Carlos Urteaga Reyesvera

Career Summary

Carlos is a Data Scientist and Machine Learning (ML) Engineer who has been involved in different sectors such as Financial, Education, Retail, and Telecommunication for more than nine years. He has a strong foundation in computer science and machine learning, and he has played a crucial role as MLOps engineer, where he deployed pipelines for integration, training and prediction using different architecture approaches to follow up model performance. Finally, he leads an MLOps squad to standardize different tools and practices to create a federation of models and resources.

Summary of Qualifications

Proficient In:

- Python
- MLOps
- Machine learning
- Pandas
- SQL
- git
- Dask
- Snowflake
- CUDA
- OpenCL
- Parallel programming
- Arduino
- Robotics
- OpenCV
- Computer vision
- REST API
- ROS
- iOS
- Android
- R/R Project

Languages

- Spanish (native speaker)
- English (Intermediate)
- French (Intermediate)

Career History

Globant, Mexico City, Mexico Mar. 2022 – Today

Data Scientist Lead, Software Designer

Augmented Gatekeeping Dec. 2020 – Oct.2022

- Propose and design MLOps task to track efficiency of the model, after this I lead MLops team to follow best practices to follow up AI models using MLFlow and ArgoFlow
- Design a classification system with random forest for technical interview matchers with 90 % of accuracy.Tools used Python, sklearn and flask.
- Extract information from resumes using Natural Language Processing (NLP) and computer vision (CV), each skill repetition and appearance permit create an score
- Generate a data flow to extract, import and evaluate a candidate to suggest a studio. Use ensemble models with KNN and random forest to resilient the data inconsistencies.
- Translation of business needs into an expert system to empower the recommendation model.
- Extract information from a data lake (presto engine) and generate an exploratory data analysis using python and sql.

Data Scientist Sr

Rappi Jun. 2020 – Dec. 2020

- Augmentation staff as Data Architect to migrate a retail catalog from legacy to snowflake, each retail has their own data description and the project was to concentrate this information into Snowflake for all latinoamerica countries with 250,000 business.
- Develop a pipeline to extract, clean and substitute information according to industry product

masters and select the features.
- Deploy and connect APIs over Flask to connect different sources and complement the information.
- Python and Dask workshop, show pros and cons of dask to use in the different projects.

Garrigan Lyman Group Mar. 2020 – May 2021
- Implement a data pipeline to enhance the quality of data extraction using regular expression. The quality was improved by 80% when the data was incomplete. Tools used, pandas, sklearn, python and docker.
- Create an ETL to add information from different sources into the data lake, some supported formats were XML, JSON and CSV.
- Identify the industry product master field to match with a regular expression.
- Using a VM (virtual machine) with Compute Engine on GCP, configure a docker to have a productive environment to extract the current data.
- Generate a demographic report from each PoS dynamically.

Deloitte Aug. 2019 – Feb 2020
- Develop an AI solution for face recognition and people counting using Deep learning using only a few pictures. The solution involves data augmentation, face recognition and identification, object tracking, those must bear over a web service with real-time streaming (video and text).The tools used were Python, TensorFlow, Keras and Flask.
- Data driven discovery to minimize cost for the IT area, we proposed a cost optimization for mobile phone plans and a Dashboard to generate reports and have enough visibility to enable agile decision-making.

IDB Apr. 2019 – June 2019
- Implemented a xgboost model in Python + Docker and some NLP to detect domestic violence.
- Generation of profile risk using feature engineering in order to predict the recidivism for LATAM mainly in Brazil, Uruguay and Honduras.

Compartamos Banco, Mexico City, Mexico Sept. 2018 – Mar. 2019
Data Scientist Manager
- Developed ML model for churn and default with Python, SQL,H2O, and Flask
- Architect and Build machine learning software products for our core from local computer to use and hybrid setup for big data (AWS).
- Lead all aspects of ML automation including model training and development, feature selection and model tuning with DM-CRISP Implementation.
- Evaluated the performance of projects using propensity score. This allows us to evaluate the impact compared without it. Generation of dashboard using Tableau.
- Perform and speed up the business as usual in SQL query and batch script.
- Propose ROI method with a genetic algorithm to identify the best option between risk for each loan

Telcel, City, Mexico City, Mexico May. 2018 – Ago. 2018
Data Scientist Sr.
- Analyzed and cleaned IVR data to recognize temporality and types of request through Python, Hadoop and SQL. This focus marketing activity invites users to install the Graphic IVR.
- Crafted a dashboard using Tableau to evaluate machine learning models. Dashboards were used for the main board to take action to analyze marketing activities.
- Redesign an IVR mobile application considers the user and client experience, this decreases the number of incoming calls to the IVR and the cost of the call center.
- Evaluation of ML models for Upsell, Cross-sell and Churn through a dashboard

KPMG, Mexico City, Mexico Oct. 2017 – May. 2018
Information Data Security Sr.
- Data analyst for compliance and monthly report of computer status to report unauthorized behavior.

