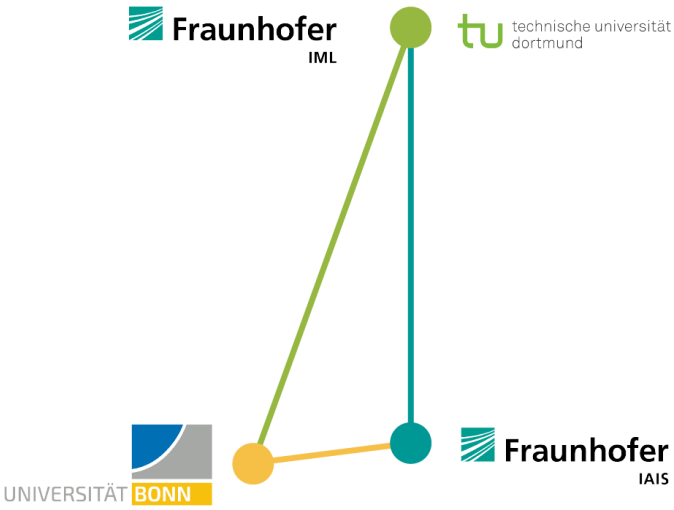
ML2R
Kompetenzzentrum
Maschinelles Lernen
Rhein-Ruhr

Report
08/2020

A summary report on the activities of the Competence Center Machine Learning Rhine-Ruhr (ML2R)

The ML2R connects pioneer research institutions in Germany:



Content Summary

Content

ML2R informs

Preface

Machine Learning plays a central role in the field of Artificial Intelligence (AI). We may consider Alan Turing the founder of AI and he stated early on that computer should be rather taught than programmed. In his famous paper “Computing Machinery and Intelligence” he wrote in 1950 that a teacher of a machine “…will often be very largely ignorant of quite what is going on inside, although he may still be able to some extent to predict his pupil's behavior. … Most of the programmes which we can put into the machine will result in its doing something that we cannot make sense (if at all), or which we regard as completely random behaviour.”[1] This early statement touches the issue of trustworthiness, which is now driving research and development of AI. We study Machine Learning in order to understand its models. We prove (statistical) bounds of correctness, fairness, robustness, and resource consumption of the algorithms and their implementations. This is still challenging because it moves beyond the abstract level of statistical theory, investigating the level of algorithms and, even further, the level of their implementations on diverse computing architectures. Hence, there are experimental methods to the same goal of trustworthiness: benchmark testing and interactive inspection or explanations of learned models. ML2R focuses on human-oriented Machine Learning, works on guarantees for classes of learning models, e.g., probabilistic graphical models, and investigates the theoretical and empirical bounds of resource consumption on diverse hardware, even on quantum computers.

As a competence center, we are commited to the upbuilding of Machine Learning in research, education, and applications. We invite colleagues to a research stay with us – in the year of the Covid-19 pandemia, this failed much to our regret. We offer an online summer school and hackathon and have informed several companies about their specific opportunities in Machine Learning applications. Networking with the other competence centers and funded German projects on Machine Learning as well as with the French competence centers on AI supports the rise of a Machine Learning community.

This report might arouse interest in the reader to learn more about the research, to participate in the summer school, to visit our campus as a guest researcher, or to benefit from a closer collaboration in terms of an Enterprise Innovation Campus, where joint teams of employees of companies and researchers of the competence center work on specific practical problems.

In any case – enjoy and stay healthy!

Prof. Dr. Katharina Morik
Speaker ML2R
17 August 2020

Prof. Dr. Stefan Wrobel
Speaker ML2R
17 August 2020

[1] Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind, 59*(236), 433-460. https://academic.oup.com/mind/article/LIX/236/433/986238

ML2R informs

Research

THE PARTNERS

The Competence Center Machine Learning Rhine-Ruhr (ML2R) brings together four leading research institutions in the field of Machine Learning (ML): TU Dortmund University (TU Dortmund), the Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS, the Fraunhofer Institute for Material Flow and Logistics IML and the Rheinische Friedrich-Wilhelms-University of Bonn (University of Bonn). They work with the common goal of establishing top-level research, fostering young scientists and strengthening the technology transfer into companies.

TU Dortmund University (TU Dortmund). The Faculty of Computer Science at the TU Dortmund University is one of the largest computer science faculties in Germany and is among the strongest faculties in terms of research output. In addition to the Competence Center Machine Learning Rhine-Ruhr, the Faculty is also home to the DFG-funded Collaborative Research Center 876 "Providing Information by Resource-Constrained Data Analysis". The spokesperson of the ML2R, Prof. Dr. Katharina Morik, holds the professorship for Artificial Intelligence (AI) at the TU Dortmund University and is director of the Collaborative Research Center 876.

Fraunhofer Institute for Intelligent Analysis and Information Systems (Fraunhofer IAIS). As part of the largest organization for applied research in Europe, Fraunhofer IAIS is one of the leading scientific institutes in the fields of Artificial Intelligence, Machine Learning and Big Data in Germany and Europe. The institute supports companies in optimizing products, services, processes and structures as well as in developing new digital business models. The ML2R spokesperson, Prof. Dr. Stefan Wrobel, is head of the Fraunhofer IAIS Institute as well as leading professor of the Information Systems and Artificial Intelligence department at the Institute for Computer Science at the University of Bonn.

Fraunhofer Institute for Material Flow and Logistics (Fraunhofer IML). The Fraunhofer Institute for Material Flow and Logistics IML is the partner of choice for integrated logistics research. It works in all fields of internal and external logistics. The institute develops solutions to problems for immediate use in businesses and conducts initial research. The representative of the Fraunhofer IML at the ML2R, Prof. Dr. Dr. h. c. Michael ten Hompel, is the managing director of the Institute and professor for Material Flow and Logistics Management at the TU Dortmund University.

Rheinische Friedrich-Wilhelms-University of Bonn (University of Bonn). The Institute for Computer Science at the University of Bonn was one of the first university institutes for computer science in Germany. The research areas at the institute are divided into four areas: Algorithms, Graphics-Vision-Audio, Information and Communication Management as well as Intelligent Systems. The representative of the University of Bonn at ML2R, Prof. Dr. Christian Bauckhage, is leading professor of the research group Machine Learning and Artificial Intelligence.

