Yasufumi Moriya

Speech and Language Processing Scientist

yasufumi1006@gmail.com

About me

My expertise is natural language processing (NLP) and speech processing using deep learning with a particular emphasis on Automatic Speech Recognition (ASR) and Information Retrieval (IR). Combining academic and industry experience, my strength is to bring research outcomes to improve production speech and language processing systems.

Experience

November 2021 - PRESENT

T-Pro, Dublin – *Speech Scientist*

- Improved the word error rate (WER) of the Kaldi ASR system for the medical domain from 15.5% to 13.2% with data augmentation.
- Improved accuracy of voice typing commands from 80% to 92%.
- Conducted research on the CTC and RNN-T based ASR systems for production use.

September 2018 - March 2021

Amazon Ireland, Dublin – *Japanese Language Engineer (part-time)*

- Focused on improvement of pronunciation accuracy of the production Japanese text-to-speech system.

November 2015 - March 2017

Ivona Software Amazon Development Center, Gdansk – *Japanese Computational Linguist*

- Supervised quality of data annotation performed by contractors.
- Improved text expansion modules for the production Japanese text-to-speech system using C language.
- Worked with software engineers and speech scientists to improve overall quality of the Japanese text-to-speech system.

Education

April 2017 - August 2022

Dublin City University, Dublin – *PhD in Computing*

Supervisor: Gareth Jones

Thesis: Augmenting Automatic Speech Recognition and Spoken Content Retrieval for User-generated Video Collection

- Explored use of multimodal features including visual and textual information for the ASR system.
- Improved the ASR system using the semi-supervised technique for noisy user-generated content with search applications.
- Improved the search system for spoken documents using a transformer-based natural language processing model.

September 2014 - September 2015

University of Edinburgh, Edinburgh – *MSc in Speech and Language Processing with Distinction*

Thesis: Stacked Denoising Autoencoder for the Front-end of DNN-based Speech Synthesis

- Gained a wide variety of knowledge about speech and natural language processing.
- Conducted a research project on use of auto-encoder for the front-end of text-to-speech systems.

April 2009 - March 2014

Sophia University, Tokyo – *BSc in Russian Studies with Linguistics Concentration*

- Learned theoretical approaches to linguistics including generative grammar and statistical approaches to analysing linguistic phenomenon.

Selected publications

- Y. Moriya, and G. J. F. Jones, "Improving Noise Robustness for Spoken Content Retrieval using Semi-supervised ASR and N-best Transcripts for BERT-based Ranking Models", In Spoken Language Technology 2022.
- Y. Moriya and G. J. F. Jones "DCU-ADAPT at the TREC 2020 Podcasts Track", In Text REtrieval Conference (TREC), 2020.
- Y. Moriya, R. Sanabria, F. Metze and G. J. F. Jones, "Grounding Object Detections with Transcriptions", In ICML 2019 Workshop, The How2 Challenge.
- Y. Moriya, and G. J. F. Jones, "Multimodal Speaker Adaptation of Acoustic Model and Language Model for ASR using Speaker Face Embedding" In International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019.

Skills

Programming languages: Python, C/C++, bash

Deep learning, ASR and NLP libraries: PyTorch, Hugging Face, scikit-learn, Kaldi, K2, NeMo, Lucene, Terrier, Spacy, NLTK

Software Development: JIRA, Git, Agile, Kanban

Languages: [Fluent] Japanese, English, [Beginner] Russian, Polish, Italian