

JOEL SUSHANTH MARKAPUDI

Boston, MA 📍 617-510-2051 ✉ markapudi.j@northeastern.edu [🌐 linkedin.com/in/joemjs/](https://www.linkedin.com/in/joemjs/) [🐙 mjsushanth.github.io](https://github.com/mjsushanth)

EDUCATION

Northeastern University

Boston, MA.

M.S. in Artificial Intelligence (G.P.A - 3.9)

Relevant Coursework: MLOps, Deep Learning, Computer Vision, Machine Learning.

Jan 2024 – Present

KL University

Vijayawada, India.

Bachelor's in Computer Science

Jun 2016 – Jun 2020

SKILLS

Programming & ML/AI: Python (PyTorch, Pandas, Polars, NumPy, Scikit-learn, OpenCV), SQL, Java;

MLOps, LLM & Cloud: Docker, MLflow, Optuna; FastAPI, Flask, Streamlit; Airflow; ELK Stack.
AWS (S3, Lambda, Bedrock, S3 Vectors, ECS); GCP; GitHub Actions CI/CD.

Data & BI: DuckDB, PostgreSQL, Power BI, Talend; ETL architecture, Data modeling.

ACADEMIC PROJECTS

FinSights: LLM RAG Financial Data Intelligence System

Sep 2025 – Dec 2025

- Built E2E **RAG system** for analyst-style queries over **SEC Form 10-K data** with **hybrid retrieval architecture**. Combines structured KPI tables and sentence-level vector search, powered by fuzzy entity extraction.
- Built token-aware, **embedding pipeline** processing 1,850 vectors/min with **intelligent batching (96 texts/call, 15K token budget)** and real-time cost tracking. Utilizes **S3 Vector store** with metadata filtering.
- Structured retrieval pipeline with **query variant generation**, triple retrieval regimes (filtered/global/union), sentence window expansion, deduplication, and stratified distance-based selection (0.35 cutoff). Includes provenance tracking.
- Deployed **3-tier production system** (LLM orchestrator, FastAPI + Streamlit) with cost modeling that achieves **\$0.017/query** through S3 Vectors. Derived scaling formulas and example scenarios for cost analysis.
- Achieved **0.82-0.86 BERTScore F1** on business-realistic gold test. Validated retrieval quality through deterministic neighbor tests (**98%+ Self@1, 70-92.5% Hit@5**). Evaluation with multiple metrics (**BLEURT-20, ROUGE-L**).

Pose Synthesis With Cross-Modal Diffusion Architecture (PyTorch, CLIP)

Jan 2025 – Apr 2025

- Devised Conditioned UNet diffusion architecture processing 23.3K+ training samples with **8-head cross-attention & 4 resolution levels**, generating 3D poses from 66-dimensional skeletal representations.
- Modeled **hierarchical cross-attention** with first-action syntactic segmentation and cluster-aware sampling, obtaining body-part and action-aware semantic conditioning. Achieved **0.69 final loss** through 3-phase progressive training.
- Enforced **biomechanical structure awareness**, anatomy validation, forward kinematics constraints, tanh-bounded outputs (1.5× scaling) and CFG guidance scale 7.0. Achieved **70% visual alignment** and topological consistency.

Late-interaction Image-Text Model: (ViT+CLIP, MaxSim retrieval, LoRA)

Nov 2025 – Dec 2025

- Implemented **ColBERT-style late-interaction retrieval** with ViT-base encoder (196 patch tokens, 768-d), CLIP, and learned projection head, using **MaxSim scoring** for fine-grained token-patch alignment over **Flickr8k** data.
- Applied **LoRA parameter-efficient fine-tuning** to ViT attention layers, training **1.3M/87M parameters (1.05%)** with symmetric **InfoNCE contrastive loss**. Achieved **48% loss reduction** with **82% retrieval accuracy**.

EXPERIENCE

FP&A and SG&A Reporting

Mar 2020 – Nov 2023

- Led BI delivery for non-technical finance and operations stakeholders across multiple business units.
- Built scalable **ETL pipelines (Talend, PL/SQL)** and **Power BI models** supporting 11+ programs including Producer Commission, QBR dashboards, DSO Reports, Hiring Metrics, and SG&A reporting.
- Shipped **25+ dashboards** covering **500+ metrics**. Automated margin/audit workflows, cutting **70+ hours/month** of manual Excel work. Implemented CI/CD upgrades across projects that **reduced deployment time 50%**.
- Developed forecasting logic for YTD actuals and projections using advanced DAX with multi-level attribution.
- Designed dynamic eligibility engine with **multi-factor scoring** (GP projections, new placements, total starts) and threshold-based qualification logic for a margin-based **ranking dashboard**.
- Partnered with data science and business teams on analytics improvements. Conducted **35+ interviews** and mentored sub-teams of **2-4** to accelerate delivery. Earned the **Innova Idol Award** for being a top performer.

Operational Reporting

May 2022 – Nov 2023

- Architected enterprise data warehouse integrating Bullhorn/HGC/JobDiva/Oracle EBS with **25+ PL/SQL procedures** across **40+ star-schema tables**. Implemented **SCD Type 2 historization**.
- Engineered **entity resolution** using 6+ fuzzy patterns and dual-path PnL attribution, achieving **88% accuracy**.
- Built a **data governance & reliability layer** with automated validation, duplicate detection, **cross-system ID reconciliation**, and role classification with tenure/designation mapping logic, reducing manual verification workflows.
- Delivered operational reporting dashboards serving **300+ KPIs** with forecasting, actuals-vs-baseline tracking, and rollups across individual, team, and division hierarchies for executive/producer analytics.
- Enabled **matrix-style drill-down** for faster exploratory analytics, RCA and improved reporting accuracy.