

Automatisiertes KI-Feedbacksystem zur Unterstützung individueller Lernprozesse

Konzeption und Anwendung im Videoanalysetool ViviAn

Marc Bastian Rieger, Jürgen Roth

Abstract: *Im Rahmen des Videoanalysetools ViviAn der Rheinland-Pfälzischen Technischen Universität Kaiserslautern-Landau (RPTU) soll Künstliche Intelligenz (KI) in der Lehrkräftebildung eingesetzt werden, um diagnostische Kompetenzen bei Lehramtsstudierenden (weiter) zu entwickeln. Das zentrale Element ist ein gestuftes Feedbacksystem, das auf dem Response to Intervention (RTI) Modell basiert und den Studierenden strukturiertes Feedback auf verschiedenen Interventionsstufen bietet. Der Einsatz von Cascading AI (CAI) ermöglicht eine präzise und transparente Analyse von Freitextantworten, die personalisiertes Feedback liefert und die Erkennung sowie Korrektur von Fehlern unterstützt, wodurch der Lernprozess individuell gefördert und verbessert werden soll.*

As part of the video analysis tool ViviAn at the Rhineland-Palatinate Technical University of Kaiserslautern-Landau (RPTU), Artificial Intelligence (AI) is to be used in teacher training to (further) develop diagnostic skills in pre-service teachers. The central element is a tiered feedback system based on the Response to Intervention (RTI) model, which will provide students with structured feedback at different levels of intervention. The use of Cascading AI (CAI) enables precise and transparent analysis of free-text responses, delivering personalized feedback and supporting the detection and correction of errors, thereby fostering and improving individual learning processes.

Keywords: *Videovignetten; KI; Didaktik; Feedback; Lernplattform; ViviAn; video vignettes; AI; didactics; feedback; learning platform*

1. Einleitung

In der dynamischen Landschaft der Hochschulbildung ist personalisiertes Feedback von großer Bedeutung, um Lernende in digitalen Bildungsumgebungen zu unterstützen und deren akademische Entwicklung zu begünstigen (Cavalcanti et

al., 2021). Die Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau (RPTU) nutzt in diesem Zusammenhang das digitale Lernwerkzeug *ViviAn* (<http://vivian.projects.rptu.de>) in der Lehrkräftebildung, um durch den Einsatz von Videovignetten die Fähigkeiten, die zur Diagnose von Lernprozessen von Schülerinnen und Schülern notwendig sind, zu entwickeln und zu fördern (Scherb et al., 2023).

ViviAn ist eine webbasierte Lernplattform, auf der zahlreiche Videovignetten zur Verfügung stehen. Die Vignetten sind videographierte Sequenzen von Schülergruppenarbeitsphasen aus dem Mathematik-Labor »Mathe ist mehr«, in denen echte Lernsituationen gezeigt werden. Die Wirksamkeit des Tools wurde in diversen Projekten und Dissertationen evaluiert und bestätigt (Enenkiel, 2022). Diagnosekompetenz ist eine wesentliche Voraussetzung für eine gezielte und individuelle Förderung von Lernenden (Enenkiel et al., 2022). Um die verschiedenen Fähigkeiten, die zur Ausbildung diagnostischer Kompetenz notwendig sind, zu trainieren, werden videographierte Lernsequenzen analysiert und spezielle Videosequenzen extrahiert sowie aufbereitet (Enenkiel et al., 2022). Diese Sequenzen werden mit zusätzlichen Materialien zur Lernsituation (Schülerbearbeitungen, Arbeitsaufträge, Arbeitsmaterialien, ...) angereichert und zusammen mit passgenauen Diagnoseaufträgen in einzelnen Vignetten auf ViviAn zur Bearbeitung durch Lehramtsstudierende zur Verfügung gestellt. Um die Ergebnisse und individuellen Lernfortschritte von Studierenden in großen Veranstaltungen und Vorlesungen zu evaluieren und individuelles Feedback zu ermöglichen, kann Künstliche Intelligenz (KI) eingesetzt werden.

