

# Anni Bamwenda

562-328-7736 | Raleigh, NC | [annibamwenda@gmail.com](mailto:annibamwenda@gmail.com) | [LinkedIn](#) | [GitHub](#)

## PROFESSIONAL SUMMARY

Software Engineer II with 3 years of experience building ML-backed systems, automation, and infrastructure validation tools. Strong Python engineer experienced in productionizing ML workflows, building automation, and improving reliability of large-scale systems.

## EXPERIENCE

### Bank of America

#### *Software Engineer II*

Charlotte, NC

Mar 2025 - Present

- Led CI/CD infrastructure enablement by consolidating scripts into version control, securing sensitive scripts, and deploying workflows using Jenkins and Ansible to ensure consistent execution across development and lower-level environments.
- Owned and built a Python-based analytics pipeline to audit Dynatrace API token usage and life-cycle, supporting secure API modernization and improving reliability of enterprise monitoring infrastructure.
- Built and productionized shell-based automation for on-demand server log analysis, enabling faster root-cause analysis during infrastructure downtime.
- Built a Python tool to automate server setup validation for Dynatrace deployments, reducing manual checks and improving deployment consistency.
- Designed an end-to-end ML pipeline for multi-output classification using multi-modal clinical and imaging data, with emphasis on feature engineering, evaluation tradeoffs, and model reliability in real-world data.

#### *Data Scientist, Data Management Analyst*

Feb 2023 – Mar 2025

- Designed an unsupervised machine learning pipeline to classify a large dataset of 80,000+ bank servers by function and design, achieving an 85% classification accuracy and improving incident response time.
- Led a team of 5 interns to build an email summarization tool using Python, achieving 90% parsing accuracy and improving executive communication workflows.

## PROJECTS

### [Metastatic Cancer Diagnosis](#) | *Python, Pandas, Scikit-Learn*

June 2024

- Engaged in a Women in Data Science datathon to develop a random forest model predicting metastatic cancer diagnosis durations to detect healthcare disparities.
- Performed extensive data cleaning and feature engineering for 20,000+ rows of data to handle missing values, outliers, and categorical data using Python and pandas.
- Leveraged GitHub to manage code versioning and document the project, fostering teamwork during the datathon.

### [Industry Classification and Analysis](#) | *Python, NLP*

Feb 2022

- Created regular expressions and implemented lemmatization to preprocess text data, leading to an increase in data quality.
- Implemented word vectorization and tokenization to convert text into numerical form to optimize the machine learning process.
- Optimized cosine similarity metric through NumPy library to improve and evaluate the unsupervised learning model.

## TECHNICAL SKILLS

**Languages:** Python, SQL, and Shell

**ML & Data:** scikit-learn, Pandas, NumPy, feature engineering, model evaluation

**Systems & Tools:** CI/CD, Jenkins, Bitbucket, Ansible Tower, Dynatrace, Bash, Subprocess, Argparse, automation scripting

**Soft Skills:** Collaboration, Problem-Solving, Communication, Adaptability

## EDUCATION

---

<b>Whittier College</b>	Whittier, CA
<i>Bachelor of Arts in Mathematics, Minor in Computer Science</i>	<i>Dec 2022</i>
<b>Cornell University</b>	New York, NY
<i>Graduate Certificate in Machine Learning</i>	<i>Aug 2022</i>
<b>Contra Costa College</b>	San Pablo, CA
<i>Associate of Science in Mathematics</i>	<i>May 2020</i>