Wang Haocheng

Mobile: +8618339942921 | Email: haocheng.wang@inf.ethz.ch | Github homepage: https://github.com/hcWang942

Add.: No.9 Nongke Road, Jinshui District, Zhengzhou, Henan Province, China

EDUCATIONAL EXPERIENCES

Xiamen University Sep. 2020 - Jun. 2025

Major: Mathematics and Applied Mathematics (Honours)

Degree: Bachelor of Science

GPA: 3.26/4.00

Relevant Modules: Mathematical Analysis, Complex Analysis, Linear Algebra, Abstract Algebra, Mathematical Theory of Games, Mathematical Graphics, ODE, Financial Mathematics, Python, C++, Statistics, Stochastic Process, Time Series

PROFESSIONAL EXPERIENCES

ETH Zürich Zürich, Switzerland

Scientific Assistant I - D-INFK (Supervisor: Rasmus Kyng) *Aug*. 2025 -

- Focusing on developing Multi-gLLM Agent for autoformalizing and formal reasoning in Theoretical Computer Science.
- Teaching Assistant at course *263-4512-00L Formalizing Analysis of Algorithms.*

ByteDance Co., Ltd. Shanghai, China

LLM & RL Research Intern - ByteDance Seed Jan. 2025 - Jul. 2025

- Focusing on developing LLMs for formal mathematical reasoning through Supervised Fine-Tuning approaches.
- Built a Scoring & Self-refinement agent pipeline for Natural Language Proof, generating reliable informal proofs to provide informal-formal data pairs for model training and enhancing the reasoning capabilities.
- Propose an innovative sketch-incorporated long Chain-of-Thought formal reasoning method. Designed data annotation template, completed pilot annotation protocols, and led quality control efforts for cold-start data annotation initiatives.
- Conducted quality reviews of multiple mainstream mathematical reasoning benchmarks including minif2f, FIMO, and Putnam.

DeepSeek AI Co., Ltd. Beijing, China

AGI Research Assistant June. 2024 - Sep. 2024

- Built a Multimodal agent *LeanAgent* for Lean 4 autoformalization tasks, facilitating data preparation and development for the Supervised Fine-Tuning of DeepSeek-Prover-V1.5 & DeepSeek-Prover-V2 models.
- Developed *ProverBench*, a domain-categorized benchmark for evaluate LLM in automatic theorem proving, sized 325 for across college-level and high-school competition level mathematics.
- Manually annotated the benchmark of Lean 4 code as demonstration data, including MINIF2F, FIMO, and IMO problems. Enhance the performance of LLMs to do recursive theorem proving.
- Adjusting truncate-point positions to monitor the performance of diversified informal proofs by using a single tactic.
- Relative results:

  arXiv Preprint: 2504.21801: ***DeepSeek-Prover-V2: Advancing Formal Mathematical Reasoning via Reinforcement Learning. for Subgoal Decomposition*** **[https://arxiv.org/abs/2504.21801]**

  arXiv Preprint: 2501.12948: ***DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning*** **[https://arxiv.org/abs/2501.12948]**

  arXiv Preprint: 2412.19437: ***DeepSeek-V3 Technical Report*** **[https://arxiv.org/abs/2412.19437]**

  arXiv Preprint: 2408.08152: *DeepSeek-Prover-V1.5: Harnessing Proof Assistant Feedback for Reinforcement Learning and Monte-Carlo Tree Search* [https://doi.org/10.48550/arXiv.2408.08152]

Huabao Securities Co., Ltd. Shanghai, China

Research Assistant Jan. 2024 – Apr. 2024

- Developed a Quantitative Fund Selection and Back-testing System using Python, capable of autonomously filtering through a vast array of funds to identify those meeting specific criteria. Performs historical back-tests to evaluate the effectiveness.
- Implemented and compared nine regression models in Python for quantitative position estimation, encompassing WLS and OLS methods, windowed LASSO, and ridge regression. Enhanced precision by integrating *optimize* for matrix-based optimization
- Systematically analyzing the portfolio rebalancing and trend analyses by quantitative strategies for to micro-stock collapses.

Wind Information Co., Ltd. Shanghai, China

Product Manager Aug. 2023 - Sep. 2023

- Utilized Wind economic database (WindEDB) to perform statistical analysis and visualization of A-share listed companies with restricted share reduction by the China Securities Regulatory Commission.
- Designed and proposed *Wind Mobile* interface and optimized UI/UX interaction for WindESG and Wind Global Asset using Axure RP. Collaborated with the development team to implement and update the app's user interface.
- Conducted competitive analysis of similar financial data products, proposed optimizations for visual enhancement and AI plugins.

Beijing Shiguang Study Culture Media Co., Ltd. Beijing, China

Intern Lecturer Mar. 2022 - May 2022

- Served as the lecturer for 48 high-school mathematics tutorial videos, totaling over 960 minutes.
- Contributed to the improvement of course content, revision of exercise answers, and gathering of user feedback.

RESEARCH EXPERIENCE

DeepSeek-Prover-V2: Harnessing Proof Assistant Feedback for Reinforcement Learning and Monte-Carlo Tree Search,

*Under review as a conference paper at **ICLR 2025** (Finished Project)* June 2023 - July 2023

- Achieved SOTA results on miniF2F (88.9%) and PutnamBench (47 out of 658), surpassing previous by significant margins.
- Proposed recursive theorem proving pipeline powered by DeepSeek-V3.
- Releasing ProverBench, the benchmark for evaluating automated reasoning.

