Christos Hadjinikolis

Pod Lead | Lead ML Engineer

contact

+44(0)7580 635 502
christos.
hadjinikolis@gmail.com
Flat 1, 74 North Hill,
N64RL, London, UK
⚑ **My-website**

in LinkedIn
○ **github**

languages
bilingual greek/english

programming
Python	●●●● ○
Java	●●●● ○
Matlab	●●● ○ ○
SQL	●●● ○ ○
Cypher	●● ○ ○ ○

author
dynamicio: python
library

developer
TDD, git + CI/CD
(Circleci)

cloud-platforms
△ *GCP*	●●●● ○
aws *AWS*	●●●● ○

Technologies & Frameworks
pySpark, Spark, Flink,
Kafka, Elasticsearch,
Neo4J, SQL, Docker,
Kubelfow, Airflow

Python Libs
pandas, numpy

visualisations
Gephi, Seaborn, plotly

about me

Lead ML Engineer with excellent presentation and communication skills. A passionate researcher with exposure to deep learning & evolutionary AI. A competent problem-solver with a proven academic record in the field of AI and with practical experience in ML-Ops, ML-engineering & advanced Data Science.

education

2010–2014   **Ph.D.** in Computer Science   King's College London, UK
Thesis: *Persuasion Dialogues & Opponent Modelling.*
Developed a prediction mechanism applied on Knowledge Graphs. Involved the application of graph-analytics on large-scale graphs & statistical inference.

2004–2010   **Diploma (BEng)** in Computer Engineering   University of Thessaly, GR
A five year-long polytechnic degree with a solid mathematical background (10 courses) ranging from pure (theoretical) computer science/engineering aspects to networking and telecommunications. *Majored in Artificial Intelligence*.

2002–2004   **Military Officer: Second Lieutenant**   National Guard, CY
Obligatory army service; position acquired through competitive IQ & athletic exams.

work-experience

since 12/2020   **Vortexa, London, UK**   Pod Lead | Lead Python ML-Engineer
Member of the Data Production Team; responsible for generating high-value forecasts by implementing ETL-/ML- pipelines to predict the operations of thousands of tankers we track. Responsible for designing, implementing & testing code, allowing historically niche research to be elevated to the real-world scale; working with data-scientists, analysts, engineers & domain experts to bridge the gap between scientific experiments & commercial products; increasing the fault-tolerance of our processing pipelines; working closely with team members on best practices and key technologies. I am also a **member of the Architecture Guild**, responsible for running a decentralised conversational architecture practice within the R&D department.

since 01/2021   **ISO - International Organization for Standardization**   Committee Expert Member for JTC21
JTC21 produces standardisation deliverables in the field of AI, related use of data to address European market and societal needs. This work aims to advance EU legislation, policies, principles and values through the adoption of international standards.

since 01/2015   **Kings College London, Dept. of Informatics, London, UK**   Visiting Researcher
Interacting with members of the Agents & Intelligent Systems group on AI related topics (my current interests orient around AI standardisation and AI ethics). Occasionally working on the development of next generation Data Mining techniques.

04/16-11/20   **DataReply, London, UK**   Lead ML Engineer
Started as a Data Scientist and gradually evolved into a ML-engineer. Led a practice of Data-Scientists/Engineers on relevant projects, with main responsibilities around translating business problems into analytic frameworks; communicating & controlling the overall plan and timeline; mentoring less experienced members, and; driving innovation.

02–03/2016   **Greenwich School of Management**   Associate Lecturer
Taught an introductory course on Computing.

2012–2016   **David Game College, London, UK**   Programming Teacher
Taught MATLAB & delivered the AS Computing Course (Java, Python, html, CSS).

2010–2014   **King's College London & University College London, UK**   Teaching Assistant
Programming (Java), Data Structures, Computer Systems, Semantic Web, DBMS-SQL & AI.

Consulting Roles with Data Reply

Vodafone Group London [07/2020 - 11/2020]   Senior ML Engineer Role
Infinity Platform
Description: Part of the Advanced Analytics Team, responsible for driving the implementation of CI/CD pipelines to support best practices around the development of ML pipelines built within an internally developed platform based on Apache Kubeflow.
Technologies: *Python, Jenkins, KubeFlow, Docker, Kubernetes, GCP*

Vodafone Group London [02/2020 - 07/2020]   Lead ML Engineer Role
Developing Cloud-native Data Analytics Competencies on GCP
Description: Operating as a technical lead as part of a new team, responsible for delivering cloud native data analytics projects in the post-migration era for Vodafone. Involved in the requirements gathering process, gap analysis and impact assessment estimates. Responsible for the delivery of Big Data related projects guided by business drives related to marketing, recommendations, automation and resource optimisation.
Technologies: *Python, pySpark, Jenkins, AirFlow, Kubernetes, GCP*

