

Luke Acosta

☎ (303) 555-0186 ✉ luke.acosta.ai@example.com

🌐 linkedin.com/example/lukeacosta 📍 Denver, CO 80206



SUMMARY

Vice President and Data Scientist Lead with eight years of applied experience delivering machine learning, **natural language processing**, and large language model solutions across financial services operations, analytics, and controls. Advanced computer science training supports hands-on leadership across Amazon Bedrock, Cortex Analyst, Amazon SageMaker, Databricks, Python, PyTorch, pandas, NumPy, and scikit-learn. Leads production programs spanning **retrieval-augmented generation**, semantic search, **content extraction**, question answering, summarization, recommendations, and agentic workflows. Builds governed APIs, pipelines, evaluation frameworks, and **machine learning operations** practices that address accuracy, faithfulness, robustness, latency, cost, safety, and rollback readiness. Trusted by senior leaders, engineers, product teams, risk partners, and commercial banking stakeholders; brings practical judgment, care for users, and enthusiasm for building reliable capabilities that improve service, efficiency, planning, and controls.

EXPERIENCE

Vice President, Data Scientist Lead

September 2023 - Present

HarborPoint Financial Group

Denver, CO

Lead a nine-person data science and applied research practice delivering governed **generative artificial intelligence** capabilities through Amazon Bedrock and Cortex Analyst for commercial banking planning, service, and control workflows.

- Led Bedrock and Cortex Analyst programs for commercial banking planning, service, and control workflows.
- Rolled out retrieval-augmented **question answering** across policy content using embeddings, vector search, and citations.
- Established evaluation gates for accuracy, faithfulness, robustness, latency, cost, **red-teaming**, and **safety checks**.
- Partnered with engineering on FastAPI services, pipelines, containers, testing, CI/CD, and operational runbooks.
- Mentored five data scientists on **prompt engineering**, tool use, **structured outputs**, reasoning, and responsible AI.

Senior Data Scientist

January 2021 - August 2023

Blue Ridge Capital Markets

Charlotte, NC

Owned machine learning and natural language processing initiatives supporting fixed-income research, portfolio reporting, and service operations across fourteen internal business functions.

- Built Bedrock extraction services for obligations, dates, and risk language across 1.2 million records.
- Designed Python retrieval workflows with pandas, NumPy, **scikit-learn**, indexing, and **vector search**.
- Implemented **prompt templates**, model routing, **tool use**, and structured JSON outputs for internal APIs.
- Created SageMaker **experiment tracking**, **model registry**, monitoring, drift review, incidents, and rollback procedures.
- Presented model tradeoffs, risks, latency, cost, and results to investment, engineering, and compliance leaders.

Data Scientist

July 2019 - December 2020

Crescent Peak Analytics

New York, NY

Delivered supervised machine learning and **natural language processing** solutions for asset-management analytics, document intelligence, reporting applications, and client-facing research initiatives.

- Developed **Python** pipelines for transaction cleansing, document parsing, embeddings, and indexed retrieval corpora.
- Compared **PyTorch**, **TensorFlow**, pandas, NumPy, and scikit-learn approaches against analyst-labeled evaluation sets.
- Exposed predictions through tested APIs and scheduled batch processes for reporting applications.
- Introduced evaluation notebooks measuring precision, recall, **faithfulness**, **latency**, and error categories for stakeholders.
- Studied research on prompting, fine-tuning, retrieval, and agent design; converted findings into practical experiments.

Machine Learning Associate

August 2018 - June 2019

LatticeWorks Solutions

Boston, MA

Supported supervised delivery of early machine learning and natural language processing engagements for **financial services** clients under senior data scientist guidance.

- Prepared account and market datasets with SQL, Python, **pandas**, and **NumPy** validation routines.
- Built scikit-learn classifiers and similarity searches against curated samples for nontechnical client reviewers.
- Contributed reusable preprocessing modules, **Git** workflows, unit tests, and **continuous integration** checks.
- Tested chunking, **embeddings**, retrieval thresholds, and source links within a controlled vector-index experiment.
- Presented findings, risks, and next steps in concise readouts for project and engineering leaders.

PROJECTS

Governed Commercial Banking Search 📅 2025

Built a governed search and question-answering capability that connected policy content with clear citations, access boundaries, and measurable evaluation criteria.

Agentic Research Assistant 📅 2024

Created a practical agent workflow using structured outputs, tool use, and multi-step reasoning for research and recommendation tasks.

LEADERSHIP & AWARDS

- Responsible AI Implementation Award, HarborPoint Financial Group, 2025
- Applied Analytics Innovation Award, Blue Ridge Capital Markets, 2022
- Graduate Research Fellowship, Georgia Institute of Technology, 2017

EDUCATION

Master of Science in Computer Science

Georgia Institute of Technology GPA: 4.0

2018

Atlanta, GA

Coursework: Machine Learning, Natural Language Processing, Distributed Systems, Data Mining

Bachelor of Science in Statistics

University of Massachusetts Amherst GPA: 3.8

2016

Amherst, MA

Coursework: Statistical Modeling, Probability, Regression Analysis, Experimental Design

CERTIFICATIONS

- AWS Certified Machine Learning - Specialty 📅 2024
- Databricks Generative AI Fundamentals 📅 2025
- Snowflake Hands-On Essentials: Cortex 📅 2025

TECHNICAL SKILLS

- **Generative AI Platforms:** Amazon Bedrock, Cortex Analyst, Databricks
- **Machine Learning Cloud:** Amazon SageMaker, AWS, Snowflake
- **Programming Languages:** Python, SQL, JSON
- **ML Frameworks:** PyTorch, TensorFlow, scikit-learn
- **Data Science Libraries:** pandas, NumPy, SciPy
- **Retrieval Systems:** Embeddings, vector search, indexing
- **LLM Methods:** Prompt engineering, fine-tuning, agents
- **Evaluation Methods:** Accuracy, faithfulness, robustness
- **Software Interfaces:** FastAPI, APIs, structured outputs
- **Data Pipelines:** Batch jobs, preprocessing, ETL
- **Engineering Practices:** Git, testing, code review
- **Delivery Automation:** CI/CD, containers, Kubernetes
- **MLOps Controls:** Model registry, drift monitoring, rollback
- **Safety Controls:** Guardrails, red-teaming, safety checks
- **Financial Domains:** Commercial banking, markets, asset management

SKILLS

- | | | | |
|-------------------------------|----------------------------------|----------------------|-------------------------------|
| • Generative AI | • Retrieval-Augmented Generation | • Question Answering | • Amazon Bedrock |
| • Large Language Models | • Agentic Workflows | • Summarization | • Cortex Analyst |
| • Natural Language Processing | • Content Extraction | • Recommendations | • Machine Learning Operations |
| • Prompt Engineering | • Semantic Search | • Evaluation Methods | |

PROFESSIONAL AFFILIATIONS

- Mentor, Financial Services Data Science Fellowship, September 2024 - Present
- Co-lead, Responsible AI Practitioners Forum, September 2023 - Present
- Reviewer, Internal Machine Learning Standards Council, January 2022 - Present

LANGUAGES

- English (Native)
- Spanish (Intermediate)

ADDITIONAL INFORMATION

Work Status : Authorized to work in United States. No sponsorship required.

REFERENCES

AVAILABLE ON REQUEST