

Alexey Dronov

alexey.dronov@outlook.com • 07728 211527 • LinkedIn • GitHub • alexeydronov.com

MSc Data Science graduate with hands-on experience building multimodal ML systems and distributed Spark pipelines. Developed an LLM-augmented financial forecasting system for predicting hourly ETF prices, reducing forecast MAE by 16%. Engineered a large-scale analysis pipeline using Apache Spark APIs with performance optimisations achieving 95%.

SKILLS

Programming: Python (proficient), SQL (MySQL, BigQuery, SQLite), Scala (intermediate)

Machine Learning & AI: PyTorch, Scikit-learn, Hugging Face Transformers, Optuna, Numpy, Pandas

Infrastructure: Google Cloud Platform, Databricks, Apache Spark / PySpark, FastAPI, Docker, Git

Methods: Time series forecasting, sentiment analysis, model optimisation, distributed training, A/B testing

PROJECTS

Financial Market Prediction using LLMs (MSc Thesis) May 2025 – Aug 2025

- Built a multimodal prediction system combining ETF time series with LLM-derived sentiment (FinBERT, FinGPT), achieving a **16% reduction** in next-hour MAE over a baseline.
- Automated distributed model training on Google Cloud Platform; designed a custom 6-class financial sentiment schema for explainable forecasting.
- **GitHub:** github.com/AlexeyDronov/financial-market-prediction-llm

IMDb Large-Scale Data Analysis in Spark Mar 2026

- Developed a distributed Apache Spark pipeline to analyse genre trends, director ratings and crew activity across 120 years of IMDb data; applied partitioning, early filtering and aggregation optimisations, achieving a **95% grade**.
- Extended pipeline to the full 2.28 GB (41 M+ records) dataset; re-implemented in PySpark SQL API, **reducing average runtime by 75.7%**; packaged as a reproducible Docker container with CLI benchmarking.
- **GitHub:** github.com/AlexeyDronov/IMDb-Analysis-Scala

Solar Output Prediction - FLW Data Challenge (Winner) Feb 2023

- Built an end-to-end pipeline to predict solar panel output from weather data, achieving 0.94 R² on test data.
- Co-presented findings to academic and industry stakeholders; awarded **first place**.

EDUCATION

The University of Edinburgh

MSc Data Science | Upper Second-Class Honours (2:1) Sep 2024 – Aug 2025

- *Relevant Modules:* Natural Language Processing, Data Science at Scale with Scala, Machine Learning, Distributed Systems, Computational Neuroscience

BSc Mathematics | Upper Second-Class Honours (2:1) Sep 2020 – May 2024

- *Thesis:* "Image Analysis" (79%) – Utilised PyTorch framework and YOLO architecture to detect and classify objects in large datasets. Achieved 94% accuracy on a traffic light detection task.
- *Relevant Modules:* Statistics, Operational Research, Time Series Analysis, Multivariate Data Analysis, Number Theory

LEADERSHIP EXPERIENCE

Business Analyst Oct 2025 – Jan 2026

Blackmont Consulting Aberdeen Branch

- Led AI-assisted legal and market analysis across multiple jurisdictions to support international expansion into U.S. and Indian markets.
- Produced regulatory compliance guidelines and core business documentation and identified key regulatory risks; findings informed two market-entry strategies.

Coordinator & Mathematics Tutor Nov 2021 – Jul 2022

Informatics Tutoring Scheme, UoE

- Led a diverse team of 13 tutors supporting 30+ students, receiving the Edinburgh Award for leadership and community contributions.
- Delivered 30+ hours of personalised Mathematics tuition to a group of 5 pupils, with all achieving grades of 65%+ on Scottish National Exams.