

Xiwen Huang

linkedin.com/in/huangxiwen/
github.com/xiwenhuang123

Email: xiwen.huang23@imperial.ac.uk | xiwenhuang123@gmail.com

Mobile: +44 7544 576516

London, United Kingdom

PhD researcher at Imperial College London specialising in statistical machine learning and sequential decision-making under uncertainty. My research develops data-driven algorithms using online learning, Thompson Sampling, combinatorial bandits and predictive modelling, with extensive experience in Python, large-scale simulation, forecasting, model evaluation and out-of-sample validation. Interested in applying statistical and machine-learning methods to quantitative research and systematic investment problems. Expected PhD completion: 2027.

PUBLICATIONS

- **How to Purchase Labels? A Cost-Effective Approach Using Active Learning Markets:** *INFORMS Journal on Data Science*. DOI: 10.1287/ijds.2025.0093; arXiv:2511.20605. Proposes a cost-aware active learning market framework for predictive modelling under label acquisition constraints. Validated through large-scale simulations and real-world forecasting case studies in real estate pricing and building energy forecasting.
- **QueryMarket: Cost-Aware Online Active Learning in Data Markets:** Manuscript under revision for resubmission to *IEEE Transactions on Neural Networks and Learning Systems*. arXiv:2606.17805. Studies online active learning and sequential data acquisition under budget constraints, with applications to dynamic data markets and solar power forecasting.
- **CATS: Combinatorial Allocation via Thompson Sampling for Multi-Buyer–Multi-Seller Data Markets:** Developing a combinatorial semi-bandit framework for sequential data allocation across multiple buyers and sellers, using Thompson Sampling under exclusivity and budget constraints. Accepted as a poster at the Women in Machine Learning (WiML) Workshop at NeurIPS 2026.
- **Ongoing Research – Sparse Optimisation for Cost-aware Data Markets:** Studying L0-based market mechanisms for cost-aware feature or data selection, with comparisons against Shapley-value-based allocation methods in forecasting tasks.
- **Recovering Efficient Supply Function Equilibrium with a Tax-Subsidy Mechanism:** Manuscript under major revision at *Production and Operations Management*. Studies strategic electricity market modelling, equilibrium computation, and tax-subsidy mechanism design.
- **Open-Source Research Code (xiwenhuang123):** github.com/xiwenhuang123. Implements active-learning algorithms, multi-agent data market simulations, equilibrium solvers, and reproducible experimental pipelines.

INVITED TALKS & CONFERENCE PRESENTATIONS

- **CATS: Combinatorial Allocation via Thompson Sampling for Multi-Buyer–Multi-Seller Data Markets:** Accepted poster, Women in Machine Learning (WiML) Workshop at NeurIPS 2026, 2026.
- **Markets for Data and Analytics Applications:** IFORS 2026, Vienna, Austria, 2026.
- **Generalised Data Markets: Learning, Agents, and Applications:** Invited seminar talk, Johns Hopkins University, EHE Seminar, 2026.
- **Active Learning Markets for Forecasting:** Best Student Presentation Award, 45th International Symposium on Forecasting (ISF), Beijing, 2025.
- **Strategic Data Acquisition & Active Learning Markets:** Technical University of Munich (TUM), 2025; ENBIS Young Statisticians Session, Greece (Knowledge Fund Award), 2025.
- **Information Incentives & Strategic Data Use:** Invited Speaker, Tsinghua University, 2025.
- **Designing Efficient Data Markets:** Invited Speaker, Peking University, 2025.
- **Cost-Aware Active Learning in Dynamic Data Markets:** Invited Speaker, Shanghai Jiao Tong University (SJTU), 2025.

EDUCATION

- **Imperial College London** London, United Kingdom
PhD in Design Engineering (Data Markets & Algorithmic Economics) Mar 2024 – Dec 2027
 - **Supervisor:** Prof. Pierre Pinson, Imperial College London; Editor-in-Chief, *International Journal of Forecasting*.
 - **Scholarship:** ICL–TUM Fully Funded PhD Scholarship (£170k+, <8% acceptance).

- **Research:** Data markets, mechanism design, strategic ML, multi-agent learning, optimisation.
- **Teaching Assistant:** From Data to Product, Taught 240+ MSc students.
- **ECRI Student Partner:** Co-designed 2 career courses and interviewed 10+ PhD alumni; moderated PhD welcome Q&A for 200+ students.

- **Harbin Institute of Technology** Harbin, China
Master of Engineering 2020 – 2022
 - **Academic Standing:** Graduated Top 10%.
 - **International Academic Symposium:** Led with 145 PhD speakers across 5 sessions.
 - **Award – English Debating:** First Prize, English debating competition.
- **Xiangtan University** Xiangtan, China
Bachelor of Engineering 2016 – 2020
 - **Academic Standing:** Rank 1/325.
 - **Award – Mathematical Modelling:** Honorable Mention, ICM Mathematical Modelling Competition.
 - **Award – English Speech:** First Prize, English speech competition.

RESEARCH EXPERIENCE

- **Peking University** Beijing, China
Research Assistant 2022 – 2024
 - **Strategic Market Modelling (Nash Equilibrium):** Studied strategic behaviour in electricity markets by modelling agents' interactions as a game and computing Nash equilibria via best-response dynamics. Developed equilibrium-solving algorithms and large-scale simulation pipelines (MATLAB/Python) to analyse convergence and market outcomes.
 - **Coordination – International Seminars:** Coordinated 5 international seminars involving 30+ researchers worldwide.
- **Columbia University** New York, USA
Research Group Member 2019
 - **Fieldwork – Cross-Cultural research:** Conducted 100+ behavioural interviews and performed qualitative analysis for cross-cultural decision-making research.

AWARDS & HONOURS

- **ICL–TUM Fully Funded PhD Scholarship:** £170k+ funding; <8% acceptance.
- **Best Student Presentation Award:** ISF Conference.
- **ENBIS Knowledge Fund Award:** Invited speaker at ENBIS conference.

TECHNICAL SKILLS

- **Programming:** Python (NumPy, SciPy, scikit-learn, pandas), R, MATLAB, C, Fortran, Git
- **Machine Learning:** active learning, online learning, bandit-based modelling, Thompson Sampling prototypes, predictive modelling, time-series forecasting, statistical learning
- **Experimentation:** reproducible ML pipelines, ablation studies, Monte Carlo simulation, benchmarking, out-of-sample validation
- **Mathematics:** optimisation, game theory, stochastic modelling, equilibrium modelling