Taras Sereda. ML Researcher

👋 Hi, I'm a **Machine Learning researcher** based in San Francisco, USA. I have over **10 years of experience** in tech. My research focus is ML for audio. I'm currently a visiting scholar at UCU in Lviv, Ukraine, where I give lectures on ML for audio and supervise graduate and undergraduate students.

Contacts and useful links

📧 taras.y.sereda@proton.me

✍️ https://taras-sereda.github.io

🔗 linkedin.com/in/tarassereda

👾 github.com/taras-sereda

Experience

Visiting Scholar

Ukrainian Catholic University *(Mar 2022 - present)*

- Giving lectures on speech and audio processing. Focusing on text-to-speech(TTS), automatic speech recognition(ASR) and source separation(SS).
- Providing mentoring and advising to students, guiding them through academic challenges.

deep learning for audio course: https://github.com/taras-sereda/deep-learning-for-audio

TTS lecture notes: lecture1, lecture2

ML Consultant

Self-employed *(Sep 2022 - present)*

Providing consultancy services for projects focused on developing text-to-speech (TTS) and automatic speech recognition (ASR) systems.

Featured Projects:

- Real-time Automatic Speech Recognition (ASR) for low SNR (-10 to 0 dB) audio. The developed system runs on edge devices, performing simultaneous voice activity detection and speech recognition. We are surpassing open-source state-of-the-art ASR models by a significant margin on in-house domain-specific data.
- Conversational text-to-speech system for call center automation, developed in collaboration with PolyAI. Our model features zero-shot speech synthesis on discrete semantic and acoustic tokens, running faster than real-time on A100 GPUs. Both text-based and acoustic-based prompting are supported to control prosody and voice characteristics of generated. speech.

paper:
https://arxiv.org/abs/2401.02839v1
code:
https://github.com/PolyAI-LDN/pheme

Co-founder, Director of Research

Vocalics.ai *Kyiv, Ukraine – (Feb 2019 - Sep 2022)*

I co-founded Vocalics.ai with the idea of multilingual speech synthesis in mind when a speaker's identity and unique speaking style are preserved with voice translation to world languages.

As a Director of research:

- Developed a multi-lingual speech synthesis system that preserves speaker identity and unique speaking style together with methodology for quality evaluation of generated speech in dimensions of intelligibility, speaker and prosody similarity.

- Led and managed a team of 4 researchers in the development of cutting-edge speech synthesis technologies.

- Reimplemented, tested, and customized multiple SOTA papers from the speech synthesis field, including autoregressive and parallel TTS approaches; Flow/Diffusion based generative modelling approaches for audio.

As a Co-founder:

- Successfully recruited and managed a team of experienced professionals to build out the technical infrastructure and expand the company's capabilities.

- Conducted customer interviews and market research to validate product-market fit and inform product development.

ML Engineer

Whisper.ai *San Francisco, CA, USA – (Feb 2018 - Jul 2018)*

Developed advanced audio source separation models based on neural networks(PIT, Chimera network). Eventually, models were deployed on ARM devices, to improve listening experience for those with mild hearing loss to hear better in noisy environments (restaurants/public transport/public spaces)

ML Researcher

Ring Labs *Kyiv, Ukraine – (Sep 2016 - Apr 2017)*

First employee in Ukrainian R&D office, hired to establish ML department, was leading a research team. Working on object detection algorithms based on R-CNN, YOLO, with highly customized in-house implementations.

ML Engineer (Part time)

IPGraphy *Kyiv, Ukraine – (Oct 2015 - Jun 2016)*

Developed core algorithm for visual object similarity search applied for images of trade marks. IP rights NN based tool, aimed to help attorneys achieve higher productivity.

ML Engineer

DepositPhotos *Kyiv, Ukraine — (Dec 2014 - Dec 2015)*

Developed core algorithms for virtual stylist app.

- Recommender system, suggests how to combine clothes of your wardrobe
- Neural networks for automatic clothes classification and wardrobe categorization

Skills

- Python
- Rust
- Java
- Databases and SQL

- CUDA programming
- Pytorch
- Tensorflow
- Algorithm and data structures

- Performance tuning
- Project Management
- Problem solving

Languages

Ukrainian 🇺🇦

Native speaker

English 🇺🇸

Proficient speaker

Courses and Talks

- Course **Deep Learning for Audio** Projector 2019, UCU Summer Schoold 2020
- **"VocalicsAI. Artificial intelligence removes language barriers"** Talk at FW days, 2021
- **"Waveglow. Generative modeling for audio synthesis"** Talk at FW days, 2019

Education

Masters's Degree in Mathematical Modelling

2011 - 2012

KNEU - Kyiv, Ukraine

Bachelor's Degree in Cybernetics

2007 - 2011

KNEU - Kyiv, Ukraine