Formalization Auction Theory using Lean and Mathlib4, *(Final thesis)* July 2023 - Aug.2024

Department of Mathematics, Xiamen University Malaysia Advisor: Dr. Ma Jiajun

- Formalize Auction Theory in Game Theory in Lean4. [GameTheoryRepo]
- Contributing the formalized proof of Second price auction DSIC and Myerson's lemma to Mathlib4. [#PR13248]

DeepSeek-Prover-V1.5: Harnessing Proof Assistant Feedback for Reinforcement Learning and Monte-Carlo Tree Search,

*Under review as a conference paper at **ICLR 2025** (Finished Project)* June 2023 - July 2023

- Developed a novel hybrid approach combining LLMs and Monte-Carlo tree search for automated theorem proving.
- Designed a truncate-and-resume mechanism for proof search, integrating single-pass and multi-pass generation strategies.
- Implemented reinforcement learning from proof assistant feedback (RLPAF) using GRPO algorithm to enhance performance.
- Achieved SOTA results on benchmark miniF2F (63.5%) and ProofNet (25.3%), surpassing previous by significant margins.

Regression analysis and ESG Rating Methodological Frameworks in the Energy Sector, Nov. 2023 - Present

*Working for **Energy Economics** Special Issue (Ongoing project)* Advisor: Dr. Woon Kok Sin

- Developed advanced machine learning models to predict carbon emissions (Scope 1, 2, and cumulative) using:
  1. Statistical methods (Multiple Linear Regression, LASSO, Ridge Regression)
  2. Deep learning approaches (Stochastic Neural Networks)
  3. Ensemble techniques (Hyperparameter-optimized Random Forest, XGBoost)
- Performed quantitative analysis utilizing time series and regression techniques to investigate the correlation between Scope 1 and Scope 2 carbon emissions of energy sector listed companies and various socioeconomic factors.

EXTRACURRICULAR ACTIVITIES

Frontier of Formal Theorem Proving with Large Language Models *(Research Talk)* Dec. 2024

Xiamen University Malaysia

- Presented research project on LLMs for automated theorem proving and mathematical reasoning including training process of state-of-the-art models DeepSeek-Prover series.
- Explored autoformalization workflows and LeanAgent pipeline for building college-level formal benchmarks
- Discussed practical LLM-powered tools in education-level (Calculus Game from NUS, Game for K12 from SJTU), and research-level (LLMLean, LeanSearch).

AI for Mathematics Workshop on Lean: Automated Reasoning and Beyond Aug. 2024

BICMR, Peking University

- Present detailed plan to raise Pull Request for create new branch *Game Theory* in Mathlib, including topics: Auction Theory, Myerson's Lemma, Minimax Theorem, Nash Equilibrim and Brouwer Fixed-point Theorem. [#PR13248]

AI For Math: Formalization Mathematics Theories and Proofs June 2024 - Aug. 2024

BICMR, Peking University

- Assisted in teaching formalization using Lean coding exercises. Graded workshop exercises on Abstract Algebra.

Workshop on Formal Proofs and Lean Apr. 2024

Department of Mathematics, NUS

- Assist in formalization of Coxeter groups, proved Abstract Simplicial Complex and Abstract Simplicial Complex Shelling. [Ref.]

AI For Math: Formalization Mathematics Theories and Proofs Jan. 2024

BICMR, Peking University

- Presentation on formalization of theorems in Second-price and First-price Auction and Myerson's lemma.
- Formalize Myerson's lemma in Game Theory in Lean4.

Red Cross Society of China, *Member* July 2021 - Present

China Marrow Donor Program, *Member* July 2021 - Present

7.20 Zhengzhou Flood Relief Frontline Commando Team, *Vice Director* July 2021 - Aug. 2021

- [News-reports-XMUM-official-account]

Read Sparks Community Service Group, *Vice President* Apr. 2021 - July 2021

AWARDS

Intel Young Talent Award, 34th China Adolescents Science & Technology Innovation Contest Aug. 2019

Students' Projects Third Award, 34th China Adolescents Science & Technology Innovation Contest Aug. 2019

The First Place Award, 33rd Henan Province Science and Technology Innovation Competition Mar. 2019

The First Place of Grand Award, 3rd International Youth Science Fair for Discovery Videos Sep. 2018

"Scientific Kaleidoscope" Best Photography Award, 9th China Adolescents Science Video Festival Oct. 2018

First Prize for Scientific Inquiry Documentary, 9th China Adolescents Science Video Festival Oct. 2018

"Hantec" Best Cinematography Award, 2nd International Youth Science Fair for Discovery Videos Sep. 2017

The First Place of Grand Award, 2nd International Youth Science Fair for Discovery Videos Sep. 2017

SKILLS

Skills: Proficient in LEAN4, R , Python, C++, Matlab, Axure RP, WFT, Microsoft Office(PPT, Excel, Word).

Language Proficiency: Chinese Mandarin (Native), English (Fluent, IELTS: 6.5/ GRE: 318).