Nikiforos Pittaras, PhD

Machine Learning researcher / engineer

Machine Learning, AI and Data Science enthusiast, focused on research and development of deep learning methods for text & multimedia, and investigations into AI safety and alignment

🌐 https://npit.github.io
✉ pittarasnikif@gmail.com

</> python, java, matlab, C++
🚀 sklearn, pytorch, tensorflow
🗄 mysql, postgres, cassandra
>_ linux, bash, docker, git

EDUCATION 🎓

- **PhD in Machine Learning**: (University of Athens / NCSR Demokritos)   2017-2021
  4-year doctorate, awarded with an SNF Industrial Scholarship, Grade 'Excellent'.
  Specialized in augmenting Deep Neural Representations with external information for ML tasks.
  Developed a Hate Speech Detection system from research findings for industry deploymen.t
  Thesis: Beyond Deep Learning: Enriching Data Representations for Machine Learning Tasks
  `>` `python, tensorflow, keras, pytorch, DNN, nltk, Wordnet, ray, mlflow, restx`

- **MSc in Signal/Information Processing & Learning**: (University of Athens)   2015-2018
  2-year Master's degree, focused on Signal Processing and Machine Learning. Grade 8.78.
  Specialized in Machine Learning, Signal and Multimedia Analysis and Processing, Ontology Engineering
  Thesis: Multimodal video classification with deep neural networks (rg)
  `>` courses: `{matlab, java, C++}`, thesis: `{python, tensorflow, keras, DCNN, LSTM}`

- **BSc in Computer Science**: (University of Ioannina)   2005-2013
  4-year Bachelor's degree (Ptychion), with a focus on Machine Learning, AI and VLSI. Grade 7.08.
  Specialized in Machine Learning and Artificial Intelligence, Multimedia Processing, Hardware Design
  Thesis: Image registration using unified particle swarm optimization (pdf)
  `>` courses: `{C, C++, java, matlab, opengl, vhdl}`, thesis: `{matlab, UPSO}`

PROFESSIONAL EXPERIENCE ★

- **Teaching Assistant**: MLSS program - Athens, Greece   June-August 2022
  - **TA**: Theoretical / technical support on advanced Machine Learning, Deep Learning, AI Safety and Alignment.

- **AI Alignment Researcher**: Independent / EA-LTFF - Athens, Greece   June - October 2022
  - **R&D on Multiagent RL**: Reinforcement Learning environment design and engineering. Agent communication and betrayal pattern analysis for cooperative tasks on empirical experiments.
  - `>` `python, pytorch, gym, stable-baselines3, docker, AWS`

- **Postdoctoral Research Engineer**: IIT-NCSRD/ICS-FORTH - Athens, Greece   Nov 2021 - present
  - **R&D / Text ML**: Research, ML and backend engineering on argumentation mining on big crawled article data, utilizing active learning, clustering, topic modelling and token / text classification with transformer models.
  - `>` `python, pytorch, transformers, BERT, flair, elasticsearch, fastapi, docker`

- **Senior ML Engineer**: LANGaware - Athens, Greece   June 2021-present
  - **ML Engineering**: Pipeline develpoment and maintenance for multimodal neurodegenerative disease prediction. Design and implementation of ETL, classification and automation procedures.
  - **ML Interpretability**: Work on feature / instance-based classification interpretation and presentation.
  - **ML Prototyping**: Work with the Data Science team to implement new solutions, experiments and prototypes.
  - **Technical Writing**: Authorship of proposals, reports and applications related to scientific/research projects.
  - `>` `python, sklearn, RF, numpy, spacy, nltk, librosa, gcp`

- **Lead ML Engineer**: SciFY- Athens, Greece   Nov. 2020-August 2021
  - **ML Engineering**: Development, maintenance, evaluation of text classification pipelines in the legal domain.

> `python, sklearn, numpy, nltk, pytorch, transformers, BERT, flask`

- **Research Associate/ ML & Backend Engineer**: ILSP-ATHENA - Athens, Greece   Feb.-Dec. 2019
  - **R&D on Text ML**: Lead developer on multimodal topic modelling on scientific publications for the D4I research project. Pipeline maintenance and extension, administration and deployment.
  - **Result analysis and report**: Meta-analysis, data aggregation, visualization and presentation; indicator engineering and impact assessment. Project report compilation and authoring
  - **System / backend administration**: VM / system management and setup, database extension / optimization.
  - > `java, mallet, LDA, spring, postgres, python, sklearn, django, flask`

- **Research Associate / Lead ML Engineer**: IIT-NCSRD - Athens, Greece   June 2016- Sep. 2021
  - **R&D on Text ML**: Text & social media mining and processing for event detection, clustering, summarization, recommendation, entity resolution, semantic search and personalization, for the BDE and Storybot research projects. Research, development, maintenance, extension, scaling to big data, optimization and deployment.
  - **System administration & engineering**: VM and software virtualization containers design, deployment and maintenance. RESTful API frontend/backend development, website deployment and migration.
  - **Academic Writing**: Publication authorship, experiment planning, execution, automation and analysis.
  - **Community management**: Organization of the internal REPR and CAKT research groups
  - > `java, jax-rs, pyinsect, NGG, docker, cassandra, spark, bash, glassfish`

- **Research Associate / Lead ML Engineer**: ITI-CERTH - Thessaloniki, Greece   Feb. 2014 - Sep. 2015
  - **R&D on Computer Vision**: Visual descriptors engineering, normalization and clustering, for the LinkedTV and ForgetIT projects; Lead development / design of a full video concept detection, image retrieval, search and visualization pipelines. Maintenance, extension, space / time optimization on the CPU and GPU.
  - **R&D on Deep Computer Vision**: Research on deep learning methods for video concept detection with DCNNs, network design, training and transfer learning. Lead design, development, maintenance and extension of video concept detection / retrieval / search DCNN pipelines for real-time execution.
  - **Academic Writing**: Publication authorship, experiment planning, execution, automation and analysis.
  - > `C++, matlab, opencv, libsvm/linear, caffee, matconvnet, DCNN, openmp, cuda`