Zur Implementierung von KI wird die Konzeption und Implementation eines gestuften Feedbacksystems notwendig, das sich am *Response to Intervention* (RTI) Modell orientiert und eine dreistufige Struktur aufweist: (1) universeller Unterricht, (2) gezielte Intervention und (3) intensive Intervention (Hartmann & Müller, 2009). Im Folgenden wird dieses Konzept vorgestellt.

2. KI in Bildungsprozessen

Automatisiertes Feedback hat das Potenzial, die Leistung von Studierenden in 65 % der Fälle zu steigern (Cavalcanti et al., 2021) und erlaubt es, tiefere Einblicke in das Verständnis und die kognitiven Prozesse der Lernenden zu erhalten. Die Metaanalyse zeigt über verschiedene Beispiele hinweg, dass automatisiertes Feedback in verschiedenen Anwendungsbereichen und Disziplinen zu positiven Effekten führt. Krusche & Seitz (2018) berichten davon, dass die Studienteilnehmenden automatisiertes Feedback während der Aufgabenbearbeitung zu 90 % als hilfreich empfanden. Eine systematische Literaturanalyse offenbarte, dass die Anwendung von solchen datenbasierten und empirischen lernanalytischen Interventionen

derzeit noch begrenzt ist (Wong & Li, 2020). Wang et al. (2022) argumentieren, dass der Einsatz von maschinellem Lernen in Feedbacksystemen die Lernleistung verbessern kann, doch müssen diese in automatisierte Systeme überführt werden, um Forschung mit umfangreichen Stichprobengrößen zu ermöglichen. Weiterhin existiert ein Mangel an Feedbacksystemen die nicht nur auf die Identifizierung von Fehlern, sondern auf die Behebung von Defiziten ausgerichtet sind (Keuning et al., 2019).

Bei Betrachtung der Forschungslandschaft im Bereich des automatisierten Feedbacks wird deutlich, dass die Analyse von Freitextantworten bei umfangreicheren Studierendengruppen, wie sie beispielsweise in Vorlesungen anzutreffen sind, auch in der heutigen Zeit eine signifikante Herausforderung darstellt (Bernius et al., 2022). Trotz fortschreitender Entwicklungen in der Künstlichen Intelligenz stellt die automatisierte Evaluation von Freitextfragen weiterhin ein komplexes Problemfeld dar (Qiao & Hu, 2023). Verschiedenartige Ansätze existieren zwar, doch eine vollständig automatisierte und präzise Analyse umfangreicher Texte bleibt bislang ein unerreichtes Ziel (Bernius et al., 2022). Dies liegt insbesondere daran, dass holistische Systeme, die fähig sind, sämtliche Antwortvarianten korrekt zu erfassen und zu bewerten, konzeptionell komplex und technisch herausfordernd sind (Bai & Stede, 2022). Nichtsdestotrotz wurde die Bedeutung von unmittelbarem Feedback im Bildungskontext wiederholt hervorgehoben (Marwan et al., 2020). Der zunehmende Einsatz von Natural Language Processing (NLP) könnte in diesem Problemkontext als Schlüsseltechnologie fungieren (Bai & Stede, 2022).

