Cohen Gress

cohengress@outlook.com • (717) 404-4118 • https://www.linkedin.com/in/cohengress001

EDUCATION

Pennsylvania State University

Expected Bachelor of Science in Data Analytics | **GPA:** 3.7/4.0 — May 2027

Community College of Allegheny County

Associate of Science in Information Systems | **GPA:** 3.6/4.0 — May 2024

Relevant Coursework: Data Structures, Algorithms, Database Systems, System Design, OOP, Discrete Math, Stats, Calculus, R, Project Management

RELEVANT EXPERIENCE

Oshkosh Corporation | McConnellsburg, PA — May 2025 - August 2025

AI Engineer Intern

- Built ML and LLM-driven automation for email classification, intelligent reply generation with internal data retrieval, and predictive modeling; presented results to executive leadership and built relationships directly with end users.
- Developed an internal LLM-powered chatbot leveraging vector search and clustering to enable department-specific, role-restricted knowledge retrieval, reducing query latency while maintaining data isolation and contextual accuracy.

TECHNICAL PROJECTS

Automated Email Categorization & Response System | *Internship Project @ Oshkosh Corp.*

Tools: Python, Azure API, Outlook, Excel, scikit-learn, pandas, Databricks, Azure, MLflow

-Implemented Python/Azure API script to extract Outlook emails into a database based on a trigger system solving a lack of automation in support workflows.

-Built a 13-class email classifier (**scikit-learn**) with **83%** accuracy; routed emails to Outlook folders based on predictions.

-Integrated **gpt-4o** to draft reply emails by querying an internal knowledge base, enabling agents to respond faster and more accurately.

-Delivered a complete ML pipeline—from data ingestion to deployment—presented outcomes to executive leadership, and achieved estimated cost savings of **$60,000** annually and plans to scale Oshkosh-wide.

Enterprise Data Analysis and Predictive Modeling | *Internship Project @ Oshkosh Corp.*

Tools: pandas, matplotlib, scikit-learn, Python, Excel

-Analyzed 11K-row dataset (84 features); cleaned and visualized data with **pandas** & **matplotlib**.

-Built **Random Forest** model (R^2: **99%** train, **96%** test; low error) on **feature importance**, driving recommendations.

Book Recommendation System with Retrieval Augmented Generation (**RAG**) | Project

Tools: Python, Qdrant, Sentence Transformers, pandas, OpenAI API, VS Code

- Built a book recommendation system using **RAG**, integrating **Qdrant** as the vector database for efficient data storage and retrieval.

- Processed and vectorized book data using Sentence Transformers for semantic similarity searches.

- Integrated OpenAI's API into a chatbot, showcasing skills in NLP, vector databases, and LLM applications.

SKILLS & CERTIFICATIONS

- **Programming Languages:** Python, Java, SQL, R
- **Tools and Technologies:** Python, Excel, Tableau, PowerBI, RStudio, VS Code, Jupyter, Matplotlib, pandas, NumPy, TensorFlow, scikit-learn, SQL SMS, Qdrant, Ollama, OpenAI API, Azure API, DSPy, Agile, Databricks, Azure Government, MLflow
- **Certifications**: ITF+, Intro to RAG Certificate, Google Data Analytics, Google AI Essentials, ML Specialization, Intelligent Swarming Fundamentals, KCS v6 Fundamentals