

Feedback durch Machine Learning – Automatische Rückmeldung zu Concept Maps

Theoretischer Hintergrund

Feedback ist wohl das prominenteste Element des formativen Assessmentprozesses und wird in vielen Untersuchungen als zentrale Schlüsselstrategie betrachtet (Black & Wiliam, 2009; Bürgermeister et al., 2014; Heritage, 2007). Feedback kann dabei durch Lehrkräfte genutzt werden, indem sie die gesammelten Informationen verwenden, um den nachfolgenden Unterricht an das Lernniveau der Lernenden anzupassen, z. B. durch unterschiedliche Aufgabenblätter oder Unterrichtsanfänge. Eine andere Möglichkeit ist, dass die Lernenden ein direktes individuelles Feedback erhalten, mit dem sie ihren Lernprozess selbstständig optimieren können (Schütze et al., 2018).

Je nach Fachgebiet gibt es unterschiedliche Klassifizierungen von Feedback (Müller & Ditton, 2014). Eine Möglichkeit ist die Differenzierung auf der inhaltlichen Ebene. Dabei wird zwischen einfachem und elaborem Feedback unterschieden (Narciss, 2006). Die einfache Variante des Feedbacks bezieht sich nur auf die Information, ob die Lösung der Aufgabe richtig oder falsch war (*knowledge of results*). Nach Narciss (2006) können zusätzliche Informationen zu inhaltlichen Aspekten, wie z.B. Informationen über die Art der Aufgabe (*knowledge on task constrains*) oder über Erklärungen von Fachbegriffen (*knowledge about concepts*), zu den elaborierten Formen des Feedbacks gezählt werden.

Damit die Lernenden ein individuelles Feedback erhalten können muss in der Regel eine große Menge unterschiedlichen Daten gesammelt, analysiert und zurückgemeldet werden. In Klassenzimmern mit oft mehr als 20 Lernenden ist ein individuelles Feedback für jeden Lernenden jedoch weder realistisch noch machbar (Hunt & Pellegrino, 2002). Und da kurz- oder mittelfristiges Feedback den größten Einfluss auf die Leistungen der Lernenden hat (Maier, 2010), stellt die formative Beurteilung für viele Lehrkräfte eine große Herausforderung dar. Um dieses Problem zu lösen, wird in der fachdidaktischen Forschung eine neue Art der Datenanalyse mit Techniken des maschinellen Lernens eingesetzt (Zhai et al., 2020), so auch für die Erstellung von automatischem Feedback (z. B. Yik et al., 2021).

Ein zentraler Aspekt, der nicht nur für automatisch generiertes Feedback wichtig ist, ist die Qualität des Feedbacks (Wolf, 2014). Ein automatisches Feedbacksystem kann den Lehrkräften viel Arbeit in Form von Bewertungen und Analysen abnehmen und sie in ihrem Schulalltag unterstützen. In den meisten Fällen wird eine Lehrkraft jedoch in der Lage sein, qualitativ hochwertigeres Feedback zu geben, als ein automatisches Feedbackmodell (Herding et al., 2010). Insbesondere für die formative Beurteilung ist ein transparentes und klar verständliches Feedback ein notwendiges Kriterium. Wenn eine Lehrkraft ein automatisches Feedback verwenden möchte, muss es daher nicht nur schnell und einfach verfügbar sein, sondern auch eine hohe Qualität aufweisen, damit die Lehrkräfte nicht durch zusätzliche

Kontrollen mehr Arbeit haben. Voraussetzung für die Nutzung eines Machine Learning Modells im Schulalltag als automatisches Feedbacktool ist eine konkrete Ausarbeitung des Feedbacks. In dieser Arbeit soll deshalb die automatisch generierte Rückmeldung für Lehrkräfte und Lernenden vorgestellt werden. Die Ergebnisse der Studie können dazu beitragen, ein besseres Verständnis dafür zu vermitteln, wie maschinell generiertes Feedback zur Verbesserung des formativen Assessments in der Schule eingesetzt werden kann.

Concept Map und ML-Modell

Als formative Assessmentmethode wurde eine Concept Map gewählt. Concept Maps sind gerichtete Graphen, in denen die Knoten klar definierte Konzepte und die Kanten ihre Beziehungen darstellen (Novak & Cañas, 2008). Sie eignen sich gut für die formative Beurteilung und als Grundlage für Feedback (Ruiz-Primo & Shavelson, 1996). Das Thema für die Concept Map war Mechanik der 11 Jahrgangsstufe. Es wurde ein geschlossenes Format gewählt, um vergleichbare Concept Maps zu erhalten und um die Lernenden nicht zu überfordern. Konzepte und ihre räumliche Anordnung waren daher vorgegeben. Die Aufgabe der Lernenden bestand darin, die 19 Propositionen selbst zu generieren. Der Ansatz zur Analyse der Concept Maps konzentriert sich auf die Auswertung der einzelnen Propositionen. Dies hat den Vorteil, dass die Propositionen im Detail untersucht werden können und so Wissenslücken sichtbar werden (Bak Kibar et al., 2013).

