Yun Cheng Wang

(323) 578-8258 | yunchenw@usc.edu | Website | Github | Linkedin

Research Interests

Machine Learning Representation Learning, Graph Neural Networks, Explainable Machine Learning

Data Mining Knowledge Graphs, Social Network Analysis, Knowledge Base Completion

Education

University of Southern California

Los Angeles, CA

Ph.D. in Electrical and Computer Engineering *January 2021 - Present*

- Incoming Ph.D. student at USC MCL.

University of Southern California

Los Angeles, CA

M.S. in Electrical and Computer Engineering, GPA: 4.0/4.0 *August 2018 - December 2019*

- **Relevant Courses**: Machine Learning from Signals, Mathematical Pattern Recognition, Introduction to Digital Image Processing, Multimedia Data Compression

National Taiwan University

Taipei, Taiwan

B.S. in Electrical Engineering, GPA: 3.8/4.3 (A+: 4.3) *September 2014 - May 2018*

- **Relevant Courses**: Introduction to Digital Speech Processing, Signals and Systems, Introduction to Artificial Intelligence and Machine Learning, Machine Learning Fundamentals, Machine Learning, Algorithms

Research and Industrial Experience

USC Media Communication Lab

Los Angeles, CA

Member *August 2018 - PRESENT*

- Obtained high-level representation for relational data by capturing correlations between entities and propagating features among neighborhood. Develop time-efficient and explainable subspace learning framework for entity classification.
- Drew insights by surveying and testing various kinds of state-of-the-art graph embedding models and provided explanations and interpretations on experimental results based on model structure and graph statistics.

CKIP Lab, Academic Sinica

Taipei, Taiwan

Research Assistant *September 2020 - December 2020*

- CKIP lab is dedicated to establish a fundamental research environment for Chinese natural language processing. Research problems in our lab contain but not confine to knowledge acquisition, knowledge representation, and knowledge utilization.
- Responsible for constructing sememe-based word embeddings from the eHowNet ontology, and making opensourced to the public.

Taiwan Semiconductor Manufacturing Company (TSMC)

Hsinchu, Taiwan

CIM Engineer *February 2020 - May 2020*

- Responsible for reducing wafer defect ratio by anaylzing the RD reports on defect cases. Root causes are automatically extracted and prevented in the future wafer manufacturing processes.
- Construct a wafer-centric knowledge base with stations, particles, and process parameters as entities and relations to accumulate wafer manufacturing knowledge.

Taboola

Los Angeles, CA

Data Science Intern *June 2019 - August 2019*

- Led a group of three interns on a "Knowledge Graph" project. Handled over 20 thousand news articles per day by extracting key terms and converting them into structured graph data that was especially efficient for indexing and retrieval.
- Developed scalable data pipelines by using Apache Spark to manipulate large volumes of data. Ran pipelines on kubernetes machines, and ingested processed data into neo4j servers on machines on a daily basis.

Publications

- B Wang, F Chen, Y Wang, CCJ Kuo. **Efficient Sentence Embedding via Semantic Subspace Analysis**. arXiv preprint (2020) ➢ pdf
- F Chen, Y Wang, B Wang, CCJ Kuo. **Graph representation learning: A survey.** APSIPA Transactions on Signal and Information Processing 9 (2020) ➢ pdf
- B Wang, A Wang, F Chen, Y Wang, CCJ Kuo. **Evaluating word embedding models: Methods and experimental results.** APSIPA transactions on signal and information processing 8 (2019) ➢ pdf

Honors, Teachings, and Members

2019 **IEEE Student Member in Signal Processing Society**
2019 **Teaching Assistant for Graduate-level Probability (EE503)**
2019 **USC EE Master Honors Program (Highly Selective)**
2012 **Asia Pacific Mathematics Olympiad Training Camp**
2011 **Selected into KSHS Science Specialty Class (Acceptance Ratio: 0.3%)**

Skills

Natural Languages Chinese (Mandarin), English (TOEFL iBT:102, TOEIC:890)
Programming Languages Python, C++, Java, Matlab
Frameworks and APIs Docker, Apache Spark, Neo4j, Kubernetes, Google Cloud Platform, IBM Watson

Projects

Video Content Searching System

LA Hacks *Spring 2019*

- Developed an YouTube video retrieval system that could extract video segments containing certain keywords or objects.
- Sent requests to Natural Language Processing and Vision APIs on Google Cloud Platform.

Video Summarization System

KKStream Ltd. *Summer 2017 - Summer 2018*

- Developed a Video Summarization System for TV shows and TV series. The system was evaluated by subjective testing with over twenty attendances.
- Held regular meetings with data scientists in KKStream Ltd. to exchange ideas and discuss the progress.
- The project was presented to over fifty employeers in KKStream Ltd. in an annual meeting, and received many positive feedbacks.

Extracurricular Activities

2014, 2015 NTUEE Basketball Team

- Represented department to play against teams from different departments and different schools.
- Won 2015 championship in the secondary level basketball tournament.

2015 NTUEE Summer Camp Volunteer

- NTUEE summer camps are held every year to educate high school students with fundamental science and engineering knowledge.
- Volunteers helped coordinate activities and provided course materials.