# Lang-Chi Yu

(323) 334-5044 | langchi.yu@gmail.com | OckhamsRazor | OckhamsRazor

## Education

**University of Southern California** *Los Angeles, California*

M.S. in Computer Science *Aug. 2018 — Expected May 2020*

- Cumulative GPA: 3.64/4.00

**National Taiwan University (NTU)** *Taipei, Taiwan*

M.S. in Communication Engineering *Sep. 2015 –– Jun. 2017*

- Master thesis: "Abstractive Headline Generation for Spoken Content with ASR Error Modeling[1]."

B.S. in Electrical Engineering *Sep. 2010 –– Jun. 2014*

- Cumulative GPA: 3.87/4.30
- Dean's List (top 5%) in Spring 2012

## Experience

**Amazon Web Services (AWS)** *Bellevue, Washington*

Software Development Engineer Intern *May 2019 –– Aug. 2019*

- Develop new website modules with React, Node.js, and DynamoDB for new functionalities of AWS Envy Project, e.g., waste storage and management interface for AWS data centers around the globe.
- Develop unit tests for website modules with Jest.

**Research Center for Information Technology Innovation, Academia Sinica** *Taipei, Taiwan*

Research Assistant *Aug. 2017 –– Jul. 2018*

- "Hit Song Prediction for Pop Music by Siamese CNN with Ranking Loss[2]."

**BoniO Inc.** *Taipei, Taiwan*

Founding Developer & Web Developer *Mar. 2013 –– Sep. 2013*

- Developed PaGamO, the world's first MOOC-based multi-student educational gaming platform used by MOOC platform Coursera.
- Developed web front-end (realtime player status, messaging, announcement systems) with Ruby on Rails, JQuery, AJAX, and MySQL.

## Projects

### Hit Song Prediction for Pop Music by Siamese CNN with Ranking Loss

- Devised a new approach to hit song prediction problem, i.e., predicting song's popularity score in a popular music market based on raw audio (mel-spectrograms), by incorporating song ranking model into previous regression model as Siamese convolutional neural networks.
- Collaborated with Machine Learning Research Team of KKBOX Inc., Taiwan, implemented the Siamese CNN models with Python TensorFlow, which made great improvement in Kendall's tau (performance metric for ranking songs) from 0.1080 to 0.2421 compared to previous simple deep models on a commercial song dataset of KKBOX.
- Constructed a standardized and publicly-available hit song prediction benchmark with Million Song Dataset and a YouTube crawler.

### Abstractive Headline Generation for Spoken Content with ASR Error Modeling

- Proposed and implemented novel attentive RNN architecture with Lua Torch, which incorporated automatic speech recognition (ASR) error modeling mechanism to learn from text data instead of from large number of spoken documents with reference headlines.
- Generated the finding that proposed architecture outperformed RNN and ARNN models that were trained on pure text data with an increase in ROUGE-1 (correctness metric for summarization) from 21.87 to 22.89.

## Skills

### Programming

- Python: Django, Keras, TensorFlow, scikit-learn, Pandas
- JavaScript: Node.js, React, Jest, Semantic UI
- Ruby: Ruby on Rails (RoR)
- Java, C++, Lua (Torch), Shell Script
- Version control: Git

### System Management

- Cloud Services: Amazon Web Services (AWS) Elastic Compute Cloud (EC2)
- Database: Amazon DynamoDB, MongoDB, PostgreSQL, MySQL, Oracle SQL
- Operating Systems: CentOS, Debian, Ubuntu
- Web servers: Apache, NGINX

---

[1] Lang-Chi Yu, Hung-yi Lee, and Lin-shan Lee, "Abstractive Headline Generation for Spoken Content by Attentive Recurrent Neural Networks with ASR Error Modeling," IEEE Workshop on Spoken Language Technology, Dec. 2016.

[2] Lang-Chi Yu, Yi-Hsuan Yang, Yun-Ning Hung, and Yi-An Chen, "Hit Song Prediction for Pop Music by Siamese CNN with Ranking Loss," arXiv preprint arXiv:1710.10814 (2017).