Poulakis Yannis
Data Scientist

✉ giannispoy@gmail.com

📞 +30 6949212113

@ linkedin.com/in/yannis-poulakis

Education and qualifications

- Bachelor's degree in Statistics and Insurance science, University of Piraeus

- MSc, Big Data & Analytics, University of Piraeus

- PhD Candidate, Automated Machine Learning for Unsupervised Learning, University of Piraeus

- **Languages**: Greek (Native), English (Proficient)

Publications

For the complete list of publications, current h-index and number of citations please visit my google [scholar page](#).

Notable Articles

- AutoClust: A Framework for Automated Clustering based on Cluster Validity Indices, IEEE - ICDM 2020.

- HOCC: An ontology for the holistic description of cluster settings, GECON 2022

- A Survey on AutoML Methods and Systems for Clustering, TKDD 2023

Employment History

- **Research Associate UPRC (GreenDatAI Horizon2020).** (2023-Ongoing) (Part Time)
  - Application of AutoML methods in time series forecasting and optimization methods for the minimum cost flow optimization in vehicle redistribution.

- ***Research Associate BYTE Computer* (PHYSICS Horizon2020).** (2021-2023)
  - Research and development of a semantic schema to capture different functional and non-functional properties of a Kubernetes cluster setting.
  - Development of a REST service that leverages Kubernetes API and Prometheus to extract cluster information and transform it according to the appropriate schema. This service was containerized and deployed for testing on both Azure and AWS testbed clusters.

- **Lab Instructor, University of Piraeus** (2021–Ongoing)

  (Limited Hours in 2nd curriculum)
  - Designed and delivered hands-on instructional material for the Information Theory and Data Analytics courses for undergraduate Students

- **Technical Coordinator UPRC (DIASTEMA ESPA).** (2021-2023)

  - Architecture design and technical coordination of the associated developers/researchers for the realization of the included components.

  - Requirement analysis and coordination of operations to align use case pilots with platform functionalities. Realization of two different scenarios based on healthcare analytics and digitization in the maritime industry.

- **Research Associate UPRC (BigDataStack Horizon2020).** (2019-2020)

  - Research and development of a service that is capable of selecting the best performing algorithm for a dataset at hand in a data driven manner.

- **Data Science Consultant OneDealer.** (2019)

  - Design and development of a machine learning model that predicts how probable it is for a customer to make a purchase according to a sequence of actions.

- **Accountant & Insurance Intermediary, self-employed.** (2018-2020)

Skills

- **Programming Languages:** Python, R, SQL

- **ML & Data Science:** Keras, Tensorflow, Pytorch, Flask, SPSS,

- **Databases & Big Data:** MongoDB, Apache Spark, MySQL, Kafka, ElasticSearch, AllegroGraph

- **Deployment & IT:** Docker, Kubernetes, Jenkins, REST

- **Tools & Management:** GitHub, GitLab, Trello, Scrum

Ongoing Personal Projects

- **Discord Bot (Various Functionalities)**

  Developed a discord bot, using discord API through python, to accommodate functionalities like web scrapping and a novel mini-game for the users.

- **YouTube Frame Object Identification (Google Chrome Addon)**

  Development of a JavaScript based google chrome addon that retrieves the current frame of a YouTube video and calls a pre-trained YoloV5 model (in-browser) to identify existing object in the frame.

- **Twitch Bot (NLP & Chat Interaction)**

  Developed a twitch chat bot that identifies word trends, frequent chatters and chat sentiment over time in live video-streams to provide insights through a custom made html dashboard.