RIGVED THORAT

+1 (303) 435-7650 | rigvedthorat@gmail.com | linkedin.com/in/rigved-thorat | github.com/rigvedthorat
Work Authorization: F-1 Visa | OPT Eligible Jun 2026 | STEM OPT Eligible (24 months)

PROFESSIONAL SUMMARY

AI/ML Engineer with production experience building RAG infrastructure, multimodal AI systems, and LLM-powered products. Built managed RAG pipelines on AWS Bedrock and OpenSearch at Rule4 and shipped GenAI features at The Builder Market spanning image generation, smart interior editing, and multimodal repair analysis. Strong in Python, PyTorch, TensorFlow, and TypeScript, with hands-on experience deploying LLM and vision workflows on AWS and Azure. IEEE-published in computer vision for plant disease classification.

TECHNICAL SKILLS

Generative AI & LLMs: RAG, Prompt Engineering, LLM Fine-tuning, Vector Embeddings, Semantic Search, Multimodal AI, Agents, AWS Bedrock, OpenSearch, Azure OpenAI, GPT-4o
Machine Learning & Deep Learning: TensorFlow, PyTorch, Scikit-learn, OpenCV, Pandas, NumPy, CNNs, Transfer Learning, Recommendation Systems, Model Evaluation, Precision/Recall
Programming Languages: Python, JavaScript, TypeScript, SQL, HTML, CSS
Backend & Web Development: Django, Flask, Node.js, NestJS, REST APIs, React, Angular
Cloud & Developer Tools: AWS (ECS, RDS, Bedrock, S3, Lambda), Azure AI Foundry, GCP, Docker, Git, GitHub Actions, GitLab CI
Databases & ORMs: PostgreSQL, MySQL, Neo4j, vector search (OpenSearch), Prisma ORM

PROFESSIONAL EXPERIENCE

AI Engineer | Capstone Project Oct 2025 – Apr 2026
Rule4, Inc., Boulder, CO (Hybrid)

- Built an AI-assisted ticket drafting feature on Rule4's TFox helpdesk platform that generates context-aware ticket titles and descriptions from user input and attachments, using Python, Django, AWS Bedrock, and OpenSearch.
- Designed end-to-end RAG pipelines using AWS Bedrock Titan embeddings and OpenSearch kNN search to retrieve top-**5** client- and queue-scoped similar tickets for prompt orchestration.
- Developed backfill and on-save ETL workflows that vectorize historical TFox tickets from MySQL into an OpenSearch index, improving draft relevance and semantic search across past tickets.
- Shipped feedback and pipeline analytics dashboards tracking retrieval behavior and **P95** latency across embedding, search, and generation stages, with end-to-end latency of **6** seconds.

AI Product & Prototyping Intern Jun 2025 – Sep 2025
The Builder Market, Los Angeles, CA (Remote)

- Built Angular image generation workflows backed by NestJS REST APIs, integrating Azure OpenAI GPT Image 1 and Black Forest Labs Flux for text- and image-based home design generation.
- Engineered **3** smart interior editing tools for paint, object replacement, and surface reskinning using GPT-4o image analysis, reference-image inputs, and Sharp-based masking, bounding boxes, and feathering logic.
- Developed a multimodal DIY repair analysis flow with GPT-4o Vision to convert uploaded home images into **4-part** repair guidance covering tools, materials, repair steps, and cost/time estimates.

PROJECTS

LLMs as Movie Recommendation Systems | *Python, GPT-4o-mini, Pandas, PyTorch* Apr 2025 – May 2025

- Fine-tuned a GPT-4o-mini model on the Netflix Prize dataset with **50,000+** viewers and **18,000+** movies to build a recommendation engine for complex customer preference understanding.
- Constructed custom training datasets by sampling **10,000** viewers and identifying "ground truth" movies rated 5 stars by over **50%** of similar viewer groups.
- Achieved a **20x increase** in recommendation relevance, improving Precision@5 from **0.02 to 0.42** compared to the baseline zero-shot model.

Image-Based Early Classification of Tomato Leaf Diseases | *Python, TensorFlow, CNN, OpenCV* Nov 2023 – Mar 2024

- First-authored an IEEE conference paper on a CNN for multiclass tomato leaf disease classification, benchmarked against SVM, ResNet, MobileNet, DenseNet, and InceptionV3. DOI: 10.1109/INCOS59338.2024.10527655
- Trained a 4-layer CNN (16/32/64/128 filters, ReLU, global average pooling) on PlantVillage images with OpenCV preprocessing and augmentation, reaching **95.8%** accuracy on 10-class tomato leaf classification after **100** epochs.

EDUCATION

University of Colorado Boulder | M.S. Computer Science, GPA: 3.70/4.00 Aug 2024 – May 2026, Boulder, CO, USA
SRM University | B.Tech. Computer Science & Engineering (AI & ML), GPA: 8.99/10.00 Sep 2020 – May 2024, Chennai, TN, India