Efthimios Vlahos

 LinkedIn | 📱 516-606-0942 | 🌐 Portfolio | ✉ vlahos89@gmail.com | ○ GitHub

Skills

- Python | SQL | R | Java | C++ | Bash | TensorFlow | Keras | PyTorch | PySpark | JAX | NLTK | Scikit-Learn | Tableau | PowerBI | Git
- AWS | GCP | Azure | Databricks | Apache Spark | Vertex AI | SageMaker | Cloud Computing | CI/CD | ETL | Snowflake | PostgreSQL | OOP
- Kubernetes | Docker | Kubeflow | TFX | MLflow | Airflow | Terraform | Distributed Systems | DevOps | MLOps | Full-Stack | Microservices

Experience

| **Data Analyst** | **Microsoft** | *Remote, NY, USA* | **11/2023 - Present** |

- **Engineered data processing pipelines using PostgreSQL and Python,** optimizing data flow and **improving chatbot response accuracy by over 40%,** contributing significantly to Microsoft's AI-driven solutions.
- **Deployed NLP models and managed lifecycle within Kubernetes environments, utilizing Docker containers** to ensure scalable and reliable model performance, **reducing operational issues by 25%**.
- **Implemented hallucination detection algorithms within GPT models**, achieving an **99.5% accuracy** in content delivery, and ensuring reliability in AI-powered chatbot applications.
- **Designed and executed SQL and Python scripts for NLP data preprocessing**, turning unstructured data into valuable business intelligence, and significantly improving data system functionalities for ongoing project needs.
- **Streamlined CI/CD pipelines for continuous integration and deployment of AI models using Docker and Kubernetes,** which improved deployment efficiency and **shortened release cycles by 30%**.

| **Data Analyst Intern** | **KPMG** | *Remote, NY, USA* | **06/2023 - 10/2023** |

- **Fine-tuned predictive models using Python and SQL for financial market analysis**, enhancing client portfolio performance by 35% through strategic data-driven decisions.
- **Utilized NLP techniques with Python libraries such as NLTK and spaCy to extract and analyze key insights from economic reports** and news, leading to a **20% increase in client engagement** by refining investment strategies.
- Streamlined machine learning model training processes with **automated hyperparameter tuning using scikit-learn and MLflow**, achieving a **50% reduction in deployment times** and boosting client service efficiency.
- **Supported data collection and management using cloud-based SQL databases on Azure**, ensuring robust data quality for model accuracy and reliability in financial forecasting.
- **Led data cleaning and preparation efforts, employing Python's Pandas and NumPy** libraries to enhance data integrity and structure for analysis and actionable financial insights.

Projects

- **Retail Sales Forecasting with XGBoost | Github** : Deployed an **XGBoost model on AWS SageMaker** for retail sales forecasting, achieving an R-squared of 0.965 and RMSE of 4202.395, **through effective hyperparameter optimization using boto3 SDK**.
- **Alzheimer's Generative AI Project | Github: Leveraged transfer learning and NLP methods**, incorporating n-gram feature representation, to enhance Alzheimer's research, **achieving an accuracy of 88.1% using LSTM, GRU, and hybrid architectures**.
- **LLM-Based Financial ChatBot Tool | Github: Built a financial analysis ChatBot using LangChain, OpenAI GPT, and RAG**, optimized with FAISS for high-accuracy, real-time insights and decision-making support.
- **Flight Delay Prediction | Github:** Developed and deployed a **machine learning solution on GCP with Vertex AI**, achieving 80% accuracy and **75% recall** for flight delay prediction, **utilizing BigQuery for data preprocessing and Docker for deployment**.

Education

| **Master of Science** | **Stony Brook University** | *Stonybrook, NY, USA* | **08/2021 - 05/2023** |

- Major in Applied Mathematics & Statistics, Concentration in Data Science

| **Bachelor of Science** | **CUNY Hunter College** | *Manhattan, NY, USA* | **08/2016 - 05/2020** |

- Major in Mathematics & Physics, Minor in Computer Science

Certificates

- **AWS Certified Cloud Practitioner |** Certificate
- **Advanced ML on Google Cloud |** Certificate
- **Machine Learning Engineering for Production |** Certificate
- **Natural Language Processing |** Certificate
- **Deep Learning |** Certificate