

# Anusha Guruprasad

Jersey City, NY (Open to Relocation) | anusha.guruprasad14@gmail.com | (347) 216-2098 | Portfolio | LinkedIn | GitHub

## Summary

AI/ML Developer with 5+ years of experience in machine learning, NLP, generative AI, data engineering, analytics, and cloud automation. Skilled in Python, SQL, scikit-learn, XGBoost, Hugging Face, RAG, OpenAI API, Tableau, Power BI, AWS, and data-quality pipelines, with experience building forecasting models, NLP systems, synthetic data workflows, and AI-assisted decision tools.

## Education

Pace University, New York City, NY – M.S. in Data Science, GPA: 3.66 / 4.0

May 2024

## Technical Skills

**AI/ML & GenAI:** Machine Learning, Generative AI, RAG, Hugging Face, OpenAI API, NLP, BERT, Prompt Engineering, Model Evaluation  
**Modeling & Statistics:** XGBoost, scikit-learn, Regression, Classification, Clustering, Forecasting, A/B Testing, Hypothesis Testing  
**Programming & Data:** Python, Pandas, NumPy, Seaborn, R, SQL, PostgreSQL, MySQL, BigQuery, Redshift, Data Cleaning, ETL  
**Cloud, BI & Tools:** AWS, SageMaker, IAM, Airflow, Git, JIRA, Django, Tableau, Power BI, Excel, Agile/Scrum

## AI/ML Projects

### Resume Scoring ATS Bot – Hugging Face, RAG, NLP, Python

- Built an AI-powered resume scoring system using Hugging Face models and RAG to compare resumes with job descriptions and generate role-specific match insights.
- Designed NLP workflow for resume parsing, skill extraction, semantic matching, and ATS-style scoring to support targeted resume optimization.

### Personalized Drug Treatment Analysis with Generative AI – OpenAI API, Jupyter, ADReCS

- Developed a GenAI notebook using OpenAI API calls and ADReCS data to explore personalized drug treatment recommendations for patients with pre-existing conditions.
- Implemented synthetic health-profile generation, AI-driven treatment customization, and impact assessment for patient-safety and clinical-practice outcomes.

## Experience

### Data Engineer

Feb 2025 – Sep 2025

Arco Data Design Inc. – New York, NY

- Designed automated data-quality and validation pipelines in Django and SQL, improving reliability of audit-ready data products serving 10K+ users.
- Integrated external datasets through scraping and transformation workflows, improving ad targeting accuracy by 36% and reducing latency from 800ms to 440ms.

### Senior Data Science Developer

Sep 2024 – Apr 2025

Shoptaki – New York, NY

- Developed Python/XGBoost financial forecasting models, improving forecast accuracy by 27% and reducing prediction error by 35%.
- Built Tableau dashboards and automated analytics workflows, reducing executive decision-making time by 40% and maintaining 99.2% pipeline uptime.

### Research Assistant

Nov 2023 – Jul 2024

Pace University – New York City, NY

- Analyzed 10,000+ genomic and metabolomic records using Python, R, and NIST data, identifying predictive biological markers with 92% model accuracy.
- Automated LCMS/XCMS workflows for biomarker detection, metabolite profiling, reproducible reporting, and statistical model development.

### Machine Learning Developer

Jun 2023 – Sep 2023

Shoptaki – New York, NY

- Analyzed real estate personalization data and guided a 10-member ML team, contributing to a 25% ROI improvement.
- Deployed BERT-based NLP model for lease document analysis with Microsoft SQL Graph integration, reducing review time by 35%.

### Associate Application Developer

Sep 2020 – Jul 2022

Unisys – Bangalore, India

- Automated AWS EC2 microservice deployments using CodePipeline and CodeBuild for real-time fare processing systems.
- Led Scrum workflows and mentored engineers on CI/CD, ELK, Grafana, and Oracle IAM, improving team efficiency by 40%.

### Middle School Mathematics Teacher, Data-Driven Instruction

Present

Success Academy – New York, NY

- Built performance trackers across 5 cohorts and 100+ students; analyzed skill gaps and improved proficiency rates from 62% to 79%.

## Certifications, Publication & Awards

- **AWS Certified AI Practitioner** – [View badge](#); AWS ML Associate in progress.
- **Publication:** [Microbes and Metabolites of Tetrastigma](#) – Integrated metagenomic and metabolomic datasets to identify microbial and chemical patterns supporting Rafflesia propagation.
- **Morgan Stanley Hackathon Winner:** Built AI-driven food redistribution analytics system; automated storage analysis and reallocation recommendations, reducing projected food waste by 28%.