

Sertan Şentürk

Data Scientist & Machine Learning Engineer

London, UK

Website: sertansenturk.com
Email: sertansenturk@gmail.com
LinkedIn: [sertansenturk](https://www.linkedin.com/in/sertansenturk)
GitHub: github.com/sertansenturk
Phone: 07848 884 649

I am pursuing a career in data science & machine learning engineering, specializing in MLOps, information retrieval and audio analysis. I have extensive knowledge and experience in entity recognition, digital signal processing, and machine learning. As part of my endeavors, I aim to design and build scalable & robust solutions for the automatic description, retrieval and discovery of big data.

EXPERIENCE

Deliveroo

Senior Machine Learning Engineer - MLOps

London, UK

Aug 2022 —ongoing

- Capture the use cases and requirements of an ML observability service for all ML models across the organization, and lead the scoping and design. **Proposed Technologies: TensorFlow serving, kafka, seldon alibi-detect, Datadog**
- Design and implement a GitOps-based CI/CD framework to deploy experiments parallel to existing ML models in staging and production. The framework is adopted by more than 10 teams within a month, covering $\approx 40\%$ of the ML pipelines and significantly improved productivity by standardizing and automating experiment deployment processes across the organization. **Technologies: Argo CD, Argo workflows, GitHub Actions, CircleCI, kubernetes**
- Assist and advise teams to troubleshoot and improve their data processing and ML training pipelines.

Kobalt Music Group

Lead Data Scientist

London, UK

Jul 2019 —May 2022

- As the technical lead of the data science team, designed & built an entity matching service, which could scale to **100s of millions** of instances of metadata. **Technologies: Scala/Spark, ElasticSearch, sklearn, Docker, AWS Sagemaker, EMR, step functions, terraform, lambda**
- Investigated attention-based deep learning models and locality sensitive hashing to match large-scale semi-structured and unstructured metadata **Technologies: PyTorch, EMR, Sagemaker, Scala/Spark**
- With the aid of product and business, identified underreported royalties using metadata matching, which is estimated to recover **\$1M** revenue for the clients
- With product, captured metadata matching requirements in 20+ use cases across the business
- Technical supervision of data engineers, data scientists and interns in 5+ projects
- Evaluated external solutions for metadata matching, audio autotagging, version identification & fingerprinting
- Oversaw internship, learn & dev, and academic outreach efforts

Senior Data Scientist – R&D

Jan 2018 —Jun 2019

- Developed & deployed a matching app in Python, which achieved a MAP of **>0.93** on more than 8 tasks and reduced process time from 2 days to **<2 hours**. **Technologies: ElasticSearch, AWS batch, terraform**
- Deployed a serverless machine learning lifecycle management tool to AWS. **Technologies: mlflow, AWS, ECS, terraform**
- Embedded into 4 different teams and assisted them to achieve their data-driven objectives
- Applied research on state-of-the-art metadata matching methods

New York University, Abu Dhabi

Data Science Consultant

Abu Dhabi, UAE (remote)

Jun 2019 —ongoing

- Provide technical support and training (e.g. Docker, TDD, experiment tracking, machine learning model versioning) to the [Music and Sound Cultures Research Group](#)
- Develop music information retrieval methods for the computational analysis of various music traditions

Universitat Pompeu Fabra

Doctoral Researcher

Barcelona, Spain

Sep 2011 —May 2017

- Developed state-of-the-art signal processing and machine learning methods to describe audio recordings, music scores and metadata
- Implemented the methods in Python in a modular, easy-to-use and open source toolbox ([link](#)) and helped to deploy it to a web-based music discovery app ([link](#)).

SKILLS

- **Professional Interests:** Data Science, Machine Learning, Big Data, MLOps, Signal Processing, Deep Learning
- **Programming Languages:** Python, Scala, C/C++
- **Cloud Computing (AWS):** EMR, Sagemaker (Studio, Ground Truth, Batch), step functions, OpenSearch, lambda, ECR, ECS, EKS, Athena, Glue
- **MLOps & CI/CD Tools:** Docker, Circle CI, GitHub Actions, Argo workflows, Argo CD, mlflow, metaflow, k8s, terraform
- **ML and Big Data Libraries:** Spark, scikit-learn, Elasticsearch, XGBoost, pandas, librosa, tensorflow, pyTorch
- **Programming:** test driven development, continuous integration and deployment, code coverage, trunk-based development
- **Planning and Management:** agile, IC supervision & mentoring, requirements management, user story mapping

EDUCATION

Universitat Pompeu Fabra Ph.D. in Information, Communication and Audiovisual Technologies, Advisor: Xavier Serra	Barcelona, Spain 2011—2017
Georgia Institute of Technology MSc. in Music Technology.	Atlanta, USA 2009 —2011
Middle East Technical University B.S. in Electrical and Electronics Engineering	Ankara, Turkey 2004 —2009
Hacettepe University Ankara State Conservatory Certificate of Achievement in Classical Guitar	Ankara, Turkey 2005 —2009

ACADEMIC WORKS AND SERVICES

- **30+** papers on audio signal processing and machine learning <http://sertansenturk.com/research/publications/>
 - **500+** citations listed in [Google Scholar](#)
- Reviewer in top MIR and ML venues, e.g. **NeurIPS**, **ICML**, **ICLR**, **ISMIR**
 - Highlighted reviewer at **NeurIPS 2020** and **ICLR 2022**
- *Industry and Sponsorship Co-Chair.* **22th International Society of Music Information Retrieval Conference** 2021

SELECTED TALKS

- Queen Mary University of London Jul 2020
Working in the Industry: A Personal Experience
- Berlin Music Information Retrieval Meetup Sep 2019
Metadata problems in the music industry (and how to match music metadata)
- London Elasticsearch Meetup May 2019
Metadata Matching Using Elasticsearch

OUTREACH AND MENTORSHIP

- *Co-organizer.* London Audio and Music AI Meetup 2020 —ongoing
- *Mentor.* Women in Music Information Retrieval 2017 —ongoing
- 10+ mentees since 2014 on *machine learning, audio signal processing, software development, and sustainability*

LANGUAGES

- **English:** fluent
- **Spanish:** upper intermediate
- **Turkish:** native
- **German:** pre-intermediate

HOBBIES

- Literature
- Cooking
- Skiing
- Travelling