Pavol Mulinka

tel.num: +34 661816178 | **address:** Barcelona, Catalunya, Spain | **email:** mulinka.pavol@gmail.com | **Linkedin:** https://www.linkedin.com/in/mulinka/ | **Github:** http://github.com/5uperpalo | **Google scholar:** https://scholar.google.com/citations?user=zsJ4nfoAAAAJ

Summary

I am a Data Scientist with 7+ years of experience in AI/ML design and implementation. I specialize in agentic AI systems and data-driven design patterns. I lead cross-functional technical design discussions and have published research in telecommunications, NLP, and ML.

Skills

name	level	notes
Python	Advanced	Scikit-learn, Flask, FastAPI, Pandas, SQLAlchemy, SciPy, Pydantic etc.
Javascript	Beginner	solutions improvement, MOA scripting and usage
SQL	Intermediate	querying, data analysis, local database administration
KQL	Intermediate	querying, data analysis
Prompt engineering	Advanced	Ollama, Azure OpenAI SDK, Google GenAI SDK
Agentic AI	Intermediate	Langchain, Langsmith
Kubernetes	Intermediate	Microk8s, K3s, Helm charts, administrator, FIREMAN and SUCCESS6G projects
Docker	Intermediate	dockerization of multiple solutions
Data Analysis	Advanced	EDA, Statistical Analysis, Data Modeling
Data Mining	Intermediate	ETL
Machine Learning	Advanced	Predictive Modeling, GBM, Markov chains, Hidden Markov Models, etc.
Deep Learning	Intermediate	GAN, GAIN, transformers, PyTorch
Cloud	Intermediate	AWS, Azure, GCP, Cloud Computing, Distributed computing
Linux	Intermediate	Bash, scripting, lightweight solutions
Networking etc.	Intermediate	design, implementation, troubleshooting

Languages

- **Spanish** - C1/C2
- **English** - C2
- **Czech** - C2
- **Slovak** - native

Work Experience

- **Cybersecurity Data Analyst**, freelancer, hybrid, Forescout - NL (17/09/2024 - 31/12/2025)
  - **Description:** Analyzed and classified data from diverse security tools and systems; identified patterns, anomalies, and potential security threats; applied Agentic AI and prompt engineering solutions to improve classification accuracy and automatize Data Analyst tasks. Developed and implemented automated impact analysis of core solutions modifications.

- **Keywords**: applied research, data analysis, data science, machine learning, statistics, text processing, investigation, analytical thinking, collaboration, problem solving, proactivity, Agile
- **Technologies**: Python, SQL, KQL, NLP, LLM, AI, Agentic AI, Langchain, Langsmith, Docker, Terraform, Streamlit, AWS, GCP, Azure, Azure AI, OpenAI, Gemini, Google GenAI SDK, Javascript, Grafana, Github, RAG

- **Data scientist**, freelancer, hybrid, Assetario - SK (01/03/2021 - 26/01/2024)
  - **Description:** Designed, analyzed, and implemented machine learning models for predicting customer lifetime value in mobile apps and in-app purchase recommendations. Enhanced data processing and feature engineering for models to improve conversion rates and personalization.
  - **Keywords**: data analysis, data science, machine learning, deep learning, communication, recommendation system, In-App-Purchases, mobile phones, problem solving, Predicted Lifetime Value (pLTV), programming, statistics, Agile
  - **Technologies**: Python, Athena, SQL, AWS, LLM, Huggingface, Transformers, GBM, Weights&Biases, H20, DBSCAN, OPTICS, Scikit-learn, pytorch, MLflow, Github

- **F5 loadbalancer specialist**, freelancer, remote, Oksystems - CZ (01/12/2021 - 01/10/2022)
  - **Description:** Migrated an existing Apache XML firewall and loadbalancer solution to F5 loadbalancer for a Ministry of Agriculture project.
  - **keywords**: migration, scripting, loadbalancer, firewall
  - **technologies**: F5, XML, bash, Postman

