The Diversity of Students as a Challenge of AI Adoption in Boosting Efficiency of Study Programmes

An Empirical Study on the Case of a big Austrian University

Larissa Bartok
Center for Teaching and Learning
University of Vienna
Vienna, Austria
e-mail: larissa.bartok@univie.ac.at

René Krempkow
Research and Transfer Unit
IU International University of Applied Sciences
Berlin, Germany
rene.krempkow@iu.org

***Abstract* - Artificial Intelligence (AI) is gaining ground in everyday life and administrative data analysis at universities. Therefore, applying machine learning and other methods in modeling academic success at universities, for example, is becoming increasingly important. Our case study from a large Austrian university demonstrates how questions related to academic success, with a focus on diversity indicators, can be investigated at universities using various methods, including AI and administrative data of *N* = 2532 students from Science, Technology, Engineering, and Mathematics disciplines (STEM). This article discusses implications for the impact of applying such models on institutional decision-making at universities regarding academic success. Although it is yet difficult to grasp efficiency within the context of student success, the potential impact of the influence of applying such models, together with possible unwanted effects, is discussed.**

Keywords - Academic analytics; study success; machine learning; diversity indicators

I. INTRODUCTION

Students leave a wealth of data at universities: for example, data provided during admission or examination data. This data can be used by the university for evaluation and quality assurance purposes and processed accordingly to further develop study programs in an evidence-based manner. Well-prepared analyses embedded in processes thus represent, on the one hand, a source of information for decision-makers, and, on the other hand, the interpretation and derivation of measures can improve the quality of offered programs and may help to identify stumbling blocks for study success for certain student groups.

Research on AI-based approaches to student success and on digital study assistance systems using rule-based AI is still manageable [1]. Possible implications of using AI models for university development on academic success are not yet discussed. This gap in research motivates the current study, which presents adaptable models for application across universities to increase the efficiency and quality of study programmes. The efficiency of a study programme within this context refers to the extent to which a programme achieves its intended objectives while optimizing the use of available resources, such as time, funding, and institutional support. It encompasses factors such as the timely completion of degrees (e.g., within the standard duration), high graduation rates, effective resource allocation, alignment with labour market demands, and the overall satisfaction and success of students. According to experience and studies, predictive, AI-based, models seem to be well suited to using empirical data and results of empirical analyses for student success and teaching development, as it also makes it possible to formulate narratives from quantitative results and to involve stakeholders [2][3]. However, it is by no means easy to govern universities (especially in state universities in German-speaking countries). Similarly, this applies to identifying suitable starting points for measures to increase the efficiency and quality of degree programmes, as the reasons for higher or lower academic success can be manifold. They can lie in the characteristics of the students or teachers as well as in the characteristics of the study programmes or the organisation, as we will explain in the following section. Within this article we focus on analyses of the prediction of academic success within the first year of studying and identifying the influence of diversity indicators impacting (or not impacting) academic success based on a selected criterion. Our case study that is making use of data of *N* = 2532 students from Science, Technology, Engineering, and Mathematics programmes at a large public Austrian university. The case originates from a working group composed of university leadership, administrative staff, and study programme directors from STEM disciplines. The objective of this working group was to analyse study programmes with a focus on the entry phase and to identify potential barriers for non-traditional students and students from diverse backgrounds. Therefore, institutional research questions were first defined and answered making use of different analysis approaches, depending on the complexity of the questions.

Our paper is organized as follows. Section II provides a brief theoretical overview of student success and factors influencing student success at Higher Education Institutions (HEIs). Section III demonstrates the methodology of our case study including our institutional research questions. In Section IV our results are described. Our conclusion, discussion and ideas for further investigation close the article.

