Efthimios Vlahos

Portfolio: Evlahos.com
Github: github.com/EfthimiosVlahos

Email: vlahos89@gmail.com
Mobile: +1 516-606-0942

Education

- **Stony Brook University** — Stony Brook, NY
  Master of Science - Applied Mathematics & Statistics; GPA: 3.71/4.0 — *Aug 2021 - May 2023*
  Courses: *Machine Learning, Data Structures and Algorithms, Linear Programming, Probability Theory, Numerical Analysis*

- **CUNY Hunter College** — Manhattan, NY
  Bachelor of Science - Mathematics & Physics; GPA: 3.92/4.0 — *Aug 2016 - Jun 2020*
  Courses: *Intermediate Mechanics, Quantum Mechanics, Linear Algebra, Abstract Algebra, Vector Calculus*

Skills

- **Languages**: Python, SQL, C++, Java, R, MATLAB, Rust, Bash
- **Frameworks**: Scikit-Learn, TensorFlow, PyTorch, Keras, JAX, PySpark, LangChain, Hugging Face, FastAPI, Flask
- **Tools**: Docker, Kubernetes, MLflow, Airflow, TFX, Kubeflow, Git, Prometheus, Jenkins, Argo CD, Terraform
- **Platforms**: AWS, GCP, Azure, Apache Spark, Databricks, Snowflake, PostgreSQL, MySQL, Hadoop
- **Soft Skills**: Leadership, Project Management, Communication, Problem Solving, Team Collaboration

Experience

- **Cornerstone Building Brands** — New York, NY
  Machine Learning Engineer (Contract) — *June 2024 - Present*
  - **Inventory Forecasting Models**: Improved **inventory forecasting accuracy by 25%** by developing and refining predictive models using **Python, SQL, Azure ML, Databricks, and Scikit-learn**, optimizing production schedules to meet business demands.
  - **Data Pipeline Optimization**: Enhanced **data processing speed by 50%** by architecting and implementing data pipelines and infrastructure with **PySpark, Azure Data Factory, and Apache Spark** for a large-scale ERP system.
  - **Predictive Maintenance Workflows**: Reduced **downtime by 15%** in manufacturing processes by designing and deploying automated machine learning workflows using **TensorFlow, Keras, and MLflow** for predictive maintenance and real-time anomaly detection.
  - **Production Integration**: Ensured seamless integration of machine learning models into production by collaborating with cross-functional teams and implementing **CI/CD pipelines**.

- **Microsoft** — Remote
  Data Analyst (Full-time) — *Nov 2022 – June 2024*
  - **AI Plugin Development**: Achieved a **70% reduction in HR support queries** by developing an **AI Plugin for M365 Chat** using **Python, Azure ML, and RAG architecture**, automating responses and improving query accuracy.
  - **Hallucination Detection**: Increased system reliability by **35%** by optimizing **hallucination detection frameworks** through **RAG-based retrieval techniques** and real-time monitoring across Microsoft AI projects.
  - **Data Pipeline Design**: Improved **batch processing efficiency by 50%** by designing and optimizing data pipelines with **Azure Data Factory, Databricks, and Apache Spark**, scaling RAG model training and deployment.

- **KPMG** — Remote
  Data Analyst (Intern) — *Jun 2022 – Oct 2022*
  - **Financial Market Analysis**: Enhanced financial market analysis models using **Python and SQL** on **GCP**, achieving a **35% improvement in client portfolio performance** through strategic data-driven optimizations.
  - **NLP for Economic Insights**: Extracted key insights from economic reports using **NLP techniques** with Python libraries like **NLTK and SpaCy**, deploying models on **Vertex AI** to increase client engagement by **20%** through refined investment strategies.
  - **Hyperparameter Tuning Automation**: Streamlined the machine learning model training process by implementing automated hyperparameter tuning on **Vertex AI**, reducing deployment times by **50%**.

Projects

- **Retail Sales Forecasting with XGBoost (Machine Learning, Forecasting, AWS) GitHub**: Developed and deployed an XGBoost model on AWS SageMaker, improving weekly retail sales forecasting accuracy by 25%. Tech: Python, AWS SageMaker, XGBoost, S3, Lambda, CloudWatch (Jan '24).
- **Coupon Optimization Strategy GitHub (Machine Learning)**: Built a machine learning model to optimize coupon distribution, doubling conversion rates from 4.92% to 9.97%. Tech: Python, SQL, Scikit-learn, XGBoost (Nov '23).
- **NewsletterGen Crew with GUI GitHub (Natural Language Processing, Automation)**: Created a multi-agent AI system to automate weekly newsletter generation using web scraping and content analysis. Tech: Python, Streamlit, crewAI, EXA API, LangChain (Oct '23).
- **Chest Cancer Detection GitHub (Deep Learning, Medical Imaging)**: Implemented a CNN-based platform for chest cancer detection with 95% recall, deployed using AWS for scalable, real-time analysis. Tech: Python, FastAPI, MLflow, Docker, AWS (Sep '23).
- **E-commerce Data Analytics Platform on AWS GitHub (Data Engineering, Big Data)**: Built and deployed a scalable data analytics platform for e-commerce using Docker, Hive, and Spark on AWS, improving data ingestion and transformation. Tech: AWS, Hive, Spark, Sqoop, SQL, Docker (Aug '23).