035-042. Additional support was provided by the Cycling & Health Tech Industry R&D Center (CHC) and an unrestricted gift from Google.org. The authors gratefully acknowledge this support. Translation from the original language into English has been aided by an automatic tool.

REFERENCES

[1] I. L. Alberts, L. Mercolli, T. Pyka, G. Prenosil, K. Shi, A. Rominger, et al., "Large language models (LLM) and ChatGPT: what will the impact on nuclear medicine be?," Eur. J. Nucl. Med. Mol. Imaging, vol. 50, no. 6, pp. 1549–1552, 2023.

[2] C. Y. Kim, C. P. Lee, and B. Mutlu, "Understanding large language model (LLM) powered human robot interaction," in Proc. 2024 ACM/IEEE Int. Conf. Human Robot Interaction, 2024, pp. 371-380.

[3] J. Huang, S. S. Gu, L. Hou, Y. Wu, X. Wang, H. Yu, et al., "Large language models can self improve," arXiv preprint arXiv:2210.11610, 2022.

[4] H. Lyu, S. Jiang, H. Zeng, Y. Xia, Q. Wang, S. Zhang, et al., "LLM Rec: Personalized recommendation via prompting large language models," arXiv preprint arXiv:2307.15780, 2023.

[5] L. Regenwetter, B. Curry, and F. Ahmed, "BIKED: A dataset and machine learning benchmarks for data driven bicycle design," in Proc. Int. Des. Eng. Tech. Conf. Comput. Inf. Eng. Conf., vol. 85383, 2021, p. V03AT03A019.

[6] L. Regenwetter, B. Curry, and F. Ahmed, "BIKED: A dataset for computational bicycle design with machine learning benchmarks," J. Mech. Des., vol. 144, no. 3, p. 031706, 2022.

[7] Z. Hong, Z. Yuan, H. Chen, Q. Zhang, F. Huang, and X. Huang, "Knowledge to SQL: Enhancing SQL generation with data expert LLM," arXiv preprint arXiv:2402.11517, 2024.

[8] Z. Hong, Z. Yuan, Q. Zhang, H. Chen, J. Dong, F. Huang, and X. Huang, "Next generation database interfaces: A survey of LLM based text to SQL," arXiv preprint arXiv:2406.08426, 2024.

[9] T. Zhang, C. Chen, C. Liao, J. Wang, X. Zhao, H. Yu, et al., "SQLfuse: Enhancing text to SQL performance through comprehensive LLM synergy," arXiv preprint arXiv:2407.14568, 2024.

[10] G. M. Coelho, E. R. Nascimento, Y. T. Izquierdo, G. M. García, L. Feijó, M. Lemos, et al., "Improving the accuracy of text to SQL tools based on large language models for real world relational databases," in Int. Conf. Database Expert Syst. Appl., Cham, Switzerland: Springer, 2024, pp. 93–107.

[11] Anthropic, Claude Opus 4 & Claude Sonnet 4 System Card, May 2025. [Online]. Available:https://www.anthropic.com/claude-4-system-card

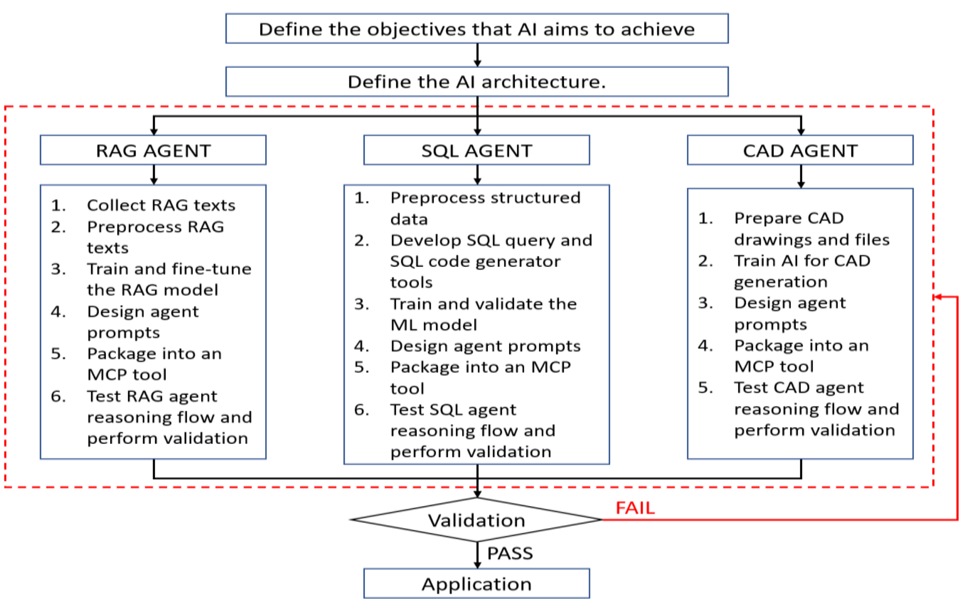

Figure 1. Research Architecture and Workflow.

