Andrej Janchevski *ML Researcher & Data Scientist*

✉ janchevski.andrej@gmail.com   📞 +41783280200   📍 Avenue de France 9, 1004 Lausanne, Switzerland

📅 19.12.1996   in linkedin.com/in/andrej-janchevski   ⬡ github.com/Bani57   📄 Swiss B Permit

👤 Profile

Experienced AI researcher with a strong background in graph machine learning, NLP, deep learning, and data science, currently pursuing a PhD, with research on the topic of scalable knowledge graph reasoning and generation. Skilled in developing and deploying machine learning models to solve complex data challenges, with over 5 years of experience in Python, PyTorch, and Pandas. Data consulting and ML/AI engineering experience in delivering AI-driven solutions to improve data infrastructure, enhance decision-making, and drive business value.

💼 Professional Experience

EPFL - Swiss Federal Technology Institute of Lausanne, *Machine Learning Researcher*   11/2021 – present
- 4 years in a full-time PhD research on-site position in a machine learning research lab   Lausanne, Switzerland
- Focused on applied AI research on improving the scalability of knowledge graph reasoning and graph generation, as well as general GenAI and deep learning, NLP & LLMs, optimization algorithms
  - Model and graph-coarsening-based approach for quadratically accelerating knowledge graph link prediction and conjunctive query answering, with an application in anomaly detection and insight mining for large knowledge bases, research funded by Swisscom AG
  - Scalable techniques for graph generation with GANs and diffusion models, with applications for drug discovery: novel single-noise-level diffusion sampling methods and hierarchical graph neural networks for generation
  - Data entry and validation of GPT models' solutions to mathematical problems
  - Supervision of 2 master thesis projects on knowledge graph abstraction and 3 student projects on solving combinatorial optimization problems with graph neural networks
- Daily full-stack ML/AI development
  - Data analysis with numpy, scipy, pandas
  - Neural network design and experimentation with PyTorch, PyG, Lightning, W&B
  - Bash-based deployment to an AI-dedicated cluster with 8 Nvidia A100 GPUs
  - Result visualization and presentation with pandas, seaborn, TeX, PowerPoint
- Research presentation at scientific conferences, company meetings, seminars, reading groups
- PhD thesis defense by the end of October 2025

Data Masters, *Data Consultant*   03/2024 – 08/2024
- 6.5 months of part-time remote work as a data consultant   Skopje, Macedonia
- Consultancy project for retail company: market basket analysis, product segmentation & customer profiling
  - Employing pandas for market data preprocessing and analysis
  - Association rule mining for product recommendation
  - Basket, customer, and product feature design and extraction
  - Customer & product clustering and feature regression with scikit-learn
  - Data visualization with seaborn

Swisscom AG, *Machine Learning & Data Analysis Intern*   02/2021 – 08/2021
- 6-month internship at the Swisscom Digital Lab, with master thesis and project presentation as deliverables   Lausanne, Switzerland
- Telco-network database analysis, modelling & prediction
  - Query design and execution on a large Neo4j NoSQL database, migrating the datasets to a Hadoop cluster
  - Network-science-based data analysis with networkx and iGraph
  - Data modelling as a large knowledge graph and learning knowledge graph embeddings to perform detection of anomalous links and answer queries with a classification model
  - Development of a new approach that yielded >700x query answering speed-up for >95% preserved accuracy

CERN, *Software Engineer*   06/2018 – 08/2018
- 2-month internship at the ISOLDE group via the Summer Student Programme   Meyrin, Switzerland
- Web app visualization of nuclear physics experiment data
  - Back-end software development with PHP for modernizing existing collider operator call log and access requests web apps, and developing a new dashboard app that will provide useful data visualizations for operators in the TI island at the CCC
  - Front-end work consisted primarily of moving from the Angular to the Vue framework and using modern Node packages and JavaScript libraries for CSS styling, data formatting, response and request management

🎓 Education

Data Science MSc, *EPFL - Swiss Federal Technology Institute of Lausanne*   09/2019 – 08/2021
- Specialization of the computer science degree to data analysis, AI, and machine learning, achieving a GPA of   Lausanne, Switzerland
  5.69 / 6.00
- Courses in machine learning, data analysis, statistics, time series, deep learning, optimization theory, intelligent agents, Python development

Computer Science & Engineering BSc, *Ss. Cyril and Methodius University*   09/2015 – 06/2019
- Backbone courses on computer science, software engineering and AI, achieving a perfect GPA of 10.00/10.00   Skopje, Macedonia
- C, C++, Java, Python, PHP, Laravel development

🧠 Skills

Programming
- Python, SQL, Golang, PHP, Java, C, C++

Software
- PyTorch, Pandas, Neo4j, Git, Kubernetes, Docker, Hadoop, AWS

Areas
- Knowledge Graphs, Deep Learning, Generative Modeling, Drug Discovery, Network Science, NLP, Data Science, Data Visualization, Statistics, Algorithms, Software Engineering, Business Intelligence (BI), Databases, Automation, MLOps, LLMs

Soft
- Research, Presentation, Teaching, Attention to Detail, Analytical Reasoning, Communication, Problem Solving, Enthusiasm, Willingness to Learn, Collaboration, Motivation, Teamwork

🌐 Languages

English ● ● ● ● ●	French ● ● ○ ○ ○
German ● ○ ○ ○ ○	Macedonian ● ● ● ● ●

📁 Projects

dnmf-python 🔗
- First-ever implementation of the state-of-the-art DNMF overlapping community detection algorithm

Assessment of Changing Collaborative Patterns Over Time 🔗
- Audio feature clustering to detect collaboration patterns between speakers, >90% accuracy in detecting collaboration state transitions with LSTM models

Tweet Classification - AIcrowd Challenge 🔗
- Competition in detection of emojis in tweets, ranked 2nd out of all teams by employing NLP features and language models

How To Ride The Shopping Cart 🔗
- Data story on retail analytics

📖 Publications

Improving Single Noise Level Denoising Samplers with Restricted Gaussian Oracles, 06/03/2025
ICLR 2025 Workshop on Deep Generative Model in Machine Learning: Theory, Principle and Efficacy 🔗

Could ChatGPT get an engineering degree? Evaluating higher education vulnerability to AI assistants, 26/11/2024
Proceedings of the National Academy of Sciences 🔗

A Study of Different Models for Subreddit Recommendation Based on User-Community Interaction, 14/10/2019
Springer, Cham 🔗

AndrejJan at SemEval-2019 Task 7: A Fusion Approach for Exploring the Key Factors pertaining to Rumour Analysis, *Association for Computational Linguistics* 🔗 06/2019

⚒ References

Andreas Loukas, *Principal Scientist & Tech Lead,* Isomorphic Labs
loukasan@gmail.com

Claudiu Musat, *Research Manager,* Google DeepMind
claudiu.musat@gmail.com

Martina Naumovska Necheska, *Partner & Project Manager,* Data Masters
martina.naumovska@datamasters.co

Volkan Cevher, *Associate Professor,* EPFL-STI-IEM-LIONS
volkan.cevher@epfl.ch