Tertuliano de Lemos Oliveira e Silva
Junior Data Scientist

Address: **London, WC2E**
Email: **leonlemos26@gmail.com**
LinkedIn: **linkedin.com/in/tertuliano-silva26**
GitHub: **github.com/LeonLemos**
Twitter: **twitter.com/Leon_theLemos**

I'm a Junior Data Scientist, with a background in Robotics and Mechatronics Engineering.
I believe data is the missing link that allows an engineer to apply their hard skills into the world and become effective in any field, because data is everywhere and in everything.
Absorbing experience in different areas has allowed me to further broaden my horizons and develop a creative thinking with new ways to solve problems and support others. I'm looking for opportunities in the tech and consulting industry with hopes of learning and growing by working with multiple sectors.

Skills
Technical skills
Programming Languages : Python, PostGreSQL.
Libraries : Numpy, Pandas, Scikit-Learn, SciPy, Statsmodels.
Big Data : Apache Spark, PySpark.
Neural Networks : TensorFlow.
Data Visualization : Seaborn, Matplotlib, Tableau.
Natural Language Processing : Tf-Idf, Countvectorizer.
Machine Learning : Supervised, Unsupervised Modelling, Timeseries ARIMA, SARIMAX.
WebScraping : Beautifulsoup, Regular Expressions, Request.

Soft Skills : Communication, Problem solver, Adaptability, Leadership, Teamworker.

Languages : English (Fluent), Portuguese (Native).

Experience
Data Science Immersive Course Student, General Assembly, London, 05/2021 - 08/2021
A full time 12 week Data Science bootcamp teaching the fundamentals of Python programming, statistics and machine learning. It was concluded with the delivery of a capstone project, in which we put machine-learning models into production to address a real-world data challenge.

Regression and Classification with the Ames Housing Data. 1 week, Solo, Conceptual
- ❖ Used Statistical Analysis to predict house prices and maximize investment profits.
- ❖ Used Regression and Classification Machine Learning techniques to estimate the relation of fixed and changeable property characteristics on house price.
- ❖ Fitted a Lasso Linear Regression Model and obtained an R2 score of 0.68.
- ❖ Concluded that fixed characteristics affect house prices more than changeable characteristics.

Web Scraping for Indeed.com and Predicting Salaries. 2 weeks, Solo, Conceptual

- ❖ Scraped the data from the job aggregation website using BeautifulSoup.
- ❖ Created a dataframe from the parsed data and a new binary class target to predict salaries.
- ❖ Fitted a Random Forest Classifier model and obtained an R2 score of 0.80.
- ❖ Concluded that jobs with the words "Senior, Administrator" would command higher salaries.

Capstone Project
Predicting Used Car Prices and the Effect of Covid-19 with Machine Learning. 4 weeks, Solo

- ❖ Obtained the Datasets using an API to extract information from live updated Covid cases. Web scraped a used car listing website, using BeautifulSoup and created a Dataframe with the respective vehicle information and all the relevant features.
- ❖ Performed Exploratory data analysis, data trend analysis, feature distributions and correlations.
- ❖ Merged the clean datasets by the common columns and times of occurrence.
- ❖ Ran Regression Machine Learning Models and Implemented Deep Neural Networks.
- ❖ Obtained an R2 score of 0.81. 81% of the data variance can be explained by our Random Forest model, meaning that it generalizes well in predicting used car prices.
- ❖ Concluded that covid cases and deaths are not significant on their own, but were still useful in improving our model score from 0.80 to 0.81.

Previous Experience

Chief Executive Officer, Banzelos Nation Record Label, Angola/Luanda, 2015-2020

- ❖ Responsible for managing artist's metadata and uploading their music catalogue.
- ❖ Devised Marketing plan for the content, and bridge the intercommunication, making sure artists and managers reached deadlines.
- ❖ Managed royalty payments.
- ❖ Kept team motivation high with good communication, support and incentives.
- ❖ Adapted the business and made it profitable during the pandemic.

Education

Data Science Immersive Course, General Assembly, London, 05/2021 - 08/2021

Robotics and Mechatronics Engineering BSC, Leeds University, Leeds, 01/2011 - 06/2014

Interests

Artificial Intelligence
With the advances of technology and increase of computational power, distant concepts like Artificial General Intelligence are getting closer to us. I'm excited to see what the future holds for the field.

Music
Music fascinates me, because of the vast amount of information that is contained inside an audio file, even though it's invisible to the average consumer. I see music as another universe inside the exciting world of data.