Srikanth Cherla

srikanth@cherla.org +46 735166693 linkedin.com/in/srikanthcherla cherla.org
Vellinge, Sweden

Professional Summary

Senior Machine Learning Engineer with 18+ years of experience building production AI systems across LLMs, agentic AI, generative models, music technology, and computer vision. Founding team member of Unity AI Assistant — an agentic LLM system embedded in the Unity Editor, serving Unity developers worldwide. Pioneered the world's first commercially deployed generative AI music composer at Jukedeck (acquired by TikTok/ByteDance, 2019). PhD in Computer Science with 20+ peer-reviewed publications at top ML and music AI venues.

Core Technical Skills

ML & Deep Learning: PyTorch, TensorFlow, Keras, Scikit-learn, HuggingFace Transformers, LoRA fine-tuning, diffusion models, RNNs, CNNs, Transformers, Restricted Boltzmann Machines
LLMs & Agentic AI: OpenAI API, LlamaIndex, prompt engineering, function calling, agentic workflows, MCP
Data & Analysis: Pandas, NumPy, Jupyter, SQL, Apache Spark, Databricks
MLOps & Infrastructure: Weights & Biases, MLflow, Docker, Kubernetes, Grafana, CI/CD
Cloud Platforms: AWS, GCP, Azure
Programming: Python, Bash
Domains: Generative AI, Large Language Models, Supervised/Unsupervised Learning, Reinforcement Learning, Computer Vision, Audio/Music AI, Recommender Systems

Professional Experience

Senior Machine Learning Engineer Nov 2020 – Present
Unity Technologies, Copenhagen, Denmark

Unity AI Assistant — Founding Team Member (Mar 2023 – Present)

- Founding member of the core team that evolved Unity AI Assistant from a web-based Q&A chatbot into an agentic AI system embedded in the Unity Editor, serving Unity developers worldwide
- Introduced tool-use and function-calling capabilities, enabling the assistant to actively perform tasks within the Unity Editor on behalf of developers
- Prototyped and deployed LlamaIndex-based agentic workflows for complex, multi-step developer request handling
- Architected MCP (Model Context Protocol) Client Integration atop the assistant's native tool-calling system
- Designed the Unity AI Assistant Skills feature, extending the assistant's capability surface to third-party integrations
- Designed a feature for AI-assisted automated testing within the Unity Editor
- Collaborated cross-functionally with product teams across Unity to AI-enable their workflows and products

AI Visual Asset Generation (Jan 2023 – Mar 2023)

- Implemented experiment tracking with Weights & Biases for a diffusion model-based visual content generation pipeline

- Introduced LoRA fine-tuning adaptor (HuggingFace) for domain-specific visual asset generation

Data Quality Monitoring — Unity Ads ML Pipeline (Jul 2022 – Jan 2023)

- Analysed training data characteristics and visualised time-varying feature statistics using Grafana to detect data drift in the Unity Ads ML pipeline
- Enabled proactive data quality management, surfacing upstream data issues before they caused silent model degradation

AI Tools for Game Designers — Research Lead (Nov 2020 – Jun 2022)

- Led the AI feature of a Unity Labs R&D initiative building ML-driven tools to assist game designers
- Developed stochastic, search-based AI agents that simulated diverse player behaviours within game economy designs, enabling designers to analyse and balance in-game economies

Senior Research Scientist Jul 2019 – Nov 2020

Moodagent A/S, Copenhagen, Denmark

- Re-architected the music recommendation data pipeline from Pandas to Apache Spark to enable production-scale data processing and model training
- Developed and improved supervised and collaborative filtering models across songs, playlists, and artists to enhance personalisation quality
- Led the adoption of software development best practices in the ML research team to improve collaboration and code quality

Machine Learning Researcher Jan 2016 – May 2019

Jukedeck Ltd., London, UK (acquired by TikTok/ByteDance, 2019)

- Co-developed the world's first commercially deployed generative AI music composer, generating music conditioned on user-specified style and mood parameters
- Designed, trained, and delivered novel deep learning (RNN) architectures for structured music generation at production scale
- Designed and implemented evaluation frameworks and quantitative metrics for assessing AI-generated content quality
- Published research on structure-aware melody generation at ISMIR 2018 (StructureNet)
- Represented the company at public conferences and events; authored research papers, blog posts, and internal reports

Earlier Research Roles 2007 – 2012

- **Research Engineer**, Siemens Corporate Technology, Bangalore (Jul 2007 – Jun 2010): Designed computer vision algorithms for human action recognition and video stitching; developed audio analytics systems for security/surveillance; published at CVPR Workshop 2008 and INTERSPEECH 2010
- **Research Assistant**, Simon Fraser University, Vancouver (Sep 2011 – May 2012): Physical modelling of the tenor saxophone using digital waveguide synthesis; published at SMC 2012
- **Research Intern**, Universitat Pompeu Fabra / PMC Technologies, Barcelona (Jan – Mar 2011): Built probabilistic models for failure prediction in semiconductor manufacturing processes (Infineon project); work published in IEEE/ASME Transactions on Mechatronics (2013)

Education

PhD, Computer Science *2012–2016*

City, University of London • Thesis: *Connectionist Models for Classification and Sequence Learning*
Fully funded studentship (fee waiver, salary & travel allowance) • Best Student Paper Award, ISMIR 2013

MSc, Sound and Music Computing *2010–2011*
Universitat Pompeu Fabra, Barcelona • GPA: 9.24/10 • Music Technology Group • 75% fee waiver scholarship

BTech, Computer Science & Engineering *2003–2007*
IIIT Hyderabad • Centre for Visual Information Technology

Selected Publications (20+ total — cherla.org/publications)

- **StructureNet: Inducing Structure in Generated Melodies** — Medeot, Cherla et al., ISMIR 2018
- **Generalising the Discriminative Restricted Boltzmann Machine** — ICANN 2017
- **Discriminative Learning and Inference in the Recurrent Temporal RBM for Melody Modelling** — IJCNN 2015
- **A Distributed Model for Multiple-viewpoint Melodic Prediction** — ISMIR 2013 *(Best Student Paper Award)*
- **Automatic Phrase Continuation from Guitar and Bass-guitar Melodies** — *Computer Music Journal* 34:3, MIT Press, 2013
- **Audio Analytics by Template Modeling and One-pass DP Based Decoding** — INTERSPEECH 2010
- **Towards Fast, View-Invariant Human Action Recognition** — CVPR Workshops 2008

Leadership & Service

Editorial: Review Editor, Frontiers in Digital Humanities (Digital Musicology) • ISMIR Reviewer (2014, 2015, 2018, 2019) • IJCAI Sub-reviewer (2015)
Mentoring: Women in MIR (WiMIR) Mentor (2018–2020, 2023–2026) • Supervised 2 graduate/undergraduate researchers (Siemens & City)
Teaching: Teaching Assistant, City University London (2013–2016) — Machine Learning, Neural Computing, C++, Computation & Reasoning
Organisation: Program Chair, City Informatics PhD Symposium (CIPS) 2014
Speaking: Panelist, CSMC: "Applying Music Patterns in Generation" • Attendee, Dagstuhl Seminar on Neural-Symbolic Learning & Reasoning (2014)

Awards & Recognition

• Best Student Paper Award, ISMIR 2013 • Fully funded PhD studentship, City University London • 75% fee waiver scholarship, Universitat Pompeu Fabra