Machine Learning: Menschen Lernen Maschinelles Lernen

Ralph Isenmann, Jürgen Prinzbach, Stephan Trahasch, Volker Sänger, Tobias Lauer, Tobias Hagen, and Klaus Dorer

Institute for Machine Learning and Analytics
Offenburg University of Applied Sciences
{ralph.isenmann/ juergen.prinzbach/ stephan.trahasch/ volker.saenger/ tobias.lauer/ tobias.hagen/ klaus.dorer}
@hs-offenburg.de

Abstract. This paper describes the concept and some results of the project "Menschen Lernen Maschinelles Lernen" (Humans Learn Machine Learning, ML2)[1] of the University of Applied Sciences Offenburg. It brings together students of different courses of study and practitioners from companies on the subject of Machine Learning. A mixture of blended learning and practical projects ensures a tight coupling of machine learning theory and application. The paper details the phases of ML2 and mentions two successful example projects.

Keywords: machine learning, data science, blended learning

1 General Description

Machine learning is a multidisciplinary field of study that is increasingly relevant to a wide range of practical applications. Unfortunately, company employees usually do not have the necessary knowledge to use machine learning in their respective areas, while computer scientists, for example, do not have the necessary domain knowledge [1]. This problem is addressed by the interdisciplinary research project ML2. It is funded by the Federal Ministry of Education and Research (BMBF) and implemented by the Institute for Machine Learning and Analytics (IMLA)[2] at Offenburg University of Applied Sciences. The central research question here is how to raise the potential of machine learning for medium-sized companies and at the same time ensure application-oriented education for the students of the university. The answer is a new teaching and learning concept, which has been specially developed for both target groups

[1] https://ml2.hs-offenburg.de
[2] https://imla.hs-offenburg.de

within the framework of this project and provides for two blended learning phases. These can optimally link theory and practice as well as the learning locations university and company as the later place of work for the students.

2 Project Objective

In addition to the basic subjects of computer science and mathematics, concrete machine learning projects always include an area of application as a third discipline. Basically all areas in which large amounts of complex data accumulate can be relevant. However, especially for the respective domain experts, who have little computer science or mathematical background, the field of machine learning is often unknown. At the same time, it is precisely these users who can identify those areas in which machine learning can offer new approaches to solving problems.

The primary goal must therefore be to resolve this dilemma by training experts from companies in machine learning, for example in industry 4.0, trade and logistics or social media. They are addressed in particular because the project was preceded by the assumption that new problems and applications of machine learning are generally not technology-driven. Rather, the users know the problems hidden in the data as well as possible solutions. It is therefore important for them to understand the thinking and approach of machine learning and to expand their understanding of its functions and effects through targeted qualification.

In addition to employees from companies, students from Bachelor's and Master's degree programs are also trained to study in various fields of application that are related to machine learning. As graduates, these students can ultimately transfer their knowledge and skills to companies and thus play an important multiplier role in solving problems through machine learning. The innovative core of the project, however, lies not only in the close integration of theory and applications, but also in the joint qualification of the target groups mentioned. This enables interdisciplinary work across subject, course and department boundaries in order to develop ideas for intelligent data products and services and implement them with the help of machine learning. A further goal is therefore imparting the professional and cross- disciplinary competences.

3 Qualification Concept

The aim of the ML2 qualification program is to qualify company employees and HSO students in the field of machine learning. The two

target groups require and enable interdisciplinary work across subject, course and department boundaries in order to develop ideas for intelligent data products and services and implement them with the help of machine learning. The program lasts nine months and is divided into a theoretical phase "Machine Learning for Practice" and a practical phase "Machine Learning in Practice" (see Figure 1). The figure reflects the general structure of the program.

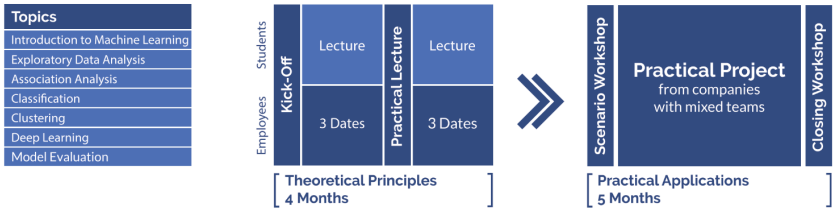

Fig. 1. Project Stages

3.1 Phase 1: Machine Learning for Practice

In the first qualification section, the basics and procedures of machine learning for practical use are taught. The aim is that at the end of the first phase, company employees and students have comparable machine learning skills and a common frame of reference and level of knowledge. Since the prerequisites and the time budget of the two target groups are heterogeneous, the first qualification phase for the users and the students takes place in different qualification paths with target group-specific learning content that takes the different prerequisites into account. In the following, the paths are referred to as student and employee tracks.

In the student track, an interdisciplinary course "Machine Learning" consisting of a lecture and laboratory/exercises is organized. In this course, students learn the same content as employees in the employee track, whereby the student track also deals with some advanced content and more strongly with theory. The students conclude this phase with an examination according to the study and examination regulations.

At the same time, a qualification on machine learning also takes place in the employee track. The qualification concept for employees is based on a blended learning approach [2] that combines the advantages of classroom-based courses with digital forms of learning that are independent of time and place. It consists of a kick-off meeting and a subsequent phase with interlinked on-line training and face-to-face meetings as well as a final meeting. The on-line phases enable participants to acquire knowledge flexibly in terms of time and place. The intermediate

presence events set the pace and support the learning processes through the professional exchange between the participants and the teachers.

