Projekt FiSK

Effekte adaptiver Feedbackbots im Simulierten Klassenraum auf prozedurales Professionswissen

Teilprojekt HFU

Adaptive Feedbackbots durch Machine Learning Methoden

Abschlussbericht

Zuwendungsempfänger:	Hochschule Furtwangen
Förderkennzeichen:	16DHB4005
Vorhabensbezeichnung:	FiSK – Effekte adaptiver Feedbackbots im Simulierten Klassenraum auf prozedurales Professionswissen – Teilprojekt Adaptive Feedbackbots durch Machine Learning Methoden
Projektleitung:	Prof. Dr. Christoph Reich
Laufzeit des Vorhabens:	01.03.2021 bis 29.02.2024, Verlängerung bis 31.07.2024

Kompakte Darstellung

Aufgabenstellung

Ziel des Projekts FiSK ist die Erforschung von adaptiven Lernumgebungen. Die dafür verwendete exemplarische Lernumgebung eignet sich für die universitäre Ausbildung von Biologielehrkräften als auch als Erhebungswerkzeug. Diese Lernumgebung enthält neben einem simulierten Klassenraum eine Dialogkomponente auf Basis von künstlicher Intelligenz. Dieses Chatbot-System soll adaptives, individuelles und formatives Feedback geben, um die Lernprozesse der Studierenden zu unterstützen und zu optimieren.

Eine Lektion in der Simulation folgt einem Frage-und-Antwort-Schema, bei dem der Lernende einer Klasse von virtuellen Schülern vordefinierte Fragen stellt. Deren Antworten sind so gestaltet, dass sie verschiedene Stufen des Lernfortschritts widerspiegeln. Drei verschiedene Themenbereiche wurden durch die manuelle Erstellung von Fragen und entsprechenden simulierten Antworten entwickelt. Die Texte folgen bestimmten Konzepten, die vom Lernenden zu identifizieren sind, wie im Fall der Evolution: Evolutionskonzepte, Schwellenwerte Konzepte und Fehlvorstellungen der Schüler. Die Lernenden nehmen an diesem interaktiven Szenario teil, indem sie Multiple-Choice-Optionen auswählen, um das Vorhandensein oder Fehlen wichtiger Lernindikatoren zu diagnostizieren. Dieser Diagnoseprozess wird dann einer summativen Bewertung unterzogen.

Das Chatbot-System soll über reaktive Fähigkeiten verfügen, zum Beispiel Fragen zum Lehrmaterial beantworten. Des Weiteren soll das System auch proaktiv einzugreifen können um den Lernweg zu lenken und so die Lernerfahrung zu verbessern. Dazu werden zwei Modellansätze für die Bot-Entwicklung kombiniert. Die Reaktivität basiert auf einem retrieval-basierten Modell, erstellt aus Expertenwissen. Die Proaktivität nutzt einen daten-getriebenen Ansatz, der auf der Anwendung von KI-Methoden auf empirische Daten basiert.

Im Forschungs-Fokus ist die automatisierte Klassifizierung von Frei-Text Inhalten. Durch Verwendung von daten-getriebenen Methoden kann ein Modell allein aufgrund der Trainingsdaten die expliziten Ordnungs-Klassen erkennen. Eine Spezifität bezüglich bestimmter Ordnungs-Klassen existiert somit nicht. Zum Beispiel erstellt ein definierter Trainingsprozess ein Modell zur Erkennung von Fehler in Schülerantworten im Fach Biologie, sofern dies in den Daten definiert ist. Derselbe Prozess kann jedoch auch ein Modell zur Erkennung der Glaubwürdigkeit von Filmbewertungen erzeugen, in Abhängigkeit der Daten. Daher steht der maschinelle Lernprozess im Vordergrund, explizit die umfassende Nutzung aller Information in den Daten. Im Bereich des bildungsnahen Einsatzes von Künstlicher Intelligenz sind Anforderung zu erfassen und geeignete Umsetzungen zu erforschen. Dies ergab hohe Qualitäts-Anforderungen, hohe Datenschutz-Anforderungen, geringe Datenmengen und Daten mit komplexen Ordnungs-Klassen. Es sollen daher Verfahren für maschinelles Lernen untersucht werden, die hohe Qualität bei geringen Datenmengen und komplexen Klassifizierungen ermöglichen, ohne etwaigen Datenschutz zu verletzen.

Das Projekt verfolgt das Ziel, durch die Entwicklung des Bots die Lernumgebung zu einem wirkungsvollen Instrument für die Hochschulausbildung weiterzuentwickeln, indem es gezielte und wirksame Rückmeldungen für den Lernfortschritt der Studierenden bietet.

Durchführungsvoraussetzungen

Das Projekt FiSK unterteilte sich in drei Teilprojekte (Teilprojekt A: Biologie, IPN; Teilprojekt B: Adaptive Feedbackbots durch Machine Learning Methoden, HFU; Teilprojekt C: TP AICoaching, AI Coaching;). Zur gemeinsamen Koordinierung von FiSK wurden regelmäßige interne Arbeitstreffen aller Projektbeteiligten durchgeführt.

Die interdisziplinäre Verzahnung lag zum einen in der Erarbeitung der Adaptiven Klassenraumsimulation, sowie in der Analyse der Erhebungs-Daten. Die Komplexität der Erkennung didaktischer Eigenschaften in Freit-Text-Antworten in Verbindung der geringen Fallzahlen eröffnete ein unerwartetes Forschungsfeld: Die Optimierung der automatisierten Text-Klassifikation auf kleine Datensätze. Dies wiederum führte zu den Forschungen bezüglich Effizienz von maschinellem Lernen.

Anknüpfung an bisherigen wissenschaftlichen Stand

Forschung zu Intelligent Tutoring Systems setzt die Arbeiten von Ashfaque et al. (2020) [1] fort. Die Nutzung lingusitischer Informationen in der Feature Selection, insbesondere für Text Klassifizierung, basiert auf den Arbeiten von Moschitti & Basili (2004) [2], Qin et al. (2015).[3], sowie Jasleen & Jatinderkumar (2016) [4]. Die Arbeiten zur Effizienz des maschinellen Lernens nutzen Erkenntnisse von Numrich et al. (2005) [5] sowie Fischer et al. (2022) [6] um Theorie und Anwendung zu fundieren.

Ablauf des Vorhabens

Zunächst wurden das Forschungswerkzeug in Form der Adaptiven Klassenraumsimulation gemeinsames entwickelt. Daraufhin begann die Durchführung von Erhebungen. Die Ergebnisse wurden dann zur Evaluierung von Technologien zur automatisierten Text Klassifizierung genutzt. Eine vertiefende Forschung im Bereich Natural Language Processing, insbesondere Feature Selection, sowie Machine Learning Effizienz setzte bereits nach der ersten Erhebung ein.

Ergebnisse

Die Ergebnisse in Projekt Fisk lassen sich in zwei Gruppen unterteilen. Technische Ergebnisse umfassen die Erstellung von zwei Simulations-Systemen, die Anbindung von Cloud-Diensten und die selbst entwickelten System zur Feedback-Generierung. Forschungsergebnisse im Bereich intelligenter Lernumgebungen sind Analyse und Konzept zur Anwendung von Künstlicher Intelligenz zur Unterstützung von Lernenden, vergleichende Untersuchung zur Eignung spezifischer Methoden maschinellen Lernens. Im Bereich Natural Language Processing sind dies eine neue Methode zur Steigerung des Informationsgehalts eines Texts, eine neue Methode zur Nutzung linguistischer Information zur Text Klassifizierung, sowie Evaluationen zur optimalen Nutzung dieser Methoden. Alle Methoden stellen dabei den Datenschutz per Design sicher. Im Bereich Machine Learning Efficiency wird eine neue Methode zur Definition von Machine Learning entwickelt, eine Metrik um diese anzuwenden, sowie Studien die die Nutzung von Machine Learning Efficiency in vergleichenden Anwendungsfällen demonstrieren.