- **Data scientist**, full-time, onsite, CTTC - ES (01/02/2020 - 30/09/2024)
  - **Description:** Designed and analyzed machine learning approaches for pattern and anomaly detection in real-world and synthetic datasets. Led projects with a focus on distributed computing and large-scale data analysis.
  - **Keywords**: research, applied research, data analysis, data science, machine learning, deep learning, project leading, pattern detection, anomaly detection, programming, statistics, analytical thinking, collaboration, research paper writing, investigation, distributed machine learning, federated learning, V2X, 5G, 6G, problem solving
  - **Technologies**: Python, SQL, Flask, Celery, Docker, Helm, Kubernetes, Redis, H20, YOLO, Transformers, GBM, LLM, Huggingface, Weights&Biases, DBSCAN, OPTICS, Scikit-learn, pytorch, Istio, Microk8s, K3s, Raspberry Pi, Flower, Kepler, Prometheus, InfluxDB, Minio, Kserve, Knative, MLflow, Github, GAN, GAIN, river, deep-river, MOA, Kafka, Airflow, KSQL, Faust, Kubeflow, Zero-to-Jupyterhub, MySQL, RAG

- **Python developer**, freelancer, remote, Slovak power plants - SK (01/06/2020 - 01/10/2020)
  - **Description:** Designed and developed a communication interface between the Slovak Electricity Hydro optimization model and a user GUI.
  - **Keywords**: containerization, in-memory database, web, multi-processing, unit testing,procedural programming, object oriented programming, code refactorization
  - **Technologies**: Docker, Redis, Flask, Celery, Python, Github

- **Data Scientist**, contract, onsite, NII Tokyo - JP (08/03/2019 - 03/09/2019)
  - **Description:** Applied unsupervised machine learning to network traces to detect and interpret unknown patterns. Improved hierarchical density-based clustering for better network measurement interpretation. Refactored code for distributed computational environments.
  - **Keywords**: MAWI, Darknet, anomaly detection, big data, machine learning, deep learning
  - **Technologies**: PySpark, Python

- **Data Scientist**, contract, onsite, O2 Telefonica - ES (21/11/2018 - 20/02/2019)

- **Description:** Analyzed relationships between socioeconomic status and network performance. Investigated potential discrimination in network deployment. Correlated public data with network measurements through geospatial analysis.

- **Keywords**: Lower-layer Super Output Areas (LSOA), data analysis, machine learning, geospatial machine learning
- **Technologies**: QGIS, ArcGIS, GeoPandas, PySpark, Python

- **Data Scientist**, contract, onsite, AIT Vienna - AT (01/03/2018 - 31/08/2018)
  - **Description:** Conducted cybersecurity and network performance analysis. Developed anomaly detection and diagnosis systems. Integrated machine learning techniques into big data platforms, BIG-DAMA project.
  - **Keywords**: stream-based machine learning, supervised machine learning, unsupervised machine learning, MAWI, Cloud latency, network performance analysis, anomaly detection, big data
  - **Technologies**: Cloudera, PySpark, Apache Pig, Hive, Kafka, Elasticsearch, Python

- **Network engineer**, full-time, onsite, CZ and SK (01/08/2007 - 01/03/2018)
  - positions in descending order:
    - Network Engineer (VSHosting, Prague, CZ)
    - Network Consulting Engineer (Verizon, Prague, CZ)
    - Senior System Engineer (ATT, Bratislava, SK)
    - HP Radia Specialist (Soitron, Bratislava, SK)
    - HP Monitoring Support Specialist (Soitron, Bratislava, SK)
    - IT VoIP support specialist (Soitron, Bratislava, SK)
  - **Description:** Designed, implemented, supported, and documented network infrastructures, with a strong focus on security, troubleshooting, and performance.

