

Josh Chang

157 W106th St #2A, New York, NY 10025

☎ 707-727-5733 ✉ sc4926@columbia.edu 🏠 joshchang1112.github.io 🌐 joshspchang 📧 joshchang1112

Education

Columbia University

MS in Computer Science

New York, NY

Sep 2021 – Expected Dec 2022

National Taiwan University (NTU)

BS in Computer Science & Information Engineering

Taipei, Taiwan

Sep 2016 – Jan 2021

Experience

Verizon Media

Data Research Intern

Taipei, Taiwan

Jul 2021 – Aug 2021

- Developed an E-Commerce recommendation system using BERT based on media news headline and abstract.
- Pre-trained and fine-tuned GPT-2 model to automatically generate slogans for recommended items.
- Worked on reading and writing large-scaled data formats on HDFS using PySpark.

Cathay Life Insurance

Machine Learning Intern

Taipei, Taiwan

Jul 2020 – Aug 2020

- Established on a fraud detection system using Neo4j, Graph Neural Networks (GNN) and Decision Tree models.
- Increased accuracy by 2% by applying graph-related features to original LightGBM model.

IBM

AI Engineering Intern

Taipei, Taiwan

Jul 2019 – Aug 2019

- Achieved image classification model accuracy to 99% and contracted model size to 1.5% using knowledge distillation techniques on detecting specific ear diseases.
- Led our intern team members in IBM Design Thinking Workshop, organized weekly meeting, and finally won 2nd prize on final presentation among 12 intern teams.

Machine Intelligence & Understanding Laboratory, NTU

Undergraduate Researcher

Taipei, Taiwan

Sep 2018 – Jun 2020

- Published a workshop paper, "**Leveraging Speaker Profile and Knowledge Enrichment for End-to-End Advising Response Ranking**", as the first author in **DSTC8, AAI-2020**.
- Researched Transfer Learning for Data to Text Generation in baseball and basketball domains.
- Presented final research project in the field of dialogue system to 200+ students and was awarded honorable mention prize in CS department.

Projects

Graph Neural Networks(GNN) in Tensorflow

Jul 2020 – Nov 2020

- Contributed to **Neural Structure Learning, Tensorflow** (800+ stars on Github).
- Implemented three state-of-the-art GNN models using Tensorflow: GCN, GAT, GIN.
- Achieved better accuracy compared to original research papers running on the Cora dataset.

Fine-tuned BERT Encoder & Graph for Node Classification

Nov 2020 – Jan 2021

- Encoded node features and bring comprehensive text information from BERT embeddings to Graph Neural Networks.
- The proposed method outperformed the 1st place in Open Graph Benchmark Leaderboard.

Dining Concierge AI Agent

Sep 2021 – Oct 2021

- Deployed a web application that gives users restaurant suggestions according to their preferences.
- Built a restaurant suggestion chatbot using AWS Lex, Lambda, SQS and integrated Lex into chat web application.
- Created a restaurant database scraped from Yelp using DynamoDB and ElasticSearch.

Awards

2021 **Ranked 21 / 852 teams (Silver Medal)**, Kaggle - MLB Player Digital Engagement Forecasting
2020 **1st prize (out of 469 teams)**, 2019 Taiwan AI Cup – Thesis Tagging Competition
2019 **1st prize (out of 244 teams)**, 2019 Cathay Big Data Competition

Technical Skills

Programming Languages: Python, C, Java, HTML/CSS, SQL

Tools/Frameworks: Pytorch, Tensorflow, Keras, Scikit-learn, AWS, Git, Flask, Docker, Spark, Hadoop, Kafka