Eingehende Darstellung

Durchgeführte Arbeiten 2021

Auswahl und Entwicklung Klassenraum-Simulation

In der Pilotierungsphase wurde eine Evaluation von Systemen zur Simulation der Lehrer-Schüler-Interaktion durchgeführt. Die Simulations-Plattform Simschool (Hersteller: SimEd LLC) wurde ausgewählt und die Entscheidung dem BMBF unterbreitet. Aufgrund von Vergaberichtlinien konnte eine weltweite Ausschreibung und Beauftragung nicht durchgeführt werden. Ersatzweise wird die Anwendung „Simulierter Klassenraum" (SKR; Cap3 GmbH Kiel) verwendet. Im Auswahlverfahren im Rahmen des Ausschreibungsprozesses bekam das Unternehmen Cap3 GmbH am 30.06.2021 den Zuschlag. Da der SKR mit Chatbot am Ende aus einzelnen Komponenten bestehen wird, die eng miteinander verbunden sind, ist die ereignisgesteuerte Aktivierung und Integration des Bot-Systems in den Event-Bus des SKRs von absoluter Notwendigkeit für die Durchführung des Projekts. Daher wurde das SKR-System auf die für den FiSK vorgesehenen Projektserver der Hochschule Furtwangen installiert. Die damit verbundenen Rechte und Datenschutzbestimmungen wurden berücksichtigt. Aufgrund von Vorlaufzeit durch die Auftragsvergabe an das Unternehmen Cap3 begann die Installation des SKR erst Mitte November 2021, sodass das SKR-System Mitte Dezember 2021 in einer ersten Basisversion auf den Servern der Hochschule Furtwangen zur Verfügung stand.

Entwicklung Adaptive Lernumgebung Gen 1

Die adaptive Lernumgebung der ersten Generation (Gen 1) umfasst eine angepasste Version der Klassenraumsimulation SKR sowie ein Chatbot-System.

Die Erstellung des Chatbot-Systems Generation 1 wurde in Kooperation mit AI Coaching GmbH im Juni 2021 abgeschlossen. Das technische Konzept zur Integration von Bot-Systemen und der Simulationsplattform wurde abgeschlossen. Die Programmierung des Chatbot-System Generation 1 wurde abgeschlossen.

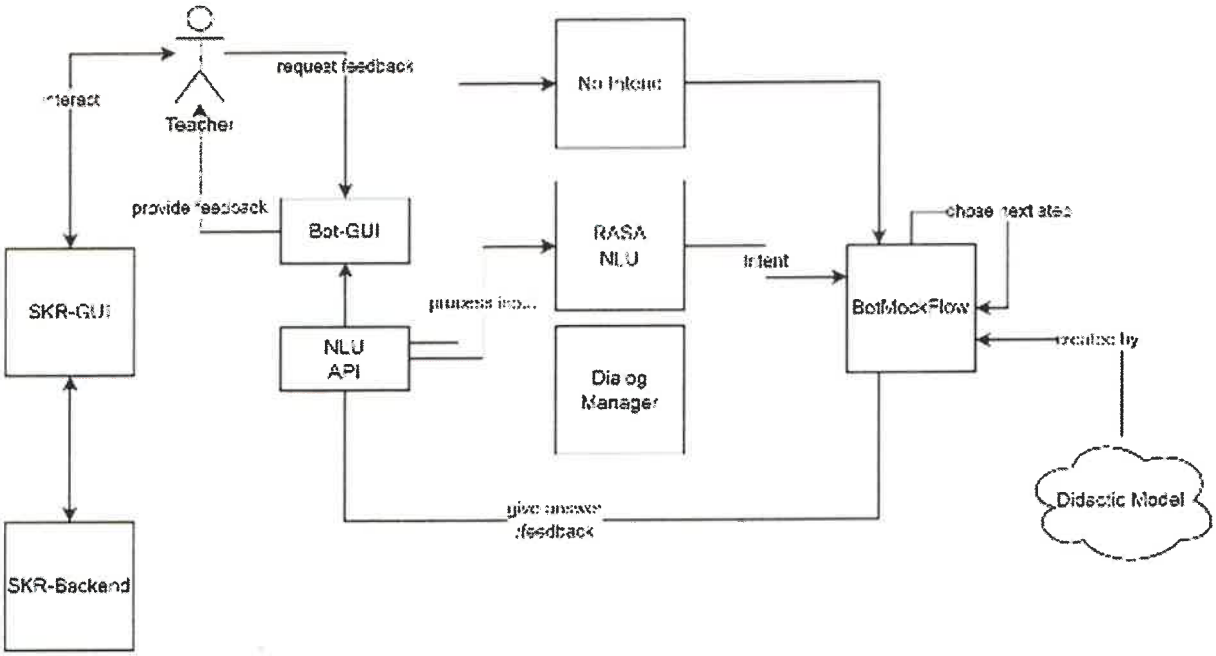

Abbildung 1 Architektur Adaptive Lernumgebung Gen 1

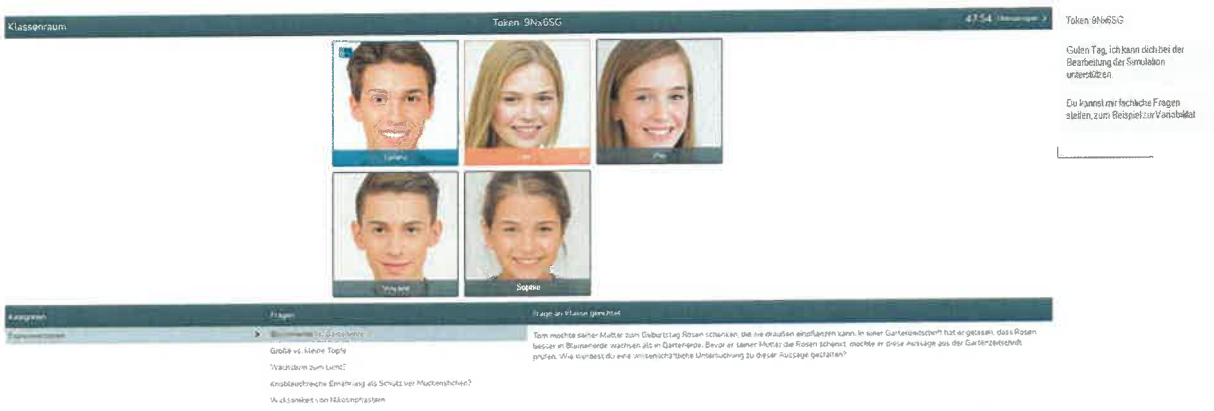

Abbildung 2: Userinterface Simulations-System Generation 1

Ausfall Chatbot-System Gen 1

Mitte Dezember 2021 ereignete sich eine unvorhergesehene Änderung zum ursprünglichen Arbeitsplan in Bezug auf das verwendete Chatbot-System. Das externe cloudbasierte Sub-System zur visuellen Erstellung des Chatverhaltens des Unternehmens Botmock wurde kurzfristig abgeschaltet und alle existierenden Projekte geschlossen. Dadurch fehlte für das Projekt ein geeignetes und insbesondere datenschutzkonformes System für den Feedback-Chatbot im SKR. Im Januar und Februar 2022 wurden deshalb Ersatzsysteme evaluiert. Die Verbundpartner konnten sich schließlich auf das System Dialogflow von Google Dublin einigen.

Entwicklung Chatbot-System Gen 2

Die Entwicklung eines Nachfolgesystem des Chatbot-Systems Generation 1 begann im November 2021 und dauert an.