Vodafone Group London [09/2019 - 02/2020]   Senior ML Engineer Role
ML Framework Development & Migration Supervision
Description: Working with a team of python developers, focusing on the migration of ML models, developed by an internal framework, to GCP. We are also managing the migration of all *CI/CD* pipelines, scheduling and orchestration jobs that employ this framework, from on-premise to cloud. The framework is based on pySpark and is mainly used for feature engineering and ML industrialisation. Additional duties concern the development of a set of pipelines to build approved base-images (for DataProc and Deep Learning with TF) to be used by the data scientist.
Technologies: *Python, pySpark, Jenkins, AirFlow, GCP*

CNHi, Chicago [05-09/2019]   Head of Data Science Role
Analysis of Agriculture Vehicle Data
Description: Led an agile team of data scientists and data engineers. Interacted directly with stakeholders to collect business requirements and translate them into data science/data engineering tasks. Guided and participated in defining and completing relevant tasks. Applied basic statistic analytics and developed time series prediction models that operate as alert triggers when relevant thresholds are excited.
Technologies: *Azure*, *Databricks*, *PowerBi*, *PySpark*, *Python (sklearn, pandas)*

UBS, London [08/2017-04/2019]   ML Engineer Role
Analysis of Unstructured and Structured Data
Description: Developed end-to-end processing pipelines to ingest and analyse data in real-time. Applied Graph Analytics for the extraction of actionable insights, challenging organisation structures and aiding process optimisation (batch analysis). Applied Process Mining techniques to aid understanding of various processes and behaviours of entities of interest in event logs (stream processing). Generated useful visuals through dashboards and python notebooks.
Technologies: *Elasticsearch*, *Kafka*, *Apache Flink*, *Python (sklearn)*, *Kibana*

Irish Road Safety Authority, London [05-07/2017]   Data Science Role
Analysis of Road Safety Data
Description: Involved analysing and extracting insights from MySQL data for predicting high-way crashes, identifying high risk nodes in the traffic network as well as categorising/profiling drivers/operators based on compliance rates, inspections and defects.
Technologies: *MySQL*, *Neo4J*, *Google Cloud*, *R*, *Python (sklearn)*

GameSys, London [03-04/2017]   Data Engineering Role
Architecture Design and Support for new Digital Platform
Description: POC work concerned with showcasing how a Kafka Schema Registry can be used for data validation purposes. Trained a team of Engineers on how to deploy it and integrate it into their existing solution.
Technologies: *Apache Kafka*, *Google Cloud*

The Mondadori Group, Milan [11–12/2016]   Data Science/Engineering Role
Development of a Recommender Engine
Description: Analysed user preferences harvested from this publisher's social commerce platform where members can interact by sharing/buying/rating books maintained in their digital "library". Applied community detection & market basket analysis techniques (ALS method) and the extraction of book association rules, all of which could then used for both user-to-user and user-to-book recommendations.
Technologies: *Scala, SQL, Apache Spark, MongoDB*

Unicredit, Milan [10–11/2016]   Data Engineering Role
Development of a Fuzzy Matching Mapping Engine
Description: Developed a batch process solution for entity identification and matching between data retrieved from distinct databases, based on a variant of Levenshtein's distance algorithm.
Technologies: *Scala, SQL, Apache Spark*

Data-Science/Engineering Projects

Python Library Author [01-04/2022]   Engineering
Main Description: Dynamicio: A pandas wrapper for abstracting I/O code
Dynamicio (or dynamic(i/o)) is a convenient wrapper around pandas I/O operations. It's a manifestation of the dependency inversion principle–a layer of indirection if you want–which keeps your code DRY and increases re-usability, effectively decoupling business logic from the I/O layer. It supports:

- seamless transition between environments;
- abstracting away from resource and data types through resource definitions;
- honouring your expectations on data through schema definitions;
- metrics auto-generation (logging) for monitoring purposes.

Github Repo: here

Side Project [01-03/2018]   Data Science Side Project
Prototyping: Feature-based Customer Review Mining- An Implementation
Description: This project, summarises and explains the approach elaborated in [Kushal Bafna & Durga Toshniwal, 2013]. Its purpose was to serve as a prototype for an internal company project I was working on. The problem I addressed concerns allowing customers to extract objective review summaries relevant to a book, without having to go through the trouble of reading all of the available reviews.
Github Repo: here
Technologies: *Python (nltk), Association Rules Mining, Python Notebooks*

DataReply UK, London [01-03/2017]   Data Science Role, Internal Project
Team R&D Project - Conference Demonstration: Building a chatbot
Description: Used a corpus of movie dialogues to train a sequence-to-sequence model for the implementation of a conversational chatbot.
Technologies: *Python (nltk), Apache Spark, s2s NNs, Tensorflow, Google Cloud*

DataReply UK, London [09–10/2016]   Data Science/Engineering Role, Internal Project
Team Project - Conference Demonstration: Facial expression recognition application
Description: Developed a Neural Network model capable of recognising facial expressions in images; then replaced faces in an image with the corresponding emoticon. Model had to be fed with a large number of categorised images. Image collection was achieved through web-scrapping (beautiful-soup) + semi-manual filtering.
Technologies: *Python, OpenCV, Keras, Tensorflow*

