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

Special Track - PGAI: Productivity Gains by AI – Myth or Measurable?
DIGITAL 2025, Advances on Societal Digital Transformation, July 06 - 10, 2025 – Venice

Larissa Bartok[1] & René Krempkow[2,3]

[1] University of Vienna, Vienna, Austria

[2] IU International University of Applied Sciences, Berlin, Germany

[3] Brandenburgische Technische Universität Cottbus-Senftenberg, Cottbus, Germany

Content

- Introduction
- Student success and influencing factors
- Methodology
- Results
- Conclusion and future work

Introduction

- Universities collect extensive student data (e.g., admission, examination) that can be leveraged for evaluation and evidence-based development of study programmes.
- Using evidence about student success and in particular making use of AI-based approaches is an emerging field [2][3].
- Aims of our study:
  - Using AI-based approaches to answer institutional research questions within a specific context (case) at a large Austrian university
  - Case: Study entrance phase of STEM fields (N = 2532)
  - Objectives: Analyze student success within the entrance phase and identify barriers for non-traditional students during the entry phase; development of measures
- Efficiency within this context: achieving program objectives while optimizing resources (e.g., timely degree completion, graduation rates, alignment with labor market demands)

Measuring student success and influencing factors

- **Common criteria**: Success rates, passed exams, grades, credit points, and graduates within the standard study period [4][5]
- **Performance-Based Funding** (PBF): models use metrics like success rates, study duration, and graduation within the standard period (+2 semesters) [7]
- **Mixed success in achieving desired outcomes** [10][16][17][18].
- **Types of factors** [7]:
  - individual factors: E.g. diversity factors (social background, prior education, school grades),
  - context factors: Employment, caring responsibilities,
  - study process factors: Performance, learning behavior, motivation [7][8]
  - institutional factors: Policies, support systems, and structural conditions.

Measuring student success and influencing factors

- **Performance data (e.g., prior academic achievements) consistently show the strongest effects** in multivariate models
- **Key findings** from German-speaking universities:
  - **Student employment** significantly impacts success [6][10].
  - **entry requirements and diversity factors** (e.g., age, social, and educational background) play crucial roles [7][12].
- **Steering and Context Factors**:
  - Effective steering requires capturing all central influencing factors on success [9].
  - **Factors outside the university's control must be modeled as context factors** to ensure meaningful evaluations and interventions [10].

Methodology

Institutional research questions

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? (gender, age at entry, university entrance, foreign language migration 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?

Methodology
Data

- Administrative and exam data of students from eight different study programs ($N = 2532$) belonging to **STEM disciplines at a large Austrian university**
- **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**
- Importance of the entrance phase: Critical for onboarding students from diverse backgrounds, fostering belonging, and ensuring continuation [6]
- Before further analysis data was retrieved from the central data warehouse, prepared and checked for inconsistencies

Methodology

AI methods modelling student success and operationalization

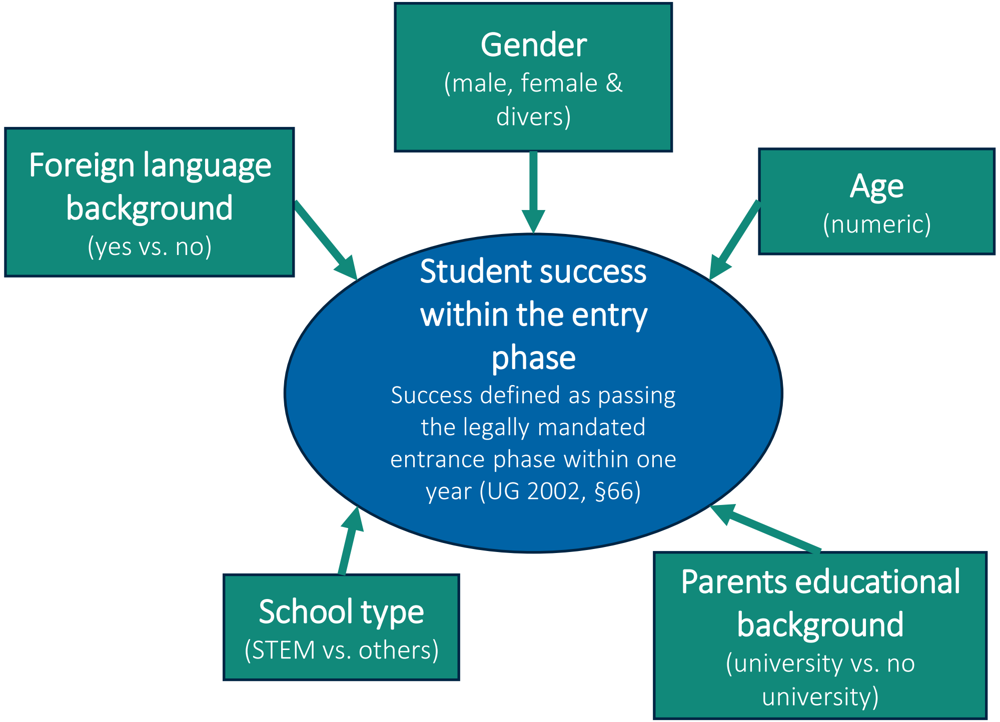
Machine Learning Models:

