HAN-YUN(HENRY) YEH

henryyeh034@gmail.com, +886-922463510

github.com/henry034/

EDUCATION

National Chiao Tung University   Hsinchu, Taiwan   *Sep. 2016 - June 2019*

- M.S. in Communications Engineering (GPA: 4.02/4.3)

National Taipei University   Taipei, Taiwan   *Sep. 2012 - June 2016*

- B.S. in Communications Engineering (Rank: 1st, GPA: 3.72/4.0)

HORNORS AND AWARDS

Research

- **6th Place** (2017) - IJCNLP Shared Task 2 [3] - Taipei, Taiwan
- **1st Place** (2015) - NTPU CE Senior Project Competition - Taipei, Taiwan
- **Best Paper Award** (2014) - Oriental COCOSDA [4] - Phuket, Thailand

Academic

- **Phi Tau Phi Award** (2016) - The Phi Tau Phi Scholastic Honor Society, Taipei Taiwan
  Award for top-ranked student in CE department among all classes
- **Dean's List** (Fall '12, Fall '15) - NTPU CE Dept. - Taipe, Taiwan
- **Scholarship** (2014) - Elytone Electronic CO., LTD. - Taipei, Taiwan
- **Honerable Mention** (2014) - Taiwan National Collegiate Programming Contest - Taipei, Taiwan

RESEARCH EXPERIENCE

National Chiao Tung University   Hsinchu, Taiwan
Graduate Student/Research Assistant, Speech Processing Lab   *Sep. 2016 - Jan. 2019*

- Research Topic: Chinese pinyin to character language model using deep learning approach [1]

  - Experimented sequence labeling (TDNN+BLSTM, joint learning with word boundary prediction) and seq2seq (Transformer) model on Chinese pinyin to character problem.
  - Preprocessed data from text corpus including wikipedia, LDC Chinese Gigaword, Sinica corpus,...through high precision CRF-based Chinese parser and rule-based G2P (character to pinyin) system.
  - Reduced character error rate to 5.6%.

- Dimensional sentiment analysis for Chinese phrases (DSAP) [3]

  - Achieved mean rank 6.5 among 24 submissions on Chinese phrases' valence and arousal prediction problem using proposed order-aware word2vec and BLSTM model on CAVT(Chinese Valence-Arousal Text) corpus.

- Child Speech Impairment Supporting System

  - Implemented corpus recording system (for speech impairment child) GUI using Java in the cooperation project with NTU Hospital Hsinchu. Collected and analyzed about 200 children's sample.

National Taipei University — Taipei, Taiwan
Research Assistant, Speech and Multimedia Signal Processing Lab — *Sep. 2012 - June 2016*

· Senior project: *"An Automatic Grade Input System via Voice"*

– Constructed speech recognition system including energy-based voice activity detection and beamforming noise cancellation module to enter student's name and score automatically. Obtained 1st place in NTPU CE Senior Project Competition.

· Mandarin prosody generation [2][4]

– Investigated the improvement of CRF-based base-phrase chunk feature and punctuation confidence in Mandarin text-to-speech system.
– Labeled base-phrase chunk feature by CRF-based base-phrase chunker.
– Generated CRF-based punctuation confidence for each lexical word boundary from input text tagged with Chinese word boundary, part of speech (POS) and base-phrase chunk, measuring the likelihood of inserting a punctuation mark (PM).
– Applied above feature in MLP-based prosody generator and confirmed that the RMSE for predicting logF0, syllable duration, energy level and pause duration are all reduced.

PUBLICATIONS

[1] **Han-Yun Yeh**. "end-to-end pinyin to character language model using self-attention mechanism". Master's thesis, National Chiao Tung University, 2019

[2] Chen-Yu Chiang, Yu-Ping Hung, **Han-Yun Yeh**, I-Bin Liao, and Chen-Ming Pan. Punctuation-generation-inspired linguistic features for Mandarin prosody generation. *EURASIP Journal on Audio, Speech, and Music Processing*, 2019(1):4, 2019

[3] Yen-Hsuan Lee, **Han-Yun Yeh**, Yih-Ru Wang, and Yuan-Fu Liao. Nctu-ntut at ijcnlp-2017 task 2: Deep phrase embedding using bi-LSTMs for valence-arousal ratings prediction of Chinese phrases. In *Proceedings of the IJCNLP 2017, Shared Tasks*, pages 124–129, 2017

[4] Yu-Ping Hung, **Han-Yun Yeh**, I-Bin Liao, Chen-Ming Pan, and Chen-Yu Chiang. An Investigation on linguistic features for Mandarin prosody generation. In *2014 17th Oriental Chapter of the International Committee for the Co-ordination and Standardization of Speech Databases and Assessment Techniques (COCOSDA)*, pages 1–5. IEEE, 2014

WORK EXPERIENCE

IBM, Inc — Taipei, Taiwan
Application Developer — *June 2018 - Oct. 2018*

· Organized health knowledge collected from internet using text processing techniques and designed rule-based health information suggestion according to user's information which includes the data from wearable device or entered by user, using Python.
· Designed backend infrastructure to collect, analyzed website user behavior data with Javascript, Python, PHP, MongoDB, MySQL and deployed service on AWS.

Application Developer Intern — *Jul. 2017 - Aug. 2017*

· Implemented automated optical inspection (AOI) algorithm of defect detection for the label on circuit board including misprint, broken label,…, using Python with OpenCV package.

TEACHING AND ADVISING EXPERIENCE

Teaching Assistant, *National Chiao Tung University* — Sep. 2016 - Jan. 2018

· "Principle of Microcomputer" course taught by Prof. Yi-Ru, Wang — Sep. 2017 - Jan. 2018
· "Logic Design and Lab" course taught by Prof. Yi-Ru, Wang — Sep. 2016 - Jan. 2017

Teaching Assistant, *National Taipei University* — Feb. 2015 - June 2016

· "Physics Lab" course taught by Prof. Cheng-Yu, Chiang — Feb. 2015 - June 2016

ACTIVITIES

2020 Formosa Grand Challenge — Taipei, Taiwan
Participant — *Nov. 2019 - Apr. 2020*

· Build model and labeled data to solve various Chinese reading comprehension task using Pytorh and build rule-based dialog system to solve conversation problem with 20 different domains.

2018 5th Pixnet Hackthon — Taipei, Taiwan
Participant — *Aug. 2018*

· Developed a device and management website for answering popular dishes using Google AIY Voice Kit and information from text mining on 200,000 Pixnet food blogs.

NTPU CE Student Association — Taipei, Taiwan
Director of Information Management — *Sep. 2013 - June 2015*

· Managed student information system and designed advertising photo for promoting the activities of student association.
· Instructor of orientation events for freshman.

Leader of Curriculum-design Section — *Feb. 2015*

· Designed curriculum including IR remoter DIY, speaker/microphone DIY and PC building of 1st NTPU CE winter camp.

The International Collegiate Programming Contest — Jakarta Indonesia
Participant — *Dec. 2014*

· Placed 29th among 70 teams in 2014 ACM-ICPC Asia Jakarta Regional Contest. Solved 4 of 11 problems.

TECHNICAL STRENGTHS

Computer Languages	Python, C, Java, Javascript, PHP
Databases	MySQL, PostgreSQL
Frameworks	TensorFlow, Pytorch
Package	OpenCV
Cloud service	AWS