Weitere Arbeiten

- Teilnahme am digitalen Kick-off-Meeting der drei Verbundpartner am 02.03.2021 mit vier Wissenschaftler*innen des IPN Kiel und der Hochschule Furtwangen sowie einem Vertreter von AI Coaching; Hauptthemen: Besprechung der Ziele und Forschungsfragen des Gesamtprojekts und der Teilprojekte, Organisation der Arbeitsebenen, Planung der nächsten Schritte.
- Organisation und Teilnahme an den monatlichen Online-Meetings der gesamten Projektgruppe

- Fertigstellung der FiSK-Projektseite im Institut für Data Science, Cloud Computing und IT-Sicherheit der HFU.
- Analyse, Design und Entwicklung eines Systems zur individuellen Nutzer-Interaktion (Chatbot Gen1)
- Anpassung einer Anwendung zur Simulation von Klassenräumen mit dem Ziel der Kommunikation mit anderen Systemen.
- Analyse und Design einer Middleware die Systeme zur Simulation von Klassenräumen, Systeme zur individuellen Nutzer-Interaktion und Systeme zur Analyse von Nutzerverhalten als Komponenten einbindet.
- Analyse eines Systems zur Analyse von Nutzerverhalten.
- Bereitstellung von Serverkapazitäten für das datenschutzkonforme Bereitstellen des SKRBIO-Feedbackbot-Systems

Durchgeführte Arbeiten 2022

Simulations-Adaption

Die Adaption der Simulations-Software durch Cap3 wurde in regelmäßigen Treffen begleitet. Die Simulation wurde erfolgreich an die Anforderungen zur Umsetzung der geplanten Erhebungen angepasst. Umfangreiche Änderungen zum automatischen Informationsaustausch zwischen der Simulations-Software (Nutzeroberfläche, Backend und Datenbank) mit der Middleware des Feedback-Bot-Systems wurden umgesetzt.

Aus technischen Gründen sowie um die Sicherheit der Software-Systeme zu gewährleisten, werden diese in einem abgesicherten Netzwerk betrieben. Dazu wurde die Simulations-Software auf Systeme der Hochschule Furtwangen übertragen.

Entwicklung Chatbot-System Gen 2

Anforderungen, Architektur und Entwurf der Middleware zur Verknüpfung der Simulations-Software. Umstellung visuellen Editors des Botmock auf die Bot-Plattform Dialogflow. Anpassen des Entwurfs an Dialogflow. Beginn der Implementierung.

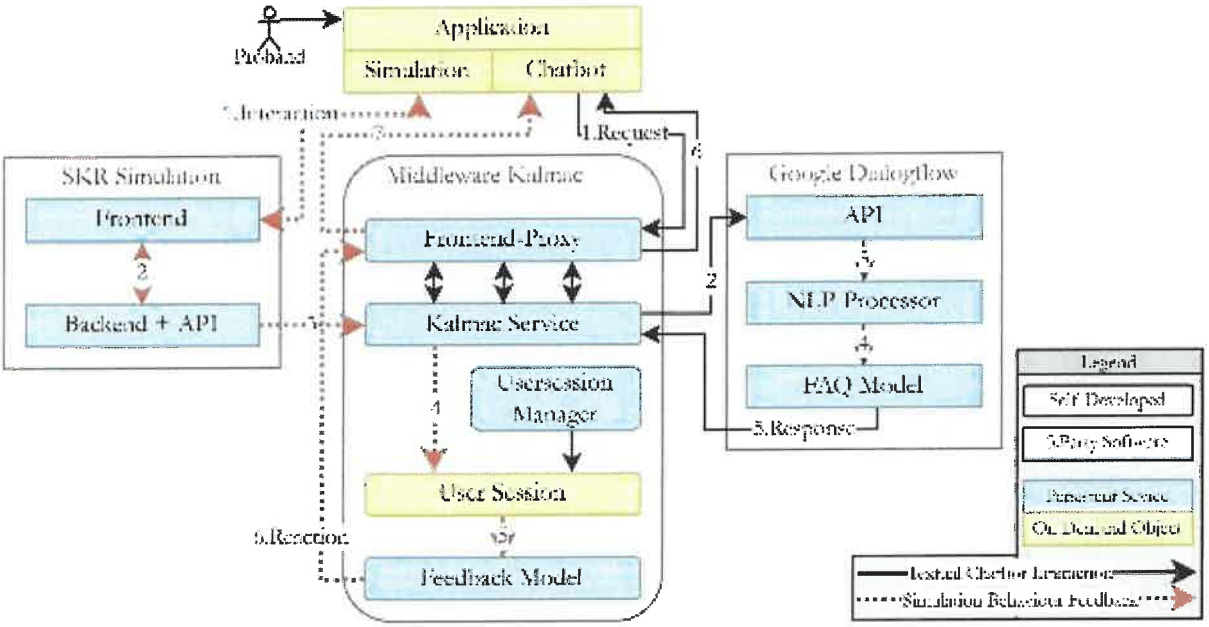

Abbildung 3 Simulations-System Generation 2

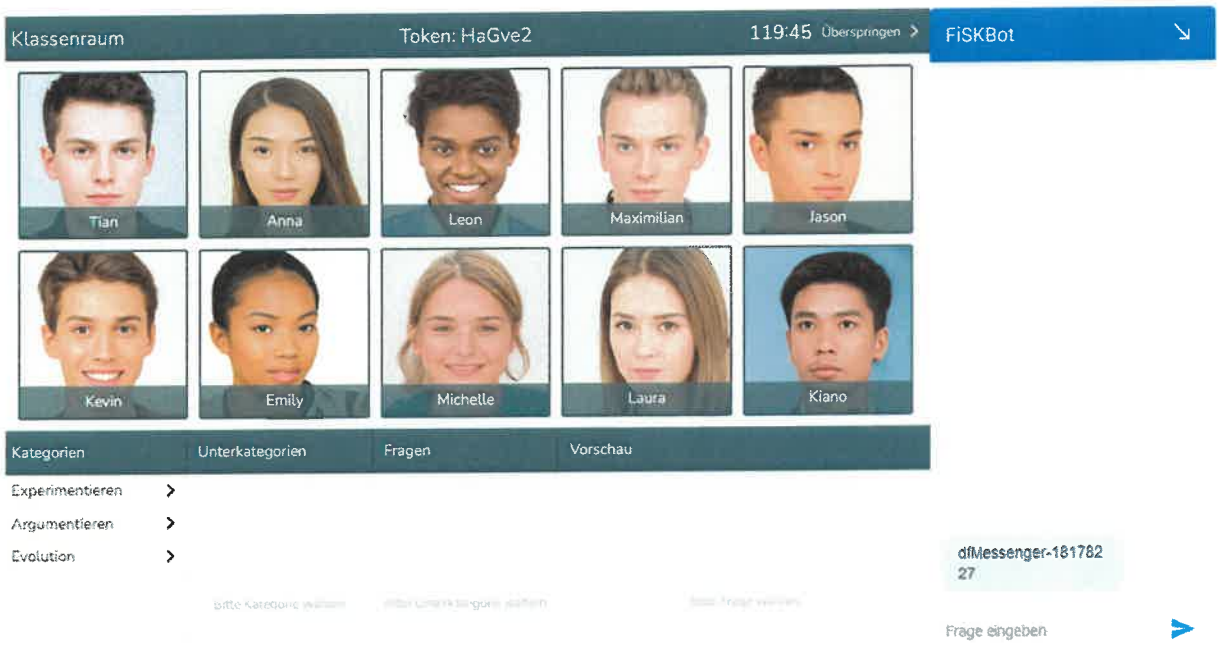

Abbildung 4 Userinterface Simulations-System Generation 2

Feedback-Katalog (HFU+IPN)

Eine Liste von Feedback-Gebungen wurde erarbeitet. Durch interdisziplinäre Zusammenarbeit konnten auslösende Situation und Eigenschaften des zu gebenden Feedbacks parametrisiert werden als Anknüpfungspunkt der technischen Umsetzung.