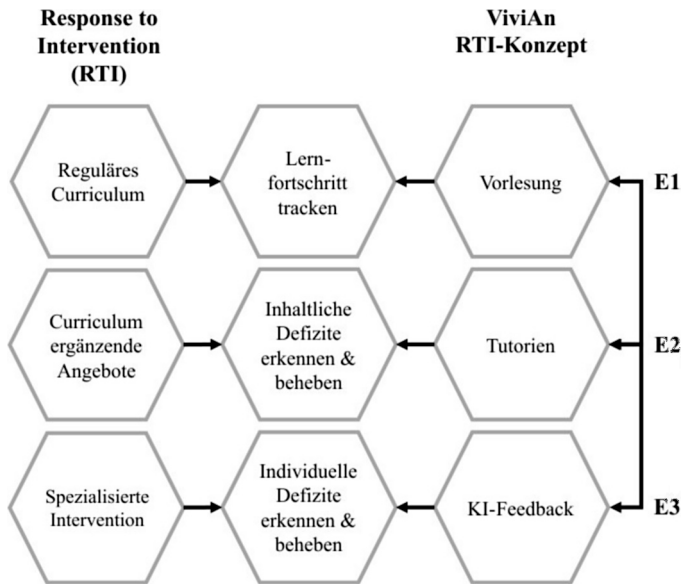
Die Integration von Künstlicher Intelligenz (KI) in Bildungsprozesse hat das Potenzial, die Art und Weise, wie wir lehren und lernen zu verändern und wirkt sich darauf bereits aus (Chen et al., 2020). KI-Technologien bieten die Möglichkeit, Lerninhalte zu personalisieren und den Lernprozess auf die Bedürfnisse einzelner Studierender zuzuschneiden. Insbesondere das so mögliche adaptive und individuelle Feedback kann zu verbesserten Lernerfahrungen und -ergebnissen führen (Gligorea et al., 2023). Diese Technologien können die Lernenden nicht nur durch individualisiertes Feedback unterstützen, sondern ermöglichen darüber hinaus Lehrenden, ihre Unterrichtsmethoden vielfältiger zu gestalten (Jaakkola et al., 2020).

3. KI-Einsatzkonzept in der Lehre mit ViviAn

Das dreistufige *Response to Intervention* Konzept (RTI-Konzept) (Hartmann & Müller, 2009) wird in folgenden drei Ebenen (E1 bis E3) auf das digitale Lernwerkzeug *ViviAn* übertragen (vgl. Abb. 1): Im Rahmen der ersten Ebene (E1) wendet *ViviAn* Curriculum-orientierte Evaluationsmethoden an, indem Studierende Videovignetten analysieren und korrespondierende Fragebögen ausfüllen, die darauf abzielen, erworbene Kompetenzen zu überprüfen und mögliche Leistungsdefizite frühzeitig zu

detektieren (Rietz et al., 2013). Die zweite Ebene (E2) umfasst tutorielle Sitzungen, innerhalb derer Lernende in Kollaboration mit den Lehrenden auffällige Antworten anonymisiert identifizieren und in Gruppendiskussionen reflektieren sowie verbessern bzw. erweitern. Hierfür wurde eine eigenständige interaktive Tutorien-Umgebung programmiert sowie pilotiert (Rieger et al., 2023). Die dritte Ebene (E3) stellt ein in der Entwicklung befindliches, KI-basiertes Feedbackwerkzeug dar, das die automatisierte Analyse von Freitextantworten ermöglichen soll. Fortschrittliche Algorithmen für Textverständnis und maschinelles Lernen ermöglichen es, die Freitextantworten der Studierenden kleinschrittig zu analysieren und unmittelbares, individualisiertes Feedback zu erstellen. Das automatisierte System soll dabei unterstützen, spezifische Lernschwächen aufzudecken, die in den vorherigen beiden RTI-Ebenen möglicherweise nicht erkannt oder nicht adressiert wurden. Durch eine nahtlose Integration in das bestehende Vignettensystem kann es in allen Arbeitsbereichen nutzbar gemacht werden.

Abb. 1: Übertrag des RTI-Konzepts auf die Lehre mit ViviAn und integriertem KI-Einsatz.



