Hari Krishna Vydana

Conversational AI Researcher; ML enthusiast

Zurich, Switzerland
(+41)779860742
harivydana@gmail.com
Linkedin
Google scholar(H-index: 14)

Scientific Interests

Agentic AI, Spoken Dialog Systems, Speech Recognition, Large Language Models (LLMs), Multimodal LLMs, Conversational AI, Speech-to-Speech Translation, Spoken and Written Language Assessment (Ed-Tech), Machine and Speech Translation, Natural Language Processing (NLP), Text-to-Speech (TTS) Models, Knowledge Distillation

Experience

Dec 2024 – Present **Tech Lead – Voice AI**, *Telepathy AI Labs*, Zurich, Switzerland.
- ***Appointment.AI*** – Developing an AI-powered system for seamless appointment booking.
- ***ASR & Dialog Management*** – Worked on ASR, VAD, and Dialog Turn Detection modules.
- ***Agentic AI & LLMs*** – Built an LLM-driven agent for automated appointment scheduling.
- ***TTS Optimization*** – Integrated TTS models and explored cost-effective voice solutions.

Sept 2023 – Nov 2024 **Research Associate – ALTA Project**, *University of Cambridge, UK*.
- ***ALTA Project*** - Contributed to Automated Language Teaching and Assessment (ALTA).
- Developed ***Automatic evaluation of spoken summaries using Large Language Models (LLMs)***.
- Researched ***Automatic spoken language assessment with Human-in-the-Loop validation***.
- ***Distillation:*** – Distilled audio foundation models to optimize inference compute.
- ***LLM's For Analytic Scoring:*** – Proposed analytic scoring flaws and benchmarked on spoken assessment task.
- ***Multi-View Error Mitigation:*** – Leveraged LLMs for multi-view error correction in language assessment.

Oct 2022 – Aug 2023 **Senior Researcher**, *Cerence, Aachen, Germany*.
- Developed ***On-Device and Cloud ASR models*** for automotive applications.
- Worked on ***Training/Evaluating/Benchmarking*** ASR models targeted for Automotive Domains.

Aug 2021 – Sept 2022 **Senior ASR Researcher**, *Huawei Technologies Oy, R&D Center, Finland*.
- Designed ***On-Device ASR models*** for Huawei mobiles.
- Built ***Voice Search models*** for ***Petal Search*** in Huawei devices.

Education

Jan 2019 – July 2021 **Postdoctoral Researcher**, *Brno University of Technology (BUT), Brno, Czech Republic*.
Research Focus: Speech Translation Systems, End-to-End Speech Recognition
Advisors: Doc. Ing. Lukáš Burget & Prof. Doc. Ing. Jan Černocký

Dec 2014 – Dec 2018 **Ph.D. in Multilingual Speech Recognition**, *IIIT Hyderabad, India*.
Thesis: Multilingual Speech Recognition for Indian Scenarios
CGPA: 9.5/10 **Advisor:** Dr. Anil Kumar Vuppala

Jan – June 2016 **Teaching Assistant**, *International Institute of Information Technology Hyderabad, India*.
Course: Speech Systems **Advisor:** Dr. Anil Kumar Vuppala

Research Projects

- **Leveraging LLMs for Feedback and Error Mitigation**

  "Using LLMs for developing Feedback systems for language assessment."

  - Used **Text/Multimodal LLMs** for multi-level feedback in spoken/written language assessment.
  - **Stefano Bano, Hari Krishna Vydana, Kate Knill, Mark J.F. Gales.** ***"Can GPT-4 Perform Analytic Assessment of L2 Writing?"*** In Proc. Workshop on Innovative Use of NLP for Building Educational Applications, ACL 2024.
  - Developed **multi-view error mitigation systems** using LLMs.
  - Built **neural text graders (BERT-based models)** for language assessment.

- **Transformer-based Speech Translation Systems**

  "Explored approaches to integrate ASR with NLP modules, developing models with an end-to-end differentiable path between ASR and NLP modules (MT). These models reduce error propagation across the pipeline."