- General Linear Model (GLM) with logistic regression
- Random Forests (RF)
- Boosted Logistic Regression (LogitBoost)
- Support Vector Machine (SVM)
- Gradient Boosting Machine (GBM)

Advantage: Greater predictive power compared to classical regression methods.

Stakeholder Communication:

- Results require explanation and moderation for stakeholders
- Graphical visualizations developed for easier interpretation
- Influence of variables reported using Odds Ratios for better clarity



universität wien | iu INTERNATIONALE HOCHSCHULE | b-tu Brandenburgische Technische Universität Cottbus - Senftenberg

Methodology
Model training and validation procedure

- **Data Splitting**:
  - Dataset split into training set (70%) and validation set (30%) using non-replacement sampling
  - Cases with missing values removed for complete-case analysis
- **Cross-Validation**:
  - Repeated 10-fold cross-validation with 3 repetitions
  - 25 stratified resamples generated during training based on the outcome variable
- **Model Training**:
  - Models trained using the caretList() function from the *caretEnsemble* package [15]
  - Features standardized (centered and scaled) before fitting

Results

Research questions 1 and 2

- **Research questions 1 and 2** answered by descriptive methods:
  - In-depth analysis by visualizing different patterns shows e.g. a descriptive gender gap:

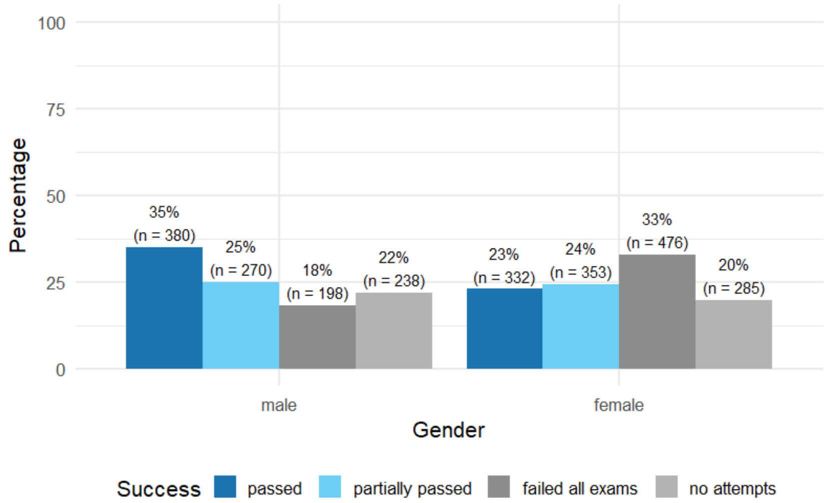


Figure 1 – Example of descriptive analysis

universität wien

INTERNATIONALE HOCHSCHULE

Brandenburgische Technische Universität Cottbus - Senftenberg

Results

Research questions 3 and 4 – Full model

- **Comparison of Predictive Power**:
  - Fit indices and ROC curves used to evaluate model performance.
  - ROC curves: Preferred model has the curve closest to the top-left corner (highest AUC).

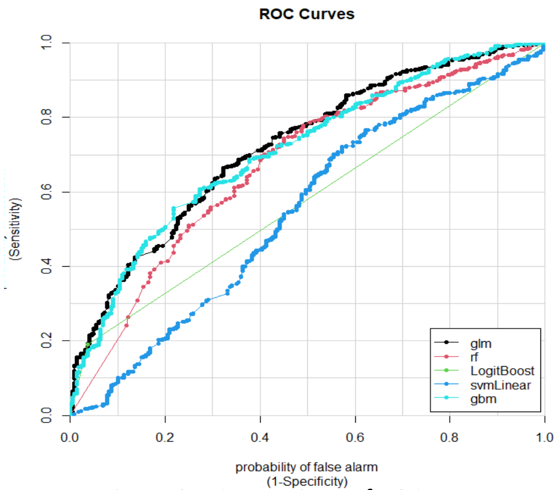


Figure 2 – Comparison of ROC curves

Results

Research questions 3 and 4 – Full model

- **Comparison of Predictive Power**:
  - Fit indices and ROC curves used to evaluate model performance.
  - ROC curves: Preferred model has the curve closest to the top-left corner (highest AUC).

