






# Efthimios Vlahos

Machine Learning Engineer

 Website  Github  LinkedIn  vlahos89@gmail.com

 516-606-0942

## EXPERIENCE

### MICROSOFT | DATA ANALYST

Nov 2023 - Present | Remote

- Led the development and implementation of NLP techniques to detect hallucinations in AI-generated text, enhancing the accuracy and reliability of machine-generated information.
- Guided a team of 5 data analysts to collaboratively improve the Hallucination Detection Project from 30% to 87% accuracy in three iterations using Excel and updated project guidelines with input from stakeholders.

### KPMG | DATA ANALYST INTERN

Jun 2023 - Oct 2023 | Manhattan, NY

- Spearheaded data-driven customer targeting strategies by assessing data quality, conducting in-depth demographic analysis, and leveraging machine learning techniques, resulting in highly effective targeted outreach campaigns.
- Utilized advanced data analytics to identify and prioritize high-value customers through a comprehensive analysis of demographics and key attributes, contributing to increased ROI for clients.

## PROJECTS

### ALZHEIMERS GENERATIVE AI PROJECT | PYTHON, TENSORFLOW, KERAS, LSTM, GRU

Nov 2023 - Dec 2023

- Led a project using LSTM and GRU neural networks to accelerate Alzheimer's disease inhibitor discovery, achieving an 88.1% accuracy with the advanced model, contributing to the fight against Alzheimer's.
- Generated 18 new molecules from the Alzheimer's dataset, showcasing their potential to inhibit BACE-1 and potentially slow down Alzheimer's disease, while implementing a coded stop character for valid chemical structures, paving the way for further research and development.

### CHEST CANCER DETECTION | PYTHON, TENSORFLOW, MLFLOW, DAGSHUB, DOCKER, AWS

Oct 2023 - Nov 2023

- Integrated MLflow and DagsHub to manage the machine learning lifecycle, enhancing experiment tracking and model performance evaluation, resulting in a 20% boost in model accuracy and a 30% reduction in development cycle time.
- Spearheaded a deployment strategy, utilizing Docker for containerization and AWS for cloud hosting, resulting in a scalable web service.
- Leveraged the VGG16 CNN model for the classification of chest cancer from CT scan images, achieving an 85% accuracy rate.

### E-COMMERCE CUSTOMER SATISFACTION | PYTHON, ZENML, OPTUNA

Aug 2023 - Dec 2023

- Successfully integrated MLflow with ZenML, enhancing the machine learning lifecycle from data ingestion to model deployment, resulting in increased robustness and scalability of models in real-world applications.
- Enhanced RandomForest model performance by 15% in predictive accuracy, as measured by comprehensive testing, by implementing Optuna-based hyperparameter tuning and refining model parameters.

## SKILLS

### PROGRAMMING

Python • SQL • R  
TensorFlow • Keras  
PyTorch • Scikit-Learn  
NLTK • JAX • Trax  
AWS • GCP • PostgreSQL  
Vertex AI • Apache Hadoop

### MACHINE LEARNING

Deep Learning(CNNs, RNNs, GRUs, LSTM, Transformers)  
NLP • Predictive Modeling  
Time Series Analysis • Feature Engineering • Model Evaluation  
Model Tuning

### DEVOPS, MLOPS, & VIZ

Docker • Kubernetes • MLflow  
Kubeflow • Git • GitHub  
Tableau • Power BI • Looker

## EDUCATION

### STONY BROOK UNIVERSITY

MASTER'S IN APPLIED MATHEMATICS  
& STATISTICS, DATA SCIENCE  
CONCENTRATION

Aug 2021 - May 2023 | Stony Brook, NY

Cum. GPA: 3.45 / 4.0

Relevant Coursework:

Machine Learning • Big Data  
Analysis • Data Structures and  
Algorithms • Linear  
Programming • Probability  
Theory • Numerical Analysis

### CUNY HUNTER COLLEGE

BACHELOR OF SCIENCE IN  
MATHEMATICS & PHYSICS, MINOR IN  
COMPUTER SCIENCE

Aug 2016-Jun 2020 | Manhattan, NY

Cum. GPA: 3.7 / 4.0

Relevant Coursework:

Intermediate Mechanics  
Quantum Mechanics • Linear  
Algebra • Abstract Algebra  
Vector Calculus • Mathematical  
Analysis I & II