

education

UNIVERSITY OF MASSACHUSETTS AMHERST

Sept 2022 – Sept 2027

Doctor of Philosophy (PhD) – Manning College of Information and Computer Sciences

Working with Prof Peter J.Haas, Prof Alexandra Meliou, and Prof Azza Abouzied

Relevant Coursework: *Advanced Algorithms, Advanced S/W Engineering: Analysis and Evaluation, Neural Networks (fall semester)*

TECHNICAL UNIVERSITY OF CRETE

Sept 2015 – Oct 2021

Undergrad/ Integrated Master's Degree – School of Electrical and Computer Engineering

"Online Machine Learning in Distributed Environments for Big Data"

Class Rank: (Top 4%/206 students), GPA (Computer Science): 3.6 / 4

Random Forest Optimizations under Data Drifts (Thesis)

- Created a distributed ensemble learning system for binary classification, using multiple decision trees (Random Forest) with datasets over 25M+ evolving streaming data.
- Researched and proposed three scaling optimizations for:
 1. Gaussian Approximation for streamlined handling of numerical attributes in unbounded data streams, yielding to improvement on resource allocation. (95% decrease on memory consumption)
 2. Resampling Enhancement: Refined the Online Bagging approach by centralizing its function reducing data transactions (80% decrease on #packets of data)
 3. Base Learner Adaptation: Designed and implemented a dynamic accuracy monitoring method to halt/resume learner adjustments accounting for "static data" periods, optimizing performance without unnecessary growth (70% memory improvement, 90% accuracy under data drifts)
- Implemented on Apache Flink and Apache Kafka, developed with Java and Scala.

work experience

Graduate Research Assistant

Sept 2022 – Now

University of Massachusetts Amherst – DREAM Lab

"Pushing analytics close to database systems"

Dynamic Query Package Maintenance

- Researching algorithms on dynamically maintaining [PackageBuilder](#), a system that extends query engines for package generation i.e., collections of tuples with global properties, during streaming data drifts.
- Ensuring improvements on quality and runtime performance, real-time adaptability and mathematical guarantees.

projects and involvements

Online Credit Card Fraud Detection

- Built an end-to-end real-time fraud detection system for credit card transactions using adaptive Random Forest of Hoeffding Trees for datasets 10M+ tuples, addressing issues such as imbalance classes, online bagging, voting boosting and avoiding overfitting.
- Developed with Scala and Java. Deployed on Apache Spark using both HDFS and Apache Kafka source/sink. Results: 92% accuracy and 95% F1-score.

Lupus (NPSLE) Classification using ML

- Implemented and designed ML pipelines including Support Vector Machines (SVM), K-Nearest Neighbors, and Random Forest (RF) on clinical data for the diagnosis of Lupus (NPSLE) disorder, researching the impact of integrating Machine Learning techniques on such sensitive high-dimensional data.
- Analyzed resting-state connectivity fMRI data by performing feature selection, classification, and cross validation techniques. Results: RF (Acc: 77% F1: 79%), SVM (Acc: 74% F1: 73%).

ToDo AI

- Designed a PyCharm extension integrating AI for code generation in extensive software development. Released a Beta version offering AI-generated code suggestions with descriptive change descriptions.
- Developed with Python, using OpenAI's GPT-3.5-Turbo Model

Ask Question App

- Built front and back-end modules, secure blockchain authentication using Metamask wallet integration, smart contracts development to support Q&A interactions and tipping (upvoting) posts through [Ethereum Attestation Service](#) using custom [schema](#).
- Implemented on Rinkeby (Ethereum TestNet) decentralized internet, developed using Solidity, Truffle Suite, Node.js and React.js

School Dashboard App

- Full stack website development for managing classes, students and grades by authorized teachers, personnel registration and authentication.
- Developed with Javascript and php on PHP-Apache server using Docker containers and MySQL

technical skills

Programming

Proficient in Python, Java, C+, PostgreSQL; Familiar with php, Javascript, Scala, Solidity, Matlab, R

Technologies

Apache Flink, Apache Spark, Hadoop, Kafka, Tensorflow, Pandas, scikit-learn, Django, phpmyadmin

honors and awards

UMass Amherst

- UMass CICS Scholarship for Best PhD Applicant 2022 (4,000\$)

IEEE - ReScan

- 1st Place Entrepreneurship Initiative Start-up Pitch Competition 2020 Greek Section