THE TEAM

Since its launch in 2018, the ML2R has established itself as a research center with international and national visibility due to the synergies generated by the joint cooperation of the partnering institutes. Its prestige is based on the excellent research performance of its continuously growing scientific staff. The team of the competence center is currently comprised of 32 researchers. In December 2019, the German Federal Ministry of Education and Research (BMBF) granted the ML2R an **additional 8 million Euros in federal funding** until the end of 2022. This will enable the ML2R to **add up to 25 new scientific staff members** to the team.

ML2R informs

ASSOCIATE PARTNERS

Together with its associate partners, the ML2R **advances interdisciplinary** and **cross-institutional research** projects. Within these cooperations, the partners gain access to the competence center's expertise and resources, e.g. algorithms. Prof. Dr. Dr. Wolfgang Rhode from astroparticle physics at TU Dortmund University adds challenging data with deep research questions to ML2R. With his specialization in autonomous intelligent systems, Prof. Dr. Sven Behnke from the University of Bonn contributes his expertise in cognitive robotics and computer vision. Through his work in action recognition, human pose estimation as well as object detection and segmentation Prof. Dr. Jürgen Gall from the University of Bonn will further advance the ML2R's research endeavors.

RESEARCH PROFILE

The joint research efforts within the competence center are based on the research and development of **modular Machine Learning (ML) applications**, in which systems are built from individual components and linked to form complex architectures. This approach covers the special aspects of **human-oriented Machine Learning** methods, **ML under resource constraints,** and **ML with complex knowledge**. Researchers at the ML2R have been working on **trustworthy Machine Learning** methods and **Machine Learning on quantum computers**.

RESEARCH RESULTS (EXCERPT)

Explainable and Trustworthy AI

The acceptance of Machine Learning is an important basis for the future dissemination and use of ML applications in industrial production and high-risk areas of use. The results and processes of ML must be **interpretable** and **explainable**. This is now a hot topic in almost all current AI or ML conferences. Since the pioneering Ph D thesis of Dr. Stefan Rüping on "Learning interpretable models" at the TU Dortmund University in 2006[2] investigated the three dimensions of understandability, accuracy, and efficiency with contributions on the levels of the optimization of the interpretability of a learner with and without knowledge of its internals (white box and black box approach), this topic was recognized in what are now the sites of ML2R.

A recent contribution detects explainable latent factors in data that ease human understanding. Recurrent Neural Networks filter prototypes within minimum enclosing balls out of the data. The respective paper by Prof. Dr. Christian Bauckhage, Dr. Tiansi Dong, and Rafet Sifa received one of the two best paper awards at the International Conference on Artificial Neural Networks 2019.[3]

It has often been claimed that the models learned by Inductive Logic Programming (ILP) are particularly easy to understand. The recent paper on "Mining Tree Patterns with Partially Injective Homomorphisms" bridges the gap between ILP and graph mining.[4] It builds upon the success of the preceding paper on probabilistic

[2] Rüping, S. (2006). *Learning interpretable models* (Doctoral dissertation). TU Dortmund University, Dortmund, Germany. https://eldorado.tu-dortmund.de/bitstream/2003/23008/1/dissertation_rueping.pdf

[3] Bauckhage, C., Sifa, R., Dong, T. (2019). Prototypes Within Minimum Enclosing Balls. *ICANN 2019*, pp. 365-376. Cham, Switzerland: Springer. https://link.springer.com/chapter/10.1007/978-3-030-30493-5_36

[4] Schulz, T. H., Horvarth, T., Welke, P., Wrobel, S. (2018). Mining Tree Patterns with Partially Injective Homomorphisms. *Proc. ECML PKDD,* pp. 585-601. http://publica.fraunhofer.de/eprints/urn_nbn_de_0011-n-5523328.pdf

frequent subtrees for graph classification and retrieval.[5] There, a simple probabilistic technique restricts the pattern language to trees and relaxes the demand on the completeness of the mining algorithm, as well as on the correctness of the pattern matching operator by replacing transaction and query graphs with small random samples of their spanning trees. This allows to efficiently handle chemical graph data sets, even those with 42,687 compounds, of which 39,337 are connected (the NCI-HIV data set) or the ZINC data set with 8,946,757 (8,946,755 connected) unlabeled, so called 'Lead-Like' molecules from the zinc database of purchasable chemical compounds. The theoretical foundation, empirical tests, and understandable patterns allow for an effective support of science through Machine Learning.

Within the network KI.NRW, the Fraunhofer IAIS team under the guidance of Dr. Maximilian Poretschkin prepared a white paper on "Trustworthy Use of Artificial Intelligence – Priorities from a Philosophical, Ethical, Legal, and Technological Viewpoint as a Basis for Certification of Artificial Intelligence" which offers a reference framework for our further work.[6]

In ML2R, we have organized a workshop together with Behrooz Omidvar-Tehrani and Jean-Michel Renders at the Extended Database Conference 2020 with the focus on the **overall pipeline of Machine Learning**.[7] Due to the Covid-19 pandemic, the workshop itself went online, but the panel became a common paper.[8] ML2R will continue its work on optimizing overall pipelines of learning processes and collecting and structuring quality and resource consumption guarantees of learning.

Human Oriented ML

When applied in practical applications, Machine Learning systems and solutions typically interact with human users. This suggests to harness the interaction and to apply user feedback in order to continuously improve a system's capabilities. In ML2R, we therefore investigate approaches towards **interactive Machine Learning**, **learning from implicit and explicit user feedback**, as well as **learning of interpretable or explainable models** (see above).