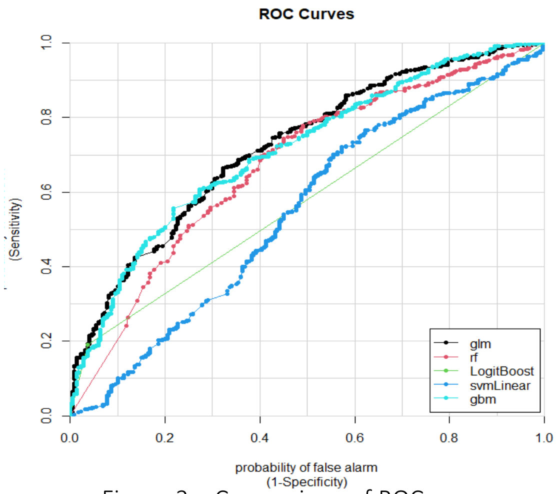


Figure 2 – Comparison of ROC curves

- **Key Findings**:
  - Logistic regression (GLM) and Gradient Boosting Machine (GBM) had the best overall fit (highest AUC and Kappa values)
  - Overall model performance ranged from poor to medium fit

Table 1 – Comparison of model indices

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

universität wien

INTERNATIONALE HOCHSCHULE

Brandenburgische Technische Universität Cottbus - Senftenberg

Results

Research questions 3 and 4 – Full model

- Influence of Diversity Indicators investigated by further analysis of logistic regression (glm)
  - Model Explanatory Power:
    - McFadden's R^2 = 0.11; Likelihood-ratio R^2_{ML} = 0.118; Nagelkerke R^2: 0.169 (modest explanatory power)
  - Key Findings (Odds Ratios):
  - Programme Effects: Differences between study programs had a larger impact on success than diversity indicators
  - Diversity Indicators:
    - Lower success probability: Older students, female students, students from non-science/math school backgrounds.
    - Higher success probability: Students with university-educated parents (FiF: No), students with no migration background
    - Interaction Effects: No significant interaction between school type and gender

Results

Research questions 3 and 4 – all programmes

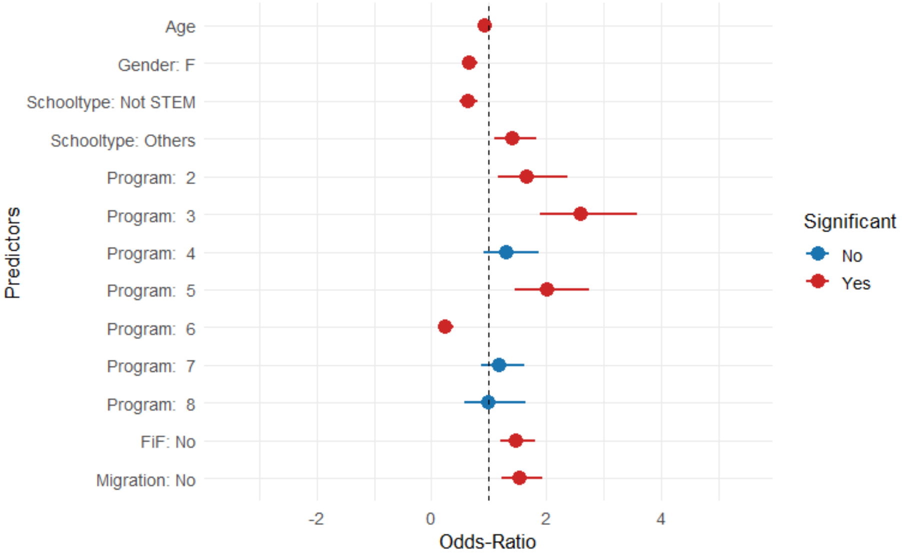


Figure 3 - Odds-Ratio – All programmes

universität wien
iu INTERNATIONALE HOCHSCHULE
b-tu Brandenburgische Technische Universität Cottbus - Senftenberg

