Taekyoon Choi

tgchoi03@gmail.com

EDUCATION

B.S. in Computer Science, Seokyung University, Seoul, South Korea — Mar 2013 – Feb 2018
- Cumulative GPA: 4.24 / 4.5

University of Southern California, English Language Program — Sep 2016 – Apr 2017

TECHNICAL SKILLS

Experties Natural Language Processing, Deep Learning, Machine Learning, Full-stack
Languages Python (Strong), C#, C, Java, HTML, CSS, JavaScript
Frameworks Tensorflow & Pytorch (Strong), Flask, Scipy, Numpy, Pandas, AngularJS, OpenCV

WORK EXPERIENCE

Mr.Mind Inc. | Dialogue Engine Developer & Founding Member

- **Toy Robot "Pororot" Dialogue Engine**(youtu.be/VQrfT53MMVM) — Apr 2018 – Jul 2018
  - Analyzed over 300K user query data by using Word2Vec to extract feasible queries for constructing dialogue dataset
  - Team Lead for designing the Natural Language Understanding engine using Word2Vec for query extraction and Deep CNN Text Model for semantic modelling. Achieved 93 % on Pororot design dataset.
  - Tested dialogue engine in play situation with 5-6 years children by developing engine API for smartphone application
- **Hotel FAQ Dialogue Engine**(youtu.be/ZdMrGo3M5Vc) — Feb 2018 – Apr 2018
  - Settled strategies for extracting Korean entity by using LSTM-CRF model and customizing entity tag for Hotel FAQ
  - Designed query filter module using idea of discriminator module in SeqGAN to decide whether engine make response
  - Achieved 4 of 5 scores from user case test by 5 adult users
- **AI Capsule Device for Voice User Interface** — Jun 2018 – Oct 2018
  - Researched AI speaker user experience to design basic architecture of device software and experienced collaborating senior engineers from AI speaker hardware solution company
  - Designed and implemented application architecture by event driven programming using pthread and flag constants
  - Developed server communication modules by websocket to build flexible message sending with server and anticipated device can communicate among people & machine remotely

Power Voice | Natural Language Processing Researcher

- **Korean Word Segmentation Model** — Mar 2019 – Apr 2019
  - Developed Word Segmentation Model to process spaced text by unspaced or wrongly spaced Korean text
  - Designed self-supervised model with 'BIE'(Begin, Inside, End) tagset with over 15M refined Korean text resources
  - Implemented model using CNN-BiLSTM model and achieved about 90% WER score for test dataset
- **Spoken Language Understanding Model for Dialogue Management System** — Apr 2019 –
  - Developed Spoken Language Understanding(SLU) Model to extract intent and entity information by users query
  - Implemented multi-task model as baseline model by using Bi-GRU with attention and slot-gate modules and achieved 97% accuracy and 88.8% F1-score for intent and NER scores by SNIPS dataset
  - Developing SLU model with Bidirectional Encoder Representations from Transformers(BERT) for better performance

RESEARCH EXPERIENCE

Data Visual Analytic Lab., Korea University | Undergraduate NLP Researcher

- **SQuAD Dataset Challenge**(youtu.be/IsQ92ahTh9Y) — Jun 2017 – Nov 2017
  - Implemented baseline model, BiDAF using Pytorch Framework by analyzing original Tensorflow project
  - Designed Sentence & Chunk Attention ideas to baseline to guide focusing on answer likely sentence or chunk area in context based on Text Similarity and Sentence Entailment models
  - Anticipated our model select answer near the most probable area in context by reducing unnecessary information

PUBLICATION

- Moonsoo Chang, Taekyoon Choi, and Minho Yoo, "Development of Reporting Tool for Supporting Language Disorder Diagnosis," in *SCIS&ISIS*, Sapporo, Hokkaido, Japan, Aug 2016.
- Changwook Jun, Taekyoon Choi, and Junghyun Cho, "Learning Natural Language Processing with Deep Learning and Tensorflow" , Released on Feb 2019. (https://github.com/NLP-kr/tensorflow-ml-nlp)

PROJECTS

- **Korean Child Language Assessment Reporting Tool** — Feb 2016 – May 2016
  - Participated in projects of analytic reports for child speech assessment
  - Designed report views for analyzing Korean spoken language ability, and representing data for assessment
  - Migrated child speech text dataset to database from raw text files
- **Sudoku Solver Application** (github.com/kai3n/AlphaKu) — Feb 2017 – Mar 2017
  - Designed the sudoku solver by recognizing numbers in Sudoku table given natural picture and finding the solution
  - Implemented cropping a sudoku table algorithm and recognizing each number letters
  - Improved MNIST data set up to 99.58% accuracy by convolution neural network algorithm