Among others, interactive learning allows for the addressing of problems due to the scarcity of labeled training data in certain application domains. An example is the problem of relation extraction in text mining where we could show that distant supervision combined with active learning and expert feedback during runtime allows neural networks to improve relation extraction results.[9]

---

[5] Welke, P., Horváth, T., Wrobel, S. (2018). Probabilistic frequent subtrees for efficient graph classification and retrieval. *Machine Learning*, 107, 1847–1873. https://link.springer.com/content/pdf/10.1007/s10994-017-5688-7.pdf

[6] Cremers, A. et al. (2019). *Trustworthy Use of Artificial Intelligence. Priorities from a Philosophical, Ethical, Legal, and Technological Viewpoint as a Basis for Certification of Artificial Intelligence*. Sankt Augustin, Germany: Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS. https://www.iais.fraunhofer.de/content/dam/iais/KINRW/Whitepaper_Thrustworthy_AI.pdf

[7] Morik, K., Omidvar-Tehrani, B., Renders, J.-M. (2020). Explainability for Trustworthy ML Pipelines. *Proc. Workshop ETMLP*. https://europe.naverlabs.com/etmlp/

[8] Giannotti, F., Kotthaus, H., Morik, K., Piatkowski, N., Schlunder, P. (2020). Explainability for Trustworthy ML Pipelines: A discussion. B. Omidvar-Tehrani/J.-M. Renders (Mod.), *Panel of the International Workshop on Explainability for Trustworthy ML Pipelines of the ETMLP*.
https://drive.google.com/file/d/1J3Ng8c9kX_ixwHrCBPbjQoqgEx9R7b5X/view

[9] Adilova, L., Giesselbach, S., Rüping, S. (2018). *Making Efficient Use of a Domain Expert's Time in Relation Extraction.* https://arxiv.org/pdf/1807.04687.pdf

Another practical example arises in social media mining for community policing and emergency response. Here, ML2R developed a first prototype of a system that analyzes geospatial distributions and sentiments expressed in posts. Since language used on social media platforms is subject to frequent changes, the system integrates various Machine Learning techniques and, crucially, frequently queries for user feedback.[10]

Quantum Computing

ML2R's **Quantum Machine Learning (QML)** related research can be factored into two main areas: **hardware** and **algorithms**. In the QML hardware area, initial studies are conducted on an Adiabatic Quantum Computer (AQC), more precisely on NASA's D-Wave 2000Q system. The Fraunhofer Center for Quantum Computing (14 Fraunhofer institutes are already involved) together with IBM cooperatively run an IBM Q System One.

The **adiabatic theorem of quantum mechanics** asserts that by changing a system's Hamiltonian gradually, the system will stay in it's n-th Eigenstate (given that we initialize the system in that state). Adiabatic quantum computers rely on this principle for solving **Quadratic Binary Unconstrained Optimization Problems (QUBOs)**. AQCs start at the solution of a built-in trivial QUBO problem H and gradually change H into a user specified QUBO instance H* - if the magnitude of the changes was small enough, the system's state will eventually be an optimizer of H* (asserted by the adiabatic theorem). A QUBO instance can be represented via it's coefficient matrix or, equivalently, as the parameters of an Ising model. The latter fact makes adiabatic quantum computing interesting from a Machine Learning (ML) point-of-view because many highly computational demanding ML problems can be (re-)formulated as an Ising model, among them: Maximum-a-posterior inference in probabilistic models like Markov Random Fields, Boltzmann machines, and Hopfield networks.

Our initial study reveals that many relevant probabilistic models cannot be optimized by real world AQCs. Beyond 64 dimensions, the hardware-architecture requires a specific sparsity pattern in the model's coefficient matrix which excludes large-scale ML models. Motivated by this fact, we designed our own quantum-inspired hardware solver that allows for dense coefficient matrices.[11] Our solver is an **FPGA**-based hardware implementation of an evolutionary algorithm with mutation-rate annealing, specifically designed for the QUBO objective function. A study on commodity FPGA hardware shows that our system can solve 2048-dimensional QUBO instances in a few seconds while consuming less than 10W. The cooling alone of D-Wave's 2000Q system consumes 25kW.[12]

In parallel, we investigate quantum gate computers. While AQCs are specialized QUBO solvers, gate-based quantum computers allow the user to directly manipulate quantum bits with a user-specified quantum circuit. Initial studies on various IBM Q System devices (from 5 to 53 quantum bits) show that the applicability is restricted by the system's quantum bit connectivity and the supported set of gates. That is, not every theoretically valid quantum circuit can be executed by real-world quantum gate machines. In the QML context, we try to overcome this limitation by devising a randomized search heuristic that learns quantum circuits for

[10] Kirsch, B., Giesselbach, S., Knodt, D., Rüping, S. (2018). Robust End-User-Driven Social Media Monitoring for Law Enforcement and Emergency Monitoring. In G. Leventakis/M. Haberfeld (eds.), *Community-Oriented Policing and Technological Innovations* (pp. 29-36). Cham, Switzerland: Springer. https://library.oapen.org/bitstream/handle/20.500.12657/27840/1002165.pdf?sequence=1#page=41

[11] Piatkowski, N., Mücke, S. (2020). *System zur evolutionären Optimierung pseudoboolescher Polynome.* Munich, Germany: Deutsches Patent- und Markenamt [Patent: DE 10 2020 206 127.0].

[12] Mücke, S., Piatkowski, N., Morik, K. (2019). Hardware Acceleration of Machine Learning Beyond Linear Algebra. *ECML PKDD Workshops*, 1, pp. 342-347. https://link.springer.com/chapter/10.1007/978-3-030-43823-4_29

classification. Given a dataset of a classification problem, the method outputs a quantum circuit which can classify data by only using those quantum gates which are physically realizable.

In the QML algorithms area, we investigate the re-formulation of known ML problems into Ising or QUBO forms, which makes them amenable for AQC and our FPGA-based solver. This line of research includes clustering problems[13], max-sum dispersion[14], which occurs in the context of web search or text summarization, as well as Markov Random Fields[15], and a version of the support vector machine.[16] In a second line of research, we investigate the performance of Machine Learning models that are already in QUBO form, like Boltzmann machines.[17] There, the focus lies on the model's conditional independence structure. The long-term goal of this work aims at a shallow embedding of a model into the physical quantum bit connectivity, since the depth of quantum circuits induces a high level of noise into the computation.

The working group on "Physics, Modern IT and Artificial Intelligence" of the German Physical Society (DPG) organized its Fall Meeting 2019 around the topic of quantum technology. Katharina Morik, being in the board of the working group which currently has about 900 members, gave an invited talk on Machine Learning at the Fall Meeting in Freiburg and socialized with international colleagues who work on ML and Quantum Computing.

