

Karen Lorena Figueroa Chilito

Data Science - Machine Learning – Bioinformatics - Cloud computing

CONTACT

Email 1: klch@unlv.nevada.edu
Email 2: k.figueroac.17@gmail.com
GitHub: <https://github.com/oikosfigueroa>
Profile: [LinkedIn Profile](#)
Personal Website: <https://www.klfc.me/>

EDUCATION

University of Nevada, Las Vegas - UNLV
Ph.D. in Public Health, Global and Environmental Health
Las Vegas, NV
July 2023 - Current
GPA: 4.0

University of Texas at Austin
Master of Science in Data Science
Austin, Texas
Jan 2022 - May 2023
GPA: 3.84

University of Texas at Austin – McCombs School of Business
Postgraduate Program in Data Science and Business Analytics.
Austin, Texas
Jun 2021 – Jan 2022
[PGP in Data Science and BA E-portfolio.](#)

Universidad Nacional de Colombia
Bachelor of Arts in Economics
Bogotá, Colombia
2011 - 2016
Classes: Game Theory, Econometrics, Data Analytics, Macroeconomics, Microeconomics

SKILLS

Programming Languages:

Python: Advanced skills in data analysis, machine learning, and geospatial modeling using libraries such as NumPy, Pandas, Matplotlib, seaborn, scikit-learn, imbalanced-learn, TensorFlow, PyTorch, NLTK, and SpaCy.

R: Expertise in statistical modeling and data visualization with tidyverse, tidymodels, ggplot2, broom, tidyr, and dplyr.

SQL: Proficient in relational database management and querying with SQLServer, PostgreSQL, MySQL, Snowflake, IBM Netezza, and SQLite.

Machine Learning and Data Science Tools:

Experience with supervised and unsupervised learning, neural networks, and deep learning models using TensorFlow and PyTorch.

Skilled in developing predictive models, handling imbalanced datasets, and performing natural language processing (NLP).

Expertise in designing and optimizing pipelines for geospatial and genomic data.

Geospatial Analysis and Tools:

Advanced proficiency in geospatial data processing and visualization using Google Earth Engine API, GeoPandas, Folium, ArcGIS, and QGIS.

Experience with high-resolution raster datasets for environmental modeling and spatial analysis.
Cloud Computing and Big Data:

I am a highly motivated and results-driven data scientist with a Master’s degree in Data Science from the University of Texas at Austin and a Bachelor’s degree in Economics from the National University of Colombia. Currently, I am pursuing a Ph.D. where my research leverages machine learning, advanced analytics, and geospatial modeling to analyze large-scale datasets and uncover patterns in complex systems.

With expertise in data analysis, predictive modeling, and neural networks, I specialize in building robust pipelines for processing and interpreting diverse datasets, including genomic, spatial, and environmental data. My passion lies in solving real-world problems by developing data-driven solutions that drive actionable insights and innovation.

PROFESSIONAL EXPERIENCE

University of Nevada, Las Vegas – UNLV, [PARAVEC Laboratory](#)
(July 2024 – Present)
Graduate Research Assistant

- Designing and implementing bioinformatics pipelines in Python and Bash to process and analyze large-scale genomic datasets, enabling efficient population genomics research.
- Applying machine learning algorithms to model mosquito population dynamics, including predictive analytics for vector adaptation and spread in new environments.
- Extracting and analyzing high-resolution raster data using Google Earth Engine and Python libraries to identify environmental drivers of mosquito-borne disease outbreaks.
- Utilizing tools like GeoPandas, Folium, Earth Engine API, and ArcGIS for advanced spatial data visualization, mapping, and geospatial analysis.
- Conducting genomics research to investigate genetic adaptation patterns in mosquito populations, providing insights into vector ecology and control strategies.
- Integrating geospatial data with genetic analyses to build predictive models for understanding the spread and resilience of mosquito populations in diverse ecosystems.

Correlation One, New York City, NY, Remote
(May 2020 – July, 2023)
Teaching Assistant, Instructor and Data Science Consultant

- Courses: Data Science for All - Colombia (2020 – 2021), DS4A Women (2022, 2023), Data Analytics (2023)**
- After completing the Data Science for All program (first iteration in Colombia), earned the possibility to work for Correlation One as a Teaching Assistant for the second iteration.
 - Teaching assistant for multiple versions of the Data Science for All program: DS4A-2.0, DS4A 3.0, DS4A 4.0, DS4A 5.0. (2020 - 2021), helping over 3,000 people to get their certificate.
 - Led and managed around 100 successful Data Science projects with immediate impact on the public and private sector. Some of the best projects: [Project 80](#), [Project 56](#), [Project 51](#), [Project 81](#).
 - Taught live remote classes to hundreds of students on:
 - NLP covering basic and intermediate concepts: Tokenization, stop words, n-grams; using NLTK, spaCy, wordcloud, scikit-Learn, data visualization, EDA, and data cleaning in Python; using pandas, Matplotlib, and seaborn. As well, databases and SQL language.

