

Victoria Puck-Karam

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

EDUCATION

The Pennsylvania State University

Bachelor of Science in Computational Data Science

College of Electrical Engineering and Computer Science (EECS)

Minors in Mathematics and Engineering Leadership Development

University Park, PA

Aug. 2020 – May 2024

EXPERIENCE

JPMorgan Chase & Co.

Software Engineer II

Wilmington, DE

January 2026 – Present

- Recognized in the top 5% of global early-career software engineers and fast-tracked for promotion based on performance and impact.
- Architected 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.
- Modernized a legacy monolithic automation system into a scalable microservices architecture, improving maintainability and reliability while preserving production stability and user experience.
- Rebranded and scaled the internal automation platform, driving 922% user growth and expanding infrastructure to support significantly higher workload volumes without performance degradation.
- Architected 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.

Software Engineer I

August 2024 – January 2026

- Delivered full-stack internal automation 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

State College, PA

Out2Win

- 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 SQL 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

JPMorgan Chase & Co.- Data Technology

June 2023 – August 2023

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

MAXAR Technologies

May 2022 – August 2022

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 <i>R, MatLab, MCMC, GARCH, Kalman Filtering</i>	January 2023 – June 2024
<ul style="list-style-type: none"> 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 Accessibility Project <i>Python, BeautifulSoup, Web Scraping</i>	November 2020 – May 2022
<ul style="list-style-type: none"> 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 <i>Python, Scikit-Learn, PyTorch, TensorFlow</i>	April 2021 – August 2021
<ul style="list-style-type: none"> 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 <i>Python, TensorFlow, Keras</i>	August 2019 – July 2020
<ul style="list-style-type: none"> 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