  - **Keywords**: networking, voip, wireless, vpn, security, firewall, design, troubleshooting, implementation
  - **Technologies**: bash, IOS

Teaching Experience

- **"Sakura Science Plan" Mentor**, NII, Tokyo - JP (2019)
- **"Network Operating Systems" Teaching assistant**, FEE, CTU, Prague - CZ (2015, 2016 winter semester)
- **"Digital Engineering" Teaching assistant**, FEE, CTU, Prague - CZ (2014 winter semester)
- **"Communication Processes Control" Teaching assistant**, FEE, CTU, Prague - CZ (2014 and 2015 summer semester)

Projects

Research

- SUCCESS-6G: Towards robust, secure and computationally efficient vehicular services in 6G, github, Co-investigator, CTTC - ES (2020-2024)
  - **Description:** Researched and developed secure and efficient real-time vehicle condition monitoring and fault provisioning systems for 6G networks using ML and IoT technologies in the context of vehicle-to-everything (V2X) communication.
- FIREMAN (Framework for the Identification of Rare Events via MAchine learning and IoT Networks), github, Co-investigator, CTTC - ES (2020-2024)
  - **Description:** Designend, developed and implemented a novel big-data based framework that encompasses all steps from sensing and data acquisition to statistical analysis and operational decisions, to accurately identify, detect, forecast and prevent rare events in a industrial physical processes.

- **Practical Privacy-Preserving Data Collection and Utilization using Provable Cryptographic Tools**, Principal investigator, FEE, CTU, Prague - CZ (2019)
  - **Description:** Researched and implemented privacy-preserving data collection methods using cryptographic tools.
- **Privacy Protection and Machine Learning Utilization of IoT Data in Cloud**, Principal investigator, FEE, CTU, Prague - CZ (2018)
  - **Description:** Engineered cloud-based IoT data analysis pipelines with a focus on privacy and machine learning.
- **Smart-home IoT and Cloud Telemetry Datamining**, Principal investigator, FEE, CTU, Prague - CZ (2017)
  - **Description:** Developed data mining tools for smart-home IoT devices and cloud telemetry systems.
- **Cloud Performance Analysis and Improvement**, Principal investigator, FEE, CTU, Prague - CZ (2015 - 2016)
  - **Description:** Analyzed cloud performance metrics and proposed optimization strategies for cloud computing environments.
- **Methods Enhancing Work with Cloud Data**, Principal investigator, FEE, CTU, Prague - CZ (2014)
  - **Description:** Designed methodologies for managing and analyzing cloud datasets efficiently.
- **Metrics for Automated Detection of Cloud Anomalous Behavior**, Principal investigator, Cisco Systems - CZ (2013)
  - **Description:** Developed automated metrics and detection systems for identifying anomalies in cloud environments.

Open Source

- **pytorch-widedeep**, Collaborator (2021-2024)
  - **main collaborations**: Deep Imbalanced Regression, New loss functions, Custom Imbalanced DataLoader
  - **Description:** Implemented methods and loss functions for imbalanced dataset processing.
- **Wikimedia Scoring platform team**, External collaborator (2020-2021)
  - **main collaborations under** Pavol86: Compress Gensim models, python-mwtext, Tokenization of "word" things for CJK, deltas, revscoring, editquality
  - **Description:** Led research and implementation efforts for CJK language tokenization. Developed comprehensive benchmarks comparing state-of-the-art segmenter. Contributed to model compression strategies for Gensim utilities.

Education

- **PhD**, Telecommunications, Faculty of Electrical Engineering, Czech Technical University, Prague - CZ (2013 - 2021)
  - **Thesis**: "Hierarchical density-based clustering and interpretation for network measurements"
  - **Description:** Developed hierarchical density-based clustering machine learning pipelines for analyzing network data and detecting unknown patterns.

- **MSc**, Telecommunications, Faculty of electrical engineering, Slovak University of Technology, Bratislava - SK (2007 - 2009)
  - **Thesis**: "Classificators for identification of the speaker"
  - **Description:** Designed and implemented speaker classification models using machine learning techniques.
- **Bc**, Telecommunications, Faculty of electrical engineering, Slovak University of Technology, Bratislava - SK (2004 - 2007)
  - **Thesis**: "Measurement of glottal period of human voice"
  - **Description:** Conducted detailed research on measuring and identifying the glottal period for voice signal analysis.

