# PERAKIS GEORGIOS

## DATA SCIENTIST/MACHINE LEARNING ENGINEER

Seasoned machine learning engineer with a robust 10-year background in developing intelligent software solutions through advanced data analytics and machine learning techniques. Expert in extracting actionable insights from complex datasets, with a strong foundation in applied mathematics and statistics. Proficient in both traditional machine learning and cutting-edge deep learning methodologies. Presently spearheading projects in natural language understanding, demonstrating a keen aptitude for evolving AI technologies. Passionate about start-up culture, entrepreneurship, and driving innovation in tech.

**perakisgeorgios[at]gmail.com**
**+306972053994**
**Athens, Greece**
**linkedin.com/in/georgeperakis**
**twitter.com/GeoPerakis**
**github.com/gperakis**

---

## Work Experience

**10/2024 – Now**          **AUTOMATTIC (TUMBRL), SENIOR MACHINE LEARNING ENGINEER**

- Enhanced end-user search experiences by integrating Elasticsearch results with top-ranking posts.
  - Conducted A/B experiments and tracked key performance indicators (user engagement, likes, shares, etc.).
  - Increased critical KPIs by at least 2.5% and successfully reduced spam by 18%.
- Contributed to the analysis and implementation of a mature-term classification system for user queries.

**07/2022 – 10/2024**          **AISERA, STAFF MACHINE LEARNING ENGINEER**

New Intentless Architecture
- Spearheaded the development of a multi-engine search component integrating Neural Search, Public Search, and Flow Search functionalities for domain specific chatbots.

Search Service
- Pod Lead for all delivery epics related to the company's Search Service.
- Instrumental in the design and deployment of a state-of-the-art Neural Search Service, enhancing query accuracy and user experience.
- Accountable for the integration of advanced embedding techniques and optimized model architectures, markedly reducing search latency and elevating system performance.
- Conducted comprehensive code reviews and mentored junior team members.

KB-Matching (Document Matching) Service:

- Pod Lead for the entire KB-Matching service.
- Led the development and launch of a KB-Matching service, employing cutting-edge ML algorithms to optimize document-matching accuracy and scalability.
- Established and applied stringent evaluation metrics, achieving continuous enhancements in document-matching precision.

**05/2019 – Now**          **AEUB, TEACHING ASSISTANT @ MSC BUSINESS ANALYTICS**

Teacher at the post-graduate course "Machine Learning and Content Analytics". Course description: here
Labs contents: Python 101, Numpy 101, Pandas 101, Scikit-Learn 101 | Text pre-processing: Bag of words approach | Embeddings: Gensim (word2vec), Glove, FastText, ELMO. | Image pre-processing & data augmentation techniques | Keras (FFNs, CNNs, RNNs, Attention Mechanisms, LLMs) | Tasks - Classification, Regression, Encoders | Model deployment using FastApi, Docker

**08/2022 – Now**          **UNIVERSITY OF PERAEUS, MACHINE LEARNING ENGINEER**

- Actively providing consulting for various machine learning-related projects.
- Integral member of the team working on the Slothub project, contributing to innovative services including:
  - Developing a predictive model to ascertain if an airline is intentionally delaying flights, leveraging advanced machine learning techniques for accuracy and reliability.
  - Crafting a recommendation system that suggests optimal time slots to airlines, combining data analysis and machine learning algorithms to enhance operational efficiency.

**01/2021 – 09/2022**          **DOCTORANYTIME, MACHINE LEARNING ENGINEER**

- Headed the development initiatives of the Data Science team, driving significant advancements in project outcomes.

- Offered consultancy on a range of data science-related projects, contributing to strategic decision-making and technical direction.
- Conducted thorough code reviews, ensuring high standards of code quality and adherence to best practices in software development.


07/2019 – 07/2022          **SKROUTZ, MACHINE LEARNING ENGINEER**

- An integral member of the Research and Development team at Greece's top e-commerce company.
- Focused on researching and developing working prototypes for visual search applications, pushing the boundaries of AI in e-commerce.
- Collaborated in the implementation of a visual clustering solution to identify and merge duplicate product listings based on their images, enhancing the user experience and platform efficiency.
- Actively contributed to the development of all recommendation engines for the company, employing advanced machine-learning techniques to personalize user experiences and drive sales.

07/2015 – 07/2019          **SOCITAL, JUNIOR DATA SCIENTIST – MACHINE LEARNING ENGINEER**

- Worked on the development of algorithms for extracting key demographic and psychographic attributes (gender, age, base location, interests, brand affiliations, and spending capacity) from social media profiles (Facebook, Twitter, Google+, LinkedIn, and email) for targeted segmentation and personalized recommendations.
- Designed and executed experiments, created prototypes and performed comprehensive statistical analyses to refine data-driven strategies.
- Responsible for the construction, deployment, and maintenance of predictive models using both machine learning and deep learning techniques. Diligently monitored, tested, and assessed model performance.
- Managed the integration and harmonization of data from diverse sources, optimizing data warehouse structures, and automating ETL processes. Skilled in interfacing with various REST APIs.
- Conducted detailed textual data analysis from multiple social networks, applying Natural Language Processing (NLP) techniques to develop robust datasets for advanced analytics.
- Utilized social network analysis methods to compute network centralities, identify cliques, and execute community detection.
- Developed OLAP reporting and KPI visualizations for clients. Innovated an automation framework for dynamic reporting, significantly reducing processing time by approximately 2/3 through the establishment of a data warehouse and a Dash-based web application.
- Regularly prepared and presented insightful reports to the business team at Socital, informing key decisions and addressing critical business issues.
- Authored comprehensive business and developer documentation covering all aspects of the data analytics pipeline, enhancing knowledge sharing and process transparency.

