Tzu-Hsin (Cynthia) Yang

+886 972619777 zixinyang92@gmail.com
https://github.com/ZixinYang https://zixinyang.github.io/

RESEARCH INTERESTS

Explainable AI, Causal Inference, Probabilistic Models, Uncertainty Estimation, Network Analysis

EDUCATION

National Cheng Kung University (NCKU), Tainan, Taiwan Jul. 2018

- Master of Science in Computer and Communication Engineering
  - Overall GPA: 4.0/4.0
  - Highlighted Courses: Internet of Things and Urban Computing, Data Mining and Social Network Analysis, Multilingual and Crosslingual Information System, Graph Theory

National Chiao Tung University (NCTU), Hsinchu, Taiwan Jun. 2016

- Bachelor of Science in Electrical and Computer Engineering
  - Last 60 GPA: 3.27/4.0
  - Highlighted Courses: Data Structure, Algorithms, Operating Systems, Computer Organization, Linear Algebra, Discrete Mathematics, Probability, Intelligent Data Analysis

WORK EXPERIENCE

Research Assistant, Academia Sinica, Taipei, Taiwan Oct. 2020 – Present

- Developing an Insurance Fraud Detection System based on **GNN** models
- Implementing a graph visualization system with **Node.js** and **D3.js**

Data Scientist, KKBOX, Taipei, Taiwan Jun. 2019 – Aug. 2020

- User Behavior Analysis
  - Developed **XGBoost** and **LightGBM** that achieved 90% accuracy to predict churn users
  - Introduced **SHAP** to discover key factors of churn behaviors, strengthened the direction of the marketing strategy with the customer relationship management team
  - Modeled user journeys on the application via semantic embedding with **fastText** that achieved 80% accuracy to predict trial-to-paid conversion
  - Visualized data on **Redash** dashboards with **PrestoSQL**, supported project managers to evaluate the effectiveness of new product releases
- Music Personalized Recommendations
  - Developed **Spark** ML Pipelines for playlists recommendation, facilitated the training process
  - Revised data preprocessed methods and excluded bias of data, attempted to increase user retention, conducted **Firebase** A/B testing to measure how the change affects user's behaviors
- Public Opinion System (Microsoft Taiwan)
  - Developed a crawler CI/CD pipeline with **GitLab** and **Jenkins**, collected online documents, implemented a NER system with **CkipTagger** to infer public opinion about Microsoft products, provided results to Microsoft Taiwan

Deep Learning Scientist and Bioinformatician, Insilico Medicine, Taipei, Taiwan Aug. 2018 – May 2019

- Molecule Generation: developed conditional generative models such as **GAN**, **JT-VAE**, provided a list of generated molecules to drug synthesis laboratories, accelerated drug discovery time
- MRI Brain Image Analysis: implemented the **Unet** model to segment images

iOS Mobile App Developer, NCKU Library, Tainan, Taiwan Aug. 2017 – Jun. 2018

- Developed a mobile library application for twenty thousand students and the faculty

Teaching Assistant, Department of Electrical Engineering, NCKU, Tainan, Taiwan Sept. 2016 – Jun. 2017

- Directed a teaching assistant team for Introduction to Computers (C++)
- Won Teaching Assistant Awards for two consecutive semesters

PUBLICATIONS

T. Yang, H. Ma and J. Huang, "DNA: General Deterministic Network Adaptive Framework for Multi-Round Multi-Party Influence Maximization", 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA) (Oral Presentation)

- Developed an **MCTS**-based framework that adapts to network changes in the long run
- Developed influence maximization algorithms from game theory perspectives
- Designed a network similarity estimation method to reduce the time for inferring new data

H. Hsieh, J. Jiang, T. Yang and R. Hu, "LSTMEnsembler: A LSTM-based Ensemble Framework for Predicting the Success of Mediation Requests Using Case Properties and Textual Information", submitted to 2020 ACM Digital Government: Research and Practice

- Conducted the first research on predicting the success of real-world mediation cases
- Developed an **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 or not

COMPETITIONS

COVID19 Global Forecasting, held by The White House OSTP Mar. 2020

- Forecasted confirmed cases and fatalities between March 25 and April 22 by region
  - Ranked in the top 13% out of 500+ competitors
  - Developed a **VARIMA** model for multivariate time series prediction

MolHack: Apply deep learning to speedup drug validation, held by Insilico Medicine May 2018

- Predicted the stability of the complex given ligand-pharmacophore pairs
  - Won 2^{nd} place out of 13 teams
  - Performed data normalization, concatenated features of pairs, predicted with neural networks

Mei-Chu Hackathon, held by NCTU and National Tsing Hua University Dec. 2014

- Developed an Automated Jumpshot Photo System
  - Won 1^{st} place out of 30 teams
  - Designed an iOS camera application connecting to an Arduino module, developed the Arduino module with accelerometer sensors plugin

SELECTED PROJECTS

Statistics with R Specialization, Duke University (Coursera) Dec. 2019

- Analyzed and visualized data in R, performed frequentist and Bayesian statistical inference

Air Quality Analysis & Prediction, KKBOX Screening Question May. 2019

- Analyzed time series data and predicted the air quality in the upcoming days
  - Performed exploratory data analysis and developed **RNN** models

Energy Consumption Analysis & Prediction for Household Planning, NCKU Jan. 2017

- Designed an algorithm to predict household electricity consumption
  - Adopted random forest to select important features and developed linear regression models

Mining Geo-Social Services for Optimal Location Placement, NCKU Nov. 2016

- Designed an algorithm to rank locations for hotel and theater placement
  - Optimized by hill climbing algorithm with **NDCG** ranking scores

Social Relationship inference from Urban Footprint, NCKU Oct. 2016

- Designed an algorithm to predict friendship on social media with check-in data
  - Selected important features, adopted cosine similarity to measure the similarity between pairs

SKILLS

Programming Languages

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

Certificate

- Microsoft Certified: Azure Fundamentals