Foundations for applied artificial intelligence: enabling and supporting AI teaching and research

Ricardo Knauer[1], Raphael Wallsberger[2], Marvin Grimm[3], Marian Bookhahn[4], Frank Neumann[5], Stephan Matzka[6] and Erik Rodner[7]

Abstract: The field of artificial intelligence (AI) has experienced tremendous advances and will undoubtedly shape a range of vertical industries over the next years. Understanding and applying AI technologies are key skills for the future and of utmost importance for virtually all disciplines. At the University of Applied Sciences Berlin, we are designing an infrastructure and community called KI-Werkstatt that allows us to tackle the challenges of applied research and teaching across vertical domains. The following paper outlines our current efforts and principles.

Keywords: AI education, applied AI research, hardware infrastructure

1 Introduction

Data is sometimes regarded as the oil of the 21st century. Indeed, algorithms with the ability to learn from data are ubiquitous in our daily lives - from voice assistants and chatbots to new medical products [DD22]. The range of services that derive insights from data is vast, and therefore it is increasingly important that students already understand at an early stage how machines can leverage data to do things "intelligently" and apply algorithms that can learn from data. Undergraduate courses do not typically cover core concepts about artificial intelligence (AI), especially in non-technical fields. Furthermore, the computational resources for tackling complex economic and societal problems with AI tools are often scarce for students and even researchers. In the KI-Werkstatt project at the University of Applied Sciences Berlin (HTW Berlin), we aim to tackle each of these issues. In this paper, we first describe how we set up a state-of-the-art infrastructure that acts as an enabler for AI applications in teaching and research, then illustrate our approach to anchor AI literacy as a multidisciplinary core competency set and finally report on the use of our infrastructure for an academic research project in the field of 3D feature extraction. Each of these topics is essential to enable students, lecturers, and researchers.

[1] University of Applied Sciences Berlin, KI-Werkstatt, 10313 Berlin, ricardo.knauer@htw-berlin.de
[2] University of Applied Sciences Berlin, KI-Werkstatt, 10313 Berlin, raphael.wallsberger@htw-berlin.de
[3] University of Applied Sciences Berlin, KI-Werkstatt, 10313 Berlin, marvin.grimm@htw-berlin.de
[4] University of Applied Sciences Berlin, KI-Werkstatt, 10313 Berlin, marian.bookhahn@htw-berlin.de
[5] University of Applied Sciences Berlin, KI-Werkstatt, 10313 Berlin, frank.neumann@htw-berlin.de
[6] University of Applied Sciences Berlin, KI-Werkstatt, 10313 Berlin, stephan.matzka@htw-berlin.de
[7] University of Applied Sciences Berlin, KI-Werkstatt, 10313 Berlin, erik.rodner@htw-berlin.de

2 Technical prerequisites for AI teaching and research

In machine learning, finding complex associations between input and output and using current state-of-the-art models is extremely resource-intensive. It is therefore crucial to provide sufficient computational resources for applying AI in teaching and research. In the KI-Werkstatt project, we logically split our computing infrastructure into 3 tiers, solving 3 different problems: (1) prototyping, (2) exploration and training as well as (3) inference. This ensures that our AI infrastructure is both scalable, allowing educators and researchers to scale their computational resources as needed, and consistently utilized.

We facilitate rapid prototyping with individualized GPU-powered virtual machines (VMs) and mobile workstations. For exploration and training, we set up a high-performance computing (HPC) cluster with a head node running SLURM, scheduling jobs across multiple Nvidia A100 GPUs. SLURM allows for full flexibility on the server side without being tied to specific machine learning frameworks. However, this also requires our users to use external or self-hosted systems for experiments and dataset tracking. For inference, we rely on a multi-node Kubernetes cluster with A30 GPUs to allow for a more granular control of GPU compute time scheduling. Finally, we deploy a large quantity of small, standardized GPU VMs for students to gain hands-on experience in applying AI while keeping the barrier of entry as low as possible. As can be seen, our infrastructure was designed with flexibility and diversity in mind without the need to achieve maximum compute performance.

3 Enabling and supporting students and lecturers

The ubiquity of AI applications in our personal and professional lives, combined with common misunderstandings about the underlying technology, makes AI literacy a core competency set of the 21st century [LM20]. We believe that the groundwork for AI literacy should be laid at an early stage, and that students of all backgrounds should become familiar with basic AI concepts. Based on the literature [LAP23, LM20, MRS23] and 25 semi-structured interviews with teaching staff and students at HTW Berlin as well as enterprises, we have therefore developed a range of competency-oriented, one- to two-hour-long teaching bricks that can be integrated into and adapted for existing lectures:

- Basics of data understanding: In this module, we address how data can be used for decision-making, why the quality of data is important (algorithmic bias) and how it can be evaluated with minimal coding. For example, students interactively explore data using pandas' Plotly backend.
- Basics of AI: In "Basics of AI", we discuss domain-specific applications of AI, buzzwords around the technology as well as the business and data understanding

phases of a typical AI lifecycle. Students cooperatively build an exemplary AI algorithm in a no-code setting to classify real-world or virtual playing cards [Mi23].

