

HAN-YUN(HENRY) YEH

henryyeh034@gmail.com

github.com/henry034/

EDUCATION

- | | | |
|---|-----------------|------------------------------|
| National Chiao Tung University | Hsinchu, Taiwan | <i>Sep. 2016 - June 2019</i> |
| <ul style="list-style-type: none">• M.S. in Communications Engineering (GPA: 4.02/4.3) | | |
| National Taipei University | Taipei, Taiwan | <i>Sep. 2012 - June 2016</i> |
| <ul style="list-style-type: none">• B.S. in Communications Engineering (Rank: 1st, GPA: 3.72/4.0) | | |

HONORS AND AWARDS

Research

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

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. - Taipei, Taiwan
- **Scholarship** (2014) - Elytone Electronic CO., LTD. - Taipei, Taiwan
- **Honorable Mention** (2014) - Taiwan National Collegiate Programming Contest - Taipei, Taiwan

RESEARCH EXPERIENCE

- | | |
|--|------------------------------|
| National Chiao Tung University | Hsinchu, Taiwan |
| <i>Graduate Student/Research Assistant, Speech Processing Lab</i> | <i>Sep. 2016 - Jan. 2019</i> |
| <ul style="list-style-type: none">• 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. The training data come from text corpus including wikipedia, LDC Chinese Gigaword, Sinica corpus,...through high precision 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-awared 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 |
| <i>Research Assistant, Speech and Multimedia Signal Processing Lab</i> | <i>Sep. 2012 - June 2016</i> |
| <ul style="list-style-type: none">• Senior project: "An Automatic Grade Input System via Voice"
Constructed speech recognition system including energy-based voice activity detection and beam-forming noise cancelation 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. The base-phrase chunk feature is labeled by CRF-based base-phrase chunker and the punctuation confidence is 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). These features are applied in MLP-based prosody generator and the result shows 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

Application Developer

Taipei, Taiwan

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.
- Implemented backend processing for statistical analysis on collected 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
- "Logic Design and Lab" course taught by Prof. Yi-Ru, Wang

Sep. 2017 - Jan. 2018

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 Pytorch 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.

TECHNICAL STRENGTHS

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