GEORGE R. CUBAS
+1.346.453.8515
georgeraulc@gmail.com

SUMMARY

Engineer and Python Developer with more than 10 years of experience. Experienced in large scale million-dollar projects utilizing cross-functional team collaboration. Skilled in AI, ML Web development, Engineering Economics, with 4 years of experience in the Energy Sector.

TECHNICAL SKILLS

Python, OOP, Django, Pandas, NumPy, Selenium, BeautifulSoup, Docker, Flask, Matplotlib, SciPy, PyTorch, Keras, Apache Spark, REST Architecture, Ruby, SQL Server, JavaScript, Linux, Machine Learning, AI, Scikit-learn, ArcGIS, GitHub, TA-Lib, SQL Server, Jupyter Notebook, SQL, Arcpy, REST, Boto3 AWS, Asyncio, Rust, PyO3, R, Fortran, C#, GIT, Kafka, Microsoft suite programs including Excel, Word, and PowerPoint.

PROFESSIONAL EXPERIENCE

EOG Resources
March 2024 – September 2024
Python Developer (Contract)

- Developed and scaled Oil and Gas economic models, focusing on cost recovery and profit split for a new application using Polars.
- Optimized and adapted models by leveraging Polars' parallelized functions for enhanced performance.
- Created Depreciation, Depletion, and Amortization (DDA) functions to improve existing cash flow models.
- Implemented Net Present Value Index (NPVI) and Profitability Index (PVI) calculations using Polars for accurate cash flow analysis.
- Contributed to the debugging and enhancement of the Planning Economics Engine, resolving critical bugs and updating the codebase.
- Played a key role in evaluating new global assets by integrating updated income tax and tangible depreciation data for economic feasibility assessment.
- Designed and developed new FastAPI endpoints and stored procedures for advanced scenario sensitivity analysis.
-

Compbuss
Aug 2023 - Present
Python Developer

- Design and build AI services including Image Learning and Classifiers using natural evolution and computer vision methods per OPENAI research paper and deploy using FASTAPI.
- Build RAG using dolphin llama model from Ollama and use Dspy with weaviate to build a rag system for the LLM so that n sources could be added and the model make inferences based on these sources
- Design, build, and maintain efficient, reusable, reliable code.
- Constantly and iteratively improve existing API functionality performance and quality of services.
- Write and maintain code following engineering best practices using CI/CD KISS tools.
- Write unit and integration tests using test driven development methodologies.

University of Virginia
Mar 2022 - Dec 2022
Python Developer (Contract)

- Generate programs for ETL process for Office of Sponsored Programs.
- Utilize Pandas, NoSQL, and SQL alchemy to streamline the ETL process for awards, proposals and agreements for full end to end project implementation using analysis, design, modeling and testing for server side scripts and applications.
- Build a crawler to verify the quality of the frontend data and see what percentage of data needs to be corrected.
- Use Pandas to generate various reports to replace reports previously generated using QlikView.
- Use Multiprocessing (bypassing GIL) and Async IO modules when applicable for parallelization/concurrent performance in various scripts.

Kinstone Investment Properties
Jul 2019 - Mar 2022
Python Developer

- Model valuation and acquisition of new properties using Pandas, NumPy and Sci-kitlearn in Jupyter Notebook.
- Utilize BeautifulSoup and Selenium to scrape Freddie Mac Multi Family Index information and Case Shiller Housing Price index information adjusted for real inflation. Construct Bridge API's to schedule updated cost valuations of properties in certain areas of Houston using Zillow API.
- Built a DCF model in python using OOP in Jupyter Notebook to evaluate the levered and unlevered IRR for each new property being considered.
- Apply Monte Carlo Analysis using NumPy, SciKitLearn and Pandas for finding the stochastic DCF Evaluation, using the DCF OOP class built.
- Build a model given various DCF valuations and data using Zillow API for decision making using a binary classifier for a given investment area using Flask framework and SciKitLearn.

Mass Action Engineering
Jan 2019 - Jul 2019
Chemical Engineer (Contract)

- Responsible for the Basis of Design for the production of Polyvinylidene Fluoride Polymer and co polymers.
- Created customized scripts for processing all P & ID's and PFD's using Python.
- Designed a process calculation utilizing NumPy, Pandas and Jupyter Notebook. The scripts were modularized to have each unit process within its own class using OOP principles.
- Utilized data structures such as Dictionaries and key value pairs consisting of arrays, or nested dictionaries for each unit process for storing the results across each unit.
- Built a Pandas data frame using dictionary to DF and then exported to Excel for rendering. The algorithms used for balancing each unit operation were selected on a case by case basis including both linear and dynamic algorithms.
- Worked with a tech team to effectively simulate the Free Radical Emulsion Polymerization of the monomer Vinylidene Fluoride.

