Marek Wydmuch

Curriculum Vitae

Phone	(+48) 784 110 482
E-mail	mwydmuch@cs.put.poznan.pl
Homepage	mwydmuch.pl
GitHub	mwydmuch
LinkedIn	marekwydmuch

About

Marek Wydmuch is machine learning researcher at Snowflake and Machine Learning Laboratory in the Institute of Computing Science, Poznan University of Technology (PUT), Poland. His research interests include various topics in machine learning, from extreme classification (XC), reinforcement learning (RL), and large language models (LLMs). He published some of his works at the top machine learning conferences (NeurIPS, ICML, ICLR, KDD).

Personal information

Date and place of birth	April 13, 1993, Poznan, Poland
Citizenship	Polish

Education

Oct. 2017 – Dec. 2025 **Ph.D., Computer Science, Poznan University of Technology**
Thesis: *Addressing the long-tail problem in extreme multi-label classification*
Advisor: Krzysztof Dembczyński, Ph.D.
Granted scholarships for the top Ph.D. students and Rector's Scholarships for the best-achieving students multiple times.

Mar. 2016 – Sept. 2017 **M.Sc., Computer Science, Poznan University of Technology**
Specialization: Intelligent Decision Support Systems
Thesis: *Online probabilistic label trees for extreme multi-label classification*
Grade: very good with distinction (A+), GPA: 4.76/5.0

Oct. 2012 – Feb. 2016 **B.Sc., Computer Science, Poznan University of Technology**
Thesis: *VIZIA: 3D Video Game-base Environment for Research on Learning Agents from Raw Visual Information*
Grade: good plus (B+), GPA: 4.43/5.0

Sep. 2009 - Jun. 2012 **High School, Adam Mickiewicz Lyceum No. 8 in Poznan**

Work experience

Aug. 2024 – present **Research Scientist, Snowflake AI Research**, Poznan, Poland
Team: Cortex Analyst, previously Arctic
Full-time, worked on various aspects of Cortex Analyst system.

Oct. 2023 – present **Research and Teaching Assistant, Poznan University of Technology**, Poznan, Poland
Part-time employee at Machine Learning Laboratory in the Institute of Computing Science.

Jun. 2022 – Dec. 2022	**Research Intern, Yahoo Research**, Paris, Ile-de-France, France Team: Scalable Machine Learning Worked on the problem of conversion rate prediction under delayed feedback in the Real-Time Bidding system in Yahoo's Demand Site Platform.
Jan. 2021 – Jun. 2022	**Specialist, Poznan University of Technology**, Poznan, Poland Part-time, prepared teaching materials and mentored students as part of AI Tech M.Sc. program.
Oct. 2017 – Dec. 2020	**Data Scientist, OLX Group**, Poznan, Poland Team: Search Part-time, worked on learning to rank, ads tagging using extreme classification, category suggestion, and spellchecking for OLX Poland, Portugal, Ukraine, Romania, and Bulgaria.
Jul. 2017 – Sept. 2017	**Data Science Intern, OLX Group**, Poznan, Poland Team: Search Worked on automatic category tree building, quality assessment for category trees, and discovering new categories in classifieds services.
Jul. 2015 – Sept. 2015	**3D Tools Programming Intern, ProGrupa**, Poznan, Poland Developed an online 3D model viewer and a converter for commercial 3D vector file formats to Wavefront .obj file format for `archiup.com`. The same model viewer is still in use on the website.
2011 – 2015	**Freelance Front-End Web Developer** Developed small web apps and WordPress and Jomla-based websites.

Leadership Experience

May. 2022 – Dec. 2024	**IT Infrastructure Coordinator, ML in PL Association** Led the development of new websites for ML in PL Conference, MLSS, and other initiatives of the association. Managed a small team of 3 people.
Jan. 2022 – Dec. 2023	**Board Member, ML in PL Association** Supported other members of the association in organizing the best machine learning related events, recruited new members, and took care of some formalities.
Mar. 2019 – Dec. 2021	**Call for Contribution Coordinator, ML in PL Association** Coordinated Call for Contributions (Call for Talks and Posters) for ML in PL Conference 2019-2021. Managed a small team of 3 people that organized talks and poster sessions during the events.

