Eduardo Pourshahidi

Orlando, FL, United State | +1 407 783 9876 | emp.sh.0703@gmail.com

Professional Summary

AI Developer with 5 years of experience delivering scalable NLP, LLM, and machine learning solutions. Proficient in designing logging architectures, enhancing data quality, and developing AI tooling to detect anomalies and improve performance. Demonstrated success optimizing inference pipelines for sub-200ms latency while integrating Python-based solutions with cloud and embedded platforms. Adept at diagnosing and resolving complex issues, ensuring production-grade reliability and efficiency.

Core Competencies

- **AI/ML Engineering**: OpenAI API (GPT-4), LangChain, TensorFlow, Hugging Face, PyTorch, Scikit-learn
- **Web Development**: Django, FastAPI, Next.js, React.js, Vue.js, Node.js
- **Cloud & DevOps**: AWS (SageMaker, Lambda, EC2), GCP (Google Cloud Platform), Docker, Kubernetes
- **Data**: NumPy, Redis, MySQL, Pandas, PostgreSQL, MongoDB
- **Testing & Tools**: Pytest, Postman, REST/GraphQL APIs, WebSockets, Figma
- **Communication**: Agile Team Communication, Cross-Functional Collaboration, Mentorship, Technical Documentation, Product-Driven AI Design

Professional Experience

Cortex Intelligence   Jan 2021 - May 2025
AI Software Engineer   *Remote*

- Developed AI tooling by deploying GPT-4-based virtual assistants using OpenAI and LangChain, reducing support ticket volume by 35%.
- Engineered a production ML API with FastAPI and Docker on AWS Lambda, achieving a response time of less than 200ms and enhancing platform performance.
- Optimized data analysis processes with an NLP parser using spaCy and Transformers, improving data extraction efficiency by 80% and supporting anomaly detection efforts.
- Directed a team of 6 in iterative model deployment cycles, focusing on issue diagnosis and logging architecture enhancements through A/B testing and monitoring via Prometheus/Grafana.
- Enabled product managers to leverage LLM-analyzed user reviews for iterative UX improvements, reducing reported usability issues by 40% across all platforms.

Neuronetics AI Lab   Jul 2018 - Dec 2020
AI Developer   *San Francisco, CA*

- Designed and maintained ETL pipelines for ML training data using Airflow, Pandas, and AWS S3, ensuring robust data validation and logging improvements.
- Converted TensorFlow models to ONNX for performance inference, reducing latency by 47% while supporting data analysis accuracy.
- Collaborated with researchers to deploy vision and NLP models with minimal downtime, emphasizing precise issue diagnosis and tool development.
- Automated deployment processes for containerized models using Docker and Terraform, streamlining integration with CI pipelines.

SmartOps Cloud   Sep 2017 - Jun 2018
Full Stack Engineer (AI Integration)   *Austin, TX*

- Spearheaded a project that integrated Google Vision and Dialogflow APIs into new user onboarding, reducing the new user onboarding time by 2 hours and improving new user satisfaction.
- Automated data capture within the onboarding pipeline by integrating Google Vision and Dialogflow APIs, processing 500+ documents per day with 99% accuracy, which increased efficiency across the team.
- Designed React-based dashboards for real-time rendering of image classification results.
- Reduced onboarding time by 30% by automating voice-to-intent detection flows.
- Authored internal docs and held training sessions for AI service adoption across teams.

Education

University of Washington, Seattle   2017
M.S., Artificial Intelligence

University of Florida   2014
B.S., Computer Science

AI Projects

Resume Intelligence Tool

- Used GPT-4 to analyze and rewrite resumes. Reduced candidate feedback loop by 50%.
- Engineered a LangChain-based Q&A bot, empowering 500+ employees with semantic search capabilities, resulting in a 30% reduction in time spent searching for internal documentation.

MLOps Retraining Pipeline
  • Automated retraining with S3 triggers, GitHub Actions, and SageMaker. Reduced manual ops by 70%.

OpenAI Integration Layer
  • Created reusable GPT interface consumed across 6 internal services, reducing dev time by 40%.

Certifications

  • AWS Certified Machine Learning - Specialty (2022)
  • TensorFlow Developer Certificate - DeepLearning.Al
  • Microsoft Azure Al Engineer Associate
  • Certified Kubernetes Application Developer (CKAD)

Publications & Contributions

  • Fine-Tuning Transformer Models for Long-Form Legal Documents. arXiv, 2023
  • Contributor to LangChain Open Source Toolkit.