II. STUDENT SUCCESS AND FACTORS INFLUENCING IT

Student success can be measured in various ways [4][5]. Objective success criteria such as student success rates, passed exams, grades or credit points and graduates within the standard period of study are often used. We evaluate student success in terms of successfully passing the student entrance phase after one year. Since the 2011/2012 academic year, the majority of degree programs in Austria have included a legally mandated introductory and orientation phase during the first semester. This phase comprises initial examinations that must be successfully completed before students are permitted to continue their studies (see UG 2002, §66). When considering the impact diversity indicators have on success, this entrance phase is of relevance, since it is well-known that onboarding of students with diverse backgrounds plays a crucial role for belonging and continuing their studies [6]. Attempts at steering within the framework of the Performance Based Funding (PBF) pursued with the new steering model also try to achieve this using such or similar criteria. In PBF models, in addition to student success rates, the study duration or the proportion of graduates within the standard period of study (plus 2 semesters in some cases) are also used [7] – albeit with only mixed success [10][16][17][18].

When considering the factors that contribute to academic success, it is useful to distinguish between institutional and individual factors [7]. Individual factors are usually included in models at universities to explain and predict study success. Individual factors can be divided into entry conditions (e.g., diversity factors, such as social background, but also previous school education and knowledge or grades in school), context factors (e.g., such as employment or caring responsibilities) and factors of the individual study process (e.g., performance factors, learning behaviour and motivation [7][8]. Although research on the factors influencing academic success varies, all levels significantly impact outcomes. However, performance data (academic or prior school achievements) consistently show the strongest effects in multivariate models.

Empirical analyses in German-speaking universities indicate that student employment significantly affects success [6][10]. Additionally, entry requirements and higher education entrance qualifications, which include diversity factors like age, social, and educational background, also play crucial roles [7][12]. However, a desired steering effect on efficiency and quality can only be achieved if at least all the central influencing factors on the target variable to be steered are recorded [9]. In addition, a change in the central factors must be within the sphere of influence of those responsible at the university. Otherwise, they must be modelled as context factors to be considered, as in performance or quality evaluations, so that corresponding steering attempts can achieve the desired effects [10].

III. METHODOLOGY

In this section we describe our institutional research questions of our case, and the methods used to answer them and to validate the results. Furthermore we describe the sample used.

A. *Institutional research questions*

The case study we present in the following aims to better understand academic success at the beginning of studies and especially focuses on developing measures specifically designed for STEM programs. Regarding STEM, Austrian universities must increase the number of students and more specifically the number of women in those fields belonging to STEM disciplines. Therefore, the investigation of the influence of diversity indicators on student success is of increasing importance. The transition from school to university, the start of studies and academic performance in the first semester, i.e., arrival at the university, are central to the students' commitment to the university. The target audience are decision-makers at universities and, indirectly, students who are expected to benefit from the measures derived. The following questions are addressed:

1) What is the proportion of students enrolled in a given semester who, after one academic year,…

- *a)* …have successfully completed all courses suggested?
- *b)* have not taken any courses?
- *c)* have partially completed courses?

2) To what extent can differences be identified between different student groups? Characteristics include: gender, age at entry, university entrance qualification (school type), foreign language, immigrant background and parents' educational background.

3) Which characteristics are most relevant in predicting success in the HEI?

4) What role do interaction effects between combinations of characteristics play?

For investigating questions 1) and 2), common descriptive analysis may be used, whereas question 3) and 4) can only be investigated by using model-based approaches like classical frequentist analysis or AI-based approaches like machine learning models. Within this paper we focus on demonstrating how algorithm-based approaches may help in analysing more complex institutional research questions and therefore help to enhance the efficiency and quality of study programmes.

B. *AI Methods modelling student success*

We used a variety of Machine-Learning methods to predict student success: A General Linear Model (GLM) with a logistic link function (logistic regression), Random Forests (RF), Boosted Logistic Regression (LogitBoost), Support Vector Machine (SVM) and Gradient Boosting Machine (GBM). Advantages of these models are that they often have more predictive power in comparison to classical frequentist approaches like classical regression analysis.

