

HAO LI

hl674@cam.ac.uk | +44 7876239581 | [Google Scholar](#)

EDUCATION

University of Cambridge

Cambridge, UK | Expected 2027

- **Doctor of Philosophy** in Mechanics, Materials and Design;
- Research interests: focus on developing mathematically grounded machine learning models to solve complex engineering problems

University of Washington

Seattle, WA | Sep 2022 - Sep 2023

- **Master of Science** in Applied & Computational Mathematics; GPA: 3.96 / 4.00
- Courses: Scientific Computing; Computational Methods for Data Analysis; High-Performance Computing; Advanced Differential Equations; Probability and Random Process; Machine Learning for Big Data

Knox College

Galesburg, IL | Aug 2017 - June 2021

- **Bachelor of Science** in Mathematics; GPA: 3.83 / 4.00
- **Bachelor of Arts** in Physics; GPA: 3.64 / 4.00
- Honors: Magna Cum Laude, Dean's List, Victoria Legner Junod Prize in Mathematics

RESEARCH EXPERIENCE

Post-baccalaureate Scholar ([GitHub repo](#)) | Knox College

Galesburg, IL | Sep 2021 – May 2022

A service project for the college, assisting faculty with the development of course materials and research projects for the new Data Science major.

- **Neural Network Exploration:** Undertook an exhaustive study of recurrent neural networks (RNNs) and LSTMs, analyzing their architecture, decoding gradient-related anomalies, and devising strategies to address challenges tied to long-term dependencies.
- **Database Design & Curation:** Approached the textbook database creation with a researcher's lens, methodically selecting, categorizing, and processing literature from Project Gutenberg. This strategy ensured the database's potential as a foundational tool for sophisticated data science investigations in the future.
- **Language Classification:** Beyond mere development, the neural network model underwent iterative testing, refinement, and validation to ensure resilient language classification. This rigorous process involved hyperparameter adjustments, training set refinements, and a close examination of misclassifications to enhance model reliability.
- **Deep Dives into Advanced Models:** An independent study on Seq2Seq and GANs rooted in demystifying the intricate mathematical principles behind these models. This research dissected complex theories like gradient descent, transportation theory, and probability distributions, contemplating their implications for practical applications.

Document Classification with OOD detection | University of Washington

Seattle, WA | Jan 2023 – Mar 2023

Worked on a team project to enhance document classification with pre-trained Transformers by incorporating multi-modal information and optimizing with PyTorch Lightning.

- **OOD Detection Framework:** To tackle the Out-of-Distribution (OOD) challenges prevalent in document classification, a specialized framework was devised. This framework targeted the model's ability to robustly identify OOD instances while preserving efficacy in primary classification tasks.
- **Multi-modal Data Integration & Validation:** Delving deeper into the realms of multi-modal integration, rigorous experiments were conducted: Incorporating text-layout dynamics through the LayoutLM transformer. Introducing a synergized image-layout-text approach by amalgamating LayoutLM with a ResNet-based feature extractor.
- **Computational Efficiency & Proof-of-Concept:** Following the successful validation of the aforementioned strategies, emphasis shifted to ensuring the model's computational scalability. The architecture was re-engineered using PyTorch Lightning, reflecting both its promise and operational efficiency.

EXTRACURRICULAR EXPERIENCE

Short-term Visiting Scholar

Chongqing, China | July 2024 – Aug 2024

- Organized technical exchange meetings at Chongqing Fine Chemical Industry Limited Company
- Assisted the technical team in developing data-driven models to predict the outcomes of a series of chemical reactions

Independent Research Assistantship

Sichuan, China | March 2024 – Sep 2024

- Conducted machine learning research under the supervision of Prof. Zhang at Sichuan University
- Delivered research presentations and organized academic seminars to facilitate discussions among scholars on various research ideas

Chinese Club of Knox College | *treasurer*

Galesburg, IL | Sep 2020 – June 2021

- Took charge of the application and allocation of funds in Chinese Club
- Organized several school events such as: Mid-Autumn Festival, Spring Festival, and Asian Student Association Formal

On campus employment at Knox College | *grader of the Math Department*

Galesburg, IL | Jan 2020 – Mar 2020

- Took charge of grading students' homework from Calculus I, II, III
- Assisted students with homework questions

PREPRINTS & PRESENTATION

Fan, T., Li, H., Paris-Moe, W., & Xing, E. (2023). Contrastive Out-of-Distribution with Multi-modal Information for Document Classification. ResearchGate. Preprint. <https://doi.org/10.13140/RG.2.2.35271.06561>

Li, H. (2023). Proofs for the Four Fundamental Equations of the Backpropagation and Algorithms in Feedforward Neural Networks. ResearchGate. Preprint. <https://doi.org/10.13140/RG.2.2.24365.87524>

Li, H., Paris-Moe, W., Hoeger, M., & Sayani Boral. (2023). Sentience Sentiments Concept Visualization A Glimpse into AI Perceptions Over 30 Years. ResearchGate. Preprint. <https://doi.org/10.13140/RG.2.2.19005.03044>

SELECTED HONORS

Dean's List	2018, 2019
The Victoria Legner Junod Prize in Mathematics	2021
The Donald L. Benedict Student Research Award	2020

SKILLS

- **Programming:** Python (NumPy, Pandas, Keras, PyTorch, PySpark), C/C++ (Multithreading, MPI, CUDA, Distributed Memory Computing), SQL, (R/R Studio, Java) (entry level).
- **Software and Tools:** Mathematica, MATLAB, Microsoft Office Suite, Jupyter Notebook, Tableau, TensorFlow.