ANDRA IONESCU

+31629348171 | andradenis.ionescu@gmail.com | Den Haag, The Netherlands

https://andraionescu.github.io | https://www.linkedin.com/in/andra-ionescu-phd/

RELEVANT WORK EXPERIENCE

TU Delft, The Netherlands *February 2020 – May 2024*
PhD Researcher

- Assessed data management issues over 4 years, engineered methods that streamlined processes for data scientists and increased data handling efficiency by 25%.
- Demonstrated sound judgement in selecting research directions by identifying gaps in state-of-the-art research and proposing 3 novel research methods, leading to 7 published research papers at A and A* conferences.
- Researched and designed an open-source feature discovery tool and a tabular data augmentation service, enhancing dataset quality and machine learning model accuracy, which advanced data processing by 15%.
- Led 2 qualitative research projects by conducting 24 semi-structured interviews of 45-60 minutes each, gathering and analysing large datasets, and deriving statistics to assess the needs and requirements for a data science pipeline.
- Produced 10+ high-quality research outputs and presentations by quickly mastering tools for data analysis (Atlasti), and design tools for mock-ups (Figma), for posters and other visualisations (Adobe InDesign), which elevated project communication.
- Collaborated with 35 experts from diverse domains to address dataset discovery and augmentation challenges, co-developing services that enhanced data usage and discovery within a data marketplace platform.
- Mentored 9 bachelor's and 2 master's students through thesis research and development, providing technical guidance in data analysis, feature engineering and ML model development.

`Python (pandas, numpy, scikit-learn, matplotlib, seaborn) / Neo4j, MySQL, DuckDB / Docker, Atlas.ti / Adobe (Illustrator, InDesign), PowerPoint, Figma`

ING Nederland, The Netherlands *October 2019 – January 2020*
Research Intern

- Led a research initiative to improve dataset discovery in the financial enterprise setting, analysed large-scale production data to identify trends, leading to data-driven recommendations that increased search and discovery times by 15% across 2 teams.
- Developed and deployed state-of-the-art schema matching algorithms using 3 ML models and 3 research studies, improving data integration and increasing matching accuracy by 33% in the enterprise platform.
- Collaborated with 2 cross-functional teams to integrate machine learning insights into product development, aligning with business goals.

`Python (pandas, numpy, scikit-learn) / Relational Tabular Data`

Maistering BV, The Netherlands *October 2018 – September 2019*
Software Engineer

- Partnered with the UI/UX team to design and iterate wireframes for a web application, integrating research insights and stakeholder feedback. This collaboration improved usability and functionality, resulting in a more intuitive user experience.
- Collaborated with engineers to translate wireframes into a functional UI, ensuring accurate implementation and backend integration. Enhanced design refinements and technical solutions, delivering the interface on schedule and meeting design goals.
- Mentored 5 junior developers with tailored guidance and clear explanations. Fostered a feedback-rich environment, accelerating learning and boosting team productivity and project outcomes.

`Python (Flask) / TypeScript / Angular 7 / d3 / vis.js`

Onior Group BV, The Netherlands *September 2017 – September 2018*
Java Developer

- Optimized AI deployment pipelines using OCR and machine learning for data extraction, improving system accuracy by 35% and reducing processing time by 20%. Integrated design, test, and analysis, stakeholder feedback, and expert insights to ensure reliability and efficiency.
- Identified the need for improved quality control, led the adoption of automated code checks in a 10-member R&D team and boosted performance metrics by reducing technical debt by 25%, thus improving code quality.

`Java EE8 / Oracle SQL`

EDUCATION

Delft University of Technology, The Netherlands *February 2020 – February 2024*
Doctor of Philosophy (PhD), Computer Science

- Thesis: *Feature Discovery in Data-Centric AI*

Delft University of Technology, The Netherlands *September 2017 – February 2020*
Master's Degree in Computer Science (MSc), Data Science Track

- Thesis: *Reproducing State-of-the-Art Schema Matching Algorithms*

University "Politehnica" of Bucharest, Romania *October 2013 – July 2017*
Bachelor's Degree in Computer Science (BSc)

- Thesis (in Romanian): *Front-end Infrastructure for Smart Cities IoT projects*

KEY TECH SKILLS

Programming Language: `Python, Java, Javascript, SQL`

Data Analysis: `Python (pandas, polars, NumPy, Scikit-Learn), matplotlib, seaborn`

Machine Learning: `AutoML, PyTorch`, Supervised Learning, Neural Networks, Decision Trees (XGBoost, Random Forest), SVM, Ensemble Methods, Feature Selection, NLP

Databases: `Neo4j, MongoDB, PostgreSQL, DuckDB`

Other: `Docker, Elasticsearch, Git/GitHub`

PUBLICATIONS

[1] Andra Ionescu, et al. "Human-in-the-Loop Feature Discovery for Tabular Data". *Conference on Information and Knowledge Management*. 2024

[2] Andra Ionescu, et al. "Key Insights from a Feature Discovery Use-Case Study". *Human-in-the-Loop Data Analytics in conjunction with the International Conference of Management of Data*. 2024

[3] Andra Ionescu, et al. "AutoFeat: Transitive Feature Discovery over Join Paths". *International Conference on Data Engineering*. 2024

[4] Andra Ionescu, et al. "Topio: an Open-Source Web Platform for Trading Geospatial Data". *International Conference on Web Engineering*. 2023

[5] **Best Demonstration Award** – Andra Ionescu, et al. "Topio Marketplace: Search and Discovery of Geospatial Data". *International Conference on Extending Database Technology*. 2023

[6] Andra Ionescu, et al. "Join Path-Based Data Augmentation for Decision Trees." *Databases for Machine Learning Workshop in conjunction with International Conference on Data Engineering*. 2022