Courses and certifications

- **Cisco Certified Network Associate (CCNA, 640-802)**, (640-553) Implementing Cisco IOS Network Security, (640-460) Implementing Cisco IOS Unified Communications, (640-721) Implementing Cisco Unified Wireless Network Essentials
- **Cisco Certified Design Associate (CCDA)**; (640-863) Designing for Cisco Internetwork Solutions
- **Cisco Certified Network Professional (CCNP)**; (642-901) Building Scalable Cisco Internetworks, (642-812) Building Cisco Multilayer Switched Networks, (642-825) Implementing Secure Converged Wide Area Networks, (642-845) Optimizing Converged Cisco Networks
- **Cisco Certified Internetwork Professional (CCIP)**; (642-642) Quality of Service, (642-611) Multiprotocol Label Switching, (642-661) Border Gateway Protocol
- **Cisco Certified Design Professional (CCDP)**; (300-31) Designing Cisco Network Service, (642-873) Designing Cisco Network Service Architectures
- **Conducting Cisco Unified Wireless Site Survey (CUWSS, 642-731)**
- **Implementing Cisco Edge Network Security Solutions (SENSS, 300-206)**
- **F5 Certified Product Consultant for LTM**; F5-PCL, F50-531
- **F5 Certified Administrator**; (101) Application Delivery Fundamentals, (201) TMOS Administration
- **Juniper Networks Certified Internet Associate EX (JNCIA-EX, JN0-400)**
- **Information Technology Infrastructure Library Foundation in IT Service Management (ITILv3, Foundation)**
- **The Open Group Architecture Framework (TOGAF 9)**
- **ArchiMate 3**
- **driving license A+B**
- **diving license OWD**

Publications

- Pavol Mulinka and Christou, Ioannis T. and Subham Sahoo and Charalampos Kalalas and Nardell, Pedro H.J.. (Oct 2025). "Towards High-Fidelity and Trustworthy Digital Twins for Fault Diagnosis in Grid Connected Inverters". In: IEEE Transactions on Dependable and Secure Computing. doi: 10.1109/TDSC.2025.3626846.
- Pavlidis, Nikolaos and Perifanis, Vasileios and Yilmaz, Selim F. and Wilhelmi, Francesc and Miozzo, Marco and Efraimidis, Pavlos S. and Koutsiamanis, Remous-Aris and Mulinka, Pavol and Dini, Paolo. (May 2025). "Federated Learning in Mobile Networks: A Comprehensive Case Study on Traffic Forecasting". In: IEEE Transactions on Sustainable Computing. pp. 576-587. doi: 10.1109/TSUSC.2024.3504242.
- Charalampos Kalalas and Pavol Mulinka and Guillermo Candela Belmonte and Miguel Fornell and Michail Dalgitsis and Francisco Paredes Vera and Javier Santaella Sánchez and Carmen Vicente Villares and Roshan Sedar and Eftychia Datsika and Angelos Antonopoulos and Antonio Fernández Ojea and Miquel Payaro. (2025). "AI-Driven Vehicle Condition Monitoring with Cell-Aware Edge Service Migration". url: https://arxiv.org/abs/2506.02785.
- Zaurin, Javier Rodriguez and Mulinka, Pavol. (Jun 2023). "pytorch-widedeep: A flexible package for multimodal deep learning". In: Journal of Open Source Software. pp. 5027. doi: 10.21105/joss.05027. url: https://joss.theoj.org/papers/10.21105/joss.05027.
- Beattie, Alexander and Mulink, Pavol and Sahoo, Subham and Christou, Ioannis T. and Kalalas, Charalampos and Gutierrez-Rojas, Daniel and Nardelli, Pedro H. J.. (2022). "A Robust and Explainable Data-Driven Anomaly Detection Approach For Power Electronics". In: 2022 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm). pp. 296-301. doi: 10.1109/SmartGridComm52983.2022.9961002.
- Mulinka, Pavol and Sahoo, Subham and Kalalas, Charalampos and Nardelli, Pedro H. J.. (2022). "Optimizing a Digital Twin for Fault Diagnosis in Grid Connected Inverters - A Bayesian Approach". In: 2022 IEEE Energy Conversion Congress and Exposition (ECCE). pp. 1-6. doi: 10.1109/ECCE50734.2022.9947986.
- Park, Souneil and Mulinka, Pavol and Perino, Diego. (2022). "A Large-Scale Examination of"Socioeconomic" Fairness in Mobile Networks". In: Proceedings of the 5th ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies. pp. 248–256. doi: 10.1145/3530190.3534809. url: https://doi.org/10.1145/3530190.3534809.
- Mulinka, Pavol and Kalalas, Charalampos and Dzaferagic, Merim and Macaluso, Irene and Rojas, Daniel Gutierrez and Nardelli, Pedro Juliano and Marchetti, Nicola. (2021). "Information processing and data visualization in networked industrial systems". In: 2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC). pp. 1–6.
- Casas, Pedro and Mulinka, Pavol and Vanerio, Juan. (2021). "NetSEC at High-Speed: Distributed Stream Learning for Security in Big Networking Data". In: Data Science–Analytics and Applications. pp. 97–104.
- Wassermann, Sarah and Cuvelier, Thibaut and Mulinka, Pavol and Casas, Pedro. (2020). "Adaptive and Reinforcement Learning Approaches for Online Network Monitoring and Analysis". In: IEEE Transactions on Network and Service Management.
- Mulinka, Pavol and Casas, Pedro and Fukuda, Kensuke and Kencl, Lukas. (2020). "HUMAN - Hierarchical Clustering for Unsupervised Anomaly Detection & Interpretation". In: 11th international Conferece on Network of the Future.