01/2005 – 05/2014          **PRIVATE INSTRUCTOR OF MATHEMATICS**

11/2012 – 01/2013          **CENTRAL BANK OF GREECE, INTERNSHIP**

- Attended bank seminars regarding the Solvency II EU Directive, Insurance Technical Reserves, and "Cashflow matching" technique for insurance companies.
- Studied and prepared a summary of the QIS5 (fifth Quantitative Impact Study) report.
- Organized communications between the Bank of Greece and insurance companies required for the implementation of the Solvency II project.

## Education

10/2016 – 05/2018          **ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS, DEPT. OF INFORMATICS:**

                            **MSC IN DATA SCIENCE, (2-YEAR DEGREE, 75 ECTS, GRADE: 8.7, "DISTINCTION")**

10/2003 – 01/2015          **NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS, DEPT. OF MATHEMATICS**

                            **BSC IN MATHEMATICS, (4-YEAR DEGREE, 250 ECTS)**

## Certifications

**Big Data & Business Analytics** / Theory and Practice (30 hours) | AUEB (11/2014), **The Data Scientist's Toolbox** | Coursera (12/2014), **Mining Massive Datasets** | Coursera (12/ 2014), **R Programming** | Coursera, (03/2015), **Programming in Python** | Coursera, (04/2015), **Introduction to R** | DataCamp, (04/2015), **Intermediate R** | DataCamp, (06/2015), **Neural Networks and Deep Learning** | Coursera (10/2017), **Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization** | Coursera (10/2017), **Structuring Machine Learning Projects** | Coursera (11/2017)

**Skills and Tools**

Programming Languages:
- Experience with **Python** (10 years). Libraries: NumPy, SciPy, Pandas, Scikit-Learn, Requests, NLTK, Spacy, Keras, TensorFlow, Pytorch, Transformers, Matplotlib, Seaborn, Plotly, FastApi, Dash, etc.
- Some Experience with **R**, **Ruby on Rails, PHP, Next.js**
- Experience with REST APIs and gRPC protocols.

Databases: **PostgreSQL**, **MariaDB**, **SQLite3**, **MongoDB**, **ElasticSearch, OpenSearch, Redis**. Direct Queries or use of ORMs.

Software: **SPSS Modeler**, **KNIME**

Visualization: **Grafana**, **Tableau**, **PowerBI, Dash, Streamlit, Looker**

Containers and Orchestration: **Docker, Kubernetes, DAG-based workflows (Airflow)**


Languages: **English** (Proficiency), **Greek** (Native language)


Excellent Communication skills, problem-solving, teamwork and collaboration
Project management (working with JIRA), Technical Analysis
API design knowledge, code reviews, Gitflow processes

**Projects: (Most projects are located on a GitHub account)**

- **Email Gender and Name Analysis** (Socital - Python) ([project](#)):
  - Given an email or formatted name, we estimate the user's gender and extract the first and last name (if exists). Gender Acc: 95%. Name Acc: 91%. Gender Coverage: 73%
- **Gender Estimation in Social Network Profiles** (Socital - Python):
  - Estimate gender using image, formatted name, email, SNA, and behavioral data.
- **Time Series Analysis using ARIMA and (G)ARCH models** (using R) (AUEB – 02/2018) ([code](#)).
- **Social Network Analysis using SNAP** (Python) (AUEB – 05/17) ([code](#))
- **Time Deposit Propensity Modeling on Banking Sector** (SPSS modeler) (AUEB – 12/2017)
- **Telephony Segmentation** (SPSS modeler) – (AUEB – 12/2017) ([presentation](#)): Value-Based Vs Behavioral.
- **Sentiment Analysis using Neural Networks** (Python) (AUEB – 04/18) ([code](#)): Perform sentiment analysis in English reviews using deep learning algorithms including MLPs, CNNs, and RNNs.
- **Sentiment Analysis using Machine Learning Algorithms** (Python) (AUEB – 04/18) ([code](#)): Perform sentiment analysis in English reviews using various machine learning algorithms.
- **Part of Speech (POS) Tagger python implementation** (Python) (AUEB – 04/18) ([code](#)): Create an implementation of a Hidden Markov Model (HMM) and a Conditional Random Field (CRF) model to perform sequence labeling of POS Tags.
- **Predictive Keyboard Implementation** (Python) (AUEB – 04/18) ([code](#)): Create a Predictive Keyboard (Next Word Predictor) for English language implementation using Python.
- **Name (entity) extraction from usernames by performing Neural Machine Translation** (NMT) using a combination of Attention (RNN) Models and Pointer Networks. (Python)
- **Sentiment Analysis on Greek reviews** (python) (AUEB's NLP Group)

**Awards and Participation in Competitions**

- Participated in the first Greek Datathon Competition (Data analysis and visualization), winning the 2nd prize (06/2014)
- Participated in the Global Management Challenge competition at the first regional level (01/2014)

**Military services fulfilled. Recommendations available upon request**