ML with Restricted Resources

The resource runtime and memory have always been investigated, most often in the O-notation of worst case complexity, but sometimes also in the worst case of execution. With the heterogeneity of modern computing architectures and the challenge of big data and streaming data in the Internet of Things (IoT), Machine Learning research is facing research challenges. **FPGAs** in particular are well suited for Internet of Things applications, but at the same time, demand their new implementations of known ML algorithms. The resource requirements of several implementations for one algorithm need to be characterized. For example, even the well-known decision tree algorithm now offers novel research questions: Which implementation is best suited for which computing architecture and application requirements?[18] Deep Learning models can be executed on FPGAs as

[13] Mücke, S., Piatkowski, N., Morik, K. (2019). Hardware Acceleration of Machine Learning Beyond Linear Algebra. *PKDD/ECML Workshops*, 1, pp. 342-347. https://link.springer.com/chapter/10.1007/978-3-030-43823-4_29; Mücke, S., Piatkowski, N., Morik, K. (2019). Learning Bit by Bit: Extracting the Essence of Machine Learning. *Proc. LWDA*, pp. 144-155. http://ceur-ws.org/Vol-2454/paper_51.pdf; Bauckhage, C., Ojeda, C., Sifa, R., Wrobel, S. (2018). Adiabatic Quantum Computing for Kernel k=2 Means Clustering. *Proc. LWDA*. http://ceur-ws.org/Vol-2191/paper3.pdf; Bauckhage, C., Piatkowski, N., Sifa, R., Hecker, D., Wrobel, S. (2019). A QUBO Formulation of the k-Medoids Problem. *Proc. LWDA*, pp. 54-63. http://ceur-ws.org/Vol-2454/paper_39.pdf

[14] Bauckhage, C., Sifa, R., Wrobel, S. (2020). Adiabatic Quantum Computing for Max-Sum Diversification. *Proc. SIAM International Conference on Data Mining*, pp. 343-351. https://epubs.siam.org/doi/pdf/10.1137/1.9781611976236.39

[15] Mücke, S., Piatkowski, N., Morik, K. (2019). Hardware Acceleration of Machine Learning Beyond Linear Algebra. *ECML PKDD Workshops*, 1, pp. 342-347. https://link.springer.com/chapter/10.1007/978-3-030-43823-4_29

[16] Mücke, S., Piatkowski, N., Morik, K. (2019). Learning Bit by Bit: Extracting the Essence of Machine Learning. *Proc. LWDA*, pp. 144-155. http://ceur-ws.org/Vol-2454/paper_51.pdf

[17] Piatkowski, N. (2019). Hyper-Parameter-Free Generative Modelling with Deep Boltzmann Trees. *ECML PKDD*, 2, pp. 415-431. https://ecmlpkdd2019.org/downloads/paper/637.pdf

[18] Buschjäger, S., Morik, K. (2020). Decision Tree and Random Forest Implementations for Fast Filtering of Sensor Data. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 65(1), 209-222. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7962153

well and several implementations exist. A recent paper from ML2R together with the Collaborative Research Center 876 and its astrophysics project now achieved an end-to-end learning from image streams of a telescope using Deep Learning and implements the learned model at an FPGA at the telescope directly.[19]

Researchers at Fraunhofer IML use a powerful **supercomputer** in current SME projects to train models for **image recognition**. Energy-efficient processors are used for mobile robots in order to send image data from the robots to the cloud, where calculations are conducted. If scaling and data protection considerations prohibit this procedure, two approaches are possible. On the one hand, a program code optimized for the system can be generated automatically, which is independent from common software packages and therefore much more efficient.[20] On the other hand, the models designed for solving problems, such as counting objects, can be specialized so that unnecessary results do not have to be calculated.[21] Together, both approaches render the use of the models directly on the resource-limited systems possible.

Turning it the other way around, we may look at the ML algorithms and see how we can change the model such that less resources are consumed. For probabilistic graphical models, the use of integer variables all through the learning and the inference, could be shown to save resources without losing too much accuracy. The complex PGM could be run even on ultra low power devices![22]

Probabilistic Gap Filling on Spatio-Temporal Data

Spatio-temporal data sets allow for the simultaneous analysis of important developments at different spatial sites over time. However, incompleteness of data at hand can heavily impact the usability. As an example, **satellite images** provide important information about processes on the Earth's surface. They can be used to illustrate the effects of climate change, urbanization or agricultural land use. However, a central problem of remotely sensed data is the occlusion of the Earth's surface by clouds. ML2R scientists at the TU Dortmund University have therefore developed a methodical framework to fill gaps in highly incomplete spatio-temporal data sets. In contrast to similar approaches, their method does not make any assumptions over data at hand and does not rely on auxiliary data input. They fill gaps with estimates of a probabilistic model, which considers partially observed information in the spatial and temporal neighborhood. In the course of their work, the scientists successfully applied the novel method on exemplary satellite images suffering from cloud cover.[23]

Combining Expert Knowledge and Neural Networks to Model Environmental Effects in Agriculture

The endeavor to design better strategies for crop selection and planting has historically formed a central core of agriculture and mathematical modeling research. In this direction, a main line of studies addresses the

[19] Buschjäger, S., Pfahler, L., Buss, J., Morik, K., Rhode, W. (2020). On-Site Gamma-Hadron Separation with Deep Learning on FPGAs. *ECML PKDD* (accepted, to be published), Ghent, Belgium.
[20] Urbann, O. et al. (2020). *A C Code Generator for Fast Inference and Simple Deployment of Convolutional Neural Networks on Resource Constrained Systems.* https://arxiv.org/pdf/2001.05572.pdf
[21] Urbann, O., Stenzel, J. (2019). A Convolutional Neural Network that Self-Contained Counts. *Journal of Image and Graphics*, 7(4), 112-116. http://www.joig.org/uploadfile/2019/1113/20191113120905251.pdf
[22] Piatkowski, N., Lee, S., Morik, K. (2016). Integer undirected graphical models for resource-constrained systems. *Neurocomputing*, 173, 9-23.
https://reader.elsevier.com/reader/sd/pii/S0925231215010449?token=6EBE8492BD3AD0DBFB42198F5513C0A7213285115EC85DED41508C43F1E8B5452CAFB78F52FA92303BF731010EC1EC72
[23] Fischer, R., Piatkowski, N., Pelletier, C. Webb, G., Petitjean, F., Morik, K. (2020). No Cloud on the Horizon: Probabilistic Gap Filling in Satellite Image Series. *DSAA* (accepted, to be published), Sydney, Australia.