As questions 1) and 2) are primarily descriptive and serve as preliminary analysis, followed by more advanced, AI-based investigations. To investigate question 3) and 4) we operationalized success by a dichotomous outcome variable (all courses in the entrance phase passed vs. at least one course failed). Logistic regression analysis or a method from the field of machine learning such as random forest can be used to compare the strength of the factors and model interaction effects. Experience has shown that these analyses require more explanation and moderation of interpretation to stakeholders (what do the results mean and what do they not mean) than purely descriptive analyses. That is why we developed graphical visualizations which are easy to interpret. The better interpretation of Odds-Ratio in comparison to importance values led us to report the influence of the variables in terms of odds ratio [13]. All variables related to diversity collected at the university were included in the model as predictors. Which variables should be included depends on the research questions to be answered, theoretical considerations, and the data availability of the respective university. The definition and categorization of the variable values should also be justified by their content. In our case study, we were investigating STEM programs, which is why it is of particular interest to determine whether a science-oriented school type has a positive impact on academic success and, subsequently, whether this positive effect differs between genders. In addition to gender and school type, other variables included in the model were parents' educational background (university vs. no university), foreign language background (yes vs. no), and age.

C. Data

Administrative and exam data of students from eight different study programs (N = 2532) belonging to STEM disciplines at a large Austrian university were used. This analysis was carried out within a working group specifically set up to identify potential adverse conditions for specific student groups and to develop measures to support students.

D. Model Training and Validation Procedure

To evaluate the predictive performance of various classification algorithms, the dataset was first randomly split into a training set (70%) and a validation set (30%), using a non-replacement sampling approach. Cases containing missing values were subsequently removed from both subsets to ensure complete-case analysis. A repeated 10-fold cross-validation procedure with three repetitions was employed to tune and evaluate model performance. For internal resampling during training, 25 resamples were generated based on the outcome variable using stratified sampling. Class probabilities were computed, and all model predictions from resampling iterations were saved to allow for further ensemble and performance analysis.

Model training was performed using the caretList function from the caret Ensemble package [15]. The outcome variable was modeled as a function of all available predictors. Prior to model fitting, features were standardized (i.e., centered and scaled). The following classification algorithms were included in the model list: Logistic regression with a binomial link function (glm), Random forest (rf), LogitBoost (LogitBoost), Support vector machine with linear kernel (SVM) and Gradient boosting machine (GBM). The reasons for selecting these models, in addition to the availability within the respective R package, were a mix of linear and non-linear and simple and complex methods, as well as different ensemble techniques allowing for a comparison of methods with different strengths and assumptions, This ensemble approach allowed for a fair comparison of different learners under the same resampling strategy, enabling both individual model assessment and potential ensemble modelling in subsequent steps.

IV. RESULTS

Research questions 1) and 2) are purely descriptive questions and can be answered without making use of machine learning models. For demonstration reasons, we are inserting a descriptive plot that was used in the communication with stakeholders for each programme separately. Here, we are displaying all programs together.

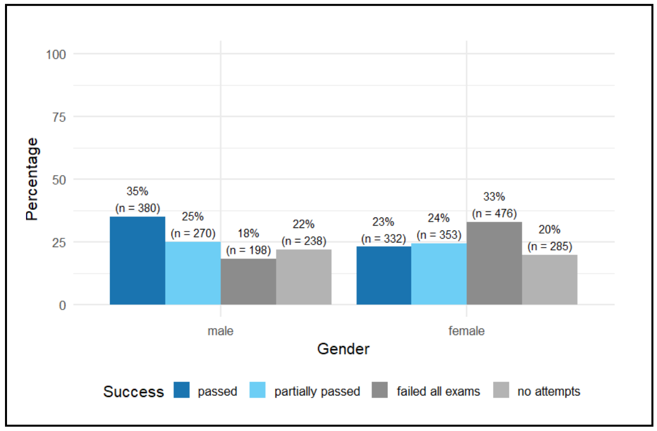

Figure 1. Descriptive differences: Gender

Descriptively speaking, only 35 percent of males and 23 percent of females have passed all entry courses within one year (Figure 1). Students of diverse genders were not represented in the graphic because there were too few observations to include them. Furthermore, it can be demonstrated, that female students more often have negative attempts only, but are attempting exams at the beginning equally often. By looking at the relative differences a gender gap exists in the eight programs analysed together.