Die Propositionen wurden zunächst von zwei menschlichen Ratern bewertet (Übereinstimmung 87 % mit einem Cohen's Kappa von 0,83). Auf dieser Grundlage wurden mehrere gängige Algorithmen für maschinelles Lernen getestet (Zhai et al., 2020). Der verwendete Ansatz entspricht einem klassischen überwachten maschinellen Lernansatz. Eine detaillierte Beschreibung des maschinellen Lernmodells und dessen Performance findet sich in Bleckmann & Friege (2023).

Automatisches Feedback

Für das automatische Feedback wurden zwei verschiedene Ansätze gewählt. Das kompakte Feedback besteht aus Bewertungen von richtig oder falsch ausgefüllten Textfeldern der Propositionen. Da die Propositionen als einzelne Items betrachtet und analysiert werden, entspricht dieses Feedback *knowledge of results*. Das elaborierte Feedback basiert auf einem vierstufigen Bewertungsschema (Fischler & Peuckert; 2000; Friege; 2001), das aus drei

Tab 1: Bewertungsschema für das elaborierte Feedback

Kategorie	Beschreibung	Beispiel-Antwort
A	Falscher Zusammenhang	... ist konstant ...
B	Einfacher Zusammenhang	... verändert sich ...
C	Gerichteter Zusammenhang	... steigt ...
D	Detaillierter Zusammenhang	... steigt linear ...

Abstufungen für richtige und einer Kategorie für falsche Beziehungen besteht (siehe Tabelle 1). Dies hat den Vorteil, dass eine viel genauere Rückmeldung gegeben werden kann.

Die automatische Auswertung ermöglicht so ein gezieltes und zeitnahes Feedback, ohne zusätzlich Zeit für die manuelle Auswertung der Concept Map aufwenden zu müssen. So können Lernende identifiziert werden, die Schwierigkeiten hatten, die Concept Map auszufüllen. Durch unterschiedliche Grafiken können schnell und einfach Informationen über den Leistungsstand der Schulklasse gewonnen werden (siehe Abbildung 1). Die Informationen können dann von den Lehrkräften für den weiteren Unterricht oder von den Lernenden direkt als Lernhilfe genutzt werden.

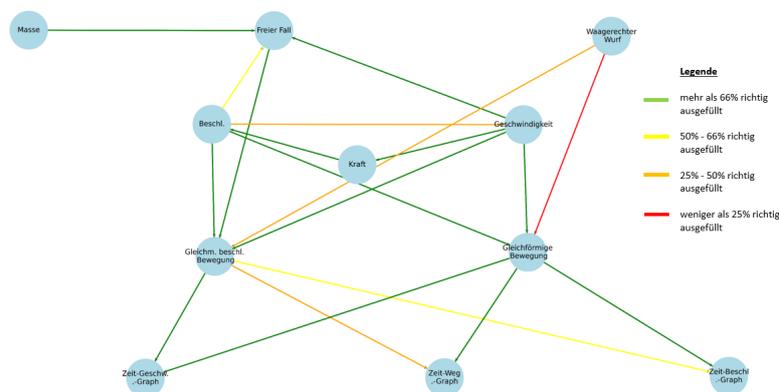


Abb. 1: Automatische Auswertung der Concept Map auf Klassenebene

Ausblick

Diese Studie legt den Grundstein für eine formative Beurteilungsmethode, die mit automatischem Feedback in Schulen eingesetzt werden kann. Aufgrund der guten Ergebnisse ist der Ansatz bereit, in der Schule getestet zu werden. Die Concept Map wird als eine Art Vorwissenstest und später, kurz vor der Prüfung, erneut eingesetzt. Die Feedbackmodelle können die Concept Maps der Lernenden schnell und einfach auswerten. Ein Beispiel für ein ausgearbeitetes Feedback für eine Proposition, welche mit Kategorie B bewertet wurde, könnte wie folgt aussehen: "Diese Proposition ist richtig, aber sie könnte noch etwas präziser formuliert werden. Du könntest mehr Details hinzufügen, z. B. wie genau sich etwas verhält, um eine umfassendere Antwort zu geben". Es werden beide Feedbacktypen getestet, um zu untersuchen, welcher Ansatz im Schulalltag die besseren Ergebnisse liefert. Lehrerinnen und Lehrer können auf diese Weise nicht nur entlastet, sondern auch in ihrer wichtigen Arbeit unterstützt werden. Sie können die automatische Auswertung der Concept Map nutzen, um ihren Lernenden ein individuelles Feedback zu geben und sie so optimal in ihrem Lernprozess zu unterstützen. Die Ergebnisse zeigen außerdem, dass der Einsatz von maschinellem Lernen die naturwissenschaftliche Bildungsforschung weiter voranbringt und die Zusammenarbeit mit anderen Disziplinen, insbesondere mit der Informatik, fördern kann.

