

# Luis Da Silva

London, UK  
+447591010534  
[luisdaniield@gmail.com](mailto:luisdaniield@gmail.com)

## Professional profile

Software developer with 4+ years of experience in software engineering and machine learning. Focused on the bigger picture. Passionate about solving exciting and complex problems with real impact. Interested in software architecture, big data, natural language processing and reinforcement learning. Happy to tackle new challenges and get out of the comfort zone. Gets a sense of accomplishment with a well-done job, thus cares about ensuring targets are met to high standards when working both within a team or own. Outside work, time is usually spent socialising, developing fun side projects or just going for a ride on a bike.

## Core competencies

- Strong IT skills. Experienced in technologies including **Python**, JavaScript, SQL, Cypher, R, Unix and Powershell. Familiar with Java and C++.
- Excellent **machine learning** skills, having worked in a wide range of professional and personal projects involving linear regressions, ensemble models, tree-based models, genetic algorithms and neural networks in order to predict an outcome, classify objects or optimise a process by finding an appropriate strategy. These projects have included time series forecasting, classification, clustering, natural language processing, reinforcement learning, recommender systems and computer vision, and have been built on top of libraries like Pandas, NumPy, SciPy, NLTK, Gensim, Scikit-learn, TensorFlow, PyTorch and Keras.
- Knowledgeable with relational, non-relational and graph **database technologies** such as Trino (Presto), MySQL, PostgreSQL, Neo4j and MongoDB.
- Experience in working with high-capacity parallel computing systems through PySpark and **Docker** in **AWS**. Additional experience working with open-source frameworks such as Airflow, Ranger, Superset and Allure.
- Design and delivery of **APIs** and dashboards using technologies like Dash, Flask, Django, React, NodeJS and Tableau to communicate results and generate insights.
- Focus on producing **clean and reusable production-level code**, with unit testing, integration testing (RobotFramework) and PEP 8 guidelines in Python along with clear documentation.
- **Team player**. Enjoys mentoring and being mentored. Equally comfortable working individually.
- Skilful at time management and able to work on different projects at the same time if necessary.
- Advance knowledge of macroeconomics, microeconomics, econometrics and financial markets.

## Experience

### **EMIS Health. Remote, UK. (full-time Employee)**

**Feb 2022 – Present: Tech Lead - Senior Software Development Engineer**

**July 2021 – Jan 2022: Senior Software Development Engineer**

**August 2020 – June 2021: Software Development Engineer (Data Engineer)**

*Support the development of an in-house petabyte-scale data-lake, delivering terabytes of data daily to partners and delivering machine learning algorithms on top of it to contribute to healthcare.*

- Improved the time frequency in which clients could get new data by a factor of 7 while only using 1.8x compute power by re-designing and re-implementing the core data processing pipeline.
- Adapted and integrated the first third party machine learning algorithm in the EMIS platform, serving thousands of queries a day and processing over 40TB of data.
- Created a local replica of each of the main components of the development cloud environment (Trino, AWS, Airflow and multiple proprietary APIs) using Docker to aid development and better testing, thus increasing the team's confidence and allowing techniques such as chaos testing.
- Development heavily focused on AWS, continually using infrastructure as code (Cloudformation) and services like IAM, EC2, S3, ECS, ECR, SQS, CloudWatch, Chalice, CodePipeline and CodeBuild.
- Support core codebase of SQL views used by the analytics division along with direct clients and partners.
- Took the initiative to guide the team whenever the team leader was unavailable, either pulled into urgent/legacy projects or on holiday.

## **August 2019 – August 2020: Thought Provoking Consulting. London, UK** **Data Scientist (full-time Employee)**

*Produce bespoke software for retail clients that allows them to optimise their internal and external processes and facilitates decision making.*

- Designed, developed, and deployed a Python framework to automate the modelling process of a machine learning model used to optimise the discount to be taken out of products during markdown periods while reaching a sales target. Reduced the Mean Squared Error by 75% and training time by 79% by restructuring the previous approach.
- Created an ML web scraping and data analysis framework that interprets a client's website and reports over 100 potential issues including pricing mistakes and missed profit opportunities. With an average of just under 2 issues per product page, clients reported increased conversion rates after acting on them.
- Managed and collaborated closely with a team of machine learning and business consultants to both understand business requirements and ensure an adequate solution is developed.
- Flexibly switched among Python, SQL and R to work on different projects as per client/team requirements.
- Performed an architectural review on all existent tools.

## **September 2017 – October 2018: Bigvene. Caracas, Venezuela.** **Co-Founder**

*Co-founded an academic research and retail consultancy company in Venezuela which grew to having 12 members and managing up to 3 projects at a time.*

- Developed a bespoke demand estimation, inventory planning and price adjusting solution for two e-commerce retailers using Python and SQL which reduced planning time from days to hours.
- Developed a competitor pricing monitoring tool for e-commerce retailers using Python and Selenium that in conjunction with the demand estimation tool was able to predict accurately the best price changes to implement to each product in a highly inflationary environment like Venezuela.
- Developed the backend in Python of a social media listening solution that used live data from Twitter and Instagram to report on customer reaction to be used when marketing campaigns and new products were launched.
- Led 2 separate teams in projects involving data cleaning, processing, and re-coding in Stata to order to run inference models on the behaviour of people against political propaganda.
- Handled customer relationships, sales, contract negotiation, finances, and marketing.

## **Formal education**

**2018 – 2019: University of Manchester, UK.**

**MSc Data Science.**

- First Class Honours. Top 5% of the class.

**2013 – 2017: Universidad Católica Andrés Bello, Venezuela.**

**BSc Economics.**

- First Class Honours. Top 5% of the class.

## **Links**

- Portfolio - <http://luisds95.github.io/>
- LinkedIn - <https://www.linkedin.com/in/luis-da-silva/>
- Github - <https://github.com/luisds95>
- Medium - <https://medium.com/@luisds95>
- StackOverflow - <https://stackoverflow.com/users/9162159/luis-da-silva>

## **Additional information**

- Languages: Spanish (native), English (fluent) and Portuguese (intermediate).
- Professional hobbies: Reading both books and papers. General investing. DeFi. Building RL agents.
- Non-work-related hobbies: Mountain biking, basketball and working out at the gym. Playing chess, go and poker. Socialising with friends.

**References available on request**