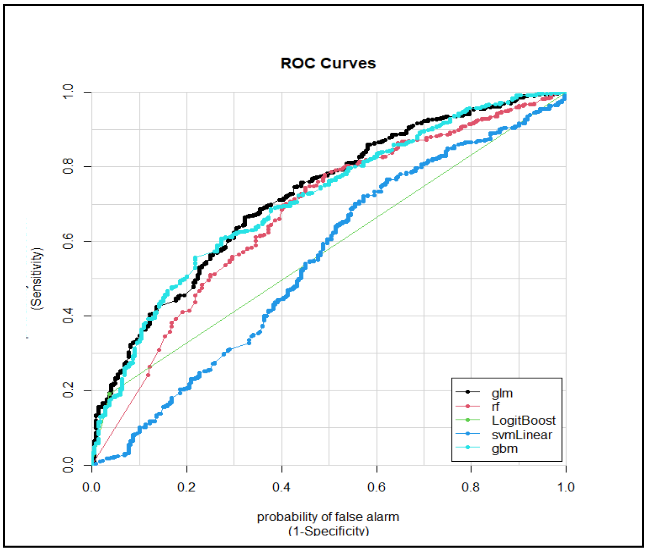

Figure 2. ROC curves

To investigate the predictive power of diversity factors fit indices of the mentioned machine learning models were compared and the ROC-curves were compared (Figure 2: ROC curves). ROC curves are often used to compare the performance of multiple models. The model with the curve closest to the top-left corner (or the highest AUC) is preferred. The visualisation shows that (boosted) logistic regression models (LogitBoost) were performing best in our specific case. A comparison between different models and their fit indices is displayed in Table 1. All the models have a poor to medium fit. The glm with the logistic link function (logistic regression) and the Gradient boosting machine (GBM) fitted best, since their AUC and Kappa values were larger.

TABLE I. MODEL FIT INDICES OF FULL MODELS

Fit of full models	Fit-Indices		
	AUC	*Mean accuracy*	*Mean Kappa*
glm	0.72	0.72	0.14
rf	0.63	0.72	0.03
LogitBoost	0.58	0.70	0.05
SVM	0.55	0.72	0.01
GBM	0.71	0.72	0.14

To investigate the influence of specific diversity indicators the results of the logistic regression (glm) were further analysed: McFadden's R2 of the full model of the logistic regression was 0.11, indicating that the model explains approximately 10.5% of the variance in the outcome compared to the null model.

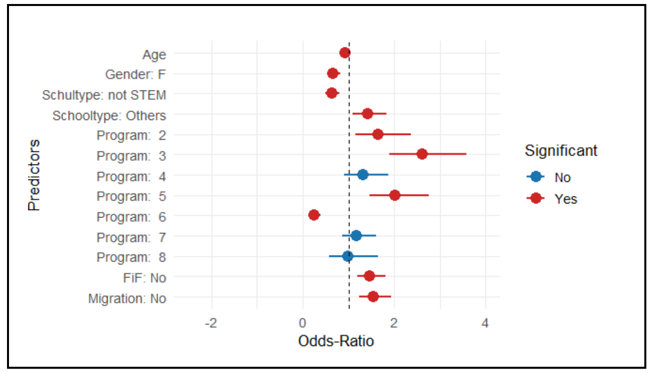

Figure 3. Odds-Ratio Logistic Regression – Full Model

The likelihood-ratio based measure $R^2_{ML}$was 0.118, and the Cragg-Uhler (Nagelkerke) R^2 was 0.169, suggesting modest explanatory power. To answer research question 2) and 3) Odds Ratio of the full model revealed differences between the study programmes on the probability to successfully complete the first courses (Figure 3). These effects were larger than the effects of the diversity indicators. Nevertheless, older students, female students and students attending schools with no science or mathematical background had a lower probability of successfully completing the entrance phase. Students with parents holding a university degree (First in Family: FiF: No) and students with no migration background had better odds to pass all courses. The interaction effect between school type and gender (research question 4) was not significant in the logistic regression model and was therefore not visualized. When modelling a specific study programme (programme 6) only, for this specific programme only age was a significant predictor with a small and positive effect: In this programme older students were having larger probabilities to successfully complete their first year (Figure 4).