relationship between environmental conditions and plant growth - in particular, the impact of certain environmental stresses (e.g. heat and drought) on crops and respective yield. Last year, ML2R experts together with colleagues from the PhenoRob Cluster of Excellence at the University of Bonn, succeeded in using Artificial Intelligence to explore some of these questions. The research lies at the intersection of sophisticated mathematical modeling and techniques from big data analysis and deep learning. More precisely, the proposed solution leverages on one hand, knowledge from classical agriculture research in terms of delicate models describing explicit environmental and plant characteristics (effects of photosynthesis, crop transpiration and soil hydrodynamics), and on the other hand, these specific pieces of domain knowledge are used to further guide and regularize Convolutional Neural Networks (CNN) designed to uncover suitable data representations and hidden connections between environmental and plant features. This results in an informed Machine Learning approach that **combines experts' knowledge with the flexibility and expressiveness of big data and neural network approaches**. This line of research ultimately provides novel algorithms for analyzing environmental stresses and identifying the most resilient plant sorts (hybrids). The contribution, submitted as part of the annual "Crop Challenge in Analytics", 2019, of the global agricultural group Syngenta, was chosen as the winner of the competition.[24]

ACCOLADES

The representatives of the competence center combine scientific expertise and excellence with an extensive network of scientific and industrial contacts. Their achievements in the field of Machine Learning methods have received numerous awards.

ML2R spokeswoman, **Prof. Dr. Katharina Morik**, was awarded with the **GI-Fellowship** by the Gesellschaft für Informatik e.V. (GI) in October 2019. Katharina Morik was honored as a pioneer of Machine Learning, as she has significantly advanced the research area on an international level. Particularly her collaboration with Yves Kodratoff from the Université Paris-Sud in the early years of the field has laid the foundation for the European research community of Machine Learning. The resulting ECML PKDD conference now covers all European countries and had 800 participants last year. In addition, the interdisciplinarity and high application relevance of her scientific work were emphasized.

Prof. Dr. Stefan Wrobel, speaker of the ML2R, was appointed as one of the **“ten influential minds” in German AI history** by the Gesellschaft für Informatik e.V. (GI) in July 2019. In the "Science Year 2019 - Artificial Intelligence" proclaimed by the German Federal Ministry of Education and Research (BMBF), the jury chose ten outstanding scientists who have made a significant contribution to shaping AI in Germany. Prof. Dr. Stefan Wrobel was chosen as an expert in the field of Machine Learning.

Prof. Dr. Dr. h. c. Michael ten Hompel, representative of the competence center’s partner Fraunhofer IML, is considered as one of the fathers of the Internet of Things and has published numerous books and over four hundred articles on logistics and IT. Michael ten Hompel was awarded the honorary distinction **"Citizen of the Ruhr Area"** in December 2018. The jury stressed his role as a pioneer for the digitalization of the logistics hub Ruhr. The Ruhr area is the biggest metropolitan area in Germany, located in the western state of North Rhine-Westphalia. In April 2019, he was also honored with the **"HERMES Verkehrs.Logistik.Preis 2019"** for his outstanding achievements in international logistics.

Prof. Dr. Christian Bauckhage, representative of the competence center’s partner University of Bonn, together with the researchers Rafet Sifa (Fraunhofer IAIS) and Dr. Tiansi Dong (Bonn-Aachen International Center for Information Technology) was awarded the **Best Paper Award** at the International Conference on Artificial Neural Networks (ICANN 2019). In the cross-institutional project under the ML2R flag, the researchers

[24] For further information, please visit: https://www.syngenta-us.com/thrive/research/winners-2019-crop-challenge.html

addressed the issue of "Prototypes within Minimum Enclosing Balls". They presented a procedure that filters out those characteristic data points that are particularly easy to interpret for human users ("explainable latent factors") from a given data set.

PROMOTING YOUNG SCIENTISTS

The Competence Center Machine Learning Rhine-Ruhr is dedicated to the promotion of young scientists in the key area of Machine Learning. The institutions involved in the competence network have a long tradition of imparting specialist knowledge to and promoting up-and-coming young scientists. ML2R offers master graduates the opportunity to obtain a doctoral degree in the research area of Machine Learning. Excellent scientists are also offered the opportunity to work as postdocs.

The outstanding achievements of ML2R's young researchers were exemplified by the awarding of a **dissertation prize** by the TU Dortmund University to ML2R researcher Dr. Nico Piatkowski. He completed his thesis entitled "Exponential Families on Resource-Constrained Systems" with distinction (summa cum laude). It is the first comprehensive investigation of resource requirements and quality bounds for probabilistic graphical models. It provides a thorough yet tutorial review of exponential families before it introduces the novel learning models of spatio-temporal models, integer random fields, and an innovative approximation of the partition function. Nico Piatkowski subsequently accepted a position as a postdoc at the competence center.

OUTLOOK

ML2R aims to **promote international research cooperations**. Based on the wide network of existing international contacts in the research community of Machine Learning, the ML2R plans to conduct joint research projects and publications. If you or your scientific team are interested in the implementation of a joint research cooperation, e.g. in the context of publications, workshops, panels etc., please contact us.