Academic / Scientific Activities   ⚗

- **Teaching / supervision**:
  - Co-supervision of BSc / MSc theses and Internships related to Machine Learning, NCSRD, 2017-present
  - Student group theoretical / technical support, coordination and organization, NCSRD, 2017-present
  - Scientific lectures and presentations on the CAKT / REPR research groups, 2017-2019
  - Course assistant (Data Structures, OOP, SoftEng for Networks Applications, Data Mining), UoA, 2015-2020

- **Conferences / Events**:
  - Workshop organizer / chair (MultiLing 2019, FNP/FNS 2020-2022)
  - Program Committee (AACL-IJCNLP2020, SLT 2021, FNP/FNS2021-2022, SDPRA 2021)
  - Paper Reviewing (ICTAI, ACL, MLj, CIj, SLSP, CSL, openreview, etc.)
  - Technical support (Multiling 2017, AI4EU Athens), research proposal evaluation (AI4Copernicus)

- **Seminar / Workshop participation**:
  - AGI Safety Fundamentals 2022 (Effective Altruism Cambridge, virtual, Jan-Mar 2022)
  - AI Safety Camp (virtual research workshop, working on Deep Multiagent RL, Jan - Jun 2022)
  - "Stanford Existential Risk Initiative" Conference, 2022, virtual (Winter 2022)
  - Reinforcement Learning / DRL courses (2) (Coursera, Huggingface, 2022)
  - Workshops (3) in open sourcing, technology transfer and utilization of AI (virtual, 2020 - 2021).
  - "Deep Learning in NLP" summer school (RANLP 2019, Fall 2019).
  - Academic CompSci seminars (5), (DIT-UoA, 2015-2018)
  - "CretaMASSS-2013 / HAISS'13-Agents" summer school on multi-agent systems & AI (TUoC, 2013).

Publications

- **Journals**:

  1. N. Pittaras, G. Giannakopoulos, G. Papadakis, V. Karkaletsis,"Text classification with semantically enriched word embeddings", JNLE Special Issue: Informing Neural Architectures for NLP with Linguistic and background Knowledge, 2020 (pdf)

  2. F. Markatopoulou, V. Mezaris, N. Pittaras, I. Patras, "Local Features and a Two-Layer Stacking Architecture for Semantic Concept Detection in Video", IEEE TETC, 2015 (ieee) (pdf).

- **Conferences / Workshops** (selected 5 out of > 20 total):

  1. N.Pittaras, V. Karkaletsis, "A study of semantic augmentation of word embeddings for extractive summarization", Multliling Workshop, RANLP2019, Varna, Bulgaria (proceedings).

  2. C. Themeli, G. Giannakopoulos, N. Pittaras "A study of text representations for Hate Speech Detection", CICLING 2019, La Rochelle, France (to appear).

  3. N. Pittaras, G. Papadakis, G. Stamoulis, G. Argyriou, E. K. Taniskidou, E. Thanos, G. Giannakopoulos, L.Tsekouras, E. Koubarakis, "GeoSensor: Semantifying Change and Event Detection over Big Data", SAC 2019, Limassol, Cyprus (pdf).

  4. N. Pittaras, G. Giannakopoulos, L. Tsekouras, I. Varlamis, "Document clustering as a record linkage problem", DocEng'18, Halifax, Nova Scotia, Canada, 2018 (ResearchGate).

  5. N.Pittaras, F.Markatopoulou, V.Mezaris, I.Patras, "Comparison of Fine-tuning and Extension Strategies for Deep Convolutional Neural Networks", 23rd International Conference on Multimedia Modeling, Reykjavik, Iceland, Jan 2017 (springer) (pdf).

- **Book chapters**:

  1. N. Pittaras, S. Montanelli, G. Giannakopoulos, A. Ferrara, V. Karkaletsis, "Crowdsourcing in Single-document Summary Evaluation: the Argo Way", Multilingual Text Analysis: Challenges, Models, and Approaches, World Scientific Publishers, 2020 (amazon).

  2. G. Giannakopoulos, G. Kiomourtzis, N. Pittaras, V. Karkaletsis, "Scaling and semantically enriching language-agnostic summarization", Trends and Applications of Text Summarization Techniques, 2019 (IGI Global).

Tech Skills

- programming: python, java, bash, LATEX, C, C++, Matlab; various IDEs / SDKs

- ml: pytorch, sklearn, transformers, tensorflow, caffee, libsvm, opencv, weka, pettingzoo, gym, stable-baselines3

- mlops, monitoring: docker, mlfow, dvc, wandb, ray-tune

- db, semantics, cloud: cassandra, spark, mysql, postgres, elasticsearch, RDF, owl, protege, sparql, GCP, mongo

- web, docs, teams: html, tomcat, hugo, flask, restx, Jira, Trello, Slack, Confluence, sphinx

- other: linux 🐧, git, maven, django, pdb, blender, godot

Language and Miscellaneous skills

- Greek (Mother tongue), English (Fluent, Cambridge CPE), German (Basic)

- Completed lifeguard training, licensed to drive cars and speedboats.

Hobbies and interests

- Long distance running and cycling, swimming, video games & indie game development.

- Musical instruments (guitar, piano, some bass and ocarina), singing, song writing.

- Popular science (philosophy of science, number theory, particle physics).