  - Developed Cascaded, End-to-End systems, and Jointly-optimized ASR-MT systems.
  - **Hari Krishna Vydana**, Martin Karafi'at, Lukáš Burget, Honza Černocký. ***"The IWSLT 2021 BUT Speech Translation Systems"***, Proceedings of the 18th International Conference on Spoken Language Translation (IWSLT 2021).
  - **Hari Krishna Vydana**, Martin Karafi'at, Katerina Zmolíková, Lukáš Burget, Honza Černocký. ***"Jointly Trained Transformer Models for Spoken Language Translation"***, Proc. ICASSP, Toronto, Canada, 2021.

- **End-to-End Automatic Speech Recognition Systems**

  - **Speech Recognition for Indian Scenarios:**

    "Developed ASR systems for code-mixing/code-switching environments. Proposed a model that handles recognition and text rendering separately to operate in code-mixed contexts."

    - **Hari Krishna Vydana**, Krishna Gurugubelli, Vishnu Vidyadhara Raju V, and Anil Kumar Vuppala. ***"An Exploration Towards Joint Acoustic Modeling for Indian Languages"***, Interspeech 2018, Hyderabad, India.
    - **Thesis:** ***"Salient Features for Multilingual Speech Recognition in Indian Scenarios"***.

  - **Combining Hybrid and End-to-End ASR Models:**

    "Trained End-to-End ASR systems and combined them with hybrid models to enhance performance."

    - Martin Karafiát, Murali Karthick Baskar, Igor Szöke, **Hari Krishna Vydana**, Karel Veselý, and Jan Černocký. ***"BUT Opensat 2019 Speech Recognition System"***, arXiv:2001.11360, 2020.
    - Katerina Zmolikova, Martin Kocour, Federico Landini, Karel Beneš, Martin Karafiát, **Hari Krishna Vydana**, Alicia Lozano-Diez, Oldrich Plchot, and Murali Karthick Baskar. ***"BUT System for CHiME-6 Challenge"***, ICASSP 2020, Barcelona, Spain.
    - **Hari Krishna Vydana**, Sivanand Achanta, and Anil Kumar Vuppala. ***"Incorporating Speaker Normalization to End-to-End ASR"***, SLTU 2018, India.
    - Ekaterina Egorova, **Hari Krishna Vydana**, Lukáš Burget, and Jan Černocký. ***"Out-of-Vocabulary Words Detection in End-to-End ASR"***, Interspeech 2021, Brno, Czech Republic.
    - Ekaterina Egorova, **Hari Krishna Vydana**, Lukáš Burget, and Jan Černocký. ***"Spelling-Aware Word-Based End-to-End ASR"***, IEEE Signal Processing Letters, 2022.

  - **Self-Supervision & End-to-End ASR Models**
    - Self-training: Using ASR to label unsupervised data and improving performance via auto-labeled data.
    - Exploring semi-supervised models for ASR (Speech-BERT, Wav2Vec, HuBERT) and MT (Mbart).

- **Language Identification systems**

  "Explored implicit and explicit models for developing Language Identification systems"

  - Tirusha Mandava, **Hari Krishna Vydana**, Ravi Kumar Vuddagiri, Hari Krishna Vydana, Anil Kumar Vuppala, ***"An Investigation of LSTM-CTC based Joint Acoustic Model for Indian Language Identification"***, ASRU 2019, Sentosa, Singapore 2019.
  - Brij Mohan Lal Srivastava, **Hari Krishna Vydana**, Anil Kumar Vuppala and Manish Shrivastava, ***"Significance of neural phonotactic models for large-scale spoken language identification"***, IJCNN, Alaska USA, May 2017.

Evaluations participated & Open-source Repositories

- **IWSLT-2020, IWSLT-2021**: Offline speech Translation
- ASR Challenge for Indian English-2021: First place
- **OPENSAT-2019 & OPENSAT-2020** : First place in ASR
- CHiME-6 Workshop (satellite of ICASSP 2020)
- Low Resource Speech Recognition Challenge for Indian Languages (Interspeech 2018 Special Session)

Open-source Repositories.

- Transformer ASR:
- Transformer-Machine Translation:
- End-to-End ASR:
- Jointly-Trained Transformers for Spoken Language Translation:

Skills

Speech-Processing	Kaldi, EESEN, ESPNET, FAIRSEQ, Nemo, Wenet, Huggingface, AZURE, OPENAI, FASTAPI
Machine-Learning	PyTorch, Numpy, Scipy, Pandas, Matplotlib, PyTorch Lightening, Databricks

Publications

Link for the complete list of papers:, *Google scholar*.