Yi-Chia Chen

(530) 648 - 7910 kvnyijia@gmail.com kvnyijia.github.io github.com/kvnyijia linkedin.com/in/kvnyijia

Education

Georgia Institute of Technology Atlanta, GA
M.S. in Computer Science *GPA 4.0/4.0* Aug 2021 – Expected May 2023
- Courses: Data & Visual Analytics, Intro to Grad Algorithms, Intro to Database Systems.

National Cheng Kung University (NCKU) Tainan, Taiwan
B.S. in Computer Science and Information Engineering (CSIE) *GPA 3.55/4.3* Sep 2015 – Jun 2020
- Courses: Data Structures, Algorithms, Operating Systems, Compiler Construction, Computer Networks, Data Mining.

Skills

Programming Languages C, C++, C#, Java, Python, R, SQL, Scala, JavaScript, HTML, CSS, Haskell.
Frameworks & Libraries Pandas, D3.js, ASP.NET, Hakyll.
Developer Tools Git, Unix/Linux, Google BigQuery, Data Studio, Matillion ETL, Apache Spark, MongoDB, Docker.

Work Experience

Institute of Information Science, Academia Sinica Taipei, Taiwan
Research Intern *Program Derivation, Functional Programming* Jul 2020 – Dec 2020
- Researched Brzozowski's derivative and parser combinators to solve parsing problems, then developed programs and derived solutions from formal specifications, using **Haskell** and functional programming paradigm.
- Built an undergraduate course website with **Hakyll** and **Pandoc** that allows the instructor to update announcements by editing Markdown files.

PIXNET Digital Media Corporation Taipei, Taiwan
Data Analyst Intern *Data Analysis* Aug 2019 – Dec 2019
- Reduced manual labor by streamlining workflow of importing data from Google Sheets to BigQuery with **Matillion ETL**.
- Retrieved **100+ GB** data from **Google BigQuery** with **SQL**, and created statistical graphics for evaluating marketing results with **Python**, **R**, and **Google Data Studio**.
- Proposed revision of website layout to improve CTR (clickthrough rate) and user experience by analyzing CTR of links.

Projects

Air Quality Data Collection, Analysis, and Prediction from Scratch at NCKU *Machine Learning* Dec 2018 – Sep 2019
- **Award: 3rd Place** of *2019 CSIE Department Research Project Competition*.
- Collaborated with teammates to collect meteorological data with **Arduino** and store them with **MongoDB** and **Docker**.
- Preprocessed **4 months** of data with **NumPy** and **pandas**, and visualized them with **matplotlib** and **seaborn** to find patterns in the periodic changes and correlations between attributes.
- Developed next-hour PM 2.5 forecasts by building a multivariate linear regression model with MAPE (mean absolute percentage error) **less than 28%** with **scikit-learn**.
- Reduced the deviation of the model by **1%** by adding other scraped weather factors into the model with **beautiful soup**.

THE ONE: Interactive Book Recommendation System *NLP, Data Visualization* Sep 2021 – Dec 2021
- Created interactive graphics with tooltip to display popular books among a country or a book category with **D3.js**.
- Produced recommendations based on scraped tweets by building bag-of-words model with **scikit-learn**.
- Calculated popularity of books or similarity between books and user input with **pandas** and **NLTK**.

Tagging of Thesis *NLP, Deep Learning* Dec 2019 – Jan 2020
- **Award:** Ranked the **top 25%** for the *AI CUP 2019: Artificial Intelligence Analysis and Classification of Thesis*.
- Conducted experiments on the dataset with various models, including **RoBERTa**, **LSTM**, **XLNet**, with **Python**.
- Built a model that predicted tags for sentences in theses by **70%** F1-score with **GRU** (Gated Recurrent Units) model.

Finance Hero: A Simple Budgeting App *Object-Oriented Programming, Relational Database* Jun 2018 – Jul 2018
- Designed a minigame in the app to make expense tracking more enjoyable with **C#**.
- Integrated the app with **Microsoft SQL Server database** to store user spending and game records with **C#** and **SQL**.
- Visualized a user's monthly spending by category in pie chart with **ASP.NET** Chart control.