Literatur

- Bak Kibar, Z., Yaman, F. & Ayas, A. (2013). Assessing prospective chemistry teachers' understanding of gases through qualitative and quantitative analyses of their concept maps. *Chemistry Education Research and Practice*, 14(4), 542–554. <https://doi.org/10.1039/c3rp00052d>
- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Bleckmann, T. & Friege, G. (2023). Concept maps for formative assessment: Creation and implementation of an automatic and intelligent evaluation method. *Knowledge Management & E-Learning: An International Journal*, 433–447. <https://doi.org/10.34105/j.kmel.2023.15.025>
- Bürgermeister, A., Klieme, E., Rakoczy, K., Harks, B. & Blum, W. (2014). Formative Leistungsbeurteilung im Unterricht: Konzepte, Praxisberichte und ein neues Diagnoseinstrument für das Fach Mathematik. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Tests und Trends: Neue Folge Band12. Lernverlaufsdiagnostik* (S. 41–60). Hogrefe.
- Fischler, H., & Peuckert, J. (2000). *Concept mapping in fachdidaktischen Forschungsprojekten der Physik und Chemie*. Logos Berlin (Verlag).
- Friege, G. (2001). *Wissen und Problemlösen: Eine empirische Untersuchung des wissenszentrierten Problemlösens im Gebiet der Elektrizitätslehre auf der Grundlage des Experten-Novizen-Vergleichs*. Logos Berlin (Verlag).
- Herding, D., Zimmermann, M., Bescherer, C. & Schroeder, U. (2010). Entwicklung eines Frameworks für semi-automatisches Feedback zur Unterstützung bei Lernprozessen. In *DeLFI 2010 - 8. Tagung der Fachgruppe E-Learning der Gesellschaft für Informatik e.V* (S. 145–156).
- Heritage, M. (2007). *Formative Assessment: What Do Teachers Need to Know and Do?* <https://doi.org/10.1177/003172170708900210>
- Hunt, E. & Pellegrino, J. (2002). Issues, Examples, and Challenges in Formative Assessment. *New Directions for Teaching and Learning*, 2002, 73–85. <https://doi.org/10.1002/tl.48>
- Maier, U. (2010). *Formative Assessment - Ein erfolgversprechendes Konzept zur Reform von Unterricht und Leistungsmessung?* *Zeitschrift für Erziehungswissenschaft*. <https://doi.org/10.1007/s11618-010-0124-9>
- Müller, A. & Ditton, H. (2014). *Feedback: Begriff, Formen und Funktionen*. In H. Ditton (Hrsg.), *Feedback und Rückmeldungen: Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 11–28). Waxmann.
- Narciss, S. (2006). *Informatives tutorielles Feedback: Entwicklungs- und Evaluationsprinzipien auf der Basis instruktionspsychologischer Erkenntnisse*. *Pädagogische Psychologie und Entwicklungspsychologie*: Bd. 56. Waxmann.
- Novak, J. D. & Cañas, A. J. (2008). *The Theory Underlying Concept Maps and How to Construct and Use Them*. Technical Report IHMC CmapTools 2006-01 Rev 01-2008, Florida. <http://cmap.ihmc.us/Publications/>
- Ruiz-Primo, M. A. & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, 33(6), 569–600.
- Schütze, B., Souvignier, E. & Hasselhorn, M. (2018). *Stichwort - Formatives Assessment*. *Zeitschrift für Erziehungswissenschaft*, 21(4), 697–715. <https://doi.org/10.1007/s11618-018-0838-7>
- Wolf, N. (2014). *Formative Leistungsmessung im naturwissenschaftlichen Unterricht*.
- Yik, B. J., Dood, A. J., Cruz-Ramírez de Arellano, D., Fields, K. B. & Raker, J. R. (2021). Development of a machine learning-based tool to evaluate correct Lewis acid–base model use in written responses to open-ended formative assessment items. <https://doi.org/10.1039/d1rp00111f>
- Zhai, X., Yin, Y., Pellegrino, J., Haudek, K. C. & Shi, L. (2020). Applying machine learning in science assessment: a systematic review. *Studies in Science Education*, 56(1), 111–151. <https://doi.org/10.1080/03057267.2020.1735757>