

SAM CONSIDINE

Stratford, London | sambconsidine@gmail.com | 07808813731 | [LinkedIn](#) | [GitHub](#)

PROFILE

Machine learning engineer with 5 years of industry experience. Passionate about solving challenging real-world problems using technology. Strong background in computer vision, neural search architecture, and language models. Seeking opportunities to increase our collective intelligence by building large-scale AI systems.

SKILLS

Python, PyTorch, Hugging Face (Transformers, Datasets, PEFT, Accelerate, Candle), NLP, PyTorch Geometric, Data Analysis tools (Pandas, Matplotlib, Seaborn, Polars), Big Data (PySpark, AWS Glue, AWS EMR, AWS Athena, Hadoop, HDFS), databases (SQL, PostgreSQL, MongoDB) NumPy, Cython, C, C++, Rust, Git.

EDUCATION

University of Cambridge, MPhil Advanced Computer Science

Oct 2021 - Jul 2022

- **Grade:** Distinction — **Thesis title:** Neural Algorithmic Reasoning for Pseudotime Trajectory Inference.
- **Awarded best Master's thesis of the year by the faculty committee.**

University of York, BSc Mathematics

Sept 2013 - July 2017

- **Grade:** First-Class Honours — **Dissertation title:** Computer Vision Based Feature Extraction for Assessing Cell Heterogeneity.

EXPERIENCE

Interim CTO, Noggin HQ

September 2023 – Present

- Designed and built our main product in PyTorch: a bidirectional encoder credit scoring model that uses personal bank transaction records to create neural representations of borrower behaviour.
- Sourced datasets totalling 7 billion bank transaction records to train the model, resulting in possibly the most accurate consumer credit scoring model in the world, with error 50% lower than Experian.
- Communication with investors, as well as full evaluation of the business feasibility and tradeoffs of various technical approaches with respect to company objectives.
- Built the company data ETL pipeline to train neural network models at scale using AWS Glue, EMR, Athena and Sagemaker, as well as custom infrastructure on EC2 to perform distributed training.
- Employed 'infrastructure as code' MLOps, using Terraform to deploy an API that could access our credit model through a Sagemaker endpoint.
- Found and assisted with hiring a new CTO and a researcher to continue work on the model.

Machine Learning Engineer, Loci

November 2022 – September 2023

- Built our main product - a multimodal search and recommendation engine for game assets, which was essential in securing our first customers, including one of the world's biggest gaming companies, a crucial component in gaining £4m in seed funding as a team of 4.
- Increased the efficiency of our existing zero-shot game asset classification system by 3 orders of magnitude while improving code simplicity by creating a custom vector store + caching system.
- Re-implemented and fine-tuned large multimodal language/image transformers (OpenCLIP-2B) on data from 3D assets, building a data pipeline, allowing us to train, benchmark and deploy models.

- Mentored a team of 6 Masters students in a collaboration with Imperial College to build a system for 3D texture synthesis from a text prompt using stable diffusion, UNet and some custom rendering.

Research Assistant, University of Cambridge

July 2022 – October 2022

- Hired to continue research on my thesis and to build a platform to be used by clinicians to use machine learning systems to assist with diagnosis.

Senior Data Scientist, Arca-Blanca

May 2021 – October 2021

- Worked for a Machine Learning and Data Science consulting firm, consulting with executives to create and assist with predictive modelling tasks throughout their enterprise.

Machine Learning Engineer, After the Off

September 2017 - January 2021

- Lead developer on our main horse racing product: using machine learning on live data feeds to generate in-play odds for horse races, establishing us as a leading provider of in-play horse racing odds.
- Designed and led a project using computer vision to derive horse positions on a race track in real time using only a live video feed of the race. Built the whole pipeline, from video feed to prediction.

Research Assistant, University of York

July 2017 - September 2017

- Hired to continue work on my dissertation, building computer vision based feature extraction to aid in cancer research.

Self-Employed, Uptodata

June 2012 – December 2016

- Started a successful marketing company as a teenager, providing marketing data to wholesalers.