



Edinburgh, United Kingdom
elena.marchetti@gmail.com
+44 7823 556 194
linkedin.com/in/elenamarchetti

Data scientist with five years of experience building ML models that ship to production, not just notebooks. Currently working on pricing optimisation at an energy company, where the models I maintain influence £290 million in annual wholesale trades. Previously did NLP work at a legal-tech startup.

Elena Marchetti

DATA SCIENTIST

EDUCATION

MSc in Data Science

University of Edinburgh, Edinburgh
Sep 2018 – Nov 2019

Distinction. Thesis on *transfer learning for low-resource text classification in legal documents*. Received the Informatics MSc Prize for best dissertation.

BSc (Hons) in Mathematics

University of Edinburgh, Edinburgh
Sep 2015 – Jun 2018

First Class Honours. Focused on probability, statistics, and linear algebra.

SKILLS

- Python (scikit-learn, XGBoost, PyTorch)
- SQL & BigQuery
- MLflow & Model Deployment
- NLP (Hugging Face Transformers, spaCy)
- Polars & pandas
- Docker & Kubernetes
- AWS SageMaker
- Tableau & Matplotlib
- A/B Testing & Causal Inference
- Git & CI/CD

CERTIFICATIONS

AWS Certified Machine Learning – Specialty

Amazon Web Services
Sep 2023 – Sep 2026

TensorFlow Developer Certificate

Google
Jan 2022 – Jan 2025

LANGUAGES

- English (native)
- Italian (fluent)

REFERENCES

Ross Campbell

Head of Data Science, Scottish Power
ross.campbell@scottishpower.com, +44 7700 900 882

Dr. Ananya Gupta

VP of Engineering, Luminance
ananya.gupta@luminance.com, +44 7700 900 109

EXPERIENCE

Data Scientist, Scottish Power, Glasgow

Jun 2022 – Present

Build and maintain ML models for energy trading and demand forecasting.

- Developed a **gradient-boosted demand forecasting model** that predicts next-day electricity demand with 96.2% accuracy — used by traders managing £290 million in annual positions
- Built a price anomaly detection system using **isolation forests** that flagged 43 suspicious trades in Q4 2024, preventing an estimated **£1.7 million** in potential losses
- Migrated 8 Jupyter notebook models into **production MLflow pipelines** with automated retraining and monitoring
- Reduced model training time by **71%** by moving from pandas to Polars for feature engineering on datasets with 180 million rows

Junior Data Scientist, Luminance, London

Sep 2020 – May 2022

Worked on NLP models for contract analysis in the legal AI platform.

- Fine-tuned a **BERT-based classifier** to identify 23 clause types across commercial contracts — achieved 91% F1 score
- Built training data pipelines that processed **340,000 contracts** from 14 law firms
- Created an entity extraction model for party names and dates that reduced manual review time by **35%** for due diligence workflows

Data Science Intern, Skimlinks (now Taboola), London

Jun 2019 – Sep 2019

Summer internship working on product recommendation and click prediction models.

- Built a click-through rate prediction model for affiliate links achieving **AUC of 0.82**
- Analysed **45 million click events** to identify seasonal patterns in product category performance

PROJECTS

Demand Forecasting Pipeline – Scottish Power

Sep 2022 – Jun 2023

Built the end-to-end pipeline for next-day electricity demand prediction.

- Trained on **8 years of half-hourly demand data** plus weather, calendar, and economic features
- Model accuracy at **96.2% MAPE** — best performing model in the trading desk's history
- Deployed on **AWS SageMaker** with automated daily retraining and drift monitoring

Contract Clause Classifier – Luminance

Jan 2021 – Feb 2022

Fine-tuned a transformer model to automatically classify clauses in commercial contracts.

- Covered **23 clause types** including termination, indemnity, non-compete, and liability caps
- Trained on **84,000 labelled clause examples** from Magic Circle law firms
- Reduced contract review time by **40%** for M&A due diligence projects

EXTRA CURRICULAR ACTIVITY

Teaching Assistant – University of Edinburgh

Sep 2022 – Jun 2024

Guest lecturer and TA for the MSc Data Science programme's machine learning module. Ran weekly tutorials for groups of 25 students and marked coursework.

Volunteer – DataKind UK

Mar 2023

Pro bono data science for charities. Completed a project for Crisis UK analysing rough sleeping data across 14 local authorities to help target outreach services.