Skills

Languages	English (full professional proficiency), Polish (native)
Skills	machine learning, deep learning, LLMs, AI agents, natural language processing, classification, recommendation, reinforcement learning, standard algorithms, big data processing, creating C++ Python bindings, CI/CD for Python and C++, visualization and 3D graphics, front-end web development, writing scientific articles, teaching
Programming languages	**Experienced: Python, C++, C, JavaScript, SQL, Bash** Basic: Julia, Lua, PHP, Java
Technologies	**Experienced: PyTorch, PyTorch Lightning, Numba, HuggingFace Transformers, vLLM, Scikit-learn, PySpark, Pandas, pybind11, Git,**

LaTeX, Snowflake
Basic: JAX, DeepSpeed, Docker, Kubernetes, AWS, Hive, Presto, Hadoop, MLflow, Airflow, Flask, aiohttp

Operating systems Linux, macOS, Windows

Other scientific activities

Software maintainer

ViZDoom – Doom-based AI Research Platform for Reinforcement Learning from Raw Visual Information
Creator and main maintainer of the project, which has over 3000 downloads from PyPI every month and a member of Farama Foundation, which maintains the largest collection of open-source reinforcement learning tools.
Website: `vizdoom.farama.org`
Farama Foundation: `farama.org`

Organizer

Helped organize the following scientific events:

ML in PL Conference 2019 – 2024, Warsaw,Poland
Member of the Organization Committee
Website: `conference.mlinpl.org`

MLSS^{S} 2023 – Machine Learning Summer School on Applications in Science 2023, Krakow, Poland
Member of the Organization Committee
Website: `mlss2023.mlinpl.org`

MLSS^{D} 2025 – Machine Learning Summer School on Drug and Materials Discovery 2025, Krakow, Poland
Member of the Organization Committee
Website: `mlss2025.mlinpl.org`

Visual Doom AI Competition 2016, 2017 and 2018 at IEEE Conference on Computational Intelligence and Games
Main Organizer
Website: `vizdoom.cs.put.edu.pl`

From Multiple Criteria Decision Aid to Preference Learning (DA2PL) 2018 Conference, Poznan, Poland
Member of the Organization Committee
Website: `da2pl.cs.put.poznan.pl`

Polish Agreement on Development of Artificial Intelligence (PP-RAI) 2018 Confernece, Poznan, Poland
Member of the Organization Committee
Website: `pp-rai.cs.put.poznan.pl`

Teaching

As Ph.D. student, specialist, and teaching assistant at Poznan University of Technology, taught the following courses:

Elements of Convex Optimization
Responsible for classes during the summer semester of 2024 and 2025.

Decision Support Systems
Responsible for classes during the summer semester of 2024 and 2025.

Introduction to Machine Learning and Neural Networks
Responsible for classes during the winter semester of 2024/2025 and 2025/2026.

Systems That Learn
Responsible for classes during the summer semester of 2024.

Advanced Methods of Computational Intelligence
Responsible for classes during the summer semesters of 2021, 2022, and 2023 and lectures during the summer semesters 2022, and 2023.

Methods of Artificial and Computational Intelligence
Responsible for classes during the summer semester of 2021.

Big Data Processing
Responsible for classes during the winter semester of 2019/2020.

Processing of Massive Datasets
Responsible for classes during the winter semester of 2018/2019.

Mining of Massive Datasets
Responsible for classes during the winter semester of 2018/2019.

Information Theory and Lossless Compression Methods
Responsible for classes during the summer semester of 2018, the course was awarded as the best new course of 2018 at the Institute of Computer Science.

Reviewer

Served as a reviewer for:

Conferences: NeurIPS 2020, **2021 (Outstanding Reviewer Award – top 8%)**, **2022 (Top Reviewer Award – top 8%)**, 2023, 2024, and 2025, ICML 2021, 2022, 2023, 2024, and 2025, ICLR 2022, 2023, and 2024, AISTATS 2021, IJCAI 2020, and 2021.

Journals: Machine Learning Journal (2 times), IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) (2 times), and Transactions on Machine Learning Research (TMLR) (3 times).

