

# TZU-HSIN YANG

 <https://www.linkedin.com/in/tzuhsinyang/>

 <https://github.com/ZixinYang>

 [zixinyang92@gmail.com](mailto:zixinyang92@gmail.com)

 <https://zixinyang.github.io/>

## Research Interests

---

Bayesian Inference, Deep Probabilistic Models, Uncertainty Estimation in Deep Learning, Reinforcement Learning

## Education

---

**National Cheng Kung University, Tainan, Taiwan**

Jul. 2018

- Master of Science in Computer and Communication Engineering
- Overall GPA: 4.0/4.0

**National Chiao Tung University, Hsinchu, Taiwan**

Jun. 2016

- Bachelor of Science in Electrical and Computer Engineering
- Last 60 GPA: 3.27/4.0

## Work Experience

---

**Research Assistant, Academia Sinica, Taipei, Taiwan**

Oct. 2020 – Present

- Developing graph visualization system
- Developing insurance fraud detection system with network data

**Data Scientist, KKBOX Inc., Taipei, Taiwan**

Jun. 2019 – Aug. 2020

- User Behavior Analysis
  - Developed **XGBoost** and **LightGBM** to predict churn users
  - Discovered key factors of churn behavior with **SHAP**
  - Developed **ARIMA** models to predict number of transactions
  - Modeled user journeys on app as a feature of users via semantic embedding with **fastText**
  - Built **Redash** dashboards to develop insights and make data-driven decisions with **PrestoSQL**
- Music Personalized Recommendations
  - Developed Spark ML Pipelines for playlists recommendation
  - Diversified playlists to increase user retention by excluding the streams caused by recommendations
  - Designed metrics of AB testing including average stream, retention and stream growth
- Public Opinion System
  - Developed a crawler pipeline for regular update of articles using GitLab and Jenkins
  - Implemented a NER system with Chinese NLP tools (**CkipTagger**)

**Deep Learning Scientist and Bioinformatician, Insilico Medicine Inc., Taipei, Taiwan**

Aug. 2018 – May. 2019

- Molecules Generation: implemented conditional generative models for generation given properties
- MRI Brain Image Analysis: implemented **Unet** model to segment images

**iOS Mobile App Developer, National Cheng Kung University Library, Tainan, Taiwan**

Aug. 2017 – Jun. 2018

- Developed a mobile library app including circulation services for twenty thousand students and faculty

**Teaching Assistant, Department of Electrical Engineering, NCKU, Tainan, Taiwan**

Sep. 2016 – Jun. 2017

- Teaching assistant for CS101 (Introduction to Computers) (C++)
- Won Teaching Assistant Awards for both fall and spring semesters

## Publications

---

**“(First author) DNA: General Deterministic Network Adaptive Framework for Multi-Round Multi-Party Influence Maximization”**, *accepted paper in IEEE International Conference on Data Science and Advanced Analytics (acceptance rate: 20%)* Oct. 2018

- Developed a *MCTS*-based framework which adapts to network change in the long run
- Developed influence maximization algorithms from game theory perspectives
- Designed a network similarity estimation method for new data prediction

**“(Third author) LSTMEnsembler: A LSTM-based Ensemble Framework for Predicting the Success of Mediation Requests Using Case Properties and Textual Information”**, *submitted to ACM Digital Government: Research and Practice* Oct. 2020

- Conducted the first research on predicting the success of real-world mediation cases
- Developed a *LSTM*-based framework based on the case information and textual descriptions
- Implemented a system for public servants and the public to decide whether to enter mediation process

## Projects

---

**(Kaggle) COVID19 Global Forecasting**, *held by The White House OSTP* Mar. 2020

- Forecasted confirmed cases and fatalities between March 25 and April 22 by region
  - Developed a vector autoregressive moving average model (*VARIMA*) to predict regional values
  - Ranked in top 13% in the competition

**(Kaggle) MolHack: Apply deep learning to speedup drug validation**, *held by Insilico Medicine Inc.* May. 2018

- Predicted the stability of the complex given ligand-pharmacophore pairs
  - Developed a regressor based on deep neural network on well-preprocessed data
  - Won 2<sup>nd</sup> place in the competition

**(Course Project) Energy Consumption Analysis and Prediction for Household Planning**, *NCKU* Jan. 2017

- Designed an algorithm to predict a household electricity consumption
  - Selected important features with random forest and predicted with linear regression models

**(Course Project) Mining Geo-Social Services for Optimal Location Placement**, *NCKU* Nov. 2016

- Designed an algorithm to rank top locations for hotels and theaters placement
  - Used hill climbing optimization algorithm with *NDCG* ranking score

**(Course Project) Social Relationship inference from Urban Footprint**, *NCKU* Oct. 2016

- Designed an algorithm to predict whether people are friends on social media with users' check-in data
  - Selected important features and used cosine similarity to measure the similarity between pairs

## Skills

---

### Programming Languages

- Python, C++, R, Scala, SQL, HTML, CSS, JavaScript, Matlab, Swift, MongoDB

### Certificate

- **TOEFL iBT Scores:** 107 (R29, L27, S23, W28)
- **Microsoft Certified:** Azure Fundamentals

### Relevant Coursework

- **University courses:** Linear Algebra, Differential Equation, Probability, Intelligent Data Analysis, Internet of Things and Urban Computing, Multilingual and Crosslingual Information System, Data Mining and Social Network Analysis
- **Online courses:** ML, DS and DL with Python (Udemy), Bayesian Inference (Coursera)