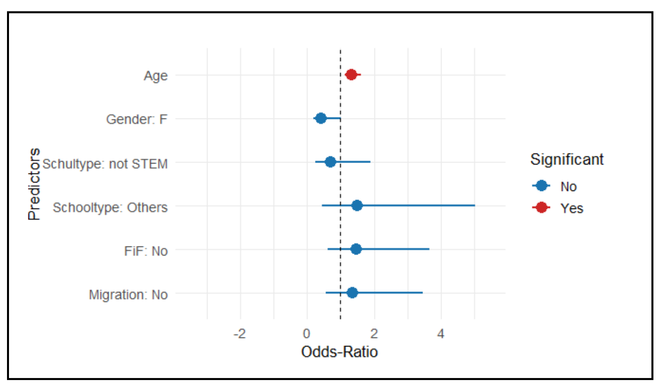

Figure 4. Odds-Ratio Logistic Regression – Full Model programme 6

 ISBN: 978-1-68558-285-2

V. CONCLUSION AND FUTURE WORK

Using the example of passing the first entrance phase, our results show that it is possible to predict study success with AI approaches, especially machine learning. Our models including diversity factors as predictors had a relatively small but measurable impact on student success. On the other hand, models without diversity indicators do have often less explanatory power than models including diversity indicators. This aligns with existing literature that performance-based indicators have strong predictive power but does not mean that the usage of diversity indicators is not relevant for HEIs, because it increases the area under the receiver operating curve and therefore the sensitivity and specificity of the model.

Which diversity indicators impact student success, seem to depend on the study programme and the diversity of the students in the sample, as we assumed based on former analyses [14]. Although data from study programmes belonging to STEM fields were used, findings were different in the respective programmes and diversity indicators were having different impacts on success. This highlights that such an analysis must be carried out individually and consequences and measures must be developed carefully for each individual study programme. AI models can help in analysing and interpreting the data, but which measures are derived are programme and context specific. Furthermore, and especially by comparing the descriptive analysis to the outcomes from models, relevant differences between different student groups exist. If those are not considered, biases by analysing the data and especially when deriving measures to support students or using the analyses for quality insurance processes may result. The analysis of the specific programme (Figure 4) revealed that older students have lower chances to pass the entry phase. A measure derived could be offering a bridging course in which relevant content can be repeated. Another measure that could be derived is the introduction of a mentoring program for older students, which might involve positive discrimination and would need to be carefully considered by experts with knowledge in higher education (research). At this point, the use of AI reaches its limits, and additional human resources are required to derive and implement measures from the results.

Since similar results are now evident in a relatively large number of such analyses, it can be assumed that ineffective forecasts and the measures derived from them to promote student success can only be reliably avoided if diversity indicators are systematically taken into account in AI approaches; and if influencing factors are examined specifically for each study programme. This is therefore a necessary prerequisite for improving the quality and efficiency of study programmes with the help of empirical analyses and AI-supported study success predictions. In addition to predicting student success and dropout, HEIs are also interested in developing study programmes which allow a diverse student body to successfully master their programs. If universities are aiming especially for the entrance phase to let a diverse student body successfully complete, the use of diversity indicators to predict and model study success may be of particularly high relevance. Therefore, investigating the specific effect of these factors, when controlling performance-based indicators is important. This can also help prevent such forecasts from being misused to attract students who are more likely to succeed in their studies from the outset, without the university having to contribute much to this - as is sometimes reported as an unintended effect of student success analyses in other countries [16]. However, a possible prioritisation of study programmes for analyses with limited resources can be derived from the experience that with greater heterogeneity or diversity of students, greater effects of diversity indicators can potentially be expected. Nevertheless, some limitations must be noted here, that are in the same time desiderata for future analyses: Contextual factors, data availability, and the choice of variables play an important role. In the social sciences, it is difficult to find perfectly uncorrelated variables and to create complete models. Therefore, it is important to interpret one's own results with caution and to consider blind spots. Furthermore, predictive performance and the identified importance of the predictors may vary depending on the operationalisation of student success. That is why this study may be repeated with other operationalisations, e.g., ECTS points [13]. In addition, further analyses would be useful to investigate the inner workings of the models, e.g., structural equation modelling (SEM) and hierarchical models; and consideration of other variables known to influence academic success (such psychological constructs as self-concept). Furthermore, the proposed analysis should be extended to further different types of degree programmes and further HEI in order to answer the question of the extent to which the results presented here can be generalised university-, state- or nationwide if necessary. Third, and finally, it would make sense not only to analyse the effects of diversity on academic success, but also to develop a monitoring system for diversity factors to observe them professionally and, if necessary, to prioritise (more or less automatically) degree programmes with significant changes for analysis.