Together with the Collaborative Research Center 876 "Providing Information by Resource-Constrained Data Analysis" at the TU Dortmund University, the ML2R will organize an **international Summer School** in the period between 31 August and 4 September 2020. The virtual event will bring together experts from the research fields of data analysis (Machine Learning, data mining, statistics) and embedded systems (cyber-physical systems). In their lectures, they will address the resource constraints of host devices used for data analysis. The participants will also be given the opportunity to present their research and network with other participants. The Summer School will be accompanied by a hackathon in the form of a Kaggle Challenge, in which participants can test their knowledge of Machine Learning and cyber-physical systems in a real-world scenario. Further information about the Summer School is available at:

https://www-ai.cs.tu-dortmund.de/summer-school-2020/

Transfer

INDUSTRIAL COOPERATIONS

The ML2R enables companies to access strategies, knowledge and data of Machine Learning (ML) processes within the framework of a cooperation. Using data provided by the companies, the staff of the competence center **analyzes the potentials of Machine Learning within the company and supports the implementation of applications into practice**. Those interested in a cooperation with the ML2R to date cover a broad spectrum of economic sectors, including the areas of logistics and industry 4.0, telecommunications providers as well as the compliance area. Cyber security, anomaly detection, quality prognoses in resource-saving processes, real-time decision support as well as the explainability and transparency of Machine Learning were identified as possible fields of application within the respective companies. The need for networking and cooperation between excellent research and innovative entrepreneurship was demonstrated by 42 talks of ML2R scientists at companies. Talks were held among others with the companies KraussMaffei, Deutsche Telekom AG/Telekom Security, Wilo SE, and targens.

The ML2R team not only supports companies in the context of cooperations through the introduction of Machine Learning into the company's own processes and products, but also helps to **build up and expand the necessary competence**. The competence center is currently developing a broad range of **seminars**. **Co-Working formats** in which companies work together with ML2R experts on innovative Machine Learning applications are also being considered. The development of know-how in the companies and the direct transfer of research into practice plays a decisive role in these endeavors.

The **generation of C source code**[25] from trained models is not only interesting for small and medium-sized German companies, which aim at integrating it into their proprietary frameworks. ML2R is further developing its already available code generator as an independent solution. At the same time, Fraunhofer IML is involved in **Google's open source project "Intermediate Representation Execution Environment" (IREE)**, which integrates a C code generation of Fraunhofer IML. Within IREE, further opportunities for an easy deployment are also being evaluated.

KION ENDOWED PROFESSORSHIP

Through the increased cooperation with industrial partners, the ML2R is a driving force of application-oriented research of Machine Learning and enables the transfer of cutting-edge research in a key area of Artificial Intelligence to the industrial sector. The collaborations prove to be mutually beneficial and enable researchers to address real-world needs of economic actors in their work.

The KION GROUP AG has established the **endowed professorship** "Machine Learning for Industrial Applications" at the TU Dortmund University. A corresponding recruitment process is currently underway. The selected scientist will be involved in ML2R's research activities, thereby promoting the expansion of the competence center. In addition, the establishment of the endowed professorship will contribute to the financing of the ML2R's intended consolidation.

FLAGSHIP-APPLICATIONS

Within the flagships, the methods researched in the competence center are evaluated in concrete scenarios. The generalizability of the results is given by the selection of application areas, which differ concerning

[25] Fraunhofer IML has published an open source Neural Network Code Generator (NNCG) (Retrieved from: https://github.com/iml130/nncg).

technical and sector-specific characteristics. Flagship projects of the ML2R are currently focused on the areas of Industry and Logistics 4.0 as well as Cognitive Process Automation.

At the Fraunhofer IML and the TU Dortmund University, there is a scientific test field based on scenarios surrounding **Industry 4.0** and **Logistics 4.0**. Demonstrators form specific logistic subsystems, where the effect of the progressive autonomization and the novel methods of Machine Learning will be most evident. The respective demonstrators will, upon completion, focus on the level of logistic processes or plant control as well as on human activity. They will thus illustrate in which way the operation and optimization of logistic systems will further develop through the application of Machine Learning. A demonstrator developed within the ML2R already enables object recognition using deep neural networks at rates of approx. 30 Hz and 0.5 W energy consumption. In research projects on indoor localization for intralogistic processes, localization approaches are also being developed using autonomous systems, standardized RGB cameras and external, non-optical sensor technology.

At Fraunhofer IML, the "**ai arena**" allows for the access to novel real-world and industry-related hardware. The research focus lies, both, on the development of code generators and on the AI autonomous swarm robotics with human-technology interaction. The latter aims at successfully coordinating and navigating swarms of robots and drones under constant optimization. Using Reinforcement Learning, scientists have tested the simulation of virtual worlds with singular robots which had to navigate through an obstructed space. This procedure will now be applied in the real world to a swarm of real robots.

***ai arena** at Fraunhofer IML:*

In addition, Fraunhofer IML and the chair for Material Flow and Logistics Management at TU Dortmund University have established "**PhyNetLab**" - a real-scale warehouse testbed, made up of cyber-physical objects (PhyNodes). It addresses the need for flexibility and scalability for emerging warehouses as well as the resulting complexity of such systems due to the multidisciplinarity in logistics, embedded systems, and wireless communications.

For the **Cognitive Process Automation** flagship, Fraunhofer IAIS builds demonstrators of Machine Learning in Natural Language Processing (NLP). They extract information from accounting documents, business reports, and court decisions, using both NLP and Machine Learning for entity recognition and anonymization. Moreover, the system supports human auditors by analyzing and annotating text modules and tables. In addition, the Flagship Demonstrators include an AI-supported tool that allows for the automated processing of checklists that serve to answer a multitude of questions, list different requirements that need to be fulfilled, like the

completeness of mandatory data, as well as to check the consistency of document content in different tables, text passages and documents.

The demonstrators for document recognition and the analysis of court decisions integrate methods from Statistical Relational Learning, which allow for the evaluation and weighing of previous knowledge of experts in the form of rules. This makes it possible to achieve good results even with smaller training quantities and is particularly helpful when evaluating semi-structured data, in this case documents with layout information.

OUTLOOK

The ML2R strives to intensify **collaborations with** partners from **the economic sector**. Based on the many visits and talks, increased cooperation with industrial partners is planned in the form of research projects and joint measures within public relations. The joint participation of topic-specific conferences is also planned. If you and your company are interested in a cooperation with the ML2R, e.g. in the context of a research project in your company or the joint participation in conferences, please contact us.

