







6+ years in developing machine learning models for audio signal processing, natural language processing, and information retrieval, and 10+ years in software development. Proficiency in machine learning, audio/text processing, data analysis, web development, and Python & Java programming.


Personal Info


 Huang Xie


 huang.xie@outlook.com

 +358 40 3219383

 Tampere, Finland

 github.com/xieh97

 linkedin.com/in/huang-xie-28b7872bb

 orcid.org/0000-0001-7609-8049

Languages

Chinese	Native
English	Professional
Finnish	Beginner

Referees

Prof. Tuomas Virtanen
@ Tampere University
✉ tuomas.virtanen@tuni.fi

Assoc. Prof. Okko Räsänen
@ Tampere University
✉ okko.rasanen@tuni.fi

Employment

Doctoral Researcher | Research Assistant 03/2019 - Present

📍 Audio Research Group, Tampere University, Finland

- Developed machine learning models (e.g., CNNs, RNNs, transformers) for multimodal learning, audio signal processing, natural language processing, and information retrieval.
- Collected, analyzed, and processed large audio and text datasets (e.g., feature extraction, semantic analysis, classification, clustering).
- Published 9+ first-authored scientific papers in prestigious journals and conferences (e.g., IEEE/ACM TASLP, SPL, ICASSP).

Software Engineer 10/2015 - 05/2018

📍 Bohai Commodity Exchange, Tianjin, China

- Developed key features for CloudBoce (🔗, an E-Commerce platform), including shopping cart, order processing, and product management.
- Collaborated with cross-functional teams (e.g., UI designers, product managers) to define, design, and ship new features.
- Troubleshoot and debugged issues related to performance, crashes, and other bugs, ensuring smooth and stable user experience.

Principal Java Developer 07/2014 - 09/2015

📍 Qingdao Industrial Software Research Institute, Qingdao, China

- Led the development and delivery of multiple enterprise-grade software solutions (e.g., BPM, CMS), overseeing the full project lifecycle from planning to deployment.
- Mentored junior developers and performed code reviews to maintain high coding standards and best practices.
- Collaborated with cross-functional teams to define requirements, troubleshoot issues, and deliver high-quality software on schedule.
- Presented technical solutions to non-technical stakeholders, and managed technical documentation and reports.

Skills

Programming: Python, Java, JavaScript, SQL, C/C++, Kotlin

Machine Learning: PyTorch, Tensorflow, scikit-learn, Ray, Spark

Data Analysis: NumPy, SciPy, Pandas, NLTK, Matplotlib

Web Development: Java EE, Spring, Hibernate, Django, Flask, MySQL, Git

Education

PhD in Signal Processing and Machine Learning 01/2021 - Present

📍 Tampere University, Tampere, Finland

M.Sc. in Data Engineering and Machine Learning 08/2018 - 11/2020

📍 Tampere University, Tampere, Finland

M.Eng. in Software Engineering 09/2010 - 06/2014

📍 University of Science and Technology of China, Hefei, China

Activities

- 👤 Reviewer for IEEE/ACM TASLP, SPL, ICASSP, WASPAA, DCASE, etc.
- 👤 Task coordinator for Language-based Audio Retrieval and Automated Audio Captioning in DCASE Challenge 2022 ([🔗](#)), 2023 ([🔗](#)), and 2024 ([🔗](#)).

Publications [\(full list 🔗\)](#)

- [1] Text-based Audio Retrieval by Learning from Similarities between Audio Captions
👤 **H. Xie**, K. Khorrami, O. Räsänen, and T. Virtanen
📅 2024 📖 IEEE Signal Processing Letters 🔗 [arXiv](#)
- [2] Integrating Continuous and Binary Relevances in Audio-Text Relevance Learning
👤 **H. Xie**, K. Khorrami, O. Räsänen, and T. Virtanen
📅 2024 📖 in Proc. Detect. Classif. Acoust. Scenes Events Work. (DCASE) 🔗 [arXiv](#)
- [3] Multi-label Zero-Shot Audio Classification with Temporal Attention
👤 D. Dogan, **H. Xie**, T. Heittola, and T. Virtanen
📅 2024 📖 in Proc. Int. Workshop Acoust. Signal Enhanc. (IWAENC) 🔗 [arXiv](#)
- [4] Crowdsourcing and Evaluating Text-Based Audio Retrieval Relevances
👤 **H. Xie**, K. Khorrami, O. Räsänen, and T. Virtanen
📅 2023 📖 in Proc. Detect. Classif. Acoust. Scenes Events Work. (DCASE) 🔗 [arXiv](#)
- [5] On Negative Sampling for Contrastive Audio-Text Retrieval
👤 **H. Xie**, O. Räsänen, and T. Virtanen
📅 2023 📖 in Proc. Int. Conf. Acoustic., Speech and Signal Process. (ICASSP) 🔗 [arXiv](#)
- [6] Language-based Audio Retrieval Task in DCASE 2022 Challenge
👤 **H. Xie**, S. Lipping, and T. Virtanen
📅 2022 📖 in Proc. Detect. Classif. Acoust. Scenes Events Work. (DCASE) 🔗 [arXiv](#)
- [7] Unsupervised Audio-Caption Aligning Learns Correspondences Between Individual Sound Events and Textual Phrases
👤 **H. Xie**, O. Räsänen, K. Drossos, and T. Virtanen
📅 2022 📖 in Proc. Int. Conf. Acoustic., Speech and Signal Process. (ICASSP) 🔗 [arXiv](#)
- [8] Zero-Shot Audio Classification using Image Embeddings
👤 D. Dogan, **H. Xie**, T. Heittola, and T. Virtanen
📅 2022 📖 in Proc. Eur. Signal Process. Conf. (EUSIPCO) 🔗 [arXiv](#)
- [9] Zero-Shot Audio Classification with Factored Linear and Nonlinear Acoustic-Semantic Projections
👤 **H. Xie**, O. Räsänen, and T. Virtanen
📅 2021 📖 in Proc. Int. Conf. Acoustic., Speech and Signal Process. (ICASSP) 🔗 [arXiv](#)
- [10] Zero-Shot Audio Classification via Semantic Embeddings
👤 **H. Xie**, and T. Virtanen
📅 2021 📖 IEEE/ACM Trans. Audio Speech Lang. Process. (TASLP) 🔗 [arXiv](#)
- [11] Zero-Shot Audio Classification Based on Class Label Embeddings
👤 **H. Xie**, and T. Virtanen
📅 2019 📖 in Proc. Work. Appl. Signal Process. Audio and Acoustic. (WASPAA) 🔗 [arXiv](#)