- Courses: Anaplan for All program, - Version 1.0 (2021), 2.0 (2021), 3.0 (2022)**
- Earned the Anaplan model builder certificate (L1, L2, The Anaplan Way) in less than three weeks to be able to be a Teaching assistant for the first iteration for the Anaplan for All program.
 - Trained and helped over 800 people to obtain their Anaplan model building certification (L1, L2, The Anaplan Way) with a deep focus on FP&A and Supply Chain modeling.
 - Built Anaplan models and created model building challenges for students.
 - Gave one-on-one sessions to students to help certify them as Anaplan Model Builders.
 - Taught live remote classes about formulas, functions, and the specific Anaplan syntax.
 - Distinction: Scored as the best Teaching Assistant for the Anaplan for All program.

Simetrik, Bogotá, Colombia
(Dec 2019 – May 2020)
Data Engineer

- Applied Data Engineering to fix multiple data sets working with SQL queries and Python scripts.
- Performed daily data validation and created unique formats for multiple datasets which had been acquired from different financial sources to finally verify daily transactions for all the different clients who wanted to verify their data.
- Created dashboards (Quicksight) based on each client to deliver main KPI’s about data validation.
- Managed the complete ETL process from PostgreSQL and Snoflake to AWS (different services including: S3, EMR, Cli, Glue, RDS, Lambda).

CredibanCo, Bogotá, Colombia
(May 2019 – Nov 2019)
Business Intelligence Analyst

<p>AWS: Experienced in cloud services such as S3, RDS, EMR, Lambda, Glue, and Quicksight for scalable data storage and analysis.</p> <p>Knowledge of parallel processing and big data tools for large-scale datasets.</p> <p>Business Intelligence and Visualization:</p> <p>Proficient in creating dashboards and interactive reports using PowerBI, Tableau, and AWS Quicksight.</p> <p>Experience in storytelling with data to support decision-making and strategic planning.</p> <p>DISCTINTIONS</p> <p>Ministry for the Information and Communications Technologies Data Science for All 1.0 Fellow Oct 2019 – Dece 2019</p> <p>Awarded merit-based fellowship (+10K applicants, <3% acceptance rate) to the first Data Science for All to train with Harvard Professor Natesh Pillai.</p> <p>Focused on advanced data science topics, including machine learning, artificial intelligence, and analytics, with practical applications in diverse fields.</p> <p>Mary Guinan and Shawn Gerstenberger Public Health Scholarship May 2024</p> <p>Awarded for academic excellence and commitment to advancing public health in Nevada, particularly in Environmental and Occupational Health.</p> <p>Recognized during the UNLV School of Public Health Honors Ceremony.</p> <p>ARPA Scholarship Award Fall 2024</p> <p>Received \$1,500 for an experiential learning project addressing public health challenges in Nevada.</p> <p>Focused on leveraging machine learning and geospatial modeling to analyze vector-borne disease patterns.</p> <p>Graduate Mentor, UNLV MENTORS Program 2024 – Present</p> <p>Officially selected as a Graduate Mentor in UNLV's MENTORS program to guide incoming graduate students.</p> <p>Recognized for leadership skills, dedication to academic success, and contributions to fostering a supportive graduate community.</p> <p>LANGUAGES</p> <p>Spanish: Native</p> <p>English: Advanced (C1) - TOEFL Certified</p> <p>PERSONAL INTERESTS</p> <p>Dog training, hiking, gardening, astronomy, cooking, sketching.</p>	<ul style="list-style-type: none">Designed and structured databases, stored procedures, reports, and data input interfaces using SQL Server and Aginity IBM Netezza (SQL language).Developed and maintained business analytics products for clients who belong to the different Industry sectors, based on the company’s strategy. (Manipulated, cleaned datasets, added statistical analysis, predictions and created dashboards for the multiple clients.Developed solutions and analyzed trends to other areas of the company transforming their data into reports which was useful to make decisions periodically. <p>TransUnion, Bogotá, Colombia (Mar 2017 – Sep 2017) IT Decisioning & Data Analyst – Latinoamerica</p> <ul style="list-style-type: none">Built scoring and risk models according to the different private financial entities’ necessities.Designed, defined, planned, and executed QA-tests using SQL and JavaScript to review the errors in the codes for every functionality that was being developed by the other part of my team according to the changes the company needed to do regarding to a transition, which were been integrated to the new applications.Trained test and client service teams the internal and external use of the new applications so they could solve any client inquiry. <p>Scotiabank-Colpatria, Bogotá, Colombia (July 2016 – Dec 2016) Junior Analyst – Credit Risk Viceprecidency</p> <ul style="list-style-type: none">Create macros in excel using Vba to automate processes that were taking more than 3 days, reducing them to only 2 hours.Processes automated: Monitoring and balance of delinquent debtors of more than one billion Colombian pesos and weekly report of goals achieved according to the different types of credit: housing credit, revolving credit, student credit, car credit and free investment credit.Built scoring and risk models according to the different private financial entities’ necessities.Reviewed accuracy of data using SQL Server for daily payments and credits of delinquent debtors of more than one billion Colombian pesos.
--	--