Feedback-Modul

Analyse der der einzelnen Feedback-Gebungen des Feedback-Katalogs und entsprechender Umsetzung als Modul in der Middleware.

Retrieval-Based Model

Erfolgreiche Umsetzung eines experten-basierten Models (retrieval-based) auf der Plattform Dialogflow.

Datengetriebenes Model

Ein datengetriebenes Model eines Chatbots verarbeitet textuelle Eingaben der Probanden. Aktuelle Verfahren zur Erkennung und Entfernung unwichtiger Textteile wurden untersucht und ein Verfahren entwickelt, das dies auf Basis linguistischer Analyse vornimmt. Die Effizienz konnte für Text-Klassifizierung zwischen 2% und 20%, abhängig von Text-Vektorisierung-Methode und Klassifizierungs-Modell, gesteigert werden. Die Ergebnisse wurden auf der internationalen Konferenz 6th International Conference on Natural Language Processing and Information Retrieval vorgestellt und im wissenschaftlichen Journal Journal of Advances in Information Technology veröffentlicht.

Learner-Model des Probanden

Ein Konzept für ein Learner-Model, das auf die Datengrundlage abgestimmt ist, wurde entwickelt. Es basiert auf Techniken zur Erstellung von Few-Shot-Modellen, da die Fallzahlen keine Big-Data-Techniken ermöglichen. Ein Konzept zu Erfassung von historischen Daten der Probanden, die eine Nutzer-Persistenz beinhalten, wurde entwickelt.

Simulations-Adaption

Die Anpassung der Simulationssoftware durch Cap3 an dynamische Anforderungen verschiedener Erhebungen ist erfolgreich umgesetzt worden.

Feedback-Modul

Die Entwicklung und Implementierung der Middleware zur Feedback-Erteilung und Integration in die Simulationssoftware wurde abgeschlossen.

Weitere Arbeiten

- Organisation und Teilnahme an den monatlichen Online-Meetings der gesamten Projektgruppe
- Organisation und Teilnahme am 1. Projekt-Treffen am 17. und 18. Mai 2022 an der Hochschule Furtwangen mit vier Wissenschaftler*innen des IPN Kiel und der Hochschule Furtwangen; Hauptthemen: Modelle und Studien im FiSK-Projekt, Design und Implementation des neuen Chatbot-Systems (Dialogflow) in den SKR, Studienplanung und nächste Projektschritte
- Teilnahme am 2. Projekt-Treffen am 26. und 27. September 2022 am IPN Kiel mit fünf Wissenschaftlerinnen und einer studentischen Hilfskraft des IPN Kiel und der Hochschule Furtwangen; Hauptthemen: Vorstellung der Ergebnisse aus der Studie " SKREVO + Knowledge Bot", Vorstellung der ausgearbeiteten Klassenräume SKREXP und SKRARG, Planung der Feedback-Bot Studien, Nächste Projektschritte

Durchgeführte Arbeiten 2023

Inbetriebnahme Adaptive Lernumgebung Gen 2 Furtwangen

Das 2. Simulationssystem, basierend auf der zweiten Generation, konnte erfolgreich in Betrieb genommen werden.

Maßnahmen Hacker-Angriff

Durch einen Hacker-Angriff wurde die gesamte IT der Hochschule Furtwangen am Standort Furtwangen zerstört. Die Auswirkungen auf das Projekt Fisk sind umfangreich: Eine Erhebung konnte nicht wie geplant durchgeführt werden. Die Systeme der SKR-Simulation und des reaktiven Chatbots wurden zerstört. Die Backups der Systeme wurden gelöscht. Eine weitere Nutzung der projektbezogenen Systeme war nicht mehr möglich. Ein Teil der Forschungsdaten ging verloren. Die Wiederherstellung der Systeme am Standort Furtwangen ist auf unabsehbare Zeit technisch nicht möglich.

Aufgrund des Cyber-Angriffs wurden Maßnahmen zur Kompensation des Schadens und einer Wiederherstellung des Betriebs der Projekt-IT entwickelt:

- Aufbau einer temporären Infrastruktur im Rechenzentrum des IPN Kiel
- Wieder-Aufbau der Simulations-Systeme
- Wiederherstellen der Google-Cloud-Verbindung

- Test der Funktionalität

Weitere Arbeiten

- Organisation und Teilnahme an den monatlichen Online-Meetings der gesamten Projektgruppe
- Teilnahme am 3. Projekt-Treffen am 11. bis 12. Juli 2023 am IPN Kiel mit sechs Wissenschaftler:innen und einer studentischen Hilfskraft des IPN Kiel und der Hochschule Furtwangen; Hauptthemen: Vorstellung bisheriger Projektergebnisse und aktuell geplanter Studien, Planung der Schritte für das letzte Projektjahr.

Durchgeführte Arbeiten 2024

Inbetriebnahme Adaptive Lernumgebung Gen 2 Kiel

Die temporäre Notbetrieb wurde erfolgreich in eine persistente Infrastruktur überführt. Dies umfasste den Wieder-Aufbau der Simulations-Systeme, das Wiederherstellen der Google-Cloud-Verbindung sowie das Testen aller Funktionalitäten. Des Weiteren wurden Maßnahmen zur IT-Sicherheit erarbeitet und umgesetzt.

Weitere Arbeiten

- Organisation und Teilnahme an den monatlichen Online-Meetings der gesamten Projektgruppe

Darstellung der erzielten Ergebnisse

Im Projekt Fisk eine Reihe von Ergebnissen erzielt. Technische Ergebnisse umfassen die Erstellung von zwei Simulations-Systemen, die Anbindung von Cloud-Diensten und die selbst entwickelten System zur Feedback-Generierung. Forschungsergebnisse im Bereich intelligenter Lernumgebungen sind eine Analyse und Konzept zur Anwendung von Künstlicher Intelligenz zur Unterstützung von Lernenden, vergleichende Untersuchung zur Eignung spezifischer Methoden maschinellen Lernens. Im Bereich Natural Language Processing sind dies eine neue Methode zur Steigerung des Informationsgehalts eines Texts, eine neue Methode zur Nutzung linguistischer Information zur Text Klassifizierung, sowie Evaluationen zur optimalen Nutzung dieser Methoden. Im Bereich Machine Learning Efficiency wird eine neue Methode zur Definition von Machine Learning entwickelt, eine Metrik um diese anzuwenden, sowie Studien die die Nutzung von Machine Learning Efficiency in vergleichenden Anwendungsfällen demonstrieren.

Intelligente Lernumgebungen

Im Fokus stand zunächst eine Erfassung von aktuellen Beweggründen und Methoden für die Integration von maschinellem Lernen in Bildungsumgebungen. Die Ergebnisse wurden in [V1] dargestellt, eine Übersicht stellt Abbildung 5 Motivation und Methoden von Machine Learning in E-Learning Systemen dar.

