ABSTRACT

The evolution of materials science has firmly entered its fourth paradigm, characterized by the data-driven methodologies of Material Informatics. Existing models, particularly those predicting wide-application fatigue strength in steels, frequently rely on restricted data covering limited alloy categories and processing parameters. This reliance impedes the development of generalized frameworks capable of driving automated alloy discovery. Furthermore, a significant research gap exists regarding the predictive modelling of complex, specialized material behaviours. For example, despite their critical utility in extreme environments like space and nuclear engineering, copper alloys often suffer from intermediate temperature embrittlement due to creep cavitation, liquid-metal, or hydrogen exposure. Yet, robust data-driven models for predicting their thermal properties remain largely unexplored. While the integration of machine learning has successfully accelerated single-target property predictions, the broader deployment of these tools in materials design remains heavily constrained by the narrow scope of available datasets.

Additionally, the design of novel materials inherently requires balancing competing mechanical properties. While multi-target regression (MTR) presents an effective solution for simultaneous property optimization, current MTR studies are yet to be explore various techniques developed for MTR. Finally, while Large Language Models (LLMs) offer immense potential for automating material discovery, standalone LLMs are hindered by inherent limitations, including knowledge hallucinations, and deficiencies in numerical reasoning.

To overcome these fundamental challenges, this thesis aims to develop comprehensive machine learning frameworks for generalized property prediction and integrates them within a tool-augmented AI agentic system. The research systematically expands the scope of fatigue strength prediction by utilizing large, diverse datasets encompassing both steels and aluminum alloys. Furthermore, it pioneers the data-driven prediction of thermal shock resistance in copper

alloys, supported by extensive feature analysis. To address complex property trade-offs, advanced MTR models are developed to simultaneously predict six mechanical properties in steels, followed by an expanded MTR framework predicting seven distinct properties across iron, aluminum, and copper alloys.

Culminating this predictive work, the thesis presents a dynamic agentic framework that bridges the gap between sophisticated regression models and accessible AI reasoning. By augmenting LLMs with direct access to the developed predictive tools, materials databases, and real-time retrieval capabilities, the framework effectively mitigates standard LLM deficiencies. Ultimately, this research provides a robust, generalized ecosystem to accelerate automated alloy search and novel materials design.