

# ASHUTOSH CHOUDHARI

4ashutosh98@gmail.com • (412) 626-9743 • [LinkedIn](#) • [GitHub](#) • Pittsburgh, PA

## EDUCATION

**CARNEGIE MELLON UNIVERSITY**, Pittsburgh, PA

**Aug 2022 - Dec 2024**

**Master of Information Systems Management**

**CGPA: 3.78/4.00**

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

**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**

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

## PROFESSIONAL EXPERIENCE

**Nexteer Automotive**

Pittsburgh, PA

**Machine Learning Engineer (Student Worker)**

**Aug 2024 – Dec 2024**

- Created a document comparison application leveraging Python to extract, compare, and analyze PDF documents using Adobe PDF Extract API and OpenAI's GPT-4o, enhancing document management efficiency thus saving 200 human hours per month.
- 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.

**Unisys**

Pittsburgh, PA

**AI Engineer (Student Worker)**

**Jan 2024 – Mar 2024**

- Improved PDF invoice data extraction accuracy from Oracle Fusion's 20% to 83% for 18 key fields across semi-structured formats by developing a Flask application powered by OpenAI GPT-4 and Azure Document Intelligence.
- Incorporated advanced prompt engineering techniques, including few-shot, chain-of-thought and zero-shot approaches, to enhance data extraction from semi-structured and unstructured invoices, reducing manual effort by 2 FTEs.

**MedSarathi LLC**

Phoenix, AZ

**Data Scientist**

**Jun 2023 - Aug 2023**

- Designed and implemented a data-driven recommendation system for medical graduates, leveraging A/B testing and regression analysis to optimize profiles and develop a weighted scoring algorithm, boosting residency acceptance rates by 14%.
- Architected a comprehensive web platform using OpenAI API, HTML, JavaScript, React, Flask, PostgreSQL, and Heroku, enabling streamlined service delivery and advanced analytics.

**Deloitte USI**

Hyderabad, India

**Solution Delivery Analyst**

**Jun 2021 - Jul 2022**

- Engineered advanced log analysis workflows using data from SIEM tools (Splunk, QRadar), EDRs, and OSINT tools identifying and mitigating suspicious activities across system, network, and endpoint logs.
- Generated over 100 detailed incident reports weekly, applying pattern recognition, anomaly detection, and OSINT tools to deliver actionable insights, driving a 32% reduction in security incidents through data-driven strategies.

## ACADEMIC PROJECTS

**Food Classification Using Transfer Learning on EfficientNet** (Pytorch, Tensorboard, Efficientnet)

**May 2024 –Jul 2024**

- Applied transfer learning on EfficientNet to classify 101 food categories from the Food101 dataset, achieving 84% accuracy through advanced data augmentation and leveraging ImageNet-trained weights.
- Optimized model performance, reducing training time by 77% and improving accuracy by 36% over a baseline VGGNet model, utilizing PyTorch and tracking metrics with TensorBoard.

**Artificial Intelligence Policy Chatbot** (Hugging Face, PyTorch, Large Language Models, Generative AI)

**Jan 2024 – Mar 2024**

- Fine-tuned Llama-2 7b model for AI policy-related inquiries using techniques like PEFT and LoRA, alongside prompt engineering and RAG, achieving a 23% improvement in BLEU, 18% in ROUGE, and 21% in METEOR scores over the baseline.
- Evaluated model performance using automated metrics (e.g., BLEU, ROUGE, METEOR, RAGAS) and human criteria, achieving a relevance score of 4.3/5 and coherence score of 4.2/5, significantly enhancing the quality of responses on AI policy queries.

**Question-Answering System Development** (Azure ML Studio, PyTorch, nltk, spaCy, SentenceTransformers)

**Sept 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, Oracle SQL, R

**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 ML studio, Vertex AI