

Alexander Koehler

978-483-8085 | Koehler.ale@northeastern.edu | [linkedin.com/in/Alexander-Koehler](https://www.linkedin.com/in/Alexander-Koehler) | github.com/AlexanderKoehler

EDUCATION

Northeastern University

MS Data Science, GPA 3.85

Boston, MA

January 2024 – December 2025

University of Massachusetts Amherst

BS Business Management, BA Psychology

Amherst, MA

September 2019 – December 2023

EXPERIENCE

Contract Data Engineer

June 2025 – Present

Tech Transformations

Cambridge, MA

- Built and deployed phone-integrated voice cloning system for cybersecurity operations, including XTTS-v2 model fine-tuning, real-time conversation orchestration, and VOIP integration for live penetration testing calls
- Developed production-grade conversational AI agents with custom-trained executive voice models, enabling security teams to conduct realistic phone-based social engineering tests through automated calling systems

Business Development Account Manager

May 2022 – September 2024

Cambridge Marketing Concepts

Cambridge, MA

- Designed and implemented a custom CRM with data collected from machine learning classification algorithms to analyze 5,000+ customer interactions, identifying key demographic segments that increased conversion rates by 20
- Served as the technical-to-business liaison for clients including Kicksta and Community Solar, securing contracts with industry-leading brands, expanding market reach by 33
- Co -lead, and delivered onboarding modules on database usage (oracle) and technical customer acquisition strategies for new team members.

PROJECTS

Radiology Classification | ML, Computer Vision, LLM Fine-tuning

October 2024 – December 2024

- Curated and processed 4,742 PA chest X-rays from the 377,110-image MIMIC-CXR dataset, implementing preprocessing pipeline with metadata merging from multiple CSV files, DICOM filtering, and binary classification labels (case/control) with 2,562 training and 285 validation samples
- Fine-tuned LLaMA-3.2-11B-Vision-Instruct for automated radiology report generation using 4-bit quantization, conversation-format instruction learning, and gradient checkpointing on NVIDIA A100 GPUs, integrating vision encoder with text generator for coherent diagnostic narratives
- Achieved strong diagnostic performance across 6 chest pathologies using ResNet-18, obtaining 84.2% F1-score and 82.8% AUC for atelectasis detection, 77.3% F1-score and 86.2% AUC for pleural effusion, and 69.1% AUC for pneumonia classification

Bluesky Sentiment Analysis | Cloud Engineering, MLOps, Business Intelligence

June 2025 – August 2025

- Built production-grade Azure streaming architecture using Functions for automated data collection, Event Hubs for message streaming, and Databricks with Delta Lake storage—processing 1,000+ posts/hour with 18-second end-to-end latency and zero data loss over deployment
- Implemented multi-dimensional sentiment analysis processing 17.5 posts/second using RoBERTa transformer models for sentiment scoring and DistilRoBERTa for 6-emotion classification, achieving 87.9% processing success rate
- Conducted temporal trend analysis across 20,000+ posts revealing hashtag trump content averaged -0.256 sentiment (65.3% negative)
- Developed automated LLM-powered trend summarization algorithm that analyzes top-engaged posts (by likes and replies) to generate real-time narrative insights, identifying prevailing opinions and thematic patterns across hashtags for comprehensive political campaign monitoring and market sentiment intelligence

TECHNICAL SKILLS

Languages: Python, Typescript, Bash/Shell Scripting, C/C++, SQL (Postgres), HTML/CSS, R

Frameworks: Git/GitHub, Docker, Jupyter Notebooks, VS Code, Azure CLI

Developer Tools: Git, Docker, TravisCI, Google Cloud Platform, VS Code, Visual Studio, PyCharm, IntelliJ, Eclipse

Libraries: NumPy, Pandas, Scikit-learn, PyTorch, Matplotlib, Seaborn, NLTK, OpenCV, Hugging Face Transformers, XGBoost