Warsaw, Poland
pytelrr@gmail.com

Rafal Pytel

https://www.linkedin.com/in/rpytel1/
https://github.com/rpytel1

I'm a driven AI Engineer with experience across agriculture, medtech, fintech, and e-commerce, applying a range of ML techniques to real-world problems. While I focus on ML and GenAI, I bring a strong backend and software architecture background, and thrive in product-oriented, client-facing environments where pragmatic solutions matter.

Work Experience (6+ years)

Senior AI Engineer (Full Time) **Zowie** **Jan 2025–Present**
Product team Remote, Poland

- Reduced client onboarding data requirements by 95% by implementing a GenAI-based intent routing system.
- Prevented nearly all rollbacks in prompt-based services by architecting and deploying a product-wide offline evaluation framework.

ML Lead, Senior ML Engineer (Full Time) **ReasonField Lab** **2022 - Present**
Remote, Poland

- Directed the development of an open-source Explainable AI library in Computer Vision, leading a team of 4 ML Engineers.
- Directed a 3-person team over 5 months to build a RAG-based internal chatbot with access to Scala documentation.
- Increased ML & GenAI-related sales by 50% by designing and executing a cross-functional strategy covering marketing, technical positioning, and sales team enablement.
- Delivered insights across agriculture, e-commerce, medtech, logistics, customer service, and fintech sectors through PoC projects using CNNs, GNNs, and LLMs.
- Shared learnings publicly through 15+ blog posts and talks at 4 national conferences (e.g., Data Science Summit, Warsaw IT Days).

ML Full Stack Engineer (Part Time) **Famateq/QEF Electronics** **2022 – 2024**
Remote, Poland

- Scaled vegetable sorting software (C++, Python, Vue.js) from a single machine to multi-unit deployments with minimal HPC infrastructure and cost.
- Enhanced ML observability in production by integrating GradCAM into live traffic with minimal processing overhead.
- ~15h/week alongside the full-time job at ReasonField Lab.

Senior ML Engineer (Full Time) **QEF Electronics** **2020–2022**
Utrecht, Netherlands

- Leading research & development of the Computer Vision side of projects in vegetable counting solutions.
- Decreased error rate of the model by 15% via R&D of model architecture(e.g. ConvNext, ResNets) and training methods (500+ models trained).
- Implemented a comprehensive TensorRT optimisation pipeline, which increased throughput by 100% for batched variable-size input.
- Improved data pipeline to decrease training time by 40%.
- Implemented various model interpretation tools(e.g. UMAP, GradCam) to understand errors in training data.
- Moved production deployment from custom code to a professionally configured Triton edge server.
- Introduced a 25% improvement in the throughput of images by implementing a custom shutter mechanism in C++ for GeniCAM cameras.

Deep Learning Fellow (Part Time) **Fellowship.ai** **Spring 2020**
Remote, Netherlands

- Implemented a self-supervised algorithm called FixMatch as an E2E service in AWS for easy and quick use for ongoing Computer Vision projects.
- ~20h/week alongside MSc studies.

Software Developer (Part Time) **Efinity** **2016-2018**
Warsaw, Poland

- Implemented various insurance offerings for clients across the World.
- ~24h/week alongside BSc studies.

Education and Certifications

- **M.Sc. Computer Science,** Delft University of Technology, Netherlands. **2018–2020**
- **B.Sc. Teleinformatics,** Warsaw University of Technology, Poland. **2014–2018**

Technologies and Languages

- Languages: Python, C++, JavaScript
- Frameworks: HuggingFace, PyTorch, Tensorflow, Ray, ZenML
- Technologies: MySQL, Postgres, Mongo, AWS, GCP, Docker

Certificates

- AWS Solution Architect Associate
- AWS ML Speciality
- HF Agents Certificate
- Tensorflow Advanced Techniques

Publications and Conference Presentations

- ICPR 2020: Tilting at windmills: Data augmentation for deep pose estimation does not help with occlusions
- Warsaw IT Days 2023: Importance and Role of ExplainableAI in developing deep neural networks
- DSS 2023 ML: What if we used tricks from Recommendation Systems to Face Recognition?
- DSS 2023: Interchangeability of Why - Insight into subjectivity of XAI for Computer Vision
- DSS 2024: Watermarking 101 - how you can detect generated data

Projects

- **Secure Agent:** Built a simple agentic workflow integrating 4 tools (web search, RAG, DB search, file editing), then systematically attacked it using techniques like prompt poisoning, injection, and misuse. Evaluated state-of-the-art defences against successful attacks, highlighting their practical strengths and limitations.
- **Prostate Cancer Metastasis Detection Challenge**: Trained multiple models using Multi-Image Learning and Teacher–Student paradigms, selectively masking easy examples to enhance learning efficiency. Utilised Python, PyTorch, MONAI, and Ray for scalable preprocessing, training, and optimisation.
- **Harvest Prediction System**: Built a state-of-the-art ML system using multimodal data (images, sensor feeds, weather) to predict harvest yield with <10% error on 2-day forecasts. Accelerated labelling with SAM2 and managed the pipeline using Python, PyTorch, GCP, and Label Studio.
- **Technical blogging** on Medium and the Softwaremill blog.

Interests

- Sports enthusiasts: windsurfing, skiing, basketball.
- Bookworm: non-fiction, technical books, but also huge LoTR nerd.