In the longer term, such models could also provide systemic benefits beyond their concrete benefits for individual study programs and universities: They are good examples of linkage possibilities with empirically informed teaching development at universities, which could contribute overall to a culture of stronger evidence orientation in decision-making at universities and thus to an even more systematic promotion of academic success. This can increase not only the quality but also the efficiency of study programmes. Future work therefore should deal with the (more) systematic derivation and development of possible recommendations for action. For example, this could apply to the fact that we have identified a gender difference and have communicated this to the study programme and responsible stakeholder. The following questions arise for the next steps: What decisions can be taken, and under what conditions and contexts? [3] Which departments do the ‘translation work’ here, which are (only) responsible for the analysis and how does this interact within the university? How can this be organized as a systematic process? AI models can, therefore, help to increase the efficiency and quality of study programmes, but must be applied systematically and indicators should not only be chosen by their predictive power, but also in terms of theory and goals defined by stakeholders. Measures must be derived and implemented by humans with experience in higher education. In addition, such models could not only provide more targeted support for certain student groups if goals and corresponding measures are adequately formulated. Rather, this could also be done for universities that (must) deal with their promotion in a special way due to their profile, their geographical location or their recruitment potential, for example.

ACKNOWLEDGMENT

We would like to thank our colleague Ingrid Purner for creating the opportunity to analyse the data and Robin Gleeson for the useful comments and ideas on our work.

This paper benefited from language refinement assistance provided by ChatGPT (April 2025 version), an AI language model developed by OpenAI. The AI was used solely for language improvement, and the authors reviewed and confirmed the final content.

REFERENCES

[1] M. Lübcke, M., J. Schrumpf, K. Schurz, F. Seyfeli-Özhizalan, T. Thelen, K. Wannemacher, & F. Weber, (eds.)"Mit digitalen Studienassistenzsystemen durchs Studium. Zeitschrift für Hochschulentwicklung", ZFHE, vol. 19, no. 4, pp. 9-13, 2024.

[2] K. R. Isett & Hicks,,. "Pathways From Research Into Public Decision Making: Intermediaries as the Third Community", Perspectives on Public Management and Governance, vol. 3, no. 1, pp. 45-58, 2020.

[3] A. Wegner, R. Krempkow, and C. Thiedig, "From evidence to impact? Practices of evidence use for quality management at German Higher Education Institutions," in *46th Annual EAIR Forum 2024, University College Cork*, Cork (Ireland), Aug. 2024.

[4] S. Falk, J. Kercher, and J. Zimmermann, "Internationale Studierende in Deutschland: Ein Überblick zu Studiensituation, spezifischen Problemlagen & Studienerfolg," *Beiträge zur Hochschulforschung*, vol. 44, no. 2-3, 2022.

[5] R. Krempkow, O. Vettori, and I. Buß, Eds., Studierbarkeit und Studienerfolg, Zeitschrift für Hochschulentwicklung – ZFHE, vol. 17, no. 4, 2021.