- Management of users at a national level to guarantee a follow of IT policies, this is accomplished using python and Machine Learning (Random Forest) to detect leaks in the process.
- Report Awareness and status to KPMG US and UK before the global system could generate the reports.

ITAM, Mexico City, Mexico Ago. 2010 – Oct. 2017

Lecturer Jan. 2016 – May. 2017

- Teaching engineering courses for mid students.
- Elaborated and coordinated teaching material and final projects involving IoT and robotics.
- Teach robotic software to develop an autonomous robot to avoid obstacles.
- Update projects and examples to real problems

Research Assistant Ago. 2013 – Oct. 2017

- Microsoft Research:
    - Analyze the effect of students' multitasking during a computer class and final score.
        - Acquired and consolidated weekly information generated along the semester.
        - Analyzed, cataloged, and prepared the information for data analysis using python; delivered with one month of anticipation.
        - Provided statistical reports; the results were used in the final analysis.
- Federal Telecommunications Institute:
    - Design a fiber network for the Federal Telecommunication Institutes to connect towns with less than 10,000 inhabitants in Mexico.
        - Proposed an architecture to transform geographic information systems (GIS) into vectors for mapping; this helped to make better decisions on the selected network.
        - Visualized the result in a map using GIS on the Internet; the maps were used to show the best distribution.
    - Audit the mobile network in Mexico to discover failures in the quality of the service for each radio station.
        - Designed and coded a mobile application to extract the relevant information of a cellular tower like Cell ID, local Signal Strength Indication (RSSI), etc.
        - Collaborated in the development of a Web server which analyzes the quality of service.
- Department of Digital Systems:
    - Deploy vision algorithm for self-driving cars using ROS and OpenCV.
    - Study human behavior in an ascending auction of the radio spectrum.
    - Fraud detection in spectrum auction by detecting unusual bids by analyzing historical binds using SQL.
    - Development of a middleman program to identify clusters by income and behavior of auctioneers.
    - Coded an auction platform using Azure cloud service with VB.net and S
    - Develop a mobile solution to differentiate the Polaroid brand in the
        - Organize a team with five other people using Scrum
        - Modified the Android operating system to add homemade mobile applications.
        - Reviewed the user experience guide of the mobile app.
    - Published IEEE conference paper about detection of scorpion using ML and CV.

IT Manager Ago. 2010 – Oct. 2017

Administration of laboratories used for academics.

- Provided software and configuration according to the class; improving resources using virtualization.
- Administration of Google Classroom and Active directory for students and lecturer.
- Advanced user in Unix/Linux, Windows, Windows Server, and OSX

Education

Columbia University, NY, USA — Jan. 2019 – May. 2019
Diploma, Applied Machine Learning

ITAM, Mexico City, Mexico — Jan. 2016 – May. 2017
MSc. Computer Science

ITAM, Mexico City, Mexico — Ago. 2007 – Dic. 2015
BS, Telematics Engineer

ITAM, Mexico City, Mexico — Ago. 2012 – Feb. 2013
Diploma, Mobile application design and development Diploma.

Courses

Coursera

Machine Learning Engineering for Production (MLOps) Specialization — Oct. 2022
Neural Networks and Deep Learning — Nov. 2017
Structuring Machine Learning Projects — Nov. 2017

Udemy
MLOps fundamentals of Continuous Integration & Continuous Delivery (CI/CD) using Azure DevOps & Azure Machine Learning — Sep. 2022

Complete MLOps Bootcamp | From Zero to Hero in Python 2022 — Sep. 2022

DataCamp — Dec. 2021
Introduction to Programming Using Python — Dec. 2021
Intermediate Python — Dec. 2021

Extra curricular

IoT and Big Data Talk Jalisco, Mexico — Jan. 2016 – Nov. 2017
Big Data and IoT Using Fiware for ANUIES

Workshop, Mexico City, Mexico — Nov. 2016
Workshop Autonomous Car imparted for Freie Universität Berlin

IEEE Conference paper, Puebla, Mexico — Sep. 2016
Urteaga-Reyesvera, J. Carlos, and Andre Possani-Espinosa. "Scorpions: Classification of poisonous species using shape features." Electronics, Communications and Computers (CONIELECOMP), 2016 International Conference on. IEEE, 2016.