- Mulinka, Pavol and Fukuda, Kensuke and Casas, Pedro and Kencl, Lukas. (2020). "WhatsThat? On the Usage of Hierarchical Clustering for Unsupervised Detection & Interpretation of Network Attacks". In: The 5th International Workshop on Traffic Measurements for Cybersecurity.
- Casas, Pedro and Mulinka, Pavol and Vanerio, Juan. (2019). "Should I (re)Learn or Should I Go(on)? Stream Machine Learning for Adaptive Defense against Network Attacks". In: The 6th ACM Workshop on Moving Target Defense (MTD 2019).
- Mulinka, Pavol and Casas, Pedro and Vanerio, Juan. (2019). "Continuous and Adaptive Learning over Big Streaming Data for Network Security". In: IEEE International Conference on Cloud Networking CLOUDNET, 2019 International Conference on.
- Wassermann, Sarah and Cuvelier, Thibaut and Mulinka, Pavol and Casas, Pedro. (2019). "ADAM & RAL: Adaptive Memory Learning and Reinforcement Active Learning for Network Monitoring". In: 15th International Conference on Network and Service Management.
- Mulinka, Pavol and Wassermann, Sarah and Marín, Gonzalo and Casas, Pedro. (2018). "Remember the Good, Forget the Bad, do it Fast: Continuous Learning over Streaming Data". In: @NeurIPS 2018 Workshops, Workshop on Continual Learning.
- Mulinka, Pavol and Casas, Pedro and Kencl, Lukas. (2018). "Hi-Clust: Unsupervised Analysis of Cloud Latency Measurements through Hierarchical Clustering". In: IEEE International Conference on Cloud Networking CLOUDNET, 2018 International Conference on.
- Mulinka, Pavol and Casas, Pedro. (2018). "Stream-based Machine Learning for Network Security and Anomaly Detection". In: Proc. of the Workshop on Big Data Analytics and ML for Data Comm. Net., Big-DAMA@SIGCOMM.
- Mulinka, Pavol and Casas, Pedro. (2018). "Adaptive Network Security through Stream Machine Learning". In: Proceedings of the ACM SIGCOMM '18 Posters and Demos.
- Tomanek, Ondrej and Mulinka, Pavol and Kencl, Lukas. (2016). "Multidimensional cloud latency monitoring and evaluation". In: Computer Networks. pp. 104–120.
- Mulinka, Pavol and Kencl, Lukas. (2015). "Learning from Cloud latency measurements". In: Communication Workshop (ICCW), 2015 IEEE International Conference on. pp. 1895–1901.
- Kacur, Juraj and Vargic, Radoslav and Mulinka, Pavol. (2011). "Speaker identification by K-Nearest Neighbors: Application of PCA and LDA prior to KNN". In: Systems, Signals and Image Processing (IWSSIP), 2011 18th International Conference on. pp. 1–4.

Interests

climbing, bouldering, motorcycles, ceramics, pottery, hiking, nature