Karsper Kwok-Him SO

sohim0419@gmail.com | Hong Kong

linkedin.com/in/karsper-so-9ba992176

EDUCATION

University of Edinburgh, United Kingdom   Sep 2020 – Sep 2021

Master of Science, Speech and Language Processing – Distinction

Core Disciplines:
- Accelerated Natural Language Processing
- Automatic Speech Recognition
- Image and Vision Computing
- Natural language Understanding, Generation and Machine Translation
- Reinforcement Learning
- Speech Processing
- Speech Synthesis
- Univariate Statistics and Methodology using R

Dissertation: "Neural Speaker Diarisation in Telephony Communication"
Industrial Collaborator: Dialpad, Kitchener, Canada
Advised by Catherine Lai

University of Tübingen, Germany   Oct 2017 – Jun 2020

Bachelor of Arts, Computational Linguistics – 1.4

Thesis: "Informing Social Media Sentiment Analysis with Cantonese Tonal Prosody:
A Recurrent Neural Network Approach" – 1.0
Advised by Detmar Meurers and Çağrı Çöltekin

RESEARCH EXPERIENCE

Fano Labs, Hong Kong   Dec 2021 – Dec 2022

Research Engineer
- Reviewed, summarised and implemented SOTA techniques from the latest academic publications to improve and productise models for end-to-end neural speaker diarisation (EEND) applications in a broad range of data-sensitive domains including finance and judiciary.
- Proposed, implemented and conducted experiments on conversational speech simulation protocols and various signal processing utilities to mitigate data paucity and privacy challenges for conformer-based EEND models.
- Collaborated closely with data, business and software engineering teams to formulate and solve challenges in building real-world ML applications with the requirement to balance research novelty and fulfil end-user requirements.
- Gave regular research talks to audience groups with non-technical and research backgrounds.

Dialpad, Kitchener, Canada   May 2021 – Aug 2021

AI Research Contractor
- Selected as a student researcher in an industry-collaborated dissertation project on DNN-based speaker diarisation topics. The research focused on leveraging deep speaker representations such as x-vectors and ECAPA-TDNN embeddings to enable robust clustering-based diarisation in telephony communication.

IBM Research and Development, Germany   Aug 2019 – Jan 2020

Research Internship, Department of IBM Cloud and Cognitive Software
- Implemented a prototype HMM-based system that generates syntax-valid texts with named entities for evaluating IBM analytic products.
- Modified and extended RegEx-based components of a data extraction system for the use of a project related to General Data Protection Regulation (GDPR).

CLARIN Centre Tübingen, Germany   Apr 2019 – Aug 2019
Research Assistant, CLARIN-D
- Implemented multiple filtering and ranking functions for an existing treebank annotator.
- Implemented a parser that generates metadata files from an XML source of a UD treebank.
- Extended a command-line application for validating and modifying metadata files with Java.

University of Tübingen, Germany   Nov 2018 – Apr 2019
Research Assistant, Department of Quantitative Linguistics
- Conducted sanity check and preprocess more than 100 hours of acoustic data for the project "CAPTuring the essence: On the value of native, non-native and speaker-specific data in computer-assisted pronunciation training".

PUBLICATIONS

Nianheng Wu, Eric DeMattos, Kwok Him So, Pin-zhen Chen, and Çağrı Çöltekin. 2019. Language Discrimination and Transfer Learning for Similar Languages: Experiments with Feature Combinations and Adaptation. In Proceedings of the Sixth Workshop on NLP for Similar Languages, Varieties and Dialects, pages 54–63, Ann Arbor, Michigan. Association for Computational Linguistics. [Paper]

COURSEWORK & PROJECTS

- Building unit-selection TTS system with my own voice and design evaluation experiments (Multilingual and Cross-lingual TTS in English and Cantonese Chinese)   Feb 2021 – Apr 2021
- Experiments on an ASR system (with WFST structures and Viterbi decoder)   Feb 2021 – Apr 2021
- RL algorithms implementation (MDPs, Continuous DeepRL, Multi-agent RL)   Feb 2021 – Apr 2021
- Implementation of multi-head attention mechanism in transformer   Feb 2021 – Mar 2021
- Implementation and evaluation of a pre-trained ResNet18 image classification model with different visual perturbations   Jan 2021 – Feb 2021
- Implementation of a vanilla recurrent neural network   Dec 2020 – Jan 2021
- Diphone-based concatenative TTS system with Python   Nov 2020 – Dec 2020
- LSTM-based sentiment analysis for Cantonese text   Feb 2020 – May 2020
- Automatic data scraping from a politically sensitive online forum in Hong Kong   Oct 2019 – Feb 2020

AWARDS/CERTIFICATES

- First place in German Dialect Identification (GDI) task and Second place in simplified variant of Discriminating between Mainland and Taiwan variation of Mandarin Chinese (DMT) task. VarDial 2019, North American Chapter of the Association for Computational Linguistics (NAACL).
- Awarded a Scholarship of €1575 for a summer exchange programme at RWTH Aachen University from the DAAD University Summer Course Scholarship 2014. Deutscher Akademischer Austausch Dienst (DAAD).
- Awarded a piano performance diploma DipABRSM. The Associated Board of the Royal Schools of Music.
- Awarded a distinction in Grade 8 piano performance. The Associated Board of the Royal Schools of Music.
- International English Language Testing System (IELTS) Overall Band 8.0. Jan 2020.

OTHERS

Language Proficiency	Cantonese (Native), Mandarin (Fluent), English (Fluent), German (B2), Russian (Elementary)
Computer Science/ML	Python (preferred), Java, C++, PyTorch, SciPy, Pandas, Git, Bash, Kaldi
Softwares	Adobe Photoshop, Adobe Premiere Pro, Adobe After Effects, Praat, Katalon
Operating Systems	MacOS (preferred), Windows, Linux