[6] V. Tinto, *Leaving College: Rethinking the Causes and Cures of Student Attrition*, 2nd ed. Chicago, IL: University of Chicago Press, 1993.

[7] R. Krempkow, "Determinanten der Studiendauer. individuelle oder institutionelle Faktoren? Sekundärdaten-analyse einer bundesweiten Absolvent(inn)enbefragung," *Zeitschrift für Evaluation*, vol. 19, no. 1, pp. 37–63, 2020.

[8] U. Heublein, J. Ebert, C. Hutzsch, S. Isleib, R. König, J. Richter, and A. Woisch, Zwischen Studienerwartungen und Studienwirklichkeit. Ursachen des Studienabbruchs, beruflicher Verbleib der Studienabbrecherinnen und Studienabbrecher und Entwicklung der Studienabbruchquote an deutschen Hochschulen. Hannover: Deutsches Zentrum für Hochschul- und Wissenschaftsforschung, 2017.

[9] E. Grande, D. Jansen, O. Jarren, A. Rip, U. Schimank, and P. Weingart, Eds., *Neue Governance der Wissenschaft: Reorganisation, Externe Anforderungen, Medialisierung*. Bielefeld: transcript, 2013.

[10] R. Krempkow, "Can Performance-based Funding enhance Diversity of Higher Education Institutions?" in *Diversity and Excellence in Higher Education: Can the Challenges be Reconciled?*, R. Pritchard, M. Klumpp, and U. Teichler, Eds. Amsterdam: Sense, 2015, pp. 231–244.

[11] R. Krempkow and R. Kamm, "Leistungsbewertung unter Berücksichtigung institutioneller Diversität deutscher Hochschulen: Ein Weg zur Förderung von Vielfalt?" in *Diversity und Hochschule. Teilhabebarrieren und Strategien zur Gestaltung von Vielfalt*, U. Klein and D. Heitzmann, Eds. Weinheim: Juventa, 2012, pp. 164–181.

[12] B. Thaler, N. Schubert, A. Kulhanek, N. Haag, and M. Unger, *Prüfungsinaktivität in Bachelor- und Diplomstudien an Universitäten*. Vienna: Institut für Höhere Studien – Institute for Advanced Studies, 2021.

[13] L. Bartok, J. Spörk, R. Gleeson, M. Krakovsky, and K. Ledermüller, "Anwendung statistischer und Machine-Learning-Methoden für Fragestellungen zu Studienerfolg - Erfahrungen in den Projekten ‚Learning Analytics - Studierende im Fokus' und ‚PASSt - Predictive Analytics Services für Studienerfolgsmanagement'," Waxmann Verlag, 2024.

[14] M. Krakovsky, R. Krempkow, L. Bartok, K. Ledermüller, and J. Spörk, "Breaking bias: Measurements, potentials and limitations for modelling study success by performance and diversity factors," *ResearchGate*, Apr. 2024. [Online]. Available: www.researchgate.net/publication/385905974

[15] Z. Deane-Mayer and J. W. Knowles, *caretEnsemble: Ensembles of caret models*, R package version 2.0.1, 2023. [Online]. Available: https://CRAN.R-project.org/package=caretEnsemble

[16] A. Matveeva, "Performance-based funding in higher education: a meta-narrative review and renewed research agenda proposal," *Tertiary Education and Management*, [Online first], 2025. Available: https://doi.org/10.1007/s11233-025-09151-y

[17] J. Castro Galvao, F. Tucker, and P. Attewell, "Bending the Curve: Institutional Factors Associated with Graduation Rates," *Higher Education Policy*, vol. 37, pp. 278–302, 2024.

[18] K. J. Dougherty, S. M. Jones, H. Lahr, R. S. Natow, L. Pheatt, and V. Reddy, "Looking inside the black box of performance funding for higher education: Policy instruments, organizational obstacles, and intended and unintended impacts," *The Russell Sage Foundation Journal of the Social Sciences*, vol. 2, no. 1, pp. 147–173, 2016, doi:10.1353/rus.2016.0007

ISBN: 978-1-68558-285-2