In addition, **seminars** and **training courses** on the ML2R's research foci are planned, placing an emphasis especially on the new topics of Quantum ML and Trustworthy ML. In addition, application-oriented topics from the two flagships are planned.

ML2R informs

Visibility

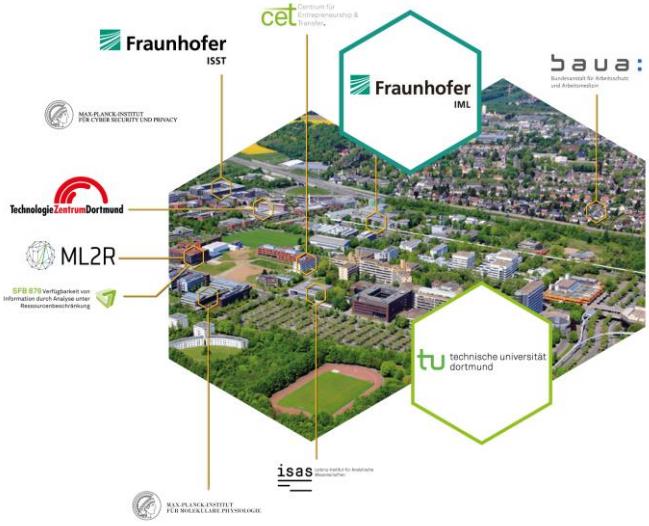
NETWORK

The competence center itself builds upon the basic-research excellence at the location of the ML2R sites, which is the Collaborative Research Center 876 on Resource-restricted Data Analysis and the Max Planck Institute for Cybersecurity on the Dortmund site, and the excellence clusters “Robotics and Phenotyping for Sustainable Crop Production” and “Matter and Light for Quantum Computing” at the Bonn site. The sites’ settings offer an ecosystem of research, transfer, and innovation, as it is shown in the figures below.

The “Centrum für Entrepreneurship & Transfer” (CET) at TU Dortmund University encourages and supports start-ups. The “Technologie Zentrum Dortmund” hosts more than 300 small and medium-sized companies with about 10 000 employees. The Leibniz “Institute for Analytic Sciences” (ISAS) at TU Dortmund University advances chemical and physical measuring, among them the recognition of viruses through an innovative telescope.[26] “BAuA” is the Federal Institute for Occupational Safety and Health. Together with Fraunhofer Gesellschaft and further institutions, the University of Bonn carries the Bonn-Aachen International Center for Information Technology (b-it). “b-it” offers three international Master Programmes in Computer Science and Applied IT and encompasses nine research groups – all on an internationally excellent level.

Fraunhofer IAIS, located in Sankt Augustin close to Bonn, hosts among others one of the sites of “AI4EU”, the European Union’s landmark Artificial Intelligence project. Fraunhofer IAIS is also part of the “International Data Spaces Association”. It offers its more than 100 participating companies and institutions of various industries and sizes from 20 countries, including several Fortune 500 companies, a virtual architecture, developed by Fraunhofer IAIS, that allows for the safe transfer of data.

*The **Innovative Environment** of the **ML2R sites**:*



[26] Shpacovitch, V. et al. (2017). Application of the PAMONO-Sensor for Quantification of Microvesicles and Determination of Nano-Particle Size Distribution. *Sensors, 17*(2), 244. https://www.mdpi.com/1424-8220/17/2/244/htm

ML2R informs

The **Rhine-Ruhr region**, located in the western state of North Rhine-Westphalia, offers innovation and tradition: internationally recognised research on Machine Learning and Big Data on the one hand, and a productive economic sector with over 20 DAX and MDAX companies on the other hand.

In its work, the ML2R benefits from its consolidated network. High-ranking scientific and economic actors consult the employees of the competence center according to their specific expertise within the framework of an advisory board. The **Steering Board**, made up of scientific advisors, and the **Industrial Board**, consisting of corporate advisors, help to shape the research agenda from both the scientific and the user perspective, thus enabling external quality assurance.

COORDINATION

In September 2018, German Federal Research Minister Anja Karliczek commissioned the spokesperson of the ML2R competence center, Prof. Dr. Katharina Morik, to **coordinate the German competence centers for Artificial Intelligence** (AI). The network of German AI competence centers has grown since then and is now comprised of six institutions: the Competence Center Machine Learning Rhine-Ruhr (ML2R), the Berlin Institute for the Foundations of Learning and Data (BIFOLD), the Munich Center for Machine Learning (MCML), the Tübingen AI Center - Competence Center for Machine Learning, the Center for Scalable Data Analytics and Artificial Intelligence Dresden/Leipzig (ScaDS.AI) and the German Research Center for Artificial Intelligence (DFKI). The **Office of Coordination** at the TU Dortmund University is responsible for the nationwide networking and coordination of the AI competence centers.

The Office of Coordination under the leadership of Prof. Dr. Katharina Morik is responsible for the coordination as well as the topical exchange between **the network of German AI competence centers and the French National AI Research Network**. The cooperation between Germany and France in the research field of Artificial Intelligence has thus been successfully intensified. For this purpose, Prof. Dr. Katharina Morik is in continuous contact with Prof. Dr. Bertrand Braunschweig, coordinator of the French national program for AI research at the French institute Inria. Together, they have been pushing forward the deepening of the German-French collaboration, discussing a declaration of intent and a call for proposals of collaboration. Germany and France have signed the Joint Declaration of Intent, which will further strengthen the countries' collaboration in the research field of Artificial Intelligence. The call for proposals of binational collaborations will soon be published.

EVENTS

Global Forum on AI for Humanity. The Global Forum on AI for Humanity was held in Paris from 28 to 30 October 2019. Prof. Dr. Katharina Morik, spokesperson of ML2R and coordinator of the German AI competence centers, organized the meeting together with Prof. Dr. Bertrand Braunschweig, coordinator of the French AI competence centers at the INRIA institute, and Prof. Dr. Malik Ghallab, as well as renowned experts from Australia, Canada, Germany, England, France, Japan, the Netherlands and the USA. The meeting in Paris with an address by the forum's host, French President Emmanuel Macron, served as a formal launch of the Global Partnership for Artificial Intelligence (GPAI). The future agenda of the GPAI working groups was formulated and a book is in preparation.

