

Genaro Garcia Pereyra | Data Scientist

genaroggp3@email.com • <https://www.linkedin.com/in/genaro-garcia-pereyra-cs/> • <https://github.com/ggarciap>

Persistent, detailed-oriented data scientist with a software engineering background. I use my analytical abilities in the fields of finance and risk assessment and apply a balanced perspective between technical and user requirements. I am passionate about machine learning, and I am always looking for new exciting challenges that allow me to grow professionally and personally.

Skills

Languages/Environments | Python, SQL, C, C++, Google Colab

Packages | Pandas, NumPy, NLTK, TensorFlow

Data Visualization | Seaborn, Plotly

Modeling | Statistical Modeling, Regression Models, Classification Models, Time Series

Analysis, Neural Networks, Image Classification

Statistics | Statistical Distributions, Bayesian Analysis, p-Values, Hypothesis Testing, Bayes Theorem

Projects

Mexico City robberies multi-class classification | GA 10/2020 - 12/2020

- Cleaned and preprocessed datasets from Mexico City crime database to create a dataset composed of 23 features describing geographical categorical and numerical attributes.
- Created new Engineered features using tools like Pandas dummies and one-hot encoding, and K-mean modeling to further improve predictive models.
- Developed XGBoost multi classification models to predict the type of robbery based on the affluence, distance from the metrobus stations and other geographical attributes of the crime.

Coexist | Web 01/2019 - 05/2019

- *Coexist* lets you plan events, create and assign tasks and split bills conveniently, thus making living together more harmonious.
- Contributed in the creation of an all-in-one app that enables users to split bills, coordinate tasks and stay in touch by developing a front end using technologies such as HTML, CSS and ReactJS; and a back end that uses DynamoDB, AWS and GraphQL.
- Built a task manager that allows manual or random generated task assignments for efficient teamwork.
- Built an interactive interface where the users can send messages and interact within a group.

Professional Experience

GLAC | Mexico City, Mexico

Software Developer

04/2017 - 03/2019

- Contributed to the creation of python scripts for the computation of mathematical models which generated 1000 daily variable estimations for the GLAC index platform.
- Upgraded and created new due diligence matrixes with tools such Pandas for optimization of the social-economic sector of the GLAC index platform.
- Collaborated with the development of the tracking of ~800 daily mentions (through the API of twitter) along with the creation of queries tailored for each sector analyzed in the GLAC index platform.

Purdue, Study Abroad and Service Learning | Dar es Salaam, Tanzania

Teacher

05/2018 - 05/2018

- Taught basic computing and coding concepts at Matumaini Primary School which serves nearly 200 children with disabilities.
- Taught instructors the fundamental use of block-based visual programming tools such as Scratch with the goal of allowing them to expand their curriculum for students.

Bulltick | Miami, FL

Analyst

04/2015 - 05/2015

- Produced 30 reports of the performance of mutual funds and delivered summaries of market openings in Spanish and English.
- Created correlation heatmaps of ETFs for the consumption of areas such as Sales & Trading.
- Tracked through bloomberg terminal the performance of ETFs of interest for the area of investment banking.

Education

Data Science Immersive, September 2020 - December 2020

General Assembly | Remote

12-week full-time immersive educational program strengthening Data Science skills including: Python, SQL, data cleaning, data visualization, regression models, classification models, web-scraping, APIs, NLP, advanced supervised learning, unsupervised learning, time series analysis, and statistics.

Bachelor of Science - BS. Computer Science, 2020

Purdue University, West Lafayette

Certificate of Entrepreneurship, 2020

Purdue University, West Lafayette