3.1 Cascading AI zur Feedbackgenerierung

Eines der vielfältigen Probleme des Einsatzes von großen KI-Modellen wie ChatGPT, die auf neuronalen Netzen basieren, ist die Undurchsichtigkeit der

Ergebnisproduktion (Liesenfeld et al., 2023). Um diesem entgegenzuwirken und Ergebnisse reproduzier- sowie erklärbar und transparent zu machen, wird sich in diesem Ansatz am Konzept der sogenannten *Cascading AI* (CAI) orientiert (Oppenlaender & Benjamin, 2020). CAI beinhaltet die Aneinanderreihung von verschiedenen KI-Modellen, um die individuelle Erklärbarkeit von KI-Systemen zu schärfen. Dabei wird die Interpretation von einzelnen Modell-Ergebnissen stärker mit den prozessualen Beziehungen der Modellreihung verwoben und nicht nur auf den letzten Output konzentriert (Oppenlaender & Benjamin, 2020). Bei CAI wird ein gegebener Input durch aufeinanderfolgende Modelle analysiert, wobei jedes Modell den Output des vorhergehenden nutzen kann. Essenziell bei der Generierung von automatisiertem Feedback auf Freitextfragen für den Anwendungsfall in ViviAn sind (1) die Mitteilung der Richtigkeit des Ergebnisses (Knowledge of Results; KOR) sowie (2) die Erklärung warum die gegebene Antwort richtig oder falsch ist (Response Contingent; RC) (Demaidi et al., 2018). Hieraus wird beispielsweise der Bedarf von Klassifizierungsmodellen definiert, die die Antworten der Lernenden in richtig oder falsch einteilen sowie zu Aussagen aus Expert:innenantworten zuordnen können. Durch CAI kann die Evaluierung von Freitextfragen kleinschrittig vorgenommen und es können bei Bedarf weitere Auswertungsmodelle eingebunden werden, um allen Anforderungen an individualisiertes Feedback gerecht zu werden. Beispielsweise muss die gegebene Antwort nicht nur dem korrekten Expert:innenfeedback zugeordnet werden, sondern auch noch auf einer passenden Metrik eingeordnet werden, um den Grad der Korrektheit festlegen zu können. Ebenso muss die vorhandene Abweichung begründet werden können.

Auf der Videoanalyseplattform ViviAn erhalten die Studierenden bereits vorgefertigte, im Rahmen von Dissertationen entstandene schriftliche Expert:innenantworten zu den jeweiligen Fragen, um ihre eigenen Antworten mit diesen abgleichen zu können. Um diesen Prozess zu individualisieren, kann die Zuordnung von Freitextantworten zu Expert:innenfeedback mit Klassifikationsmodellen vorgenommen werden. Sie eignen sich gegenüber vorgefertigten Schablonenauswertungen besonders, da sie auf einer großen Menge sowohl richtiger als auch falscher Antworten beruhen und somit eine genauere Unterscheidung sowie eine automatisierte Fehlererkennung möglich ist. Wichtig zu beachten ist, dass der Trainingsdatenkörper stetig erweitert und überarbeitet werden sollte, da die semantische Vielfalt der Sprache es nahezu unmöglich macht, alle äquivalenten Sätze zur Beantwortung einer Frage zu besitzen (Sychev et al., 2020). Die Sammlung von Freitextantworten zu allen verfügbaren Vignetten über acht Jahre hinweg ermöglicht das Training von Klassifikationsmodellen mit ausreichend großen Datensätzen zu spezifischen Fragen. Somit ist das Training von eigenen spezialisierten Modellen möglich, ohne auf generalisierte externe Modelle zurückgreifen zu müssen. Mit Klassifikationsmodellen ist es möglich, Studierendenantworten zu Expert:innenantworten zuzuordnen oder beispielsweise eine Klassifizierung in richtig oder falsch vorzu-

nehmen. Als Trainingsgrundlage für das Klassifikationsmodell dienen Datensätze, die bereits richtige Zuweisungen enthalten. Der erste Prototyp des Klassifikationsmodells mit ViviAn-Daten zur Zuordnung von gegebenen Antworten zu Expert:innenantworten erreicht bereits eine Gesamtgenauigkeit von 85 %. Das bedeutet, dass eine gegebene Antwort bereits zu 85 % zur korrekten Expert:innenantwort zugeordnet wird. Durch das Hinzufügen weiterer Informationen, die von anderen Modellen generiert werden, wie beispielsweise den Lernständen und der Klassifikation in richtig und falsch, kann am Ende der CAI-Kette ein generatives Large Language Modell (wie GPT4) alle analysierten Daten gesammelt in einem personalisierten Text als Feedback ausgeben.