The Academy Day of Artificial Intelligence (North Rhine-Westphalian Academy of Sciences, Humanities and the Arts). Prof. Dr. Katharina Morik organized the "Artificial Intelligence Day - Machine Learning as the Key to Artificial Intelligence" of the North Rhine-Westphalian Academy of Sciences, Humanities and the Arts which took place on 29 October 2019 in Düsseldorf. In the context of lectures and a poster exhibition, the breadth of the field of Machine Learning methods as well as application projects were presented. The speaker of the ML2R, Prof. Dr. Stefan Wrobel, gave the speech: "Artificial Intelligence and Machine Learning: Opportunities and Challenges".

ML2R informs

Meeting of the German and French competence centers. On 16 September 2019, representatives of the six German AI competence centers and the French AI institutes met surrounding the ECML PKDD conference in Würzburg. The meeting was organized by Prof. Dr. Katharina Morik and Prof. Dr. Bertrand Braunschweig (INRIA), coordinator of the French AI institutes. The researchers presented their current work. The aim of the meeting was to lay the groundwork for a virtual Franco-German center for the coordination of the competence centers of both countries as well as to elaborate on a joint declaration for an intensified transnational cooperation.

Summer of Research. The ML2R employees, who conduct research at the ML2R locations of Dortmund, Bonn and Sankt Augustin, exchanged their ideas and presented their current work on 24 July 2019 at the Summer of Research in Sankt Augustin. The central workshop for the PhD students, postdocs and experienced scientists focused on current challenges of the practical application of Machine Learning along their different research foci. It is planned that the Summer of Research will provide ML2R with an annual forum for collaboration.

Visit of Research Minister Anja Karliczek. During her visit to ML2R in Dortmund on 9 July 2019, German Federal Minister of Education and Research Anja Karliczek gained insights into the competence center's cutting-edge research. Accompanied by journalists, she encountered robots that make AI and ML understandable in a playful way and discovered AI systems that analyze spoken language, improve satellite images and make autonomous driving safer. In addition, a swarm of drones hummed above her. The visit gave the participants the opportunity to try out the exhibits themselves and to find out about future plans of the competence center.


Oliver Schaper / TU Dortmund


Oliver Schaper / TU Dortmund

Meeting of ML experts. Starting with an opening event at the VIP club of Dortmund's football club BVB 09 on the evening of 4 June 2019, the ML2R invited all researchers of the Machine Learning projects funded by the German Federal Ministry of Education and Research (BMBF) to Dortmund. More than 100 experts from over 50 projects in industry, business, and science took the opportunity to exchange ideas and network with each other. In around 30 lectures and a poster exhibition, the diverse areas of application of Machine Learning were demonstrated - from logistics, transport and production, meteorology, biology and a broad medical field of application to wind power plants and data security in companies.


Oliver Schaper / TU Dortmund


Oliver Schaper / TU Dortmund

ML2R informs

Kick-off event. The ML2R kick-off event on 23 January 2019 illustrated the triad of basic research, bridging competence, and economy. Standing for outstanding research, Prof. Dr. Thorsten Joachims (Cornell University, USA) gave a talk on “Learning from and for Interventions” and Prof. Dr. Michèle Sebag (Université Paris-Sud, France) presented the state of the art in “Selecting an algorithm (and hyper-parameters)”. Of course, the bridging competence center ML2R was introduced and speakers from both the federal government (Prof. Dr. Wolf-Dieter Lukas) and the government of North-Rhine Westfalia (Minister Isabel Pfeiffer-Poensgen), as well as the rector of TU Dortmund University, Prof. Dr. Dr. h. c. Ursula Gather, marked its societal importance. Success stories of Machine Learning applications, e.g. at T-Systems, Volkswagen AG, and Pricewaterhouse Coopers were accomplished by an innovation show, where the companies Comma Soft AG, Crosscan GmbH, Integrated Analytics GmbH, Old World Computing UG, Point 8 GmbH, QuinScape GmbH, RapidMiner GmbH, Roberta-Initiative of Fraunhofer IAIS, sourcewerk GmbH, and UNICEPTA Gesellschaft für Medienanalyse mbH demonstrated the ecosystem of the Ruhr area.

OUTLOOK

Over the course of the year 2020, the ML2R will set up a **virtual ML showroom**, which will bring Machine Learning demonstrators and applications to life in a virtual environment. In addition to technological demonstrators, the online platform will also offer access to online courses and teaching materials, curated data and algorithms for practical application, as well as up-to-date information on the transfer potential of ML research results into the economy. Thus, the ML2R will provide an open access to knowledge and data for an interested audience from science, economy, and society. The aim is to reduce social and economic reservations toward the technologies of Artificial Intelligence and to strengthen collaborative scientific endeavors.

In addition to the Virtual ML Showroom, the ML2R will host a **blog**, that is dedicated to the increased transfer of ML research results into the economic sector. Within the context of the blog, ML2R experts will report on their research and thus aim to strengthen the exchange on current research results, topics and applications of ML research, especially with experts from companies.

Under the leadership of the coordination office at ML2R, a **newsletter** of the nationwide network of **all AI competence centers** will be set up. The newsletter, which will address an interested audience, will include current publications and information from the competence centers.

Contact Informationen

Prof. Dr. Katharina Morik

Tel. +49 231/755-5101
Fax +49 231/755-5105
katharina.morik@tu-dortmund.de

Prof. Dr. Stefan Wrobel

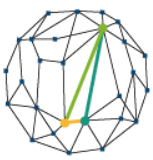
Tel. +49 2241 14-3015 (Secr. -3014)
Fax +49 2241 14-2436
stefan.wrobel@iais.fraunhofer.de

ML2R informs

Address

Competence Center Machine Learning Rhine-Ruhr
TU Dortmund University, Department of Computer Science
Otto-Hahn-Str. 12
44227 Dortmund

Competence Center Machine Learning Rhine-Ruhr
Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS
Schloss Birlinghoven
53757 Sankt Augustin



ML2R
Kompetenzzentrum
Maschinelles Lernen
Rhein-Ruhr