

HAN TANG

LinkedIn Profile: [Han Tang](#)

Github Page: [Hirah Tang](#)

PUBLICATIONS

Contextual Molecule Representation Learning from Chemical Reaction Knowledge Published in Arxiv, 2024

EDUCATION & TRAINING

Univrsity of Copenhagen, Copenhagen, Denmark *September 2023 - December 2026 (expected)*

PhD. in Machine Learning for Biology

Department of Computer Science

Relevant Subjects:

Probabilistic Machine Learning: A course covers concepts including Variational Inference (VI), Variational Auto Encoder (VAE), Diffusion Model, Sampling, Gaussian Process (GP), etc.

Technological University of Dublin, Dublin, Ireland

September 2018 - Febraury 2020

MSc. in Computing (Data Science)

School of Computing

Relevant Subjects:

Machine Learning (A⁻), Deep Learning (A), Data Mining (B⁺).

Dissertation:

[A Comparison Study on State-of-the-art Minority Class Data Oversampling Techniques for Imbalanced Learning.](#)

Dublin Institute of Technology, Kevin St., Dublin, Ireland

September 2017 - June 2018

Pre-master for MSc in computing.

Beijing University of Chemical Technology, Beijing, China

September 2013 - June 2017

BSc. in Applied Chemistry

Relevant Subjects:

Inorganic Chemistry, Organic Chemistry, Physical Chemistry, Structural Chemistry

Statistics and Probability, Linear Algebra, Multivariable calculus

Dissertation:

Study Factors Influencing the Layer Heights for Layered Double Hydroxides Using Computing Approaches.

(Published on The 13th National Conference of Quantum Chemistry (China) as the second author)

[Paper Link\(Chinese\)](#)

CAREER EXPERIENCE

Institute for AI Industry Research (AIR), Tsinghua University

December 2021 - Now

Research Engineer

- My work at the Institute for AI Industry Research (Tsinghua AIR) is composed of Artificial Intelligence for Small Molecule Drug Discovery and Protein Structure Prediction.

Technological University Dublin & AutoPlan

September 2020 - March 2021

Research Assistant

- Cooperate with AutoPlan as a research assistant to develop similarity ranking algorithms, which find precedents to assist architects' planning applications.
- Build Text classifiers from the development descriptions of planning applications. This classifier classifies the categories of planning applications. The category type becomes a feature for the similarity ranking algorithm.

Shanghai Data-turing Technology

July 2020 - September 2020

Data Scientist

- Design data mining systems to reduce the energy consumption of HVAC (Heating, Ventilation, and Air Conditioning) systems. Our main clients are Volkswagen manufacturing plants and data centres of Tencent.
- The data-based optimised control system saves €1500 of electric bills for each data centre per day.

Postal Savings Bank of China

July 2019 - September 2019

Data Mining Analyst Intern.

- [Github Link: Credit Risk Assessing \(Illustration\)](#)

State Key Laboratory of Chemical Resource Engineering

November 2015 - May 2017

Research Assistant

- [Github Link: Chemical Parameter Calculator](#)
- Review works of literature of chemical computational methods to calculate physical chemistry parameters for complex organic compounds.
- Develop programmes with user interfaces that embed chemical parameters calculating methods using Python.

LANGUAGES & SKILLS & INTERESTS

Technical Skills & Certificates

R, Python, [Java](#), [Machine Learning \(Coursera\)](#)
[Deep Learning \(Coursera\)](#), Graph Neural Networks
[Full list of certifications earned.](#)

Languages

English (Full professional proficiency) -
[Duolingo English Test: 130 \(Equivalent to 110 in TOEFL\)](#)
Mandarin (Native proficiency)

Interests

Football (Soccer), played in multiple weekend football amateur leagues.