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** — Remote
  Machine Learning Engineer (Contract) — *May 2024 - Present*
  - **Demand Forecasting Optimization**: Increased demand forecasting accuracy by **30%** by developing and refining predictive models using **Python, Databricks, and Azure Machine Learning**, enabling more efficient inventory management and reducing stockouts.
  - **Data Pipeline Enhancement**: Reduced data processing time by **50%** by designing and optimizing data pipelines with **Databricks** and **Azure Data Factory**, enabling real-time analytics for supply chain operations.
  - **Distributed Data Processing**: Leveraged **Apache Spark** on **Databricks** to process large-scale manufacturing data, improving model training speed and supporting scalable machine learning operations across multiple plants.
  - **ERP System Integration**: Collaborated with cross-functional teams to integrate ML-driven insights into **ERP systems**, improving decision-making and operational efficiency across several manufacturing facilities.
- **Microsoft** — Remote
  Data Analyst (Full-Time) — *Nov 2020 – April 2024*
  - **AI Plugin Development for Knowledge Management**: Reduced HR support queries by **70%** through the development of an **AI Plugin for M365 Chat**, leveraging **Python, Azure ML, and LLMs**, automating responses for knowledge-based management, and improving query accuracy.
  - **LLM Hallucination Detection in Business Applications**: Increased system reliability by **35%** by refining **LLM-based hallucination detection frameworks**, improving content quality in Microsoft's AI-driven business tools through real-time monitoring.
  - **Data Pipeline Optimization for LLM Training**: Boosted **batch processing efficiency by 50%** by optimizing **data pipelines** for large language model training using **Azure Data Factory, Databricks, and Apache Spark**, supporting the development of scalable AI solutions for enterprise clients.
- **KPMG** — Manhattan, NY
  Data Analyst (Intern) — *Jun 2020 – Oct 2020*
  - **Financial Market Analysis for Clients**: Improved client portfolio performance by **35%** through data-driven optimizations using **Python and SQL** on **GCP**, providing valuable insights into market trends and predictive analysis for investment strategies.
  - **NLP for Generating Economic Insights**: Increased client engagement by **20%** by deploying **NLP models** using **SpaCy and NLTK**, extracting key insights from economic reports to guide investment decisions, aligned with clients' objectives.
  - **Automated Hyperparameter Tuning for Financial Models**: Decreased model training time by **50%** by implementing automated hyperparameter tuning on **AWS SageMaker**, optimizing financial models for accuracy and speed to meet client needs.

Projects

- **2048 Game Deployment on AWS EKS-Fargate GitHub (AWS, CloudFormation, Terraform)**: Designed a scalable and reusable **model execution component** using **AWS Services**, reducing deployment time by **40%**. Streamlined production deployment with **CloudFormation** and **Terraform**, integrating into existing **CI/CD pipelines** (Aug '23).
- **Alzheimer's Generative AI Project GitHub (Observability, AWS, Monitoring)**: Designed and integrated **observability components** for a cloud-based **Generative AI application**, enhancing real-time monitoring and issue resolution, leading to a **35% improvement in application uptime** (Jul '23).
- **Retail Sales Forecasting with XGBoost GitHub (Machine Learning, Forecasting, AWS)**: Developed and deployed an XGBoost model on **AWS SageMaker**, improving weekly retail sales forecasting accuracy by **25%**, leveraging **AWS S3** for data storage and **Lambda** for data preprocessing (Jan '24).
- **Coupon Optimization Strategy GitHub (Machine Learning)**: Doubled conversion rates from **4.92% to 9.97%** by building a machine learning model for coupon optimization, leveraging **Python, SQL, Scikit-learn, and XGBoost**. Deployed the model on **AWS SageMaker** for production use (Nov '23).