VASILEIOS VITTIS

vvittis@umass.edu
vvittis.github.io
+1 (857) 706 9391

education

UNIVERSITY OF MASSACHUSETTS AMHERST Sept 2022 - Sept 2027

PhD - Manning College of Information and Computer Sciences

"Pushing analytics close to database systems"

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

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.

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

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"

Working with Prof Antonios Deligiannakis

Class Rank: Top 5%

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:

- Data Modeling: Gaussian Approximation for streamlined handling of numerical attributes in unbounded data streams, yielding to improvement on resource allocation. (95% improvement)
- Resampling Enhancement: Refined the Online Bagging approach by centralizing its function, resulting in reduced data transactions (80% improvement)
- 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.

Relevant Coursework: *Advanced Topics in Database Systems, Approximation Techniques for Massive Datasets, Artificial Intelligence, Multi-Agent Systems, Services in Computational Cloud and Fog*

projects and involvements

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

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.

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

Lupus (NPSLE) Classification using ML

– Implemented and designed ML pipelines including Support Vector Machines (SVM), K-Nearest Neighbors, and Random Forest 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: Random Forest (Acc: 77% F1-score: 79%), SVM (Acc: 74% F1-score: 73%).

RεScan

– Founded a start-up initiative, RεScan, a user-friendly app promoting correct recycling habits. By scanning the barcode of items, the integrated machine learning model determines and suggests the recyclability category (if any), guiding users to dispose of their waste properly.

– Implemented Convolutional Neural Networks to distinguish the recycling categories based on barcode images, used Firebase to maintain a dynamic repo of barcode-to-product feature mappings.

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.

IEEE TUC website

– Built a full stack website for the local IEEE student branch to support user registration and authentication, team creation for contests, workshop management, posts and news sections.

– Developed using DJango framework, Python, PostgreSQL, html, css.

volunteer activities

Chairman

– Served as chairman of the IEEE TUC student branch, organized 20+ events (hackathons, workshops and lectures), initiated specialized mentorship sub-teams; personally guided students in web development and data science.

Lecturer

– Led a 'Mathematical Shapes with Python' workshop for +40 freshmen students during TUC WelcomeDay.

Ambassador

– Oversaw IEEEXtreme competitive programming contest, managing logistics,and sponsorships.

honors and awards

UMass

– UMass CICS Scholarship for Best PhD Applicant (4,000$)

IEEE - RεScan

– 1st Place Entrepreneurship Initiative Start-up Pitch Greek Section

skills

Programming

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

Technologies

Apache Flink, Apache Spark, Hadoop, Tensorflow, Pandas, scikit-learn, DJango, phpmyadmin