

# IOANNIS PETRIDIS

DATA SCIENTIST - KLDISCOVERY



## PERSONAL DETAILS

**Date of birth:** 16-07-1989

**in** [www.linkedin.com/in/ipetridis](https://www.linkedin.com/in/ipetridis)

## CONTACT

**Kappadokias 97, 54453, Thessaloniki, Greece**

**t:** +30 693 6833 316 | **e:** [ioannis.petridis89@gmail.com](mailto:ioannis.petridis89@gmail.com)

## SUMMARY

*I seek to apply lean Data Science and Machine Learning solutions to solve business problems and deliver impactful and innovative products within an engaging and challenging environment. I am self-directed and curious data scientist dedicated to improve and grow both in a team and individually. I am reliable and hard working with a lifelong learning spirit.*

## EXPERIENCE

### Data Scientist

September 2021 - Current

#### *KLDISCOVERY*

- Engineered NLP processes such Text Cleaning, Clustering, Classification, Summarization and Topic Modelling using Transformer Models, Azure DevOps, Docker, Python, Postman, PostgreSQL and DBeaver.

### Data Scientist | Research Scientist

September 2018 - September 2021

#### *Aristotle University of Thessaloniki - EnvELab*

- Consulted in modelling a neural network to predict the chemical properties of 100 unknown chemicals achieving an R-squared of 0.87 (45% higher than the accepted value for such models).
- Data reporting and visualization using R-Markdown and ggplot, Python and matplotlib, Tableau, Power BI and MS Excel.
- Expedited process lead time by 60% by guiding the development of a time efficient framework in R and Python for preprocessing time series data.
- Facilitated the creation of physiologically based biokinetic models in ACSLX language to simulate time series tissue concentrations of chemicals and mixtures.
- Attained a position in the Greek Government's modelling team by designing a precise dynamic model in ACSLX to predict Covid-19 dynamics (cases, deaths, hospitalization).
- Awarded by a 3-year-long Marie Skłodowska-Curie Fellowship.
- Earned a 1.5-year-long scholarship among 400+ scientific teams.

### Lean Six Sigma Consultant

June 2020 - November 2020

#### *"Thriasio" General Hospital*

- Provided a cost and a time-efficient concept solution (76% decrease in the current preoperative time per patient and 81% decrease in the current financial annual cost) by optimizing the preoperative check-up procedure utilizing Project Charter, Process Mapping, Waste Analysis, Swimlane Diagrams and Financial Analysis.

### Data Science and Machine Learning Engineer

January 2020 - June 2020

#### *Enve-X Solutions*

- Combined, cleaned and visualized multimodal time data series in R.
- Engineered and directed a methodology to impute-forecast missing values of environmental time series data using Regression Analysis and Generalized Additive Models (GAM) in R and S(ARIMA)X modules in Python based on log likelihood ratio and AIC information criteria.

### Computer Engineer

December 2017 - August 2018

#### *CEDEFOP European Centre for the Development of Vocational Training*

- Integrated AppleScript programming language to resolve compatibility issues, perform automatic disk mounting and backup.
- Analysed and assessed log files and security alerts acquired from Splunk in order to detect security bridges.

### Database and Web Engineer

September 2017 - November 2017

#### *"Athena" Research and Innovation Centre*

- Utilized MySQL Workbench, Bash Scripting, PHP, Yii-2 framework and D3.js library to design and develop an interactive web search interface for visualizing RNA structures and their dependencies.

### Biomedical Research Engineer

January 2017 - June 2017

#### *Leiden University Medical Centre (LUMC)*

- Analysed and interpreted nonlinear time and frequency data series utilizing MATLAB and EEGLab in order to quantify the brain dynamics of migraine patients and healthy controls in the presence of sinusoidal light simulation.

---

## PROFESSIONAL SKILLS

- R, Python, SQL, MATLAB
- Statistical Modelling
- Data Analysis and Visualization
- Regression, Clustering, Classification
- Neural Networks
- Machine Learning
- Monte Carlo Analysis
- Natural Language Processing

## SOFT SKILLS

- Teamwork
- Hard working
- Good listener
- Multitasking
- Persistence
- Motivation to learn
- Reliable

## LANGUAGES

- Greek (Native)
  - English (Professional proficiency)
  - German (Elementary knowledge)
- 

## PROFESSIONAL TRAINING

- **Data Science and Machine Learning**

*Key Learnings:*

- Utilizing Data Science to solve Business Problems, Customer Analytics, Time Series Analysis, Credit Risk Modelling, Web Scraping, Financial Analytics
- Clustering, Classification, Regression, Neural Networks
- Tableau, Power BI, SQL, Python, R, Excel Pivot Tables
- Machine Learning, Deep Learning, Convolution and Recurrent Neural Networks, TensorFlow
- Data Cleaning, Exploratory Data Analysis, Statistical Inference, Shiny Applications, Google Cloud Platform

- **Lean Six Sigma Yellow & Green Belt**

*Key Learnings:*

- DMAIC Methodology, SIPOC, Project Charter, Process Map, Value Stream Map, Waste Analysis, Process Cycle Efficiency

- **Negotiation, Mediation and Conflict Resolution**

*Key Learnings:*

- Deeper understanding of cross-cultural behaviour and communication
- 

## EDUCATION

December 2018 – September 2021	●	<b>PhD Candidate in Environmental Health Engineering</b> Aristotle University of Thessaloniki, Greece Department of Chemical Engineering
September 2014 - June 2017	●	<b>MSc in Biomedical Engineering</b> Delft University of Technology, Netherlands Department of Mechanical, Maritime and Material Engineering
September 2014 - July 2019	●	<b>MSc in Information and Communication Systems (remote)</b> Open University of Cyprus, Cyprus Department of Applied Sciences
September 2007 - September 2013	●	<b>BSc/MSc in Computer and Electrical Engineering</b> University of Thessaly, Greece Department of Electrical and Computer Engineering

---

## VOLUNTEERING

May 2019	●	<b>AUTH on Sundays (Education)</b> - Raise public's awareness with respect to the University's research
April 2018	●	<b>TEDxThessaloniki (Networking)</b> - Support the team during the main event
April 2017	●	<b>RoboValley (Science and Technology)</b> - Audience Handling
March 2017	●	<b>IDHEM (Civil Rights and Social Action)</b> - Email Management

## HONORS AND AWARDS

January 2020	●	Operational Program, Human Resources Development, Education and Lifelong Learning, Co-financed by Greece and the European Union
September 2018	●	Marie Skłodowska-Curie Actions ESR Fellow (MSCA-ETN-ITN) funded by the European Commission's Framework Programme Horizon 2020

---

## INTERESTS