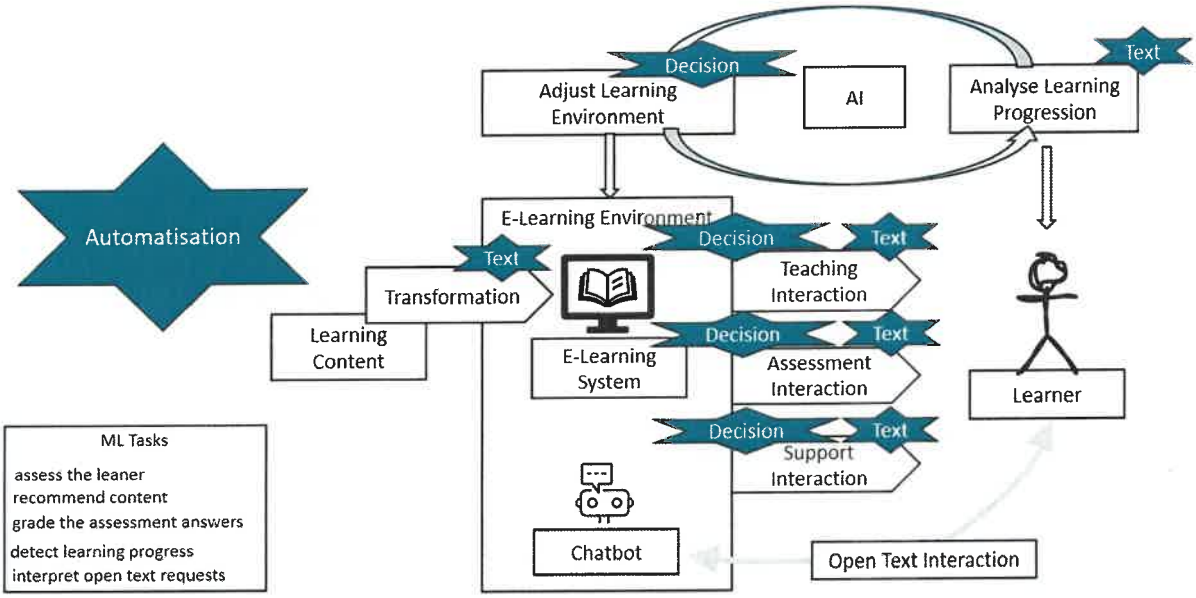

Abbildung 5 Motivation und Methoden von Machine Learning in E-Learning Systemen

Des Weiteren wurden die Herausforderungen bei der Implementierung von maschinellem Lernen analysiert und ein Framework zur Integration von ML in Lernumgebungen in [V1] dargestellt. Eine Zusammenfassung stellt Abbildung 6 Herausforderungen und Chancen bei Verwendung von Machine Learning in E-Learning Systemen dar.

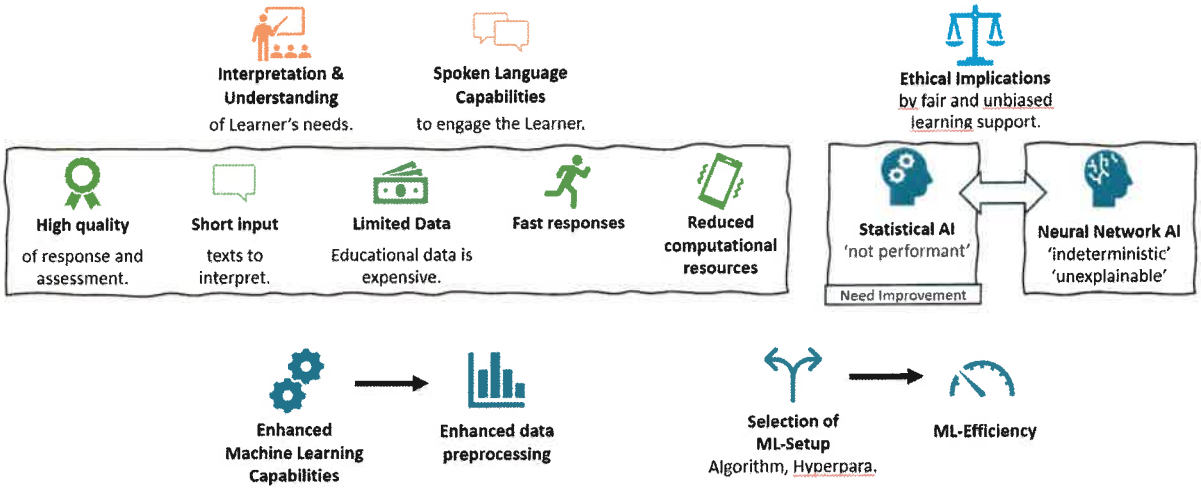

Abbildung 6 Herausforderungen und Chancen bei Verwendung von Machine Learning in E-Learning Systemen

Natural Language Processing: Feature Selection

Die gängige Methode der Merkmalsauswahl, insbesondere die Entfernung von Stoppwörtern, basiert auf der Annahme, dass hochfrequente Text-Token ohne negative Auswirkungen auf die Klassifizierungsleistung eliminiert werden können. Dabei werden jedoch wesentliche linguistische Aspekte, wie die Satzstruktur und die spezifischen Anforderungen der eingesetzten Klassifikationstechnologie, unzureichend berücksichtigt. In [2] wird das Weighted Unimportant Part-of-Speech (WUP) Model vorgestellt, das eine

alternative Herangehensweise zur Tokenentfernung während der Textvorverarbeitung bietet. Dieses Modell weist jedem Token eine gewichtete Relevanz zu, die sowohl dessen Bedeutung für die Klassifikationsergebnisse als auch dessen Gesamteinfluss auf die Klassifikatorleistung berücksichtigt. Zur Kategorisierung werden dabei Part-of-Speech-Daten (POS) herangezogen. Analog zur klassischen Stoppwortentfernung ermöglicht das Modell die Erstellung eines WUP-Instanzensatzes, um linguistisch weniger relevante POS-Daten gezielt auszusondern.

Eine Proof-of-Concept (PoC)-Evaluierung wurde durchgeführt, um den Nutzen von WUP-Instanzen für verschiedene Klassifikationsalgorithmen zu prüfen. Die Ergebnisse zeigen, dass die Verwendung des WUP-Modells im Vergleich zur traditionellen Stoppwortentfernung signifikante Vorteile bietet. Besonders ein baumbasierter Klassifikator profitierte mit einer 65%igen Steigerung der Laufzeiteffizienz und einer 25%igen Verbesserung der Klassifikationsleistung. Andere Klassifikatoren wiesen hingegen Leistungseinbußen zwischen 0,2 % und 2,4 % auf, die jedoch mit Laufzeitverbesserungen zwischen 4,4 % und 24,7 % einhergingen. Diese Ergebnisse unterstreichen die potenziellen Vorteile des WUP-Modells und seine positiven Auswirkungen auf die Effizienz und Leistung von Klassifikationssystemen.

Eine detaillierte Beschreibung wurde in [V2] und [V4] veröffentlicht, eine Übersicht stellt Abbildung 7 Beispiel für WUP-Filterung dar.

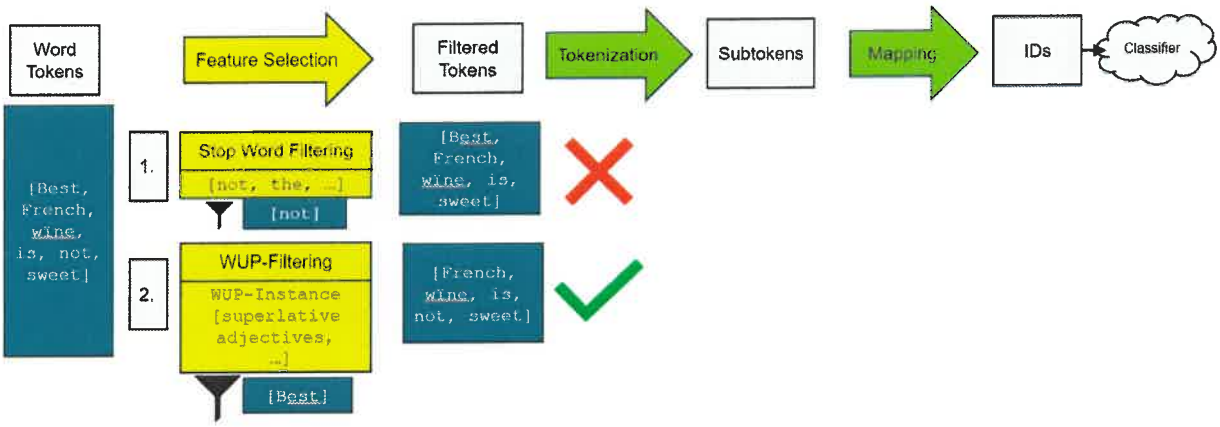

Abbildung 7 Beispiel für WUP-Filterung

Natural Language Processing: Feature Detection

Die Effektivität datengetriebener maschineller Lernverfahren (ML) ist stark von der Qualität und den Eigenschaften der zugrunde liegenden Daten abhängig. Aspekte wie Stichprobenumfang, Merkmalverteilung, Datenintegrität sowie das Vorhandensein von Anomalien, Verzerrungen und Rauschen können die Datenqualität erheblich beeinträchtigen und dadurch die Leistungsfähigkeit von ML-Modellen verringern. In verschiedenen ML-Bereichen werden spezialisierte Ansätze entwickelt und eingesetzt, um diese Herausforderungen entlang der gesamten Pipeline zu adressieren.