Results

Research questions 3 and 4 – all programmes vs. programme 6 only

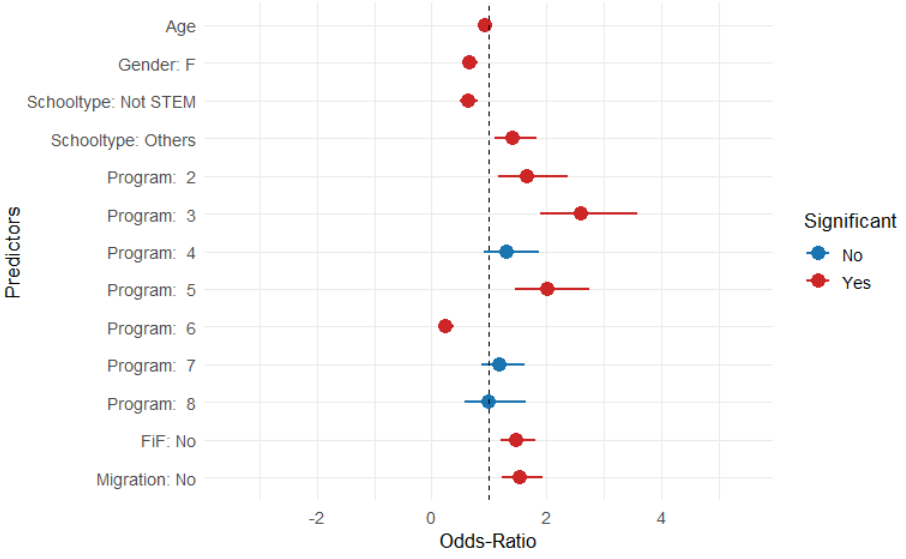


Figure 3 - Odds-Ratio – All programmes

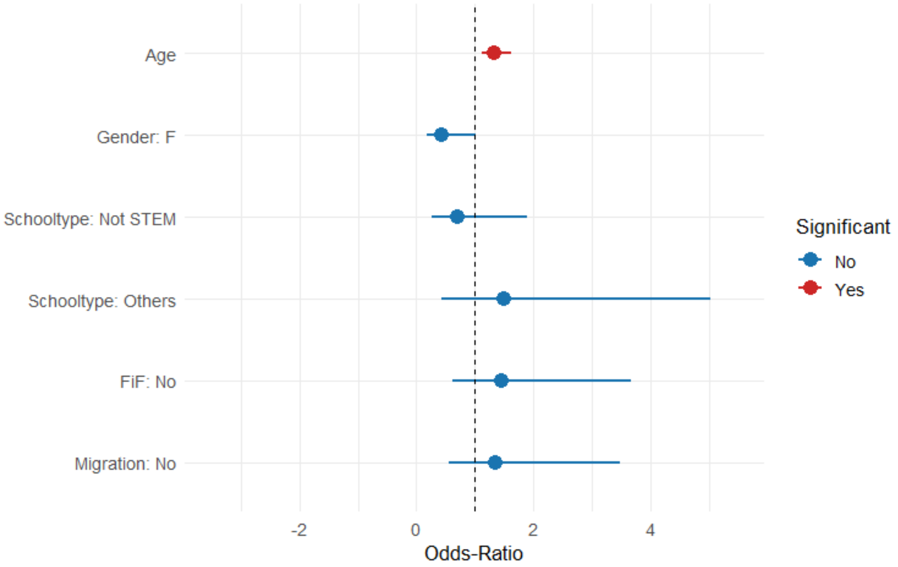


Figure 4 - Odds-Ratio – Programme 6 only

Conclusion and future work

- **Predicting study success**:
  - AI and machine learning models **can predict study success**, with diversity indicators improving model sensitivity and specificity.
  - Models without diversity indicators often have **less explanatory power**, aligning with existing literature [14].
  - **Diversity indicators: Impact varies by study program and student sample diversity**, requiring tailored analyses and measures.

  → **Example measures**: Bridging courses for older students or students from different fields; mentoring
- **Systematic use of diversity indicators**:
  - Essential for improving quality and efficiency of study programs: Ensures meaningful interventions
  - Particularly relevant for diverse student bodies and entrance-phase success

Conclusion and future work

- **Contextual factors, data availability, and variable selection affect results**
- Future analyses:
  - Extension to **other degree programs** and institutions for generalizability
  - Develop **diversity monitoring systems** for ongoing evaluation
- AI models **can support evidence-based teaching development and decision-making at universities** and therefore, can increase quality and efficiency of study programmes by: higher success rates, higher percentage of timely degree completion
- **Systematic processes are needed** to translate findings into actionable measures
- **Collaboration** between departments and stakeholders is critical
- **Models should align with theoretical goals and stakeholder-defined objectives, not just predictive power**

Thanks a lot for the attention!
We are looking forward to your feedback and questions!

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

universität wien

iu INTERNATIONALE HOCHSCHULE

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? Sekundärdatenanalyse 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.

universität wien

INTERNATIONALE HOCHSCHULE

Brandenburgische Technische Universität Cottbus - Senftenberg

References

[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

Additional material