Occidental Petroleum
Jan 2017 - Jan 2019
Software Engineer (Contract)

- Planned various Field Development and Section Development plans in locations throughout New Mexico and Texas.
- Built an Anti-Collision Risk Analysis Web application using Django Framework that calculated the risk between planned Wells and existing wells.
- Created scripts that used a pretrained neural network to extract scanned well file data to a CSV using a Linux shell script.
- Built scrapers using BeautifulSoup4 and Selenium.
- Designed a drilling web application using Django and a machine learning algorithm that modeled drawdown for production using Scikit-learn.
- Designed most programs using Python, some JavaScript, the Django framework, and the modules NumPy, Pandas and scikit-learn. All Python code was written using OOP principles and applying the necessary algorithmic approach on a case by case basis.
- Developed models using ArcGIS, Python, Arcpy, SQL, Excel, and Powerpoint.

British Petroleum (Alaska)
Jan 2013 - Dec 2016
Drilling and Completion Engineer

- Supervised and managed the daily drilling operations in various roles.
- Managed teams of 30+ people and projects in excess of $20MM.
- Planned a $10MM sidetrack well which included building a time estimate based on gathered metrics in each development hole section based on a Monte Carlo time estimate and cost estimate based on gathered rig burn rates and total equipment costs.
- Executed the technical engineering design based on the recommended practices and standard procedures for the North Slope Developmental area. The well was put on production at a rate of 100,000 bbl/per day.
- Worked in the role of Rig workover engineer in charge of replacing ESP's in dynamic pressured reservoirs in Milne Point and extensive decompletes and scab liner jobs.

PROJECTS

Generative AI (RAG using Weaviate, Dspy)   **2024**
[https://github.com/mateBarey/Rag-GEN-AI/blob/main/generalizedRagAgg.py]
A framework for integrating and retrieving information from multiple sources to enrich a LLM knowledge base and enhance its contextual accuracy

- **Implemented a Retrieval-Augmented Generation (RAG) pipeline** leveraging Weaviate for semantic search, enabling precise and contextually relevant information retrieval from large PDF datasets for enhanced model outputs.
- **Utilized Weaviate's vectorized search capabilities** to efficiently store, index, and retrieve document chunks based on contextual similarity, optimizing data access for the RAG system.
- **Integrated Dspy's Ollama Local model** within the RAG framework to generate accurate and context-driven answers, enhancing the AI's ability to respond to complex queries with technically detailed explanations.
- **Engineered parallelized PDF processing pipelines** with concurrent futures to load and chunk documents into Weaviate, significantly reducing data preprocessing time and improving overall system efficiency.

Apache Spark Prediction Pipeline   **2019**
[https://github.com/mateBarey/Apache-Spark-IOT-Prediction-Pipeline]
Weather IOT data is cleaned using PySpark SQL and vectorization, a correlation matrix built to check for independent variables and then a train and test dataframe is split by 0.8 and 0.2, the models are then built using a train_DF to predict a label called "HOURLYPressureTendency"

- Coded in Python using Jupyter Notebook.
- Utilized Apache Spark for building a machine learning pipeline and training/testing data.
- Predictors used were Linear Regression, Gradient Boosted Tree Regressor, Logistic Regression, Random Forest Classifier and a GBT Classifier.

Actor Critic Reinforcement Learning NN   **2019**
[https://github.com/mateBarey/Reinforcement-Learning]
A reinforcement algorithm that uses an actor for finding the best policy and critic which enables finding the best probability associated with each action to take in order to solve the Cart Pole Problem

- The program is written in Python using the PyTorch Module for creating Neural Networks.
- It has 2 hidden layers with 2 neural nets (NN) for both actor and critic and 512 neurons in each.
- Uses Relu Activation functions.

DB as a Rest API   **2017**
[https://github.com/mateBarey/DB-as-a-Rest-API-using-Flask]
A database as a Rest API using Flask that allows users to register and get tokens. A token is taken away when a sentence is stored and then retrieved.

- Coded in Python using the Flask framework.
- Utilized RESTful architecture.
- Integrated a NoSQL database which allowed for a dynamic schema and linear scalability.
- Utilized Docker Containerization.

LANGUAGES SPOKEN
Spanish, Portuguese, French

EDUCATION AND CERTIFICATIONS
Bachelor of Science, Chemical Engineering - University of Houston

AI Engineering Professional Certificate - IBM

Machine Learning for Traders - Google Cloud

Coding Bootcamp - Le Wagon