In lectures given by data scientists from industry and research, the participants and students can gain further experience already in the theoretical phase and combine it with learned knowledge. At the same time, they provide an opportunity for exchange and networking with the lecturers.

3.2 Phase 2: Machine Learning in Practice

In the practical phase, the linking of theory and practice as well as the linking of the learning locations university and company as the later place of work is of particular relevance for the students. The practical phase involves the transfer and deepening of the participants' machine learning expertise by applying and developing the knowledge acquired in the first phase in a professional and entrepreneurial context. It is necessary that the participants apply the Machine Learning know-how according to the situation and adapt it to the existing conditions of practice and to company goals. For example, a method that is best suited in theory might not be applicable in practice due to the amount of data, its nature or availability, or due to the time or resource budget.

At the end of the first phase, a selection of practice-relevant projects from companies of the participating employees will be available, which will be processed in this second phase. A presentation of these projects will take place at a final attendance date where both the employees and the students will be present. This allows participants and students to form preferences as to which projects they would like to work on. Thus, interdisciplinary teams are formed for the practical phase. The lecturers determine in a final vote which projects will be carried out and how the teams will be formed, taking preference into account. The practical phase starts with a classroom event in which the teams meet for the first time and discuss the course and the milestones of the project.

The practical phase is carried out agilely in four sprints of four weeks each, which are based on the CRISP-DM [3] standard. Each sprint begins with the development of the goals of the sprint with regard to the overall project goal from a business perspective (Business Understanding), followed by data acquisition (Data Understanding) and its preparation (Data Preparation). In the next step, a data modeling technique is selected (modeling), applied to the data, the results carefully evaluated (evaluation) and presented to the company (deployment).

The aim of the first sprint is to be able to decide whether the desired project objective can be achieved within the framework of the qualification program. If this is not the case, the experience gained and the

remaining three sprints will allow sufficient time to adjust the objective. The further sprints serve to broaden the database, to refine the preparation of the data or to apply and evaluate several different machine learning methods. The practical phase ends with a final workshop at which all project teams discuss their project status. This allows experience to be gained from all projects across the board.

A special aspect of the practical phase is that not only machine learning algorithms are applied, but also the typical problems during the project work occur, such as availability and quality of the data. In addition, the applicability of the algorithms can be better experienced and learned in a real project. In a real project, a statement can also be made about the economic efficiency, which is an essential success factor.

4 Practical Implementation

Already in the first round, which took place between March 2018 and January 2019, 26 employees from 17 different companies as well as 19 students of the University of Applied Sciences Offenburg took part. They were able to apply the skills they had acquired in the theoretical phase in 13 different projects in the practical phase. Even though there were some hurdles to overcome and the hoped-for goals could not always be achieved in all cases, the projects led to new insights and experiences. Two projects and their results are presented below.

4.1 Lead Classification of Test Users

During the test phase of a software product, it would be good if the users could be evaluated on the basis of their behavior and thus better understood. The question of whether the user becomes a customer is essential. The aim of the project was therefore to use machine learning methods to identify users who are likely to purchase the software and then to use individual measures such as on-board e-mails and calls as well as discount campaigns. This was achieved by creating a prototype based on Python and Amazon Web Services that first excludes the data relevant for classification from the database and then classifies different potentials from low to high using a previously learned model. The results are then made available to the employees responsible for marketing campaigns in order to support them in their work. The prototype is currently being tested and, if successful, further developed into a productive system.

4.2 Advertising Media Framing

In general, the business case is about providing the user with promotional prices from various advertising media. The project was specifically about automatically recognizing sales, prices and advertising texts in previously scanned advertising brochures. For this purpose, Deep Learning methods are used to automatically draw frames around the characteristics to be classified. In concrete terms, manually framed and classified advertising media have already been used to further train the existing MobileNet SSD network from Google with the help of the Machine Learning Framework Tensorflow. The learned network could thus be used to localize and classify the already mentioned features in new advertising media by means of object detection. During the activities, numerous additional insights were gained. These include, for example, that a lower resolution of the images can significantly accelerate the training of the networks while maintaining the same quality. Finally, it can be stated that all previously defined goals have been achieved.

5 Conclusions

In summary, it can be said that the interdisciplinary research project "Menschen Lernen Maschinelles Lernen" achieved the desired goals and enabled the close interlocking of theory and applications in a joint qualification of company employees and students. This means that interdisciplinary work is carried out across disciplines, courses and departments in order to develop ideas for intelligent products and services, including machine learning. The experience gained in this way helps to sensitize and inspire the participants to the subject area. This success is currently to be repeated in a second round of the project and will lead again to exciting projects.

This research has been funded by the Federal Ministry of Education and Research of Germany in the framework of *IKT 2020 - Forschung für Innovationen* (project number 01IS17080).

References

1. Cao, L.: Data science: Challenges and directions. Commun. ACM **60**(8) (July 2017) 59–68
2. Bonk, C.J., Graham, C.R., Cross, J., Moore, M.G.: The Handbook of Blended Learning: Global Perspectives, Local Designs. Pfeiffer & Company (2005)
3. Shearer, C.: The crisp-dm model: The new blueprint for data mining. Journal of Data Warehousing **5**(4) (2000)