

PRASOON RAJ

📍 Boston, MA 📞 +1 (857) 991-9119 ✉️ prasoon.raj.041@gmail.com 🔗 LinkedIn 🐙 GitHub 🌐 Portfolio

SUMMARY

AI/ML engineer with 4+ years of experience designing and deploying machine learning and generative AI systems for enterprise and government clients. Proven ability to translate complex challenges into impactful real-world solutions, including a GenAI procurement platform showcased at the Massachusetts State House.

EDUCATION

Northeastern University, Boston, MA

Master of Applied Machine Intelligence

Relevant Coursework: Applications of AI, Big Data and Data Management, Project Management

Jun 2025

GPA: 3.9

TECHNICAL SKILLS

Languages: Python, SQL, Typescript (familiar)

AI/ML: LLMs (OpenAI, Claude, Llama), RAG, Prompt Engineering, AI Agents, NLP, Deep Learning, RAGAS

Frameworks: PyTorch, TensorFlow, Keras, Flask, FastAPI, LangChain

Cloud/Data: AWS (Bedrock, Kendra), Azure (DevOps, ML, Foundry, Repos), PySpark, FAISS, Git, CI/CD

Tools: Power BI, Tableau, Linux, MongoDB, SQL Server, Postman, Docker

CERTIFICATIONS

Solutions Architect – Associate (AWS Certified)

Azure Fundamentals (Microsoft Certified)

WORK EXPERIENCE

Burnes Center for Social Change

Generative AI Intern

Jan 2025 – Present

Boston, MA

- Built a FedRAMP-compliant multi-agent AI document search system for Health and Human Services (HHS) using AWS Kendra, reducing document retrieval time by 60% and accelerating ticket resolution for 500+ support staff.
- Designed LLM agents with AWS Bedrock, and led development of a RAGAS-based evaluation pipeline that improved response quality scoring by 25% across critical use cases.
- Mentored 10+ co-op students in developing GenAI solutions for different state government agencies, achieving an 85% success rate in prototype delivery and enabling 5 projects to demo at institutional showcase events.

Burnes Center for Social Change

Generative AI Engineer Co-op

Jul 2024 – Dec 2024

Boston, MA

- Partnered with the Operational Services Division (OSD) to streamline procurement workflows by developing a generative AI chatbot with AWS Bedrock and Knowledge Base, reducing document lookup time by 60%.
- Implemented a metadata-enhanced RAG pipeline supporting mixed document types (images, videos, PDFs), boosting document discoverability and reducing time-to-decision for buyers by 25%.
- Spearheaded a conflict detection system, collaborating closely with compliance teams, enhancing data reliability and aligning with state directives to improve regulatory compliance by 40%.

Enquero – Genpact

Senior Data Scientist

Aug 2022 – Aug 2023

Bengaluru, India

- Drove engagement with a US-based global sportswear brand on financial modeling and inventory clustering, leveraging unsupervised machine learning and aligning cross-functional collaboration to optimize product launches, reduce overstock by 15%, and boost profitability by 8%.
- Engineered a modular analytics framework and integrated Python-based predictive models into Power BI dashboards, reducing reporting time by 90% and enabling finance teams to make faster, insight-driven decisions.
- Acted as a lead coach in Genpact's GenAI Hackathon, guiding 500+ freshers through prototype design and creation, resulting in three winning entries showcasing innovative generative AI applications.

Accenture

Data Science Analyst

Jun 2019 – Aug 2022

Bengaluru, India

- Architected a document anonymization system for a healthcare and insurance client using NLP and computer vision to automatically redact personal identifiers and brand elements, automating 80% of redactions across 200 enterprise clients.
- Enhanced image classification using CNNs (TensorFlow-Keras) and YOLO (PyTorch), with Azure ML for experimentation and deployment, improving redaction precision by 15% and speed by 40%.
- Integrated Django-based NLP APIs with real-time computer vision pipelines and used Azure Repos for version control, delivering compliance dashboards now used by over 3,000 employees.