- AI and ethics: This teaching brick tackles the question how ethics relates to data and AI. We zoom in on the ethical, legal and societal implications of trending AI technologies such as large language models, and discuss opportunities and challenges. Students actively apply web-based AI tools[8] to classify animal images and understand how low-quality data can lead to biased decisions.

- Trending topics in AI: In this teaching brick, we sketch the inner workings of selected AI models. State-of-the-art AI models can require dozens of gigabytes of memory for low-latency inferences, so students must commonly resort to using closed-source models that ought not be fed with sensitive data. Our AI infrastructure enables them to apply and evaluate open-source models for sensitive and insensitive data alike. For *the* major trending topic at present, large language models, this flexibility also allows them to experience how prompt engineering strategies differ between larger closed-source and smaller open-source models [e.g., Ko22], and how iterative prompt development is key to avoid hallucinations and biases across model architectures.

To lower the barriers for lecturers to implement these teaching bricks, they are individually supported along the way in integrating these bricks into their own teaching. Additionally, we offer workshops both at the HTW Berlin and the Berliner Zentrum für Hochschullehre (BZHL) that address teaching about and with AI[9]. By providing lecturers with multiple means to incorporate AI into their own teaching context, including an infrastructure for resource-intensive computations, we take important steps towards anchoring AI as a multidisciplinary core competency set in technical and nontechnical domains – both at the HTW Berlin and beyond. We are currently in the process of releasing our materials as open educational resources under the CC BY-SA 4.0 license[10].

4 Enabling AI research at scale

Apart from AI applications in teaching, we also support a range of AI projects in research[11]. In the following, we focus on the project SparePartAssist as a concrete example. SparePartAssist aims at finding spare parts in a huge (> 1 million) and ever-changing collection of computer-aided design (CAD) models, based on 3D scans. With a specifically developed app, service technicians can identify urgently needed components, when documentation is not available or up-to-date. To train and evaluate robust feature

[8] https://teachablemachine.withgoogle.com/
[9] https://bzhlprogramm.tu-berlin.de/e/FACHBEZOGENE-KI-INHALTE
[10] e.g., https://mediathek.htw-berlin.de/video/Introduction-to-kernel-density-estimation-/2f3c144721aadae43b8564b52519ca07
[11] https://kiwerkstatt.htw-berlin.de/alle-projekte

extractors for 3D scans, a large synthetically generated dataset was created utilizing different rendering techniques including computationally expensive raytracing algorithms. For the first prototyping, comparatively small GPU VMs were used. Having a tested container setup, it was easy to transfer the different pipelines to our HPC cluster for training. Without the HPC cluster, we performed a training run in roughly 3 days on an Nvidia RTX 3090 GPU; with the cluster, we were able to start 3 parallel jobs which finished in less than 1 day without even utilizing the entire capacity of the cluster. It is important to note that our infrastructure also allows for experiments with state-of-the-art models on a much larger scale. Next to high-performance computing, our infrastructure, in particular the mobile workstations, also enabled students to easily test the interplay of the app with external equipment like cameras or depth sensors.

5 Conclusion

The KI-Werkstatt project lays the groundwork for applied AI teaching and research at the HTW Berlin and beyond, by providing a hardware infrastructure for resource-intensive computations, designing educational resources for AI literacy across technical and nontechnical fields, supporting lecturers in implementing core AI concepts into their own teaching as well as facilitating state-of-the-art applied AI research. Open challenges remain how to best trade-off infrastructure performance against maintenance effort, as well as how to foster teaching brick integration with minimal external support.

Bibliography

[DD22] DIN, DKE: Deutsche Normungsroadmap Künstliche Intelligenz (Ausgabe 2), https://www.din.de/go/normungsroadmapki/, accessed: 31/05/2023.

[LAP23] Laupichler, M.C. et al.: Delphi study for the development and preliminary validation of an item set for the assessment of non-experts' AI literacy. Computers and Education: Artificial Intelligence 4/23, 2023.

[LM20] Long, D.; Magerko, B.: What is AI literacy? Competencies and design considerations. In (Bernhaupt, R. et al.): Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1-16, 2020.

[Ko22] Kojima, T. et al. Large language models are zero-shot reasoners. In (Oh, A.H. et al.): Advances in Neural Information Processing Systems, pp. 22199-22213, 2022.

[Mi23] Michaeli, T. et al.: Data, trees, and forests – decision tree learning in K-12 education. arXiv preprint arXiv:2305.06442, 2023.

[MRS23] Michaeli, T. et al.: What students can learn about artificial intelligence – recommendations for K-12 computing education. arXiv preprint arXiv:2305.06450, 2023.