4. Fazit

Der Einsatz von künstlicher Intelligenz (KI) in der Bildung stellt eine fortschrittliche Methodik dar, die die Lehrqualität und den Lernerfolg in der Lehrkräftebildung verbessern kann. Durch die Integration eines gestuften Feedbacksystems, das auf dem Response to Intervention (RTI) Modell basiert, wird eine effektive Evaluation und Förderung diagnostischer Fähigkeiten bei Studierenden ermöglicht. Es werden KI-gestützte Tools zur detaillierten Analyse von Freitextantworten eingesetzt, wodurch personalisiertes und präzises Feedback ermöglicht wird.

Der innovative Ansatz der Cascading AI, der eine Aneinanderreihung verschiedener KI-Modelle umfasst, adressiert die Herausforderungen der Freitextanalyse durch schrittweise Verfeinerung und Spezialisierung der Auswertung. Dies kann nicht nur zur Genauigkeit des Feedbacks beitragen, sondern erhöht auch die Transparenz und Nachvollziehbarkeit des Feedbackgenerierungsprozesses. In weiteren Konzeptions- und Erprobungsschritten werden weitere KI-Modelle ausgewählt, trainiert und evaluiert, die zur Generierung von individuellem Feedback hilfreich sind. In einem Design-Based-Research Ansatz können die Modelle in ViviAn integriert und mit Lernenden evaluiert werden.

Literaturverzeichnis

- Bai, X., & Stede, M. (2022). A Survey of Current Machine Learning Approaches to Student Free-Text Evaluation for Intelligent Tutoring. *International Journal of Artificial Intelligence in Education*, 1–39. <https://doi.org/10.1007/s40593-022-00323-0>
- Bernius, J. P., Krusche, S., & Bruegge, B. (2022). Machine learning based feedback on textual student answers in large courses. *Computers and Education: Artificial Intelligence*, 3, 100081. <https://doi.org/10.1016/j.caeai.2022.100081>

- Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y.-S., Gašević, D., & Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, 2, 100027. <https://doi.org/10.1016/j.caeai.2021.100027>
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. *IEEE Access*, 8, 75264–75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Demaidi, M. N., Gaber, M. M., & Filer, N. (2018). OntoPeFeGe: Ontology-Based Personalized Feedback Generator. *IEEE Access*, 6, 31644–31664. <https://doi.org/10.1109/ACCESS.2018.2846398>
- Enenkiel, P. (2022). *Diagnostische Fähigkeiten mit Videovignetten und Feedback fördern: Gruppenarbeitsprozesse zur Bestimmung von Längen, Flächen- und Rauminhalten*. Springer Fachmedien; Imprint Springer Spektrum.
- Enenkiel, P., Bartel, M.-E., Walz, M., & Roth, J. (2022). Diagnostische Fähigkeiten mit der videobasierten Lernumgebung ViviAn fördern. *Journal für Mathematik-Didaktik*, 43(1), 67–99. <https://doi.org/10.1007/s13138-022-00204-y>
- Gligorea, I., Cioca, M., Oancea, R., Gorski, A.-T., Gorski, H., & Tudorache, P. (2023). Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review. *Education Sciences*, 13(12), 1216. <https://doi.org/10.3390/educsci13121216>
- Hartmann, E., & Müller, C. M. (2009). Schulweite Prävention von Lernproblemen im RTI-Modell. *Schweizerische Zeitschrift für Heilpädagogik*(9), 25–33.
- Jaakkola, H., Henno, J., Lahti, A., Jarvinen, J.-P., & Makela, J. (2020). Artificial Intelligence and Education. In 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO) (S. 548–555). IEEE. <https://doi.org/10.23919/MIPRO48935.2020.9245329>
- Keuning, H., Jeuring, J., & Heeren, B. (2019). A Systematic Literature Review of Automated Feedback Generation for Programming Exercises. *ACM Transactions on Computing Education*, 19(1), 1–43. <https://doi.org/10.1145/3231711>
- Krusche, S., & Seitz, A. (2018). ArTEMiS. In T. Barnes, D. Garcia, E. K. Hawthorne & M. A. Pérez-Quinones (Hg.), *Proceedings of the 49th ACM Technical Symposium on Computer Science Education* (S. 284–289). ACM. <https://doi.org/10.1145/3159450.3159602>
- Liesenfeld, A., Lopez, A., & Dingemanse, M. (2023). Opening up ChatGPT: Tracking openness, transparency, and accountability in instruction-tuned text generators. In M. Lee, C. Munteanu, M. Porcheron, J. Trippas & S. T. Völkel (Hg.), *Proceedings of the 5th International Conference on Conversational User Interfaces* (S. 1–6). ACM. <https://doi.org/10.1145/3571884.3604316>
- Marwan, S., Gao, G., Fisk, S., Price, T. W., & Barnes, T. (2020). Adaptive Immediate Feedback Can Improve Novice Programming Engagement and Intention to Persist in Computer Science. In A. Robins, A. Moskal, A. J. Ko, & R. McCauley