Bei der Vorverarbeitung von Textdaten kommen häufig Techniken wie Lemmatisierung und Stoppwortentfernung zum Einsatz. Textdaten zeichnen sich jedoch durch ihre inhärent linguistischen Eigenschaften aus, die einzigartige Möglichkeiten zur Extraktion und Nutzung

verborgener linguistischer Informationen bieten. Diese Informationen, wie z. B. syntaktische Details in Form von Part-of-Speech-Tags (POS), können dazu genutzt werden, Merkmalssätze zu bereichern und so Anwendungen wie die Textklassifikation gezielt zu verbessern.

Der LinPair-Ansatz bietet eine innovative Methode zur Tokenisierung, die die Herausforderungen bei der Verarbeitung von Textdaten für maschinelles Lernen (ML) adressiert. Diese Methode erweitert die WordPiece-Tokenisierung durch die Integration linguistischer Erkenntnisse, indem Token mit syntaktischen Informationen in Form von Token-Tag-Paaren angereichert werden. Dadurch wird es möglich, große Sprachmodelle (LLMs) gezielt mit zusätzlichen syntaktischen Informationen zu unterstützen.

Der LinPair-Ansatz wird in zwei Varianten umgesetzt: CompleteLinPair Tokenization (CLT) und SmartLinPair Tokenization (SLT). CLT ist speziell für kleinere Datensätze konzipiert und reichert jedes Token systematisch mit Part-of-Speech (POS)-Informationen an. SLT hingegen wurde entwickelt, um die negativen Auswirkungen von Out-of-Vocabulary (OOV)-Tokens zu reduzieren, die häufig bei der Verarbeitung großer und vielfältiger Textkorpora auftreten.

Eine experimentelle Analyse zeigte, dass SLT die Leistung bei Textklassifizierungsaufgaben mit BERT-Modellen signifikant verbessert. Im Vergleich zur herkömmlichen WordPiece-Tokenisierung konnte eine Steigerung des F1-Scores um bis zu 13,9 % erzielt werden. Im Gegensatz dazu führte die CLT-Variante zu einer Verschlechterung der Klassifikationsleistung. Es ist jedoch wichtig zu betonen, dass diese Ergebnisse experimentell bedingt sind und nicht uneingeschränkt auf andere Anwendungsbereiche übertragbar sein könnten.

Das LinPair Verfahren wurden in [V6] veröffentlicht, eine Übersicht stellt Abbildung 8 Beispiel für LinPair Tokenisation dar.

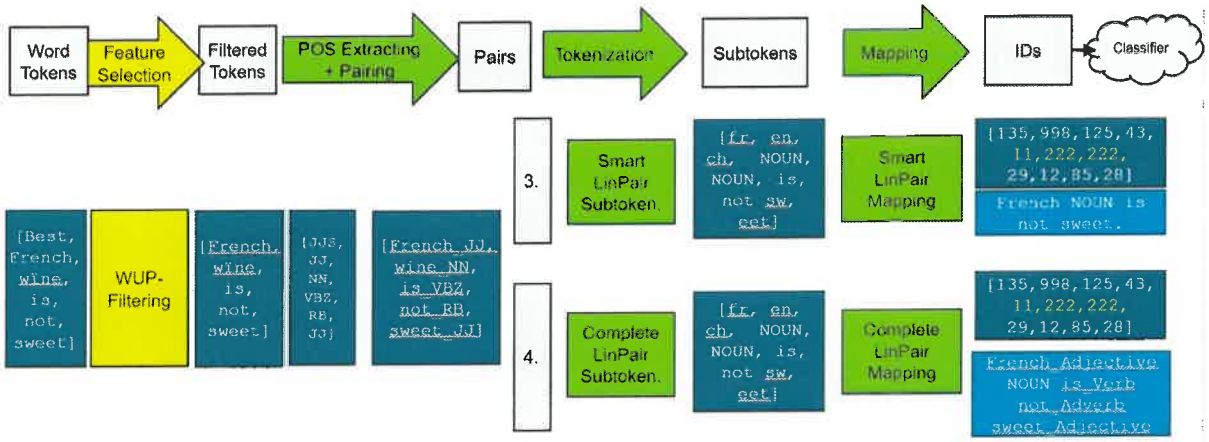

Abbildung 8 Beispiel für LinPair Tokenisation

Machine Learning Efficiency

Mit der zunehmenden Verbreitung maschineller Lernsysteme (ML) steigen auch deren Ressourcenbedarf und die damit verbundenen finanziellen Aufwendungen. Dies verdeutlicht die Notwendigkeit, den Energieverbrauch und die Kosten zu optimieren, um den Erwartungen der Stakeholder gerecht zu werden. Während Qualitätsmetriken für prädiktive

ML-Modelle gut etabliert sind, werden Effizienzmetriken vergleichsweise selten berücksichtigt.

In diesem Kapitel wird ein umfassender Rahmen für die Bewertung von Effizienzmetriken vorgestellt, der den Vergleich unterschiedlicher Effizienzdimensionen erleichtert. Dabei wird ein neuartiger Ansatz eingeführt: die Compact Efficiency Metric (CO). Diese Metrik berücksichtigt nicht nur die Vorhersagequalität eines Modells, sondern auch dessen Ressourcennutzung, Rechenaufwand und Laufzeit. Zur praktischen Umsetzung wurden spezialisierte Varianten wie die QCO-Metrik entwickelt, die auf spezifische Anwendungsfälle zugeschnitten ist.

Darüber hinaus wird eine Pareto-basierte Methodik zur Auswahl von ML-Modellen mit Fokus auf Effizienz präsentiert. Die QCO-Metrik wurde in umfangreichen Tests validiert, um ihre Eignung, Plausibilität und Anpassungsfähigkeit an unterschiedliche Datensatzgrößen und Hosting-Umgebungen zu überprüfen. In experimentellen Anwendungen wurde die QCO-Metrik auf zwei Datensätze angewandt und für verschiedene ML-Modelle berechnet. Dabei identifizierte sie eine optimale Konfiguration, insbesondere bei der Bestimmung der geeigneten Sequenzlänge für transformatorbasierte Modelle.

Die Ergebnisse der Pareto-basierten Methodik korrelierten stark mit den Resultaten der QCO-Metrik, was die Praktikabilität und Effizienz dieses Ansatzes zur Vorauswahl bevorzugter Lösungen unterstreicht.

Die theoretische Grundlage sowie die Defintion der QCO-Metrik wurden in [V5] veröffentlicht.

