+420 723 089 475
me@freedatoms.eu
freedatoms.eu

Tomáš Frýda

Svatovítská 361/4, Prague 6, Czechia, 160 00

GitHub: freedatoms, tomasfryda
LinkedIn: freedatoms

Professional Summary

Curiosity-driven Senior Software & ML Engineer with a strong background in Theoretical Computer Science. Proven track record of co-authoring production-grade AI tools (H2O Sonar) and architecting core algorithms for globally recognized open-source AutoML frameworks (H2O-3). Expert in Machine Learning Explainability (XAI), advanced statistical modeling (GLM/Tweedie), and building advanced evaluation systems for Generative AI (LLMs, RAG, and Agentic architectures).

Skills

- **Languages:** Python, Java, R, Clojure, Emacs Lisp, Common Lisp, C/C++
- **Machine Learning & AI:** AutoML, Generative AI, LLM Evaluation, RAG & Agentic Systems, Explainable AI (SHAP), Statistical Modeling (GLM, Tweedie), Conformal Prediction, Probabilistic Programming (Stan, PyMC)
- **Tools & Infrastructure:** Linux, Git, Docker, K8s, MLOps, Data Warehousing, Emacs/Org-mode

Work

2020 – now Senior Software & ML Engineer @ H2O.ai

- **Generative AI & Evaluation:** Co-authored H2O Sonar (GenAI Evaluators) and architected core ML components for H2O Eval Studio, establishing rigorous frameworks to evaluate Large Language Models (LLMs), RAG architectures, and agentic systems.
- **Open-Source AutoML:** Core developer of open-source AutoML and Stacked Ensembles within H2O-3 across `Java`, `Python`, and `R`, optimizing both algorithm performance and system scalability.
- **Explainable AI (XAI):** Researched and implemented core mathematical methods for machine learning interpretability; created the explain module and integrated Marginal/Interventional SHAP algorithms (TreeSHAP, DeepSHAP, Generalized DeepSHAP) into the open-source H2O-3 engine.
- **Advanced ML Modeling:** Implemented advanced interpretable algorithms, including AdmissibleML and a complex Tweedie family estimator for GLM with simultaneous estimation of dispersion and variance power (PR 1, PR 2).
- **Ecosystem Maintenance:** Acted as the official maintainer of the H2O-3 R package, ensuring ecosystem stability and package governance.

2018 – 2020 Freelance Data Scientist & Team Lead @ Datamole & Consulting

- **Team Leadership:** Promoted to Data Science Team Lead, managing engineering priorities and client delivery.
- **Engineering & MLOps:** Architected custom optimization routines, contributed to an internal MLOps system, and co-designed enterprise data warehouses.
- **Academic Teaching:** Taught Distributed Data Mining at the Czech Technical University (FIT CTU).
- **Consulting:** Provided independent advisory services in information theory, statistics, and graph theory.

2017 – 2018 Data Scientist @ Vendavo

- **Price Optimization:** Engineered price optimization and economic models using classical statistical methods in `R`.

- **Collaboration:** Partnered with an interdisciplinary PhD-heavy team bridging advanced Econometrics and Artificial Intelligence.

Jul 2016 – Sep 2016 Research Assistant @ FIT (VýLeT)

- **Research Focus:** Researched scalable machine learning of predictive ensembles using the FAKE GAME approach on the H2O platform, which directly laid the foundation for my award-winning Master's thesis and publication in Machine Learning.

Education

2014 – 2017 Master's in Knowledge Engineering, Faculty of Information Technology, Czech Technical University (FIT CTU)

- **Thesis:** Scalability of Predictive Modeling Algorithms – Awarded the Dean's Award (Winter 2016/2017)

2011 – 2014 Bachelor's in Theoretical Computer Science, Faculty of Information Technology, Czech Technical University (FIT CTU)

- **Thesis:** Comparison of different neuroevolution algorithms on several problem domains

2010 – 2011 Mathematical Informatics, Faculty of Nuclear Sciences and Physical Engineering, CTU

- Transferred to FIT for closer alignment with Computer Science interests.

Interests & Preferences

Open Source & Technology

Linux user since 2001. Open-source contributor, including maintaining multiple packages in the Arch User Repository (AUR) and Gentoo/Funtoo overlays for custom Emacs, MKCL, and SBCL environments (now all are unfortunately unmaintained due to lack of time). Long-time LinuxDays attendee.

Algorithmic Fairness & Research

Deeply engaged in topics of algorithmic fairness, accountability, transparency, and preventing discrimination in AI. Interested in uncertainty estimation techniques (conformal prediction) and probabilistic programming languages (Stan, PyMC).

Work Preferences

Due to a busy life outside of work — which includes exploring molecular gastronomy, studying plant-based nutrition, and managing two energetic dogs (Pixel & Okurka) — I prefer **fully remote positions** or roles with a **dog-friendly office space** in Prague.

Publications

Full registry available at ORCID and Google Scholar.

- Nykodym, Tomas and Kraljevic, Tom and Wang, Amy and Wong, Wendy and Fryda, Tomas (2025). *Generalized linear modeling with h2o*, Published by H2O.ai Inc. Available here.
- Kordík, Pavel, Černý, Jan, Frýda, Tomáš. (2018). *Discovering predictive ensembles for transfer learning and meta-learning*. Machine Learning. 107. doi: 10.1007/s10994-017-5682-0.
- Kordík, Pavel, Frýda, Tomáš (2018). *On Scalability of Predictive Ensembles and Tradeoff Between Their Training Time and Accuracy*. In: Shakhovska, N., Stepashko, V. (eds) Advances in Intelligent Systems and Computing II. CSIT 2017. Advances in Intelligent Systems and Computing, vol 689. Springer, Cham. doi: 10.1007/978-3-319-70581-1_18.
- Kordík, Pavel, Frýda, Tomáš, Šnorek, Miroslav, Čepek, Miroslav. *Scalability of predictive ensembles*, 2017 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), Lviv, Ukraine, 2017, pp. 555-560, doi: 10.1109/STC-CSIT.2017.8098848.
- Frýda, Tomáš (2017). *Scalability of Predictive Modeling Algorithms*, České vysoké učení technické v Praze. Available here.
- Frýda, Tomáš (2014). *Comparison of Different Neuroevolution Algorithms on Several Problem Domains*, České vysoké učení technické v Praze. Available here.