 ISBN: 978-1-68558-293-7

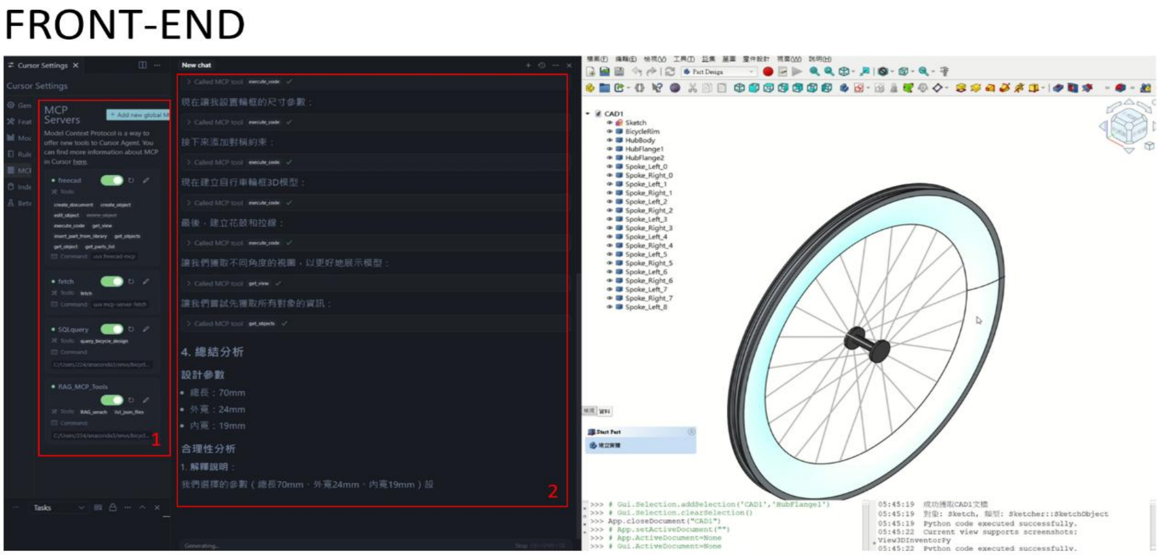

Figure 2. Frontend UI of the AI-Based Bicycle Rim Design System.

Figure 3. SQL Agent Response.

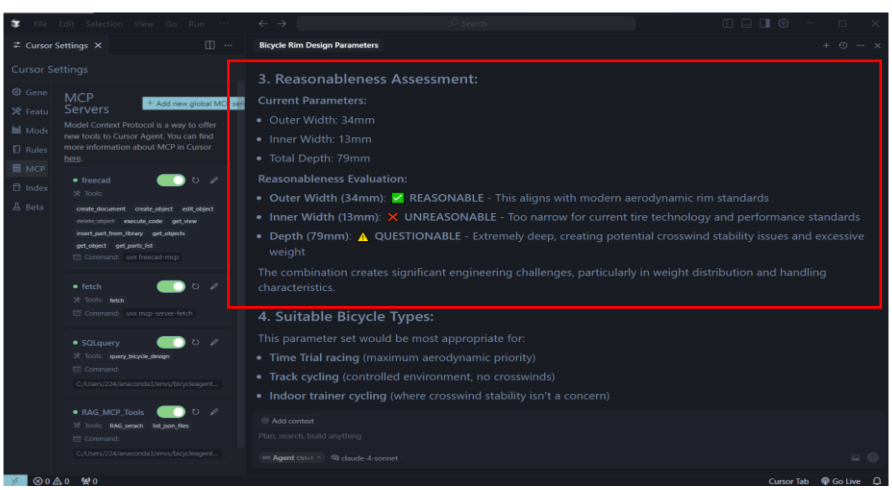

Figure 4. RAG Agent Response.

Figure 5. 2D Drawing of CAD Training Data.

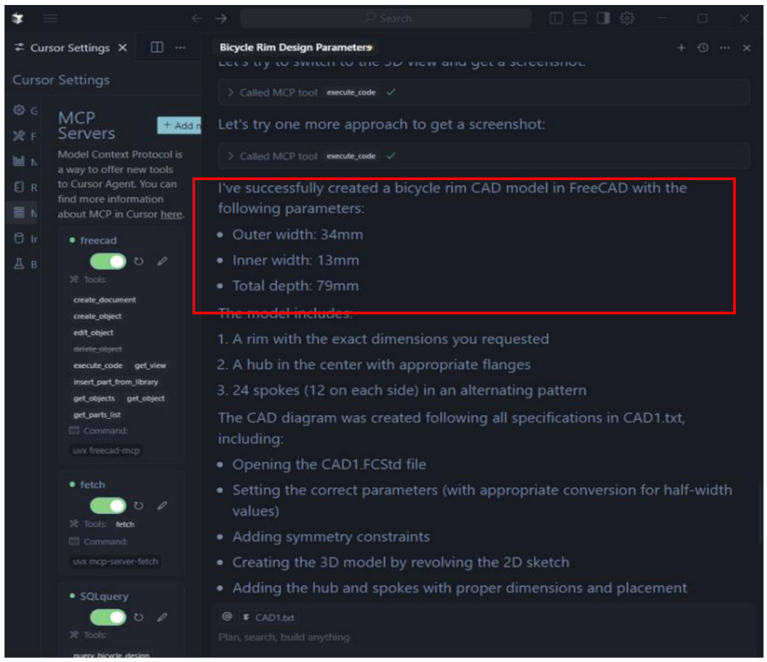

Figure 6. CAD Agent Response.