Summer schools participant

Attended the following summer schools:

Eastern European Machine Learning Summer School 2020: Deep Learning and Reinforcement Learning

MLSS^{S} 2023 — Machine Learning Summer School on Applications in Science 2023

Publications

Number of citations 1400+ (Google Scholar)

H-index 9 (Google Scholar)

1. Marek Wydmuch, Łukasz Borchmann, and Filip Graliński. Tackling prediction tasks in relational databases with LLMs. *arXiv preprint arXiv:2411.11829*, 2024.
2. Maciej Beręsewicz, Marek Wydmuch, Herman Cherniaiev, and Robert Pater. Multilingual hierarchical classification of job advertisements for job vacancy statistics. *arXiv preprint arXiv:2411.03779*, 2024.
3. Wojciech Kotlowski, Marek Wydmuch, Erik Schultheis, Rohit Babbar, and Krzysztof Dembczynski. A general online algorithm for optimizing complex performance metrics. In *Proceedings of the 41st International Conference on Machine Learning*, (**ICML '24**). PMLR, 2024.
4. Erik Schultheis, Wojciech Kotłowski, Marek Wydmuch, Rohit Babbar, Strom Borman, and Krzysztof Dembczynski. Consistent algorithms for multi-label classification with macro-at-k metrics. In *The Twelfth International Conference on Learning Representations*, (**ICLR '24**), 2024.

5. Erik Schultheis, Marek Wydmuch, Wojciech Kotłowski, Rohit Babbar, and Krzysztof Dembczynski. Generalized test utilities for long-tail performance in extreme multi-label classification. In *Advances in Neural Information Processing Systems*, volume 36 (**NeurIPS '23**). Curran Associates, Inc., 2023.
6. Erik Schultheis, Marek Wydmuch, Rohit Babbar, and Krzysztof Dembczynski. On missing labels, long-tails and propensities in extreme multi-label classification. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, (**KDD '22**), New York, NY, USA, 2022. Association for Computing Machinery.
7. Marek Wydmuch, Kalina Jasinska-Kobus, Rohit Babbar, and Krzysztof Dembczynski. Propensity-scored probabilistic label trees. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, (**SIGIR '21**), New York, NY, USA, 2021. Association for Computing Machinery.
8. Kalina Jasinska-Kobus, Marek Wydmuch, Devanathan Thiruvenkatachari, and Krzysztof Dembczynski. Online probabilistic label trees. In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 (AISTATS '21). PMLR, 2021.
9. Thomas Mortier, Marek Wydmuch, Eyke Hüllermeier, Krzysztof Dembczynski, and Willem Waegeman. Efficient algorithms for set-valued prediction in multi-class classification. *Data Mining and Knowledge Discovery*, 2021.
10. Kalina Jasinska-Kobus, Marek Wydmuch, Krzysztof Dembczynski, Mikhail Kuznetsov, and Robert Busa-Fekete. Probabilistic label trees for extreme multi-label classification. *arXiv preprint arXiv:2009.11218*, 2020.
11. Marek Wydmuch, Kalina Jasinska, Mikhail Kuznetsov, Róbert Busa-Fekete, and Krzysztof Dembczynski. A no-regret generalization of hierarchical softmax to extreme multi-label classification. In *Advances in Neural Information Processing Systems*, volume 31 (**NeurIPS '18**). Curran Associates, Inc., 2018.
12. Marek Wydmuch, Michał Kempka, and Wojciech Jaśkowski. ViZDoom Competitions: Playing Doom from Pixels. *IEEE Transactions on Games*, 11(3), 2019. **The 2022 IEEE Transactions on Games Outstanding Paper Award**.
13. Michał Kempka, Marek Wydmuch, Grzegorz Runc, Jakub Toczek, and Wojciech Jaśkowski. ViZDoom: A Doom-based AI research platform for visual reinforcement learning. In *IEEE Conference on Computational Intelligence and Games*, Santorini, Greece, 2016. IEEE. **The Best Paper Award**.

Awards

Sept. 2021	The 2022 IEEE Transaction on Games Outstanding Paper Award.
Mar. 2017	1st place in the machine learning "The Let's Roq Challenge" competition.
Mar. 2017	2nd place in the machine learning OLX Data Ninja 2017 competition.
Sept. 2016	The Best Paper Award at IEEE Conference on Computational Intelligence and Games, 2016.

Interests

Hobbies	board games, escape rooms, movies, programming, graphic design, computer graphics
Sports	hiking, bouldering

Poznan, December 14, 2025