Philadelphia, PA william.zebrowski@me.com (267) 816-8258



ARTIFICIAL INTELLIGENCE ENGINEER

PROFESSIONAL SUMMARY

Artificial Intelligence Engineer with over 2 years of experience specializing in LLM integration, model development, and generative AI. Proven track record in building and managing custom ML/AI frameworks and developing domain-specific models. Expertise in optimizing model performance, cost, latency, and quality through state-of-the-art techniques such as LoRA, RLHF, and instruction-tuning. Passionate about leveraging neural networks and deep learning to drive innovative solutions, enhance AI capabilities, and deliver impactful results in real-world applications.

EXPERIENCE

JUL 2023 - PRESENT

Al Engineer, Benefits Data Trust, Philadelphia, PA

Generative AI Development:

- · Built, sustained, and managed the organization's first customer-facing Generative AI Assistants.
- Developed fully connected backend systems to support GenAl Assistants, integrating data flow and storage mechanisms for conversational data.
- Performed data wrangling, cleansing, and preprocessing to ensure high-quality input for AI models.
- **Utilized RAG-based architecture**, vector databases (ie, *Milvus, Pinecone*), performing semantic search, and embedding models to enhance AI development.
- Specialized in LLM API calls, custom model training and deployment, and retrieval augmented generation to enhance the functionality of AI solutions.
- Participating in the full AI development lifecycle, including data preparation, design, coding, testing, deployment, and support.

Model/LLM Development:

- Developed ML/Neural Network NLP frameworks for Conversational AI.
- Fine-tuned language models with domain/partner-specific data to meet specific requirements.
- Use Fine-tuned embedding models to enhance performance of embeddings.
- Implemented continuous improvement strategies through feedback loops and iterative training processes.
- Trained and deployed custom large language models using state-of-the-art techniques (LoRA, RLHF, instruction-tuning).
- Assisting in the identification and correction of biases in datasets and AI models.

Al Agents:

- Developed, implemented, and maintained AI Agent systems using Autogen and CrewAI.
- Created tools for AI Agents tailored to their specific use cases.
- Integrated AI Agents into other AI systems, enhancing user experience.

R & D:

- Researched the latest emerging technologies in AI.
- Built prototypes to integrate the latest AI technologies into BDT's products.
- Optimized model performance, cost, latency, and quality by specializing in model optimization techniques.

Collaboration:

• Collaborated with cross-functional teams to implement AI-driven solutions for real-world applications.

- Worked with product managers & design teams to develop new product experiences using NLP & LLMs.
- Defined and improved business & product metrics to optimize the quality and cost of AI usage.
- · Advocated for, and built, new and better ways of doing things, ensuring continuous improvement in every project.

MAR 2023 - JUL 2023

Automation Engineer, Benefits Data Trust, Philadelphia, PA

- Developed and maintained automation frameworks for UI and API testing using Python Playwright, ensuring high-quality software delivery.
- · Designed automated test scripts to validate application functionality and performance, improving test accuracy.
- · Collaborated with development and QA teams to identify test requirements, achieving comprehensive test coverage.
- Implemented CI/CD pipelines integrating automated tests, enhancing development lifecycle efficiency.
- Provided detailed test reports and logs to support defect tracking and resolution, streamlining debugging processes.
- Developed new automation tools enhancing productivity and efficiency.

MAR 2022 - MAR 2023

QA Engineer, Benefits Data Trust, Philadelphia, PA

- Built API Testing Framework from scratch w/ over 4000+ JavaScript tests in Postman.
- Perform UI, functional, API and regression testing.
- Develop and execute functional and performance automation scripts using tools such as POSTMAN, Ghost Inspector, Cypress and their frameworks. Test scripts created in Python and JavaScript.
- Works closely with Development, Product Management and Architecture to understand and guide product features and testing.
- · Contributes to the development of new theories and methods for automation in OA.
- Attends and leads daily Scrum ceremonies for the QA Team.

JUN 2018 - JUL 2022

QA Analyst, Benefits Data Trust, Philadelphia

- Led daily Scrum ceremonies, ensuring QA team alignment and efficiency.
- Executed UI, functional, API, and regression testing, identifying critical issues.
- Developed 4000+ JavaScript tests in Postman, enhancing API testing framework.
- · Collaborated with cross-functional teams to refine product features and testing.
- Automated scripts in Python and JavaScript, improving testing accuracy and speed.

EDUCATION

JUN 2022 - AUG 2022

Computer Science, Harvard University

Python Programming Bootcamp. Computer Science with Python

JAN 2010 - MAY 2012

Business/Managerial Economics, The University of Alabama

AUG 2003 - MAY 2005

Economics, William Paterson University of New Jersey

SKILLS

Generative AI

Python

Neural Networks

Model Development

Natural Language Processing

Machine Learning

Deep Learning

Data Science

ARTICLES

Building a LLM: Leveraging PyTorch to Construct a Large Language Model

LLM Foundations: Constructing and Training Decoder-Only Transformers

The Algorithmic Core of Positional Encoding in Transformers

Unlock the Power of Llama3 8B Model with Apple MLX Server and Chainlit

Advanced RAG w/ Re-Ranking | Groq + Ollama + LangChain + Cohere + PineCone + Llama3-70B

Architecting Intelligent IR with Neural Networks in Python

LINKS

<u>LinkedIN</u> <u>Github</u>

45+ Licenses & Certifications

REFERENCES

References can be found in the $\it Recommendation$ section on LinkedIN:

www.linkedin.com/in/williamzebrowski