Victoria Puck-Karam

(571) 551 1500 | vbp5103@gmail.com | victoriapuckkaram.com

Education

The Pennsylvania State University — University Park, PA
Bachelor of Science in Computational Data Science — *Aug. 2020 – May 2024*
College of Electrical Engineering and Computer Science (EECS)
Minors in Mathematics and Engineering Leadership Development

Experience

Software Engineer I — August 2024 – Present
JPMorgan Chase & Co.- Loan Origination Operations — *Wilmington, DE*

- Architected and productionized a domain-specific LLM RAG pipeline integrating OpenSearch vector DB with GPT embeddings model for semantic retrieval of unstructured data, eliminating knowledge silos, and accelerating root-cause analysis.
- Delivered full-stack internal tools with Java Spring, Terraform, Python, and AWS, eliminating bottlenecks in the Home Lending loan delivery pipeline and saving hundreds of engineering hours annually.
- Translated complex regulatory and financial logic into backend automation using JS; deployed solutions that saved millions in compliance costs.
- Built and maintained resilient CI/CD infrastructure with Spinnaker and Jenkins, integrating contract, performance, and chaos testing to ensure safe, stable production releases.
- Implemented advanced observability tooling with CloudWatch, enabling real-time performance monitoring and reducing mean time to resolution (MTTR).
- Developed a privileged access management system for a SaaS platform, automating configurations and reducing deployment time across environments.

Software Engineering Consultant — September 2023 – November 2023
Out2Win — *State College, PA*

- Developed scalable Python AWS Lambda pipelines leveraging Playwright for automated web scraping, optimizing data ingestion for the startup's AI-powered athlete marketing system.
- Designed Airtable-based data schemas to store and query structured athlete information, supporting machine learning workflows and analytics.
- Integrated social media scrapers for Instagram and TikTok, collecting engagement and follower data to enrich athlete profiles and enhance predictive marketing models.

Software Engineer Intern — June 2023 – August 2023
JPMorgan Chase & Co.- Data Technology — *Wilmington, DE*

- Engineered a high-throughput Java Spring Boot microservice to decrypt, parse, and store events from multiple data streams; integrated Kafka consumers to process over 10,000 messages per minute with real-time ingestion and persistence.
- Spearheaded the development of a Kubernetes-hosted Java back-end service for data pipeline management, alleviating a critical pain point for 20+ developer teams.
- Implemented JUnit unit tests and Cucumber integration, achieving a test coverage rate of 98.2%.
- Orchestrated deployment of a production-ready Java microservice through Jenkins and Kubernetes.

Data Science Intern — May 2022 – August 2022
MAXAR Technologies — *Westminster, CO*

- Surfacing actionable insights from large-scale historical sales and geospatial data sets and visualizing data using an ESRI dashboard to advise optimization and effective data-driven decision making essential to intelligence customers and in advising the sales execution teams.
- Executing the business understanding step of the data science life cycle by facilitating interdisciplinary communication to translate client, product manager, and sales team asks into executable technical tasks.
- Automating robust ETLs (Python & SQL) using an Apache Airflow server to ingest data from S3 buckets into PostgreSQL RDS.
- Deriving data understanding through performing statistical analysis using Scikit learn, PyMC and NumPy and producing preliminary visualizations with matplotlib, Seaborn, and PyPlot, as a precursor to modeling and finalized visualization.

Projects

CMCI Index | *R, MatLab, MCMC, GARCH, Kalman Filtering* January 2023 – June 2024

- Utilized advanced methodologies: temporal and geographical expansion, Principal Component Analysis (PCA), Markov Chain Monte Carlo (MCMC) modeling with Stein Thinning, Kalman Filtering and GARCH Volatility Model
- Optimized index offers a comprehensive understanding of the U.S. economy, particularly post-COVID-19
- Contributed to literature on proxy indexes for consumer sentiment by refining factor model, assessing predictive power, and uncovering symmetries between consumer sentiment and macroeconomic factors
- Developed a front-end dashboard to display the CMCI using Apache Airflow and Apache Superset

Website Accessiblity Project | *Python, BeautifulSoup, Web Scraping* November 2020 – May 2022

- Developing Python algorithms to automate data analytics of elements in Digital Healthcare to draw insights about the accessibility of essential healthcare services.
- Utilizing web-harvesting technology to extract data from 10,000 US hospitals' user interfaces
- Normalizing and wrangling web-scraped data for analysis

Spot-Suggest | *Python, Scikit-Learn, PyTorch, TensorFlow* April 2021 – August 2021

- Deployed and trained a machine learning algorithm utilizing data from the Spotify API.
- Utilized cosine similarity functionality to determine mathematical similarity of MIDI files

X-READ | *Python, TensorFlow, Keras* August 2019 – July 2020

- Developed CNN Machine Learning algorithm in Python3 using TensorFlow and Keras libraries within an Anaconda environment.
- Automated the diagnosis of common Cardiothoracic illnesses from X-RAY scans from the NIH database.
- Achieved an average accuracy of 65-75%, with the highest precision in diagnosing X-rays with visible masses reaching 91% accuracy

Technical Skills

Languages: Java, Python, SQL, JavaScript, HTML/CSS, R, MatLab, Octave
Frameworks: AWS CLI & SDK, Terraform (IaC), Spring Boot, React, Node.js, JUnit
Developer Tools: Git, Docker, VS Code, Visual Studio, PyCharm, IntelliJ, Eclipse, Anaconda, JIRA
Libraries: pandas, NumPy, Matplotlib, pytorch, tensorflow, keras, scikit, scipi, pybrain, BeautifulSoup