Marvin Jakobs

LinkedIn: linkedin.com/in/marvinjakobs
Github: github.com/OfficialMarvin
Email: marv416@gmail.com
Mobile: +1-612-298-3926

Education

- **The Pennsylvania State University** — State College, PA
  Data Sciences *- College of Information Sciences and Technology* — *August 2020 - May 2024*
  Semester Abroad - Rochester Institute of Technology, Croatia
  Courses: Data Science 1-4, Computer Science 1-2, Calculus 1-2, Statistics 1-3, Data Integration, and Machine Learning

Skills Summary

- **Languages**: Python, R, SQL, JavaScript, Solidity, Scala
- **Libraries**: Sklearn, Matplotlib, Pandas, NumPy, Beautiful Soup, Git, PyTorch, Web3js
- **Tools**: Salesforce, RStudio, SQLite, Jupyter, Tableau
- **Soft Skills**: Teamwork, Writing, Speaking, Problem-solving

Employment History

- **Cencora (formerly AmerisourceBergen)** — Conshohocken, PA
  Customer Relationship Management Intern — *Jun 2022 - Aug 2022*
  - Worked with the IT team to export and document metadata and UI/UX for multiple permission sets in production. Gave suggestions for field and flow updates. Became proficient in **Salesforce** and used **SQL** and **Excel**.

- **College of IST, Penn State** — State College, PA
  Summer Research Assistance — *May 2021 - September 2021*
  - Researched anomalies in global wildlife trade and illegal seizure data using **Python** tools. Cleaned and spatially clustered data to recognize patterns and modus operandi of poachers/illegal wildlife traders.

- **App Developer** — Remote
  Freelance — *April 2018 - July 2020*
  - Developed and sold both iOS and Android apps. Used Unity with **C#** for Amazon, and **XCode** with **C++** for iOS. Auctioned the rights to several apps on Flippa, used Google AdSense, and ran Instagram ads to market.

Publications

- **An Analysis of Wildlife Seizure Data Distributions using Spatial Clustering**:
  Presented at Tanzania Wildlife Research Institution Conference, December 2021 (Co-authored with Faizan, Kalidindi, Mitra, and Kinyua)

Top Projects

- Ocean Trash Detection: Developed a ML system using Sentinel-2/Landsat-8 imagery and CNN models like **ResNet50** to detect marine pollution. Integrated Hugging Face APIs for enhanced detection. Implemented a web-based visualization platform.
- TNFD Research Tool: Worked with a professor to develop a Retrieval-Augmented Generation (**RAG**) system using **LangChain** and advanced LLMs. Integrated reports on the Taskforce on Nature-related Financial Disclosures for real-time and accurate data retrieval.
- Stock Trend Prediction: Utilized sentiment analysis on tweets and historical stock data to predict stock trends with a model accuracy of 51%. Implemented data processing and machine learning models in **Databricks**, employing VADER for sentiment analysis.
- Brain Tumor Classification: Developed machine learning models for grading gliomas using imaging and genomic data, emphasizing algorithm selection and model accuracy. Used datasets from UC Irvine and TCGA, tested **SVM**, **Random Forest**, and neural networks.
- CloudyAI: Developed a generative AI chatbot using **TypeScript**, Tailwind CSS, and OpenAI's API, deployed at cloudyai.fun using Vercel. Designed context-aware responses with system prompts and built responsive front-end features.
- Blockchain Battleship: Developed a decentralized game on Ethereum with **Solidity** and **JavaScript**. Features included real-time messaging and gameplay mechanics, with a focus on smart contract security and frontend interaction via **Web3.js**.