Martin Vaculik

Hustopece, Druzstevni 17 xvaculm00@stud.fit.vutbr.cz 702 204 480
https://github.com/Vacko32

About Me!

Final-year CS student at VUT FIT Brno specializing in speech processing and multimodal LLMs, with a focus on training and evaluating large-scale neural models. I work on end-to-end systems, from data preprocessing to model evaluation, and enjoy pushing current architectures beyond standard baselines.

Education

BS **VUT FIT**, Computer Science – Brno, CR Sept 2023 – present

- Placing in first top 10% of students based on GP, 3rd semester reaching top 5% for students and getting stipendium.

Experience

Conference Support, VUT FIT – Brno VUT FIT May 2025 – Sept 2025

- Worked as a support staff member at the JSALT Research Conference in Brno. My responsibilities focused on ensuring researchers and interns were comfortable and well supported. This role exposed me to many current trends, state-of-the-art research workflows and topics in the audio and NLP fields.

Projects

Extending LLM to process speech – **Bachelor's Thesis** Jan 2026 – present

Speech-LLM Alignment for End-to-End ASR

- Connected frozen Whisper Encoder to a frozen LLM via a trainable connector Trained on 7k hours with DDP. Achieved 2.76% WER on LibriSpeech (clean), 9.89% on CommonVoice, and 8.91% on How2, matching performance of whisper-small.

Machine Translation From Scratch Winter 2025

- Encoder-decoder Transformer in pure PyTorch for En→De on WMT19. Conducted systematic ablation experiments comparing RMSNorm, LayerNorm, DyT, and Derf normalization strategies, analyzing their impact on convergence and translation quality. Inspired by DyT paper.

Other Projects

- LLM pretraining from scratch — BPE tokenizer and GPT-style training pipeline using only PyTorch primitives (Stanford CS336 Inspired)
- Geospatial clustering and hypothesis testing on Czech Police dataset using GeoPandas

Skills & Knowledge

Programming: Python, Pytorch, Numpy, Huggingface (datasets, transformers), basic familiarity with C, C++ and CUDA

ML: Probability & Statistics, hypothesis testing, supervised & unsupervised learning, clustering, model evaluation metrics, DER curve, regularization, KL Divergence, optimalization

NLP: N-gram and count-based models, TF-IDF & BM25 retrieval, Transformers (Bert & GPT), RLHF & DPO, evaluation metrics (BLEU, perplexity), multimodality, BPE tokenization, decoding techniques

Audio/Speech: WavLM & Whisper encoders, neural ASR, speaker verification (x-vectors, DET curves), diarization, LPC, basic exposure to vocabulary-based ASR systems, DTW and HMM

TeamWork: Git, Weights & Biases