

# Krystian Safjan

DATA SCIENTIST · TECHNICAL LEADER · RESEARCHER

Wrocław, Poland

✉ ksafjan@gmail.com | 🌐 www.safjan.com | 📱 izikeros | 🌐 krystiansafjan

## Summary

I'm a Data Scientist with 5 years of experience in Machine Learning and 14 years of experience in research. I'm solving real business problems using Machine Learning techniques from the fields of NLP and time series prediction. In the past, while developing 4G and 5G networks, I have mastered ability to understand, model and simulate complex systems. I lead the work in collaborative, EU funded projects. Python is my programming language of choice.

## Experience

### Data Scientist & Technical Leader

NOKIA

Feb. 2019 - present

- Created system for 4G/5G Base Station log classification.
- Retrieving information from unstructured, inconsistent data.
- Working with human or machine produced text data using different approaches like: sentence clustering, text generalization, time series cross-validation, classification.
- Managing experiments, results and models in frequent iterations.
- Presentations of project outcome and plans tailored to audience.
- Quality assurance including Data Science specific aspects like notebook testing.

### System AI Architect for Cloud Base Stations

NOKIA

Sep. 2018 - Feb. 2019

- Designing AI architecture for hybrid distributed and Cloud mobile networks. Covering areas such as data storage and flows, distributed: training, models storage and updates. Coordination of work on Nokia Cloud RAN BTS Reference Architecture specification.

### Data Scientist, Web-app developer, Entrepreneur

LADATA

Apr. 2017 - Aug. 2018

- Cleaned the textual data, extracted features, selected best-performing algorithms for the NLP-based document classification service.
- Prototype of e2e ML text-classification pipeline for a client Wilabs.
- Design and implementation of web applications (Django, Flask); Python programming and testing.
- DevOps including Docker containers and their orchestration, CI (with GitLab) and QA (pytest, tox, linters).

### 5G Senior Research Engineer

NOKIA BELL LABS

2015 - 2017

- Developing mm-wave system reliability ML based enhancements based on historical data. Predict and mitigate blockages and interference.
- Leading work package in EU funded project mmMAGIC by driving technical work on architectural topics and coordinating work of 10 partners from industry, operators and academia.

### Radio Research Engineer

NOKIA SIEMENS NETWORKS

2007 - 2015

- Development, evaluation of concepts by modeling and simulations in MATLAB for interference mitigation in LTE networks.
- Supporting standardization delegates by providing results and explanations of performed studies.

### Software and System Architect

SIEMENS

2006 - 2007

- Research and development of LTE system level simulator, modeling of radio channel access methods, designing schedulers.
- Software quality assurance, work with VCS, technical documentation development, simulator performance optimization, software testing.

## Data Science Experience

SYSTEM LOG CLASSIFICATION

2019 - 2021

- Choosing efficient representation for log data ("Bag of Lines", sequence of lines).
- Building project environment for effective collaboration, experiment and model management (using JupyterText, GitLab, MLflow).
- Ensuring model reproducibility; model packaging, monitoring and retraining, handling data drift (MLOps).
- Designing models using LSTM elements and word embeddings (with TensorFlow and FastText).

DOCUMENT CLASSIFICATION - NLP SERVICE

2017

- Cleaning data, extracting features selecting best algorithms for NLP-based assessment of document creativity. Customized polish stemmer.
- Prototype of e2e ML pipeline for the client.

- Face elements localization (eyes, nose, mouth), face detection in the image, face classification. Using synthetic faces for sensitivity evaluation.

## SIDE PROJECTS

2003 - 2005

- Machine translation using sequence-to-sequence model. More of side projects are available on github: <https://github.com/izikeros>.

## KAGGLE COMPETITIONS

2016 - present

- Participation in Kaggle community, keeping eye on the trends, learning from the best Kagglers to design own solutions and compete with others e.g. recipe to cuisine mapping.

## Leadership Experience

---

### Technical Lead of ML project

NOKIA

Feb. 2019 - present

- Technical lead of 5-person project team working with OKR methodology.
- Setting research and experimenting directions, experiment and task prioritization.
- Guiding less experienced team members and students.
- Challenging methods and results to build solutions with confidence of being correct.

### Work package lead, Project Management Team Member in EU funded project

mmMAGIC, WINNER+ PROJECTS

2016 - 2017

- Leading "RAN Functions, Architecture, and Integration" work package, member of Project Management Team in mmMAGIC EU project.
- **Gained Experience:** Planning the utilization of research resources. Prompt plan revisions due to dynamic situations (e.g. partners leaving the project). Coordination of multi-company work on a public deliverable documents and journal paper. Effort estimation and workload balancing and alignment. Communicating with other task- and work package-leaders to harmonize the effort. Task delegation to task leaders.

## Education

---

### Wrocław University of Technology, Digital Signal Processing

M.SC. IN TELECOMMUNICATIONS

Oct. 1999 - Jul. 2004

- Thesis in the field of Computer Vision: "Face image based person recognition".

### University of Wrocław, Communication

M.SC. IN SOCIAL SCIENCE

Oct. 1999 - Jul. 2004

- Thesis in fields of Communication and Data Mining: "Characteristics of Internet Mediated relation establishment - study of dating service".

## Other Skills

---

- Teaching/mentoring experience: Nokia Lecturer at Wrocław University of Technology (WUT), teaching Assistant at Faculty of Electronics WUT, serving as adviser for Masters students.
- Able to look at things from various perspectives, ability to see the bigger picture.
- Eager to explore unknown.
- Effective communication skills.
- Empathy and self-awareness.

## Publications, Patents, Certifications

---

- 11 Conference papers, 4+ IPRs, Chapters in 3 books, 1 Journal Paper - details on google scholar.
- Able to look at things from various perspectives, ability to see the bigger picture.
- "Neural Networks and Deep Learning" certificate from deeplearning.ai (2TMZFJMHKZF).
- Passed Microsoft Data engineering exam DP-201: "Designing an Azure Data Solution"

## Online Writings

---

- "How to organize Data Science project based on Jupyter notebook",
- "Data Science Command Line Tools" - on data processing from Linux command line,
- "Finding the spy - post on Markov Chains and stochastic matrices",
- "Kaggle evaluation metrics used for regression problems",
- and more on [safjan.com](https://safjan.com)