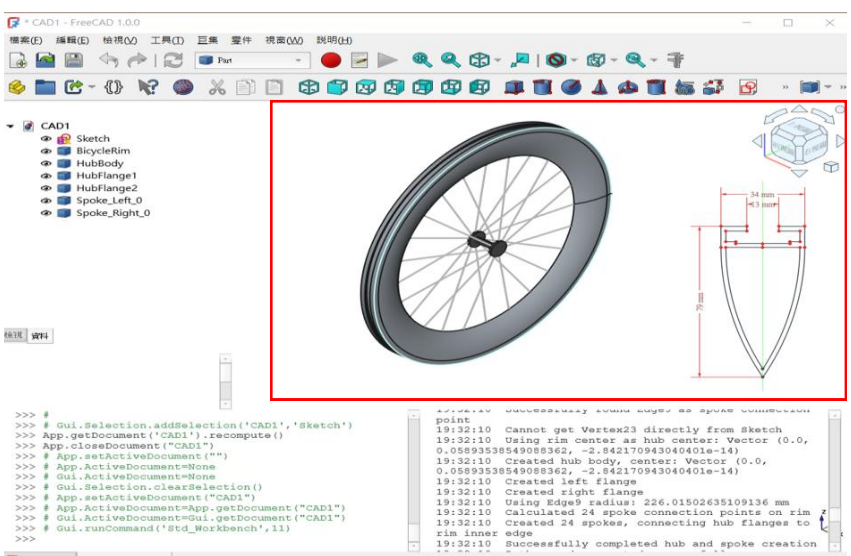

Figure 7. CAD Agent 2D and 3D Modeling.

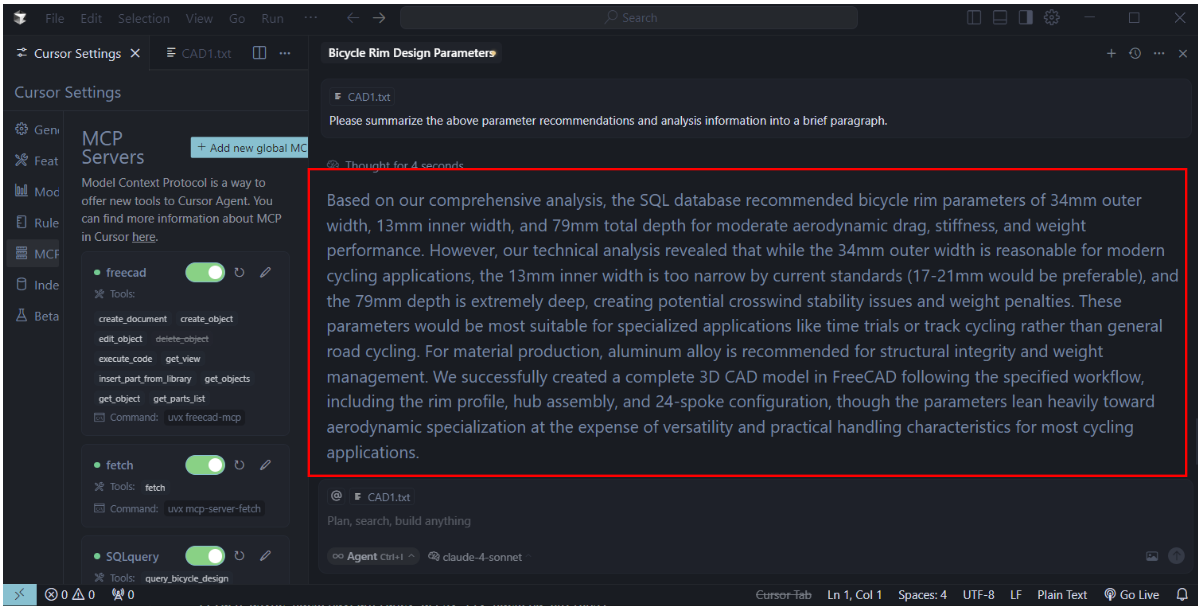

Figure 8. Main Agent Response of the AI-Based Bicycle Rim Design System.

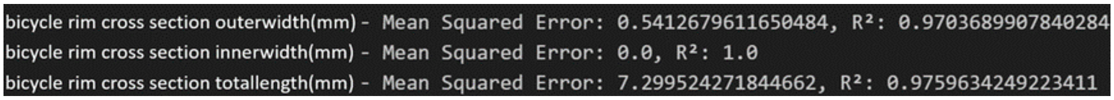

Figure 9. Training Results of the ML Multi-Output Regression Model.