# Yu-Shiang Huang

F09946004@ntu.edu.tw
https://www.linkedin.com/in/yu-shiang-huang-309108126/
3F., No. 2, Aly. 21, Ln. 501, Zhonghe Rd., Yonghe Dist., New Taipei City 234021, Taiwan (R.O.C.)
Nationality: Taiwan (R.O.C.)
Date of Birth: 04 November 1997

**Summary**

I am a Ph.D. candidate in Data Science at National Taiwan University. My research interests include retrieval-enhanced machine learning (REML), natural language processing (NLP), and recommender systems (RecSys), with a particular focus on their applications in the financial domain.

**Education**

Ph.D. **National Taiwan University** Since 2022
Data Science

M.S. **National Taiwan University** Admitted to the Ph.D. program in September 2022
Data Science

B.S. **National Taiwan University** 2020
Accounting

**Positions**

**Academia Sinica** Graduate Student Researcher
Taipei, Taiwan September 2020-present

- REML Research: Designed task formulation and an end-to-end retrieval-augmented pipeline to extract important signals from financial reports.
- RecSys Research: Designed experiments to validate the effectiveness of a new loss function compared to state-of-the-art (SOTA) recommendation methods.
- RecSys Industry Cooperation: Researched multi-behavior algorithms for banner recommendation systems and deployed the developed model in a real-world system with millions of users.
- NLP Research: Analyzed the characteristics of financial annual reports and designed a new algorithm to classify sentence relations based on these findings. Leveraging the results, we conducted transfer learning in a self-supervised setting, and our paper was accepted by ACL 2023.
- NLP Industry Cooperation: Fine-tuned BERT with a Siamese network and contrastive loss to obtain sentence embeddings in legal data, building a retrieval system aimed at finding the most relevant internal regulations.

**National Taiwan University** Head Teaching Assistant, CSIE5043 Machine Learning
Taipei, Taiwan September 2023-January 2024

- Teaching Assistance: Led a team of 10 to assist in lecturing Machine Learning at National Taiwan University.
- Project Design: Designed the final project and maintained the data streaming pipeline.
- Homework Validation: Validated the homework problem sets related to machine learning theory.

**National Taiwan University** Teaching Assistant, CSIE5043 Machine Learning
Taipei, Taiwan September 2021-January 2022

- Final Project Design: Created an in-class competition for a 300-student course by adding noise to an open dataset and testing competition metrics with various machine learning models.
- Assignment Validation and Consulting: Validated assignments to ensure accuracy in problem sets and provided answers to machine learning theory-related questions during TA hours.

**National Taiwan University** Teaching Assistant, GENEDU5010 Programming and Web Scraping
Taipei, Taiwan Febuary 2021-June 2021

Instructional Support: Possessed a background in designing programming assignments and offering instructional support, allowing me to contribute effectively to teaching endeavors.

**beBit Inc.** Part-time Quantitative Analyst
Taipei, Taiwan October 2020-March 2021

- Survey Data Analysis: Collaborated with a qualitative research team and applied statistical methods to verify business hypotheses and identify key factors in the customer experience journey.
- Methodology Establishment: Improved quantitative analysis methodology based on research papers and technical blogs, enhancing the persuasiveness of the final deliverable.

**BluePlanet Technology** Project Assistant
Taipei, Taiwan November 2019-Febuary 2021

- Product Mechanism Design: Surveyed literature related to accounting and regulation to support the product mechanism design, resulting in the successful launch of the product in January 2021.
- Communication: Bridged the gap between the finance research group and the product management team by facilitating communication and recording discussion details during meetings.

**National Taiwan University** Part-time Research Assistant
Taipei, Taiwan July 2020-September 2020

- Website Development at Department of Economics: Utilized Python to create a website for an economics experiment based on a research paper in experimental economics, catering to a participant pool of 100 individuals.

**Academia Sinica** Intern
Taipei, Taiwan July 2019-June 2020

- Publication: Completed and submitted the work "NFinBERT: A Number-Aware Language Model for Financial Disclosures" to SwissText 2021, which has been accepted.
- Textual Data Analysis: Implemented machine learning models to conduct experiments on the relationship between financial report text and financial measures.
- Experimental Data Preparation: Collected SEC 10-K filings from 1993 to 2018 and performed data cleansing on the corpus through shell scripting, making the 330 GB data available for further research.

**National Taiwan University** Teaching Material Assistant
Taipei, Taiwan November 2018-January 2019

- Teaching Material Creation at Department of Electronic Engineering: Self-studied advanced topics about kernel tricks in Support Vector Machines (SVM) in machine learning. Developed teaching slides with explainable examples beyond textbooks, aimed at making the theorems easier to understand.

## Publications

### Under Review

Yu-Shiang Huang, Shih-Yang Liu, Cheng-Wei Lin, and Chuan-Ju Wang. DAGCF: Distance-aware Graph Collaborative Filtering. 2024

### Conference Papers

Jia-Huei Ju, Yu-Shiang Huang, Cheng-Wei Lin, Che Lin, and Chuan-Ju Wang. A compare-and-contrast multistage pipeline for uncovering financial signals in financial reports. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 14307–14321, 2023

Ta-Wei Huang, Jia-Huei Ju, Yu-Shiang Huang, Cheng-Wei Lin, Yi-Shyuan Chiang, and Chuan-Ju Wang. FISH: A Financial Interactive System for Signal Highlighting. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, pages 50–56, 2023

Hao-Lun Lin, Jr-Shian Wu, Yu-Shiang Huang, Ming-Feng Tsai, and Chuan-Ju Wang. NFinBERT: A Number-Aware Language Model for Financial Disclosures (short paper). In *SwissText*, 2021

## Recognition

### Awards

Legal Compliance Special Award, Lawsnote Legal-Tech Hackathon, 2021

## Service

### Organizational Leadership

ccClub (a Python Community), Director and Teaching Assitant, 2018-2024

## Skills

- Extensive programming experience with Python, particularly for machine learning projects on Unix platforms.
- Proficient in machine learning related packages such as PyTorch, Transformers, Sentence-BERT, LangChain, Scikit-learn, NumPy, Pandas and Matplotlib.
- Fluent in English and native proficiency in Traditional Chinese.

## References

*Available on request.*