DataReply UK, London [05–09 2016]   Data Science/Engineering Role, Internal Project
R&D Project: Graph analytics & Data Augmentation
Description: Developed a graph analytics solution for company-director networks. Local analysis: developed a real-time graph neighbourhood constructor for companies of interest with multiple visualisation options. Global analysis: carried out centrality (GraphFrames & PageRank) and community detection analytics.
Github Repo: here
Technologies: *Scala, Apache Spark, CYPHER, Neo4j, Python, webAPI*

DataReply UK, London [06 2016]   Data Science/Engineering Role, Internal Project
Conference Demonstration Project: Twitter Sentiment Analysis
Description: Used the twitter API to collect MP tweets related to #BREXIT. Tweets were then categorised based on the public position of the respective MP and were analysed using NLP libraries. Finally, a naive Bayes prediction model was developed to predict any user's opinion, by leveraging their own tweets and those of people they follow.
Github Repo: here
Technologies: *Twitter API, Python (nltk, pySpark), Apache Spark*

certificates & awards

07/2020	**AWS Certified Machine Learning - Specialty** Validation Number: D0R0WC7CN1VQ13SV	AWS
01/2019	**Process Mining** License number: AFU2VTP6273V	Coursera Course Certificates
10/2017	**Google Certified Professional - Data Engineer** Credential ID: 000000400	Google Cloud Certified
07/2017	**Advanced Elasticsearch: Data modelling** Enrolment Id: 33442	Elastic

06/2016	**Graph Analytics for Big Data**	Coursera Course Certificates
	License number: NZADYAQ7N8FW	
05/2016	**Apache Spark Developer Certification**	O'Reilly Media
	License number: 1.x–0548	
04/2016	**Neo4j Certified Professional**	Neo Technology, Creators of Neo4j
	License number: here	
2013	**Best Poster Award**	IJCAI Conference, Beijing, China
	Selected amongst 413 contestants (Prestigious Conference)	
2012	**Outstanding Teaching Assistant Award**	King's College London
	Notable Mention: Department of Informatics (based on student voting)	
2011	**Graduate Certificate in Academic Practice**	King's College London
	Related with improving one's teaching abilities	
2010-2013	**PhD Tuition Fees Scholarship**	EPSRC
	Provided by EPSRC and King's College London	

talks

2022	**Open Data Science Conference**	Industry Conference
	Dynamicio, a python library for abstracting your I/O, <u>**Details:**</u> here	
2020	**Big Data Warsaw**	Industry Conference
	Monitoring & Analysing Communication and Trade Events as Graphs, <u>**Details:**</u> here	
2018	**Connected Data London**	Industry Conference
	AI tribes and interconnections: what's graph got to do with it?, <u>**Details:**</u> here	
2018	**Minds Mastering Machines**	Industry Conference
	A talk on doing Data Science the Agile Way, <u>**Details:**</u> here	
2018	**London Big Data Week**	Industry Conference
	A talk on Big Data and Graph Analytics, <u>**Details:**</u> here	
2016	**Open Data Science Conference**	Industry Conference
	A talk on the advancements in the field of graph-isomorphism, <u>**Video:**</u> here	
2013	**The Imperial College London ACM Student Chapter Seminars**	Workshop
	On anticipating opponent information in competitive contexts, <u>**Video:**</u> here (Invited)	
2011	**The 6th Joint Working Group Workshops**	Maastricht University, NL
	On Strategic Consideration in Dialogue Games (Invited)	

publications

2015	**Building Support-based Opponent Models in Persuasion Dialogues**
	C. Hadjinikolis, S. Modgil, E. Black
	Theory & Applications of Formal Argumentation (LNAI) (Workshop)
2013	**Opponent Modelling in Persuasion Dialogues**
	C. Hadjinikolis, Y. Siantos, S. Modgil, E. Black, P. McBurney
	23rd International Joint Conference on AI (AAAI)(Highly recognised)
2012	**Mechanisms for Opponent Modelling**
	C. Hadjinikolis, S. Modgil, E. Black, P. McBurney
	Imperial College Computing Student Workshop (*Schloss Dagstuhl*)(Workshop)

2012   **Investigating Strategic Considerations in Persuasion Dialogue Games**
        C. Hadjinikolis, S. Modgil, E. Black, P. McBurney, M. Luck
        Starting AI Researcher Symposium (IOS Press)(Workshop)
2012   **Minimising Processor Communication in Parallel Approximate String Matching**
        C. Hadjinikolis, C. S. Iliopoulos, S. P. Pissis, and A. Stamatakis
        Heidelberg Institute for Theoretical Studies (Tech. Report)

programming commitees

2012   **MIWAI**   PC Member
        The 10th Multi-Disciplinary International Workshop on AI
2015   **MIWAI**   PC Member
        The 9th Multi-Disciplinary International Workshop on AI
2012   **ICCSW**   PC Member
        Imperial College Computing Student Workshop
2012   **STAIRS**   PC Member
        6th European Starting AI Researcher Symposium

interests

- **Sports:** Basketball, Football, Tennis, Jogging, Swimming

- **Reading:** Comics, Philosophy, Politics