

# ASHUTOSH CHOUDHARI

4ashutosh98@gmail.com • (412) 626-9743 • [LinkedIn](#) • [GitHub](#) • [Portfolio](#) • Jacksonville, FL

## EDUCATION

**CARNEGIE MELLON UNIVERSITY**, Pittsburgh, PA

**Aug 2022 - Dec 2024**

**Master of Information Systems Management**

**CGPA: 3.78/4.00**

Coursework: Computational Data Science, Unstructured Data Analytics, Time Series Forecasting, Data Structures and Algorithms

**VELLORE INSTITUTE OF TECHNOLOGY**, Vellore, India

**Jul 2016 - Jun 2020**

**Bachelor of Technology - Electronics and Instrumentation Eng** (Minor in Computer Science and Eng)

**CGPA: 3.94/4.00**

Coursework: Database Management systems, Neural networks, Python programming, Statistics, Operating Systems

## PROFESSIONAL EXPERIENCE

**AI/ML Engineer, Pyramyd** (Remote)

**Jun 2025 - Present**

- Designed a **transformer-based ML model** in **PyTorch** to predict ratings for software product reviews achieving **87%** accuracy.
- Developed **AI-powered RFP Requirements Extractor** using **o4-mini LLM** and **Azure AI Search**, implementing parallel processing architecture to extract, categorize requirements from proposal documents with **90%+** accuracy across documents.
- Optimized performance for **5x** concurrent processing with vector search for intelligent document cross-referencing.

**Software Developer, Confirmed LLC** (Pittsburgh, PA)

**Feb 2025 - Sep 2025**

- Developed microservice using **LangChain**, **FastAPI**, **Azure Container Apps** for nametag scanning, achieving **92%** accuracy.
- Integrated **Python service** with .NET backend via **REST APIs**, optimized SQL query performance by **43%** through **NL-to-SQL**.
- Built **CI/CD pipeline** in **Azure DevOps** with **pytest** testing and **Docker** deployment; reduced deployment time by **70%**.

**Machine Learning Engineer (Student Consultant), Nexteer Automotive** (Remote)

**Aug 2024 - Dec 2024**

- Created document comparison application to extract, compare, and analyze PDF documents using **Azure AI Document Intelligence**, **Adobe PDF Extract API** and **OpenAI's GPT-4o**, enhancing document management efficiency and saving **2 FTEs**.
- Implemented user-friendly Streamlit GUI with a robust backend data handling system leveraging **MongoDB** for persistent caching, to minimize API call costs and **improve processing speed by 5 times**.

**Data Science Intern, MedSarathi LLC** (Remote)

**Jun 2023 - Aug 2023**

- Developed an A/B tested, regression-weighted recommendation engine, increasing residency acceptance by **14%**.
- Built a full-stack web app (**GPT-4**, **React**, **Flask**, **PostgreSQL**, **Heroku**) for streamlined service delivery and advanced analytics.

**Solution Delivery Analyst, Deloitte** (Hyderabad, India)

**Jun 2021 - Jul 2022**

- Developed a **Python-based IOC triage tool** to process **500+ IPs, domains, URLs, and emails** daily, classifying inputs, and integrating OSINT APIs (VirusTotal, WHOIS, AbuseIPDB), reducing average query time by **65%** with parallel execution.
- Architected a Flask REST backend with PostgreSQL and Redis**, exposing endpoints for IOC normalization, rate-limit handling, and API key management, enabling **multi-user scalability and cutting redundant OSINT calls by 40%**.
- Built a React-based frontend with secure API integration** to provide dashboards, results tables, and export workflows (Markdown/PDF/CSV), ensuring **consistent and auditable threat intelligence reporting across environments**.

## ACADEMIC PROJECTS

**Network Security Phishing Detection System** (*MongoDB, MLflow, AWS ECR, EC2, Docker, FastAPI*).

**Jan 2025 - Mar 2025**

- Developed a machine learning pipeline for phishing detection using 7 tuned classification algorithms and achieving 96% F1-score and 97% precision with the best-performing model (XGBoost), representing a 15% improvement over baseline models.
- Engineered an MLOps solution with CI/CD pipeline using GitHub Actions, AWS ECR containerization, and MLflow for experiment tracking, reducing deployment time by 78% and enabling real-time phishing detection with 99.2% uptime.

**CGM Systems Analysis and Improvement Recommendations** (*NLTK, Scikit Learn, genism, matplotlib*)

**Aug 2023 - Dec 2023**

- Analyzed 40,000+ social media posts about CGM systems using LDA topic modeling (0.78 coherence score) and VADER sentiment analysis (89% accuracy), identifying 5 key discussion topics per brand and user sentiment patterns.
- Extracted actionable insights by segmenting user sentiment across product features, revealing critical pain points around insurance coverage (37% negative) and app functionality (28% negative), informing strategic product improvement recommendations.

**Question-Answering System Development** (*Azure Machine Learning, PyTorch, nltk, spaCy, Transformers*).

**Sep 2023 - Nov 2023**

- Fine-tuned a pre-trained BERT model for question-answering tasks, achieving 68% accuracy on the SQuAD test set, a 28% improvement over the pre-trained model.
- Deployed the fine-tuned BERT model as a scalable cloud-based question-answering service on Azure Container Instances (ACI), leveraging Azure ML for deployment, achieving efficient inference on unseen data.

## SKILLS

**Programming languages:** Python, Java, SQL, R, JavaScript, HTML, CSS

**Frameworks / Libraries:** PyTorch, Tensorflow, NLTK, Hugging Face, scikit-learn, NumPy, pandas, XGBoost, flask, LangChain

**Tools:** Tableau, Jira, GitHub, Docker, MongoDB, Neo4j, FAISS, Chroma, Ollama, Sqlite, Nvidia NIM

**Certifications:** AWS Certified Cloud Practitioner, Cert. of Cloud Security Knowledge, Agile Product Owner, Agile Scrum Master

**Cloud tools:** AWS Bedrock, SageMaker, Lambda, S3, API Gateway, Azure AI Foundry, GCP Vertex AI, BigQuery