Chen-Kang Lee

+1 (858)-500-2105 Chen-Kang Lee Personal Website chl126@ucsd.edu

Education

University of California San Diego **Sep 2021 - Dec 2022 (Expected)**
M.Sc. in Computer Science and Engineering (GPA: 3.84/4.0) *La Jolla, CA, USA*

- Highlighted Coursework: Convex Optimization, Algorithms Design and Analysis, Advanced Compiler Design

National Tsing Hua University **Sep 2015 - Jun 2019**
B.S. in Computer Science (GPA: 3.82/4.0, Academic Achievement Award) *Hsinchu, Taiwan*

- Highlighted Coursework: Operating Systems, Data Structures, Databases Systems, Parallel Programming

Skills

Programming Languages Python, C/C++, Java, Javascript/Typescript, HTML, CSS, C#, SQL
Tools & Framework Pytorch, Tensorflow, Node.js, React.js, Docker, Git, GCP, AWS, Apache Beam, Apache Airflow Jira, Kubernetes, Hadoop MapReduce, MPI, CUDA, Unix/Linux

Professional Experience

Software Engineer Intern **Jun 2022 - Aug 2022**
Pyramid Systems Inc. *Fairfax, VA, USA*

- Worked on a cloud-based MLOps project running machine learning models on *Google Cloud Platform* to identify natural disaster hot zones from FEMA disaster data and classify FDIC bank data by their risks.
- Designed and implemented an automatically scaling distributed data preprocessing pipeline using *Apache Beam*.
- Implemented machine learning models on *Vertex AI* and automated the pipeline execution using *Apache Airflow*.
- Created a full stack web application using *React* front end to present the result of the pipeline and deployed the web application server on *Google App Engine*.

Research Assistant **Jun 2020 - May 2021**
Institute of Information Science, Academia Sinica *Taipei, Taiwan*

- Developed H-FND, an iterative *reinforcement learning* framework that achieved 0.81 F1-score (state-of-the-art performance) in relation extraction on the SemEval distantly supervised dataset with false negative noise.
- Utilized the H-FND framework to automatically extract new knowledge from un-annotated news articles.
- Developed and maintained the relation extraction functionality of the lab's CoreNLP pipeline.

Software Engineer Intern **Jan 2019 - Mar 2019**
SkyEyes GPS Technology Co., Ltd. *Taichung, Taiwan*

- Implemented MobileNet in *Pytorch* to detect mis-classification in the driver assessment software with an 88% accuracy.
- Preprocessed and augmented 3000+ images taken by dash cams of the driver assessment software to generate training data for the deep learning classification model.

Projects

Python to WebAssembly Compiler **Mar 2022 - Jun 2022**
Project for UCSD CSE231: Advanced Compiler Design

- A compiler written in *Typescript* that compiles a subset of Python code into *WebAssembly* and executes in-browser.
- Worked on a large-scale development project to extend other core python functionalities such as iterables, memory management, and an in-browser temporary file system to support I/O operations. [Code]

Rain-induced Landslide Prediction **Jan 2020 - Apr 2020**
Project as Research Assistant at FCU AI Research Center

- Preprocessed historical satellite landslide data, geographic data, and rainfall data to predict rain-induced landslides in the Chenyulan river watershed.
- Co-developed a *graph-based sequential model* that utilizes *generative adversarial networks (GANs)* to combat the lack of training data. The model achieved 85% accuracy (0.72 F1-score) predicting landslides in test.

Publications

Jhi-Wei Chen, Tsu-Jui Fu, **Chen-Kang Lee**, Wei-Yun Ma
"H-FND: Hierarchical False-Negative Denoising For Distant Supervision Relation Extraction" [Link]
59th Annual Meeting of the Association for Computational Linguistics (ACL), 2021 (Findings)