

Sangeet Sagar

[Github](#) | [Google Scholar](#)

Email: sasa00001@stud.uni-saarland.de

Mobile: (+49) 15257 657433

Address: Saarbrücken– 66123, Germany

EDUCATION

- **Universität des Saarlandes** Saarbrücken, Germany
Master of Science in Language Science and Technology; GPA: 2.0 (ECTS, lower is better) Oct. 2020 – Feb. 2023
Thesis: Noise Robust Speech Recognition for Search and Rescue Domain
- **The LNM Institute of Information Technology** Jaipur, India
Bachelor of Technology in Electronics and Communication Engineering; GPA: 7.13/10.0 Aug. 2015 – June. 2019
Thesis: [Analysis of Emotion Recognition using Speech Features](#)

PROGRAMMING SKILLS

- **Languages:** Python, Unix, MATLAB, R
- **Toolkits/Libraries:** PyTorch, [SpeechBrain](#), GPT, Scikit-learn, Numpy, Huggingface, Matplotlib
- **Misc.:** GIT, Advanced Linux user, Extensive experience with cluster computing like SLURM

EXPERIENCE

- **Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH** Saarbrücken, Germany
Research Assistant | HiWi (1.6 years) June 2021 – Feb 2023
 - As a part of MS thesis, working on building a noise robust automatic speech recognition system (STT) (German language) capable to work under hostile conditions like search and rescue missions. Also, train RNN based large language models for speech recognition.
 - Trained open-source attention-based BiRNN punctuation restoration system+TruCasing for the German language. The system outperformed the baseline model- Vosk model by over 14% in recall metric.
- **Institute of Formal and Applied Linguistics, Charles University** Prague, Czechia
University research assistant (1 year) Oct. 2019 – Dec 2020
 - Principle tester of the live speech-language translation (SLT) system at live test sessions for the [ELITR project](#).
 - Training and testing (with in-domain/out-of-domain data) Czech punctuator system for live ASR.
 - Perform domain adaptation for live ASR by fine-tuning LM by using speaker-specific data.
 - Set up pipelines for the [ELITR test set](#) and to get it ready for evaluation.
- **Faculty of Information Technology, Brno University of Technology** Brno, Czechia
University research assistant (8 months) Feb. 2019 – Sept. 2019
 - Developed a system for cross-lingual topic identification in low resource scenarios for Kinyarwanda, Zulu, Hindi and Reuters corpus.
 - Achieved a remarkable weighted average precision score of 0.52 (on 10 different topics) on Kinyarwanda language by learning a simple linear transformation from this language to English language space.
 - Primary task: text feature extraction, training classifier, generating document embeddings, and learning transformation from unknown language to English space.

ACADEMIC PROJECTS

- **English and Chinese poetry generation** April 2022
English poetry generation and comparison (using PPL as evaluation metric) using LSTM, encoder-decoder based models, and transformer model (GPT). We conclude that fine-tuning GPT-2 model generated the highest quality poems i.e. with the least PPL score. Also trained a topic-prediction model to study how well a machine-generated poem is interpreted by a system trained on human-written poems. [\[Report\]](#) [\[Code\]](#)
- **Span extraction based slot-filling using attention and RNNs** April 2022
Performed slot-filling (and intent recognition) using RNN on a multi-head attention mechanism. The main idea was to model the slot-filling task as a span extraction and to utilize available information about slot type for which value is to be provided. F1 score of 0.83 was achieved compared to 0.96 of the baseline model. [\[Report\]](#) [\[Code\]](#)

- **Out-of-vocabulary(OOV) word estimation using subword representation** Aug. 2021
Achieved a better OOV rate and perplexity score than the baseline for three levels of granularity (char level, small, large vocab) with appropriate hyperparameter tuning. This was done by training RNN based language model to artificially generate corpus and compute OOV rate on varying sizes of the generated corpora. [Report] [Code]
- **[Bachelors Thesis] Analysis of Emotion Recognition using Speech Features** 2018
7% improvisation over baseline on the classification of speech signals based on human emotions like angry, disgust, fear, happy, etc. Implemented on SAVEE and Emo-DB datasets using classifiers like GMM, CNN, MLPNN, we propose the use of feature extraction algorithms like S-Transform and image spectrogram of the speech signal to perform emotion recognition. [Report] [Code]

FOREIGN LANGUAGES

- **Fluent:** English
- **Beginner:** German (A1)

PUBLICATIONS

- [1] Ondřej Bojar et al. “ELITR Multilingual Live Subtitling: Demo and Strategy”. In: *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*. Online: Association for Computational Linguistics, Apr. 2021, pp. 271–277. DOI: [10.18653/v1/2021.eacl-demos.32](https://doi.org/10.18653/v1/2021.eacl-demos.32). URL: <https://aclanthology.org/2021.eacl-demos.32>.
- [2] Ondřej Bojar et al. “ELITR: European Live Translator”. In: *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation*. Lisboa, Portugal: European Association for Machine Translation, Nov. 2020, pp. 463–464. URL: <https://aclanthology.org/2020.eamt-1.53>.
- [3] Dario Franceschini et al. “Removing European Language Barriers with Innovative Machine Translation Technology”. English. In: *Proceedings of the 1st International Workshop on Language Technology Platforms*. Marseille, France: European Language Resources Association, May 2020, pp. 44–49. ISBN: 979-10-95546-64-1. URL: <https://aclanthology.org/2020.iwlt-1.7>.
- [4] Santosh Kesiraju et al. *Bayesian multilingual topic model for zero-shot cross-lingual topic identification*. 2020. DOI: [10.48550/ARXIV.2007.01359](https://doi.org/10.48550/ARXIV.2007.01359). URL: <https://arxiv.org/abs/2007.01359>.
- [5] Dominik Macháček et al. *ELITR Non-Native Speech Translation at IWSLT 2020*. 2020. DOI: [10.48550/ARXIV.2006.03331](https://doi.org/10.48550/ARXIV.2006.03331). URL: <https://arxiv.org/abs/2006.03331>.
- [6] Peter Polák et al. “CUNI Neural ASR with Phoneme-Level Intermediate Step for ~Non-Native~SLT at IWSLT 2020”. In: *Proceedings of the 17th International Conference on Spoken Language Translation*. Online: Association for Computational Linguistics, July 2020, pp. 191–199. DOI: [10.18653/v1/2020.iwslt-1.24](https://doi.org/10.18653/v1/2020.iwslt-1.24). URL: <https://aclanthology.org/2020.iwslt-1.24>.