Josh Chang

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

📞 707-727-5733 ✉ sc4926@columbia.edu 🏠 joshchang1112.github.io 🔗 joshspchang ⬛ joshchang1112

Education

Columbia University   **New York, NY**
MS in Computer Science   Sep 2021 – Expected Dec 2022

National Taiwan University   **Taipei, Taiwan**
BS in Computer Science & Information Engineering   Sep 2016 – Jan 2021

Experience

Verizon Media   **Taipei, Taiwan**
Data Research Intern   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   **Taipei, Taiwan**
Machine Learning Intern   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   **Taipei, Taiwan**
AI Engineering Intern   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   **Taipei, Taiwan**
Undergraduate Researcher   Sep 2018 – Jun 2020
- Published a workshop paper on AAAI to leverage speaker profile information for improving response ranking task.
- Researched Transfer Learning for Data to Text Generation in baseball and basketball domains.
- Presented final research project to 200+ students and acquired 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.

Miubot   **Apr 2019 – Jun 2019**
- Built an task-oriented dialogue system for developers to train models and replace their own model structures easily.
- Designed Encoder-Decoder, Transformer-based models on Language Understanding and Natural Language Generation.

Awards

2021	**Ranked 21 / 852 teams (Silver Medal)**, Kaggle - MLB Player Digital Engagement Forecasting
2019	**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