Stuttgart, Germany
+49 152 57625087
irina.nicolae@proton.me
ririnicolae.github.io
linkedin.com/in/irina-nicolae-a2251638
ririnicolae
Google Scholar

Irina Nicolae

Lead machine learning researcher

Selected work experience

2019–now **Researcher, machine learning & security**, *Bosch Research*, Stuttgart, Germany

LLM trust — Project lead—trust for LLM applications (ongoing)
- Automated the evaluation of LLM applications by developing methods, metrics and tools.
- Established "Risks of GenAI" as strategic topic for Bosch Research.

Fuzz testing — Automated software testing using ML
- Increased fault discovery in CAN protocol fuzzing by 20x by generating test data from a transformer.
- Identified fundamental limitations in neural program smoothing for fuzzing.

Planar robots — Hybrid ML control for magnetic levitation
- Improved levitation performance in a PID controller by training a model on an inverse problem.

Power tools — Embedded ML for electric screwdrivers
- Implemented an automated screw stop feature by training an ML model on time series.

2017–2019 **Researcher, adversarial ML**, *IBM Research*, Dublin, Ireland
- Created and led the IBM Adversarial Robustness Toolbox, one of the top libraries in the field.
- Developed new defenses against adversarial attacks and model robustness metrics.
- Won Darpa grant, IBM Outstanding Technical Accomplishment Award and EU Horizon projects.

Education

2013–2016 **Ph.D. in machine learning**, *Jean Monnet University*, France
- Developed new metric learning algorithms for classification on tabular data and time series, with theoretical guarantees on their performance.

2011–2013 **MSc in computer science**, *Ensimag, Grenoble INP*, France, Major in Information Systems

2007–2011 **BSc in computer science**, *Politehnica University of Bucharest*, Romania

Skills

ML paradigms	Supervised, unsupervised and self-supervised learning	**Deep learning**	CNNs, RNNs, autoencoders, GANs, transformers
ML tools	TensorFlow, Keras, PyTorch, ScikitLearn, Numpy, Pandas	**LLM tools**	LangChain, LangGraph, HuggingFace, OpenAI, LiteLLM
Programming	Python, C	**DevOps**	Docker, GitHub Actions, GitLab
Languages	English, French, Romanian	**Interpersonal**	Driven, pragmatic, honest, team player

Other accomplishments

Publications **Academic** ICML, ICLR, NeurIPS, ESEC/FSE, CCS, AAAI, ECML/PKDD.
Industrial BlackHat, RSA, Dagstuhl and over 15 patents.

Reviewer ICLR 2021–, NeurIPS 2016–, ICML 2016–, AAAI 2020–2022, IJCAI 2020–2024.

Editor `https://medium.com/security-garten`: security, privacy and safety blog.

Organizer Nemesis Workshop at ECML/PKDD 2018, IDA International Symposium 2015.

Mentoring N. Niayesh, Z. Zhuang, K. Sheng, J. Messner, A. Soutiff , V. Zantedeschi.