- (Hg.), *Proceedings of the 2020 ACM Conference on International Computing Education Research* (S. 194–203). ACM. <https://doi.org/10.1145/3372782.3406264>
- Oppenlaender, J., & Benjamin, J. J. (2020). *Towards Metaphors for Cascading AI*. <https://doi.org/10.31219/osf.io/gxt7y>
- Qiao, C., & Hu, X. (2023). Leveraging Semantic Facets for Automatic Assessment of Short Free Text Answers. *IEEE Transactions on Learning Technologies*, 16(1), 26–39. <https://doi.org/10.1109/TLT.2022.3199469>
- Rieger, M. B., Ossadnik, H., Pfaffmann, C., & Roth, J. (2023). Design und Evaluation einer interaktiven Tutorien-Umgebung für das Videoanalysetool ViviAn. In J. Aufenanger & M. Bigos (Hg.), *Digitalisierung in der Lehrer:innenbildung: Corona als Katalysator?! (S. 312–324)*. Beltz Juventa.
- Rietz, C., van Koll, S., & Franke, S. (2013). RTI in der Methodenausbildung von Sonderpädagoginnen und Sonderpädagogen. Erste Implementierungsschritte und Evaluationsergebnisse. *Empirische Sonderpädagogik*, 5, 374–384. <https://doi.org/10.25656/01:9231>
- Scherb, C. A., Rieger, M., & Roth, J. (2023). Untersuchung von Usability und Design von Online-Lernplattformen am Beispiel des Video-Analysetools ViviAn. In J. Roth, M. Baum, K. Eilerts, G. Hornung, & T. Trefzger (Hg.), *Die Zukunft des MINT-Lernens – Band 1 (S. 105–121)*. Springer. https://doi.org/10.1007/978-3-662-66131-4_6
- Sychev, O., Anikin, A., & Prokudin, A. (2020). Automatic grading and hinting in open-ended text questions. *Cognitive Systems Research*, 59, 264–272. <https://doi.org/10.1016/j.cogsys.2019.09.025>
- Wang, X., Zhang, L., & He, T. (2022). Learning Performance Prediction-Based Personalized Feedback in Online Learning via Machine Learning. *Sustainability*, 14(13), 7654. <https://doi.org/10.3390/su14137654>
- Wong, B. T., & Li, K. C. (2020). A review of learning analytics intervention in higher education (2011–2018). *Journal of Computers in Education*, 7(1), 7–28. <https://doi.org/10.1007/s40692-019-00143-7>