

Keshav Goel

US Citizen | kgoel9657@gmail.com | linkedin.com/in/goel-keshav | keshavgoel.dev | github/keshavgoel787

EDUCATION

Northeastern University Bachelor of Science in Data Science and Mathematics GPA: 3.88/4.00 Coursework: Data Structures, Algorithms, Machine Learning, Statistics, Data Science, Linear Algebra, Calculus, Probability Theory, Database Systems, Computer Vision, Natural Language Processing, A/B Testing, Data Visualization	Expected May 2027 Boston, MA
--	---------------------------------

EXPERIENCE

ProcureMate AI <i>Software Engineering Intern</i> <ul style="list-style-type: none">Engineered OCR pipeline using GPT-4o Vision and Tesseract.js to parse 100+ dental supply invoices, using fuzzy matching to map products to inventory and reduce manual entry by 90%.	Jan 2025 - Present Boston, MA
WPS Data Lab <i>Data Science Research Assistant</i> <ul style="list-style-type: none">Engineered TensorFlow regression model analyzing 20K+ EPA H-2A violations to predict high-risk agricultural areas, improving enforcement efficiency by 35% through targeted inspections.Developed Python ETL pipeline processing 500GB EPA data across 15 sources, reducing ingestion errors by 40% via automated Pandas schema validation.	Oct 2024 - Present Boston, MA
Johnson & Johnson <i>Software Engineering Co-op</i> <ul style="list-style-type: none">Developed AI/ML infrastructure for research acceleration and automation, assisting 300+ researchers across 5 clinical teams.Architected LangChain agent pipeline with fine-tuned GPT-4o autonomously summarizing 500+ biomarker papers, reducing literature review time from 2 hours to 30 minutes per paper.Engineered Neo4j knowledge graph of 100K+ disease–gene–drug links powering GPT-based GraphRAG pipeline, eliminating manual cross-referencing and increasing retrieval accuracy by 25%.Architected unified Streamlit MLOps interface for training and inference across 10 wearable health models, replacing manual deployments with one-click solution reducing model overhead costs by 42%.	Jun 2025 - Dec 2025 Raritan, NJ
VDart Inc. <i>Software Engineer Intern</i> <ul style="list-style-type: none">Reduced screening time by 86% (7 min to 1 min) per candidate by automating resume analysis and interview transcription for 200+ daily applications.Fine-tuned Ollama LLM on 20K resumes via HuggingFace Transformers and built FastAPI pipeline with MongoDB and AWS S3 for scalable candidate processing.	May 2024 - Aug 2024 Atlanta, GA

PUBLICATIONS

Co-Author, et al. "Worker Protection Enforcement: Description, Variation, and Predictors of Regulatory Vigor" (In Prep)

PROJECTS

ForgetMeNot (HackRU Winner) <i>FastAPI, Snowflake, Gemini, Google Cloud, React, OpenCV</i> <ul style="list-style-type: none">Built AI-powered memory assistance platform for dementia patients using facial recognition and multimodal LLM narration, winning HackRU against 500+ participants.Implemented vector search pipeline for content clustering with real-time audio generation, processing patient photos and medical records for personalized care.	
BreatheEasy (CSBase Hacks Winner) <i>Flutter, Dart, Google Maps API, Shelf, Docker</i> <ul style="list-style-type: none">Built Flutter mobile app optimizing navigation routes based on real-time air quality data, winning CSBase Hacks against 300+ participants.Designed Dockerized microservices backend computing optimal routes balancing air quality, distance, and travel time.	

TECHNICAL SKILLS

Languages/Frameworks: Python (FastAPI, Flask), JavaScript/TypeScript (React, Node.js), Go, Java, C++, R, SQL
Databases/Cloud: MongoDB, MySQL, PostgreSQL, Snowflake, AWS (EC2, S3, Lambda), Docker, Kubernetes, Jenkins
ML/Data Science: PyTorch, Scikit-Learn, Pandas, NumPy, HuggingFace, Ollama, OpenCV, NLTK, Matplotlib, Snowflake Cortex
Concepts: Microservices, REST APIs, ETL Pipelines, Machine Learning, Computer Vision, Data Modeling, End-to-End pipeline, Cross functional Decision Making, Workflow Orchestration Layers, Data Querying, Data-Driven Decision Making, Semantic Models, Analytics