Integration und Anwendung von Machine Learning in E-Learning-Umgebungen

Maschinelle Lerntechniken (ML) in virtuellen Lernumgebungen basieren maßgeblich auf ihrer Fähigkeit, textuelle Inhalte präzise zu verstehen und zu verarbeiten. Dieses Kapitel untersucht die Auswirkungen fortgeschrittener linguistischer Anreicherungstechniken auf die Vorverarbeitung textbasierter Benutzereingaben. Hierbei wurden Ansätze wie transformatorbasierte Datenerweiterung, Part-of-Speech (POS)-erweiterte Merkmalsauswahl und die LinPairPiece-Tokenisierung eingesetzt.

Die Evaluierung konzentrierte sich sowohl auf die Klassifizierungsqualität als auch auf die Trainingseffizienz. Dabei wurde ein synthetisch erstellter Frage-und-Antwort-Datensatz verwendet, der durch begrenzte Stichprobengrößen, eine hohe Anzahl von Klassen und die Herausforderung der Identifikation didaktischer Elemente gekennzeichnet war. Die Ergebnisse zeigten, dass SVM in den meisten Qualitätsmetriken DistilBERT übertraf. Dennoch konnte die Integration linguistischer Elemente die Leistung von DistilBERT signifikant verbessern: Der F1-Score stieg um 7,62 %, und der Hamming-Score verbesserte sich um 17,02 %.

Trotz dieser Leistungssteigerungen schnitt DistilBERT in Bezug auf Effizienzmetriken schlechter ab als SVM. Dies legt nahe, dass SVM für synthetische Daten eine robuste Wahl bleibt, während Large Language Models (LLMs) wie DistilBERT erhebliche Potenziale bei der Verarbeitung komplexer linguistischer Daten entfalten können, wenn sie gezielt mit linguistischen Informationen angereichert werden.

Die Ergebnisse unterstreichen die Eignung beider Ansätze als effektive Werkzeuge für die automatisierte Bewertung in Bildungsumgebungen und verdeutlichen den Mehrwert linguistischer Anreicherungstechniken für moderne ML-Modelle.

Die Forschung hierzu wurde in [V8] veröffentlicht. Eine Übersicht stellt Abbildung 9 ML Implementation dar.

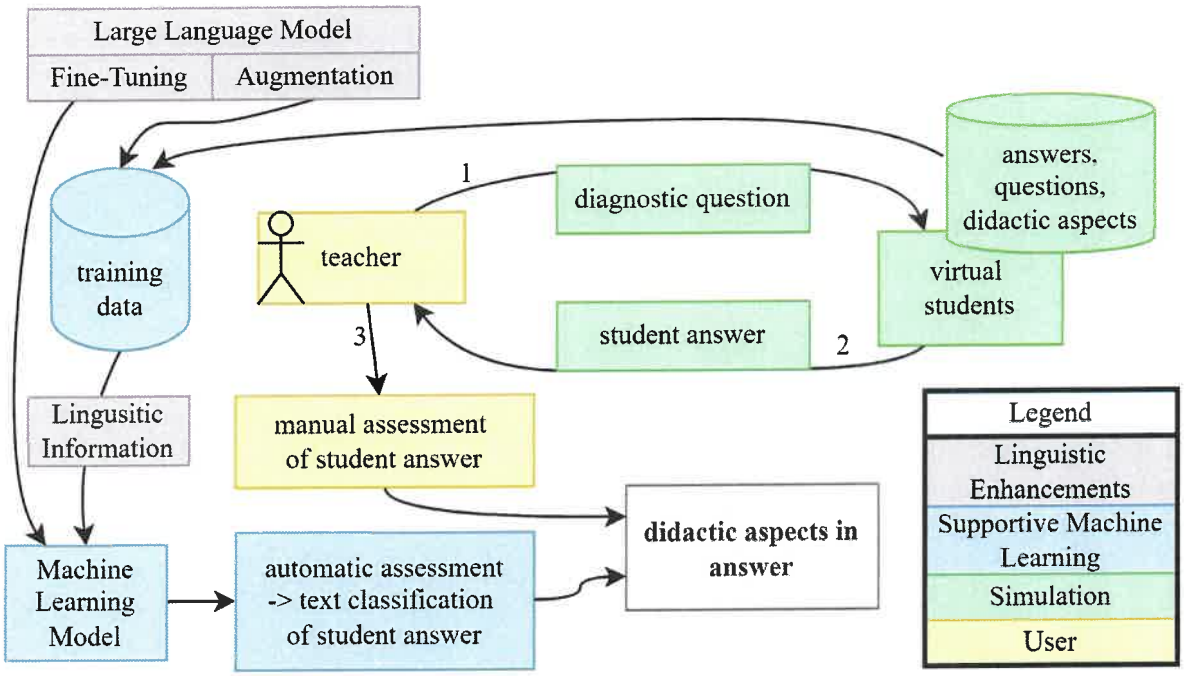

Abbildung 9 ML Implementation

Transformative Technologie: Bewertung und Akzeptanz

Im Bereich der Bildungsinnovation gewinnen E-Learning-Systeme zunehmend an Bedeutung, da sie das Potenzial bieten, den Kompetenzerwerb zu fördern und die Wissensverbreitung effizienter zu gestalten. Die Integration von Chatbot-Systemen in solche Umgebungen bringt dabei wesentliche Vorteile: Zum einen unterstützen sie Lernende während ihrer gesamten Bildungsreise, zum anderen schaffen sie einen offenen Kommunikationskanal für textbasierte Interaktionen. Fortgeschrittene Chatbots können personalisierte Antworten und proaktive Nachrichten generieren, die auf das individuelle Verhalten der Lernenden abgestimmt sind.

Eine umfassende Lernsystemarchitektur wurde entworfen, die eine intuitive Weboberfläche, ein Backend für die Klassenzimmersimulation, ein Cloud-basiertes Chatbot-Backend und eine Middleware umfasst. Letztere ermöglicht die proaktiven Funktionen des Chatbots und stellt eine nahtlose Integration der Lerninhalte sicher.

Im Rahmen der Studien wurden verschiedene Aspekte des Forschungsprojekts untersucht, darunter die Akzeptanz durch die Nutzer, die bevorzugte Interaktionsdichte und der gewünschte Informationsumfang. Darüber hinaus wurde das Chatbot-System kritisch evaluiert, wobei Attribute wie Persönlichkeit, Freundlichkeit, Benutzerfreundlichkeit und die Qualität der Interaktionen analysiert wurden.

Q-Code	Aspekt	Frage-Text
FB02_1	Interaktionshäufigkeit	Ich hätte mir mehr Interaktion mit dem Chatbot gewünscht.
FB02_2	Interaktionshäufigkeit	Der Chatbot war zu aufdringlich.
FB02_3	Tiefe der Informationen	Ich hätte mir gewünscht, dass der Chatbot kompetenter wäre.
FB02_4	Tiefe der Informationen	Der Chatbot brachte zu viel Wissen ein.
FB02_5	Persona des Chatbots	Die Persönlichkeit des Chatbots war realistisch und ansprechend.
FB02_6	Persona des Chatbots	Der Chatbot wirkte zu roboterhaft.
FB02_7	Persona des Chatbots	Der Chatbot war bei der Ersteinrichtung einladend.
FB02_8	Persona des Chatbots	Der Chatbot kam sehr unfreundlich rüber.
FB02_9	Benutzerfreundlichkeit	Der Chatbot war einfach zu navigieren.
FB02_10	Relevanz und Klarheit	Der Chatbot hat mich gut verstanden.
FB02_11	Relevanz und Klarheit	Die Antworten waren nützlich, angemessen und informativ.
FB02_12	Benutzerfreundlichkeit	Der Chatbot kam gut mit Fehlern oder Irrtümern zurecht.
FB02_13	Benutzerfreundlichkeit	Der Chatbot war sehr einfach zu bedienen

Von den 175 Teilnehmenden bevorzugte ein Großteil eine geringere Interaktionshäufigkeit (47 %) und eine moderate Informationstiefe (43 %). Die Freundlichkeit der Chatbot-Persona wurde von 48 % positiv bewertet, während die Ansprechbarkeit mit 61 % eine besonders hohe Zustimmung erhielt. Dennoch zeigten sich bei diesen Bewertungen starke Polarisierungen (Modus: 5, Standardabweichung: 1,12), was auf unterschiedliche Nutzererfahrungen hinweist. Die Relevanz und Klarheit der Antworten wurde überwiegend positiv bewertet (42 % bzw. 51 %). Indifferenz zeigte sich hingegen bei der Beurteilung der Persönlichkeit des Systems und seiner Fehlerbehandlungsfähigkeiten. Die allgemeine Benutzerfreundlichkeit erhielt eine positive Bewertung (63 %).

