

Emilio Nguyen

Data Science Intern

📞 (303) 555-0178 ✉️ emilio.nguyen@example.com

🌐 linkedin.com/in/emilionguyen 📍 Denver, CO 80202

STRENGTHS

- 🔄 **Adaptable**
Readily adjusts strategies based on project feedback, gaining insight from varying perspectives throughout research phases.
- ⚙️ **Analytical Thinker**
Breaks down complex problems methodically; peers often sought advice through challenging project hurdles.
- 👥 **Team Collaborator**
Receives high marks frequently for fostering cooperative relationships among team members across departments.
- 🎯 **Self-Motivated**
Proactively seeks additional challenges; volunteered for leadership roles in student research groups, paving way for personal growth.
- 🗣️ **Communicative Leader**
Strives for clarity in all communications; this transparency deepens trust and enhances teamwork effectiveness.

SKILLS

Python

Natural Language Processing

Data Analysis LLM APIs

Data Cleaning EHR Data Handling

LANGUAGES

English Native

Spanish Intermediate

SUMMARY

Detail-oriented Data Science Intern passionate about leveraging technology to improve healthcare outcomes. Proficient in developing and validating Large Language Model (LLM) pipelines, specializing in extracting vital clinical endpoints from Electronic Health Record (EHR) data. Currently pursuing a Master's degree in Data Science with demonstrated experience in data cleaning, extraction, and analysis tailored for clinical applications. Notable team player committed to continuous learning through cross-functional collaboration, resulting in effective communication of insights to both technical and non-technical stakeholders.

EDUCATION

Master's Degree in Data Science

University of Denver 🎓 GPA: 3.8 📅 2026 📍 Denver, CO

Coursework: Python Programming, Machine Learning, Natural Language Processing, Data Visualization

TECHNICAL SKILLS

- **Programming Languages:** Python
- **NLP Techniques:** Natural Language Processing, Text Mining, Sentiment Analysis
- **Data Visualization Tools:** Matplotlib, Seaborn, Tableau
- **Machine Learning Frameworks:** TensorFlow, PyTorch, Scikit-learn
- **Database Management Systems:** SQL, MongoDB, PostgreSQL
- **Cloud Platforms:** AWS, Google Cloud, Azure
- **Version Control Systems:** Git, GitHub, Bitbucket
- **EHR Data Standards:** HL7, FHIR, DICOM
- **Software Development Methodologies:** Agile, Scrum, Kanban
- **Data Warehousing Solutions:** Snowflake, BigQuery, Redshift

EXPERIENCE

Data Science Intern

University Research Lab 📅 May 2026 - Present 📍 Denver, CO

Facilitate innovative data projects focusing on clinical endpoint extraction and validation using LLM technology. Leverage strong coding skills in Python throughout the development of efficient workflows and documentation practices. Active participant in research initiatives across diverse teams within the lab set-up.

- Developed Python-based workflows to process unstructured clinical data, enhancing overall efficiency for health studies.
- Conducted validation efforts comparing automated outputs against manually curated datasets, ensuring heightened data accuracy.
- Collaborated actively with healthcare professionals to communicate clinical needs and translate them into actionable technical solutions.
- Built a comprehensive code repository that streamlined processes for data input and output in ongoing research projects.
- Analyzed performance metrics of different data extraction methodologies, presenting findings to enhance stakeholder decision-making.
- Contributed to the creation of an interactive interface allowing researchers to easily access processed clinical datasets.

Data Analyst

Academic Research 📅 August 2025 - April 2026 📍 Denver, CO

MY CAREER



● Data Science Intern at University Research Lab (3 Months)

● Data Analyst at Academic Research (8 Months)

- Engaged in significant academic work related to large-scale assessments of clinical data patterns and technologies. Focused primarily on optimizing various stages of dataset preprocessing and analysis in a university setting.
- Collaborated in research projects focused on applying LLM technologies to derive clinical endpoints from EHR records.
 - Utilized NLP methods for thorough data cleaning and preprocessing, successfully implementing prompt techniques for improved model results.
 - Performed detailed examination of edge cases in text data, validating results against recognized clinical scenarios.
 - Documented critical research operations and results, facilitating valuable contributions to team publications and presentations.
 - Created visualization tools effectively converting complex data results into accessible content for diverse audiences.
 - Participated in peer review sessions ensuring research deliverables were rigorous and met the required scientific standards.

LEADERSHIP & AWARDS

- Dean's List, University of Denver, 2025
- Best Presentation Award, Data Science Symposium, 2025

CERTIFICATIONS

- Python for Data Science and Machine Learning Bootcamp 📅 2025
- Natural Language Processing with Python 📅 2025

PROFESSIONAL AFFILIATIONS

- Member, Data Science Club, University of Denver
- Volunteer, Local Health Data Initiative

ADDITIONAL INFORMATION

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

REFERENCES

AVAILABLE ON REQUEST