Marinos Poiitis

DATA SCIENTIST

+306942713478
marinospoet@gmail.com
linkedin.com/in/marinos-poiitis/

B.Sc. in Computer Science and M.Sc. in Web and Data Science. Data Scientist, ML/AI Researcher, Web developer and project manager. Practical experience in various Horizon 2020 and other European projects. Theoretical experience in data and AI-related research. Seeking to contribute to large projects of social impact while advancing my technological, organizational and managerial skills.

SKILLS

ML FRAMEWORKS (PYTORCH/TENSORFLOW) | PYTHON | NO SQL | SPARK | SQL | HADOOP | KAFKA | MQTT | GIT | DOCKER | FRONT-END (HTML/CSS/JS/REACT) | BACK-END (FLASK, PHP, DJANGO) | ANDROID | JAVA

WORK EXPERIENCE

Data and Web Science Lab (Datalab)

Thessaloniki, Greece

DATA SCIENTIST – WEB DEVELOPER – RESEARCH ASSISTANT

December 2018 - November 2021

- Project manager for a small data engineering team using agile methodologies and tools such as Azure DevOps, Git and Slack
- Data scientist/data engineer in various projects under the "H2020 Research and Innovation Programme" and the "Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE"
- Experience on data engineering pipelines using Docker, Kubernetes, Kafka queues and the MQTT protocol, Edge analysis for real-time patient emergency detection

- Social network analysis, data collection, preprocessing and analysis using Python, Twitter API and MongoDB
- Natural Language Processing on Twitter and Instagram and integration with web applications
- Energy disaggregation and Non-Intrusive Load Monitoring using Graph Neural Networks
- Research on Graph Representation Learning, Graph Neural Networks and theoretical AI (publications in NeurIPS conference, TWEB journal etc.)
- Research on Graph theory and information diffusion

- Web app development (HTML5/CSS3/JS, Flask, Python) and data visualization (D3.js, Chart.js)
- Lectures regarding Web Technologies: HTML, CSS3, JavaScript, PHP, MongoDB, SQL, SPA frameworks (mostly React.js), Flask. Supervising projects on Web app and Website development using the taught technologies and stacks

TELEFONICA

Barcelona, Spain

DATA SCIENTIST – MARIE CURIE RISE SECONDEE

September 2019 - December 2019

- Social network data mining and analysis
- Development of influence maximization algorithms and cascading models. Application on online aggression diffusion modeling and minimization
- Research experimentation and paper writing

Data and Web Science Lab (Datalab)

Thessaloniki, Greece

WEB DEVELOPER– DATA ENGINEER

March 2017 – August 2017

- Front-end developer using HTML5/CSS3/JS and React.js
- Back-end developer using Flask, Python and MongoDB.
- Data visualization for monitoring European regions, and aid in resource and facility management.

EDUCATION

Aristotle University of Thessaloniki

Master of Science – M.Sc. in Web and Data Science (GPA: 9.67/10.0) Thessaloniki, Greece 2018-2020

Bachelor of Science - B.Sc. in Computer Science (GPA: 9.22/10.0) Thessaloniki, Greece 2013-2017

DATA SCIENCE PROJECTS

Representation Learning on Graphs

Sep 2020 - Nov 2021

- Python, PyTorch, Tensorflow, NetworkX
- GraphM: unification of traditional graph representation learning algorithms and datasets. Specifically:
  - DeepWalk, Node2Vec, LINE, TADW, MNMF
- iSpine: unification of graph neural network-based representation learning algorithms and datasets. Specifically:
  - GCN, GAT, VGAE, DANE, DGI, AGC, AGE, PointSpectrum

Neural Network Generalizability and Training behavior

May 2021 - Oct 2021

- PyTorch, Python, Matplotlib
- Analyze the training behavior of neural networks using Lipschitz continuity
- Investigate the neural network efficiency on seen (training) and unseen data based on the learned function's complexity
- Experiment on MLPs and CNNs using baseline image datasets (CIFAR10) and generated signal data
- Visualizations of the learned function using ReLU's properties and global Lipschitz constant
- In collaboration with EPFL researcher Andreas Loukas and MIT associate professor Stefanie Jegelka.
- Publication in NeurIPS '21

Bot and anomaly detection on Twitter

- Python mostly, minor usage of R and Matlab.
- Data collection using Twitter streaming API.
- Apply models to find whether data follows a specific distribution, e.g. power law
- Fit data using approximation techniques, e.g line fitting
- Unsupervised machine learning for cluster detection.
- In collaboration with Amazon researcher and Carnegie Mellon University professor Christos Faloutsos.
- Best paper award

Hashtag popularity prediction on Twitter

- Python, Tensorflow, some Pyspark
- Data collection using twitter streaming API
- Feature extraction from hashtags, tweets, users
- Dimensionality reduction techniques for reducing complexity
- Machine learning and Neural Networks for predictions
- Insight: Generalizes on ad popularity prediction

Automated Quasi Identifier Prediction on Open Data

- Spark for big data processing, Java
- HDFS and cluster utilization for data and results storage
- Health data anonymization
- Analysis of data features

- Protect information from triangulation attacks
- In collaboration with IBM researcher Aris Gkoulalas-Divanis

HONORS AND AWARDS

Best Paper Award
July 2020

WIMS 20 (The 10th International Conference on Web Intelligence, Mining and Semantics)

Best Master Thesis Award
March 2020

Excellence Scholarship
March 2016

Compensatory excellence scholarship given by Aristotle University of Thessaloniki (AUTh)

LANGUAGES

ENGLISH | GREEK

OTHER

- ACM AUTh Student Chapter
- Linux-centric team lead
- 3D Game development team lead
- Lectures on Network monitoring and Cybersecurity

WEBSITES

- https://github.com/mpoiitis
- https://mpoiitis.github.io/