Nicolas Sanchez Noguera

Nicosanchez0411@gmail.com · (786)-683-9388 · https://github.com/nicosanc

SUMMARY

Graduate from the University of Michigan, specializing in computational and statistical programming, with hands-on experience with Machine Learning, Web Systems, and ETL development. I learn on the fly and pride myself in my adaptability, discipline, and resilience. I'm skilled in Python, with extensive experience with data related frameworks including PyTorch, Sci-kit Learn, Numpy, and Pandas. My academic foundation in Computer Science was built on mastering Data Structures & Algorithms and Object Oriented Programming in C++, leveraging the language's Standard Template Library.

EDUCATION

University of Michigan, **August 2023**
Bachelor of Science in Engineering in Data Science, Minor in Mathematics
Relevant Coursework: Artificial Intelligence, Calculus, Data Mining, Financial Engineering, Linear Algebra, Machine Learning

SKILLS

Programming: Bash, Cairo, C++, C#, CSS, HTML, JavaScript (React), Python, R, Rust, SQL
Technologies: Apache Kafka, Apache Spark, AWS, Azure Databricks, Docker, Git, Github Actions (CI/CD), MySQL, VS Code
Operating Systems: Windows, Linux
Languages: English, Spanish, French
Certifications/Certificates: Databricks Certified Associate Data Engineer, Confluent Apache Kafka Accreditation

PROFESSIONAL EXPERIENCE

DataAnnotation – Contract Coding Prompt Engineer
October 2023 - Present

- Promoted to Reviewer, where I assess other users' performance on coding prompt conversations and determine project eligibility
- Simulating user-chatbot dialogues by carefully curating Python Data Analysis and Visualization code generation prompts
- Assessing model performance using metrics such as verbosity, harmfulness, correctness, and instruction following

Rovisys - Software Engineering Intern
June - August 2022, Aurora OH

- Collaborated with team to design, build, and launch a full-scale data reporting Windows Service for FirstEnergy in C#
- Programmed a data collector that retrieved data from governmental websites using Newtonsoft and HTTPClient
- Configured a PI Data Archive which stored and organized the service's collected data using PISDK
- Provided FirstEnergy with real-time updates on Midwest power plant metrics, boosting operational efficiency

PROJECTS

Decentralized Event Management System - StarkHack Hackathon Project (TSX, Cairo), June 2024

- Creating a decentralized event management system on Starknet using Scaffold Stark typescript stack with Cairo smart contracts
- Expanding to a stake-to-attend mechanism using OpenZepellin's ERC20 Components to minimize user no-shows for events
- Configuring the UI to make read and write operations to the Cairo Smart Contracts using Starknet React hooks

DeRisk Research - ODHack 4 Web3 Open Source Contribution (Python), May 2024

- Aggregated a deposit field to the Pandas DataFrame in the ZkLend loan state computation pipeline, enhancing data management
- Modified the models and schema Python files to include a new Deposit field
- Generated and applied a new migration file to update the database schema

Land Use and Land Cover Classification – Personal Project, October 2023

- Designed a multi-class Image Classifier using a PyTorch 2-D CNN on a Google Colab Notebook
- Trained, validated, and tested the model on remotely sensed images from the EuroSAT dataset, scoring a test accuracy of 84%
- Wrote a comprehensive report on the project exploring the data, model architecture, and training and eval processes

Simulated Search Engine – Team Project (Python, Bash), December 2022

- Constructed a scalable search engine using MapReduce to create segmented inverted index creation from Wikipedia data
- Implemented Python MapReduce programs for data processing, including text cleaning and term frequency computation
- Deployed an Index Server to process search queries, leveraging a RESTful approach to generate real-time JSON results
- Integrated PageRank and TF-IDF for enhanced search result relevance, balancing query dependent and independent factors
- Enabled concurrent processing in a Search Server, optimizing parallel Index Server queries for expedited results