


# TZU-HSIN YANG

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

 <https://github.com/ZixinYang>

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

 (886) 972619777

## Research Interests

---

Deep Learning, Machine Learning, Reinforcement Learning, Game Theory, Causal Inference, Computer Vision

## Education

---

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

Jul. 2018

*M.Sc in Computer and Communication Engineering, Supervised by Jen-Wei Huang*

*Overall GPA: 4/4.3*

**Thesis:** DNA: General Deterministic Network Adaptive Framework for Multi-Round Multi-Party Influence Maximization.

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

Jun. 2016

*B.Sc in Electrical and Computer Engineering*

*Overall GPA: 3.1/4.3*

## Work Experience

---

**Data Scientist, KKBOX**

Jun. 2019 – Present

- User Behavior Analysis
  - Churn Prediction: *User behavior insight discovery / Churn user prediction with boosting methods*
  - Subscription Prediction: *Time series analysis with ARIMA / Modeling user journeys via semantic embeddings*
- Music Recommendation system
  - Seed songs selection: *Personalized song prediction*

**Deep Learning Scientist and Bioinformatician, Insilico Medicine**

Aug. 2018 – May. 2019

- Molecules Generation: *Development of generative models to generate potential valid molecules*
- MRI Brain Image Analysis: *Development of Unet model to segment images and predict ages*

**iOS developer, National Cheng Kung University, Main Library**

Aug. 2017 – Jun. 2018

- Development of a mobile library app

**Teaching Assistant, National Cheng Kung University, Department of Electrical Engineering**

Sep. 2016 – Jun. 2017

- Teaching assistant for CS101 (Introduction to Computers) (C++)

## Publications

---

**DNA: General Deterministic Network Adaptive Framework for Multi-Round Multi-Party Influence**

**Maximization.**, accepted paper in *The 5th IEEE International Conference on Data Science and Advanced Analytics* Oct. 2018

- **First author:** generate node-selection policies to maximize influence on social network in the long term with graph mining and reinforcement learning methods

## Skills

---

### Programming Languages

- PYTHON, C++, R, SCALA, SQL, HTML, CSS, JAVASCRIPT, MATLAB, SWIFT, MONGODB

### Data Science Techniques

- **Python Packages:** Pytorch, Keras, TensorFlow, Numpy, Scipy, Pandas, Matplotlib

### Deployment Pipeline

- GITLAB, JENKINS, DOCKER, KUBERNETES

### Language Qualification

- TOEFL: 96/120 (R28, L20, S23, W25)

## Projects

---

**COVID19 Global Forecasting** , *Kaggle Competition mainly held by The White House OSTP* Mar. 2020

- Forecast confirmed cases and fatalities between March 25 and April 22 by region
  - Using vector autoregressive moving average model (VARIMA) to predict regional values simultaneously
  - Top 13% in the competition

**MolHack: Apply deep learning to speedup drug validation** , *Kaggle Competition held by Insilico Medicine* Apr. 2018 – May. 2018

- Given ligand-pharmacophore pairs, predict the stability of the complex
  - Applying a regressor based on deep neural network on well-preprocessed data
  - Won 2nd place in the competition

**KKBOX Data Game: TV Show Recommendation**, *Kaggle Competition held by KKBOX* Jun. 2017

- Design an algorithm to predict what users will watch next
  - Exploratory data analysis / Linear regression

**Social Relationship inference from Urban Footprint** , *National Cheng Kung University* Sep. 2016 – Jan. 2017

- Design an algorithm to predict whether people are friends on social media with users' check-in data
  - User and behavior similarity estimation

**Mining Geo-Social Services for Optimal Location Placement**, *National Cheng Kung University* Sep. 2016 – Jan. 2017

- Design an algorithm to rank top 20 locations for hotels and theaters placement
  - Hill climbing optimization with NDCG ranking score

**Energy Consumption Analysis and Prediction for Household Planning**, *National Cheng Kung University* Sep. 2016 – Jan. 2017

- Design an algorithm to predict a household electricity consumption
  - Feature selection with random forest and linear regression modeling

## References

---

- |                          |   |                          |
|--------------------------|---|--------------------------|
| <b>Research Advisor</b>  | <b>Jen-Wei Huang, Professor,</b><br>National Cheng Kung University, Taiwan<br><b>Email:</b> jwhuang@mail.ncku.edu.tw  | <a href="#">Homepage</a> |
| <b>Course Instructor</b> | <b>Hsun-Ping Hsieh, Professor,</b><br>National Cheng Kung University, Taiwan<br><b>Email:</b> hphsieh@amazon.com  | <a href="#">Homepage</a> |
| <b>Research Mentor</b>   | <b>Emmanuel Salawu, Research Scientist,</b><br>Amazon Web Services, Washington, D.C., USA<br><b>Email:</b> esalawu@amazon.com <b>Phone:</b> +1 202 891 9265 | <a href="#">Homepage</a> |