

# Nikiforos Pittaras, PhD

Machine Learning researcher / engineer

Machine Learning, AI and Data Science enthusiast, with a passion for research and development of deep learning approaches for text and multimedia.

<https://npit.github.io>

[pittarasnikif@gmail.com](mailto: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 [Industrial Scholarship](#), Grade 'Excellent'.
  - Specialized in augmenting Deep Neural Representations with external information for various ML tasks.
  - Developed a Hate Speech Detection system from research findings for industry deployment
  - 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}

## EXPERIENCE



- **Postdoctoral Researcher:** [IIT-NCSR](#) - Athens, Greece Nov 2021 - present
  - **R&D on Text ML:** Research and ML engineering on argumentation mining with transformer models.`>` python, pytorch, huggingface, BERT, flair, elasticsearch, fastapi, docker
- **Senior Machine Learning Engineer:** [LANGaware](#) - Athens, Greece June 2021-present
  - **ML Engineering:** Pipeline development 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 Machine Learning Engineer:** [SciFY](#) - Athens, Greece Nov. 2020-August 2021
  - **ML Engineering:** Development, maintenance and evaluation of text classification pipelines in the legal domain.`>` python, sklearn, numpy, nltk, pytorch, transformers, BERT, flask
- **Research Associate/ Lead Developer:** [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 Developer:** [IIT-NCSR](#)D - Athens, Greece Jun. 2016- Sep. 2021
  - **R&D on Text ML:** Text & social media mining and processing for event detection, clustering, summarization, recommendation, entity resolution 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 REPR and CAKT research groups, NCSR

> java, jax-rs, pyinsect, NGG, docker, cassandra, spark, bash, glassfish
- **Research Associate / Lead Developer:** [ITI-CERTH](#) - Thessaloniki, Greece Feb. 2014 - Sep. 2015
  - **R&D on Computer Vision:** Visual descriptors and colour extensions, normalization and clustering, for the LinkedTV and ForgetIT projects; Lead developer and designer of a full video concept detection and visualization pipeline. 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 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 ACTIVITIES



- **Course / Supervision support:**
  - Course assistant (Data Structures, OOP, Soft. Dev. for Networks, Data Mining), UoA
  - Co-supervision of BSc / MSc theses and Internships related to Machine Learning, NCSR
  - Theoretical / technical support and guidance, student group coordination and organization, NCSR
  - Scientific lectures and presentations on the CAKT / REPR research groups
- **Conferences / Events:**
  - Workshop organizer / chair (MultiLing 2019, FNP/FNS 2020-2022)
  - Program Committee (AAACL-IJCNLP2020, SLT 2021, FNP-FNS2021, SDPRA 2021)
  - Reviewing (ICTAI, ACL, Machine Learning, Computational Intelligence Journal, SLSP, CSL, etc.)
  - Technical support (Multiling 2017, AI4EU in Athens)
- **Seminar / Workshop attendance:**
  - AGI Safety Fundamentals Course (Effective Altruism Cambridge, virtual, Jan-Mar 2022)
  - AI Safety Camp 2022 (virtual research workshop, working on Deep Multiagent RL, Jan - Jun 2022)
  - Academic seminars (3) in Dept. of Informatics and Telecommunications, (UoA, 2015-2018)
  - “Stanford Existential Risk Initiative” Conference, 2022, virtual (Winter 2022)
  - “CretaMASSS-2013 / HAISS’13-Agents” summer school on multi-agent systems & AI (TUoC, 2013).
  - “Deep Learning in NLP” summer school (RANLP 2019, Fall 2019).
  - “Open Source AI Workshop”, workshop on AI, Eclipse Foundation (virtual, Summer 2020).
  - “AI for Multilanguage Services for Citizens and Businesses”, workshop, ELRC(virtual, Winter 2020).
  - “Prometheus Technology Transfer program”, workshop, SciFY, SRI, US Embassy(virtual, April 2021).
  - “Fundamentals of Reinforcement Learning”, virtual course, Coursera (Winter 2022).

## 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 and Workshops (selected 5 out of > 20 total):


1. N. Pittaras, V. Karkaletsis, “A study of semantic augmentation of word embeddings for extractive summarization”, Multiling 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: C, C++, Java, Matlab, python, bash, L<sup>A</sup>T<sub>E</sub>X; various IDEs / SDKs
- ml: pytorch, tensorflow, sklearn, transformers, coffee, libsvm, opencv, weka, ray-tune, pettingzoo, gym, stable-baselines3, mlflow, dvc, wandb
- databases, semantics & cloud: cassandra, spark, mysql, postgres, RDF, owl, protege, sparql, GCP, mongo
- web, docs and teams: html, tomcat, hugo, flask, restx, Jira, Trello, Slack, Confluence, sphinx
- other: linux , git, maven, docker, 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.
- Philosophy of science, AI safety , popular science (number theory, particle physics).