Zusammenfassend verdeutlicht die Umfrage eine grundsätzliche Akzeptanz von interaktivem und geführtem Lernen mit Chatbots. Gleichzeitig besteht jedoch erhebliches Verbesserungspotenzial, insbesondere bei der Vermittlung des Nutzens solcher Lernwerkzeuge. Zukünftige Arbeiten sollten sich auf die Verfeinerung der Anpassungsmechanismen, die Verbesserung der natürlichen Integration in die Lernumgebung und die Rationalisierung der Inhaltserstellung konzentrieren, um eine breitere Anwendbarkeit zu gewährleisten. Fortschritte bei der Verarbeitung von Texteingaben sowie die Integration automatisierter formativer Bewertungsmethoden könnten die Lernerfahrung zusätzlich bereichern.

Die detaillierte Vorgehensweise und Ergebnisse hierzu wurden in [V9] veröffentlicht. Ein Einblick in die Ergebnisse gibt Abbildung 10 Verteilungsmuster der Antworten.

Tabelle 1 Q-Codes

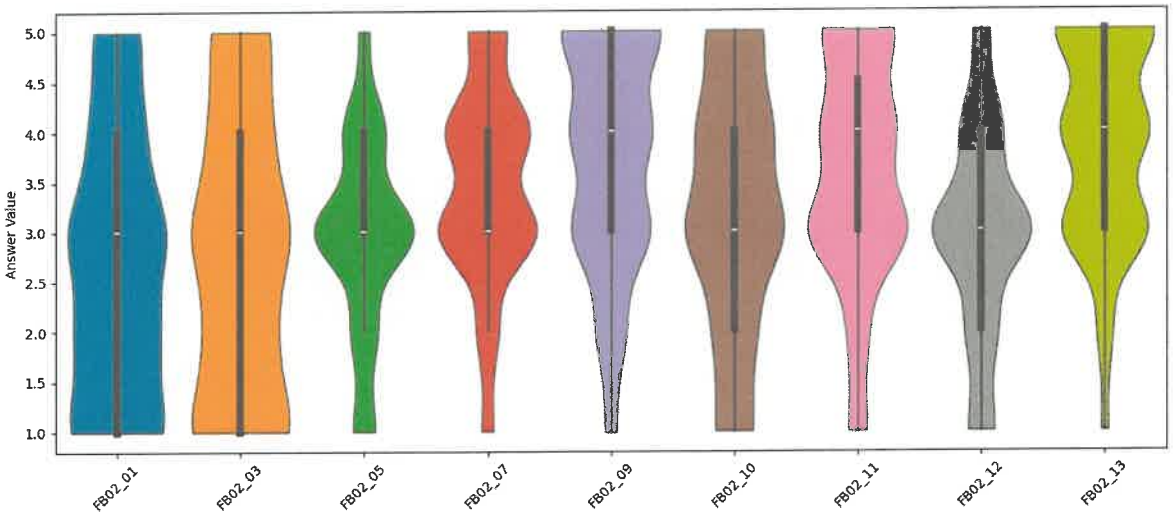

Abbildung 10 Verteilungsmuster der Antworten (Legende siehe Tabelle 1)

Präsentationen

Im Folgenden sind alle Tagungsbeiträge, Vorträge und Messestände aufgeführt, die während des Förderzeitraums von FiSK präsentiert wurden.

November 2021	4th Annual International Conference of Education, Research and Innovation. ICERI2021. Seville, Spanien. • Vortrag: Data-Driven Tutoring: challenges and prospects.
Mai 2022	connect2research Villingen-Schwenningen • Messestand
Dezember 2022	6th International Conference on Natural Language Processing and Information Retrieval. Bangkok, Thailand. • Vortrag: Linguistic driven feature selection for text classification as stop word replacement. 6th International Conference on Natural Language Processing and Information Retrieval
Juni 2023	Doctoral Day. Université de Haute-Alsace, Frankreich. • Vortrag: Feature Selection and ML Efficiency in FiSK
September 2023	Eighth International Conference on Green Communications, Computing and Technologies GREEN 2023. Porto, Portugal. • Vortrag: Comparable Machine Learning Efficiency: Metrics for NLP.

März 2024	6th International Conference on Natural Language Processing. ICNLP 2024. Xi'an, China. • Vortrag Linguistic-Aware WordPiece Tokenisation: Semantic Enrichment and OOV Mitigation
Juli 2024	Tag der Informatik Furtwangen • Messestand
Juli 2024	16th annual International Conference on Education and New Learning Technologies. EDULEARN24. Palma, Spain. • Vortrag: Transformative educational technologies: Chatbot integration and teacher trainee reception in simulated classrooms.
Juli 2024	The 2024 IARIA Annual Congress on Frontiers in Science, Technology, Services, and Applications. IARIA Congress 2024.Porto, Portugal. • Vortrag: Automatic assessment of student answers using large language models: Decoding didactic concepts.

Veröffentlichungen

Im Folgenden sind alle veröffentlichten Beiträge in Fachzeitschriften und Konferenz-Proceedings aufgeführt, die während des Förderzeitraums von FiSK entstanden sind.

[V1] Schönle, D., Reich, C., & Abdeslam, D. (2021). Data-Driven Tutoring: challenges and prospects. In ICERI2021 Proceedings: 4th Annual International Conference of Education, Research and Innovation, 8-9 November, 2021, Online (pp. 4584-4594).

[V2] Schönle D., Reich C. & Abdeslam D. (2022). Linguistic driven feature selection for text classification as stop word replacement. 6th International Conference on Natural Language Processing and Information Retrieval

[V3] Fiedler, D., Schönle, D., Reich, C. & Harms, U. (2022). Diagnosing evolutionary explanations: Effects of a chatbot support system on teacher students' diagnostic competence [Vortrag]. 13th Conference of European Researchers in Didactics of Biology (ERIDOB), Nicosia, Cyprus.

[V4] Schönle D., Reich C. & Abdeslam D. (2023). Linguistic driven feature selection for text classification as stop word replacement. Journal of Advances in Information Technology, vol. 14, no. 4, pp. 796–802, 2023. DOI: 10.12720/jait.14.4.796-802

[V5] Schönle, D., Reich, C., & Abdeslam, D. (2023). Comparable Machine Learning Efficiency: Metrics for NLP. In Eighth International Conference on Green Communications, Computing and Technologies GREEN 2023.

[V6] Schönle, D., Reich, C., & Abdeslam, D. (2024). Linguistic-Aware WordPiece Tokenisation: Semantic Enrichment and OOV Mitigation. 6th International Conference on Natural Language Processing (ICNLP 2024).

[V7] Schönle, D., Reich, C., Abdeslam, D. (2024). Streamlining AI: Techniques for Efficient Machine Learning Model Selection. The International Journal on Advances in Intelligent Systems 2024 Volume 17 Nr. 12.

[V8] Schönle, D., Fiedler, D., Harms, U., Reich, C. & Abdeslam, D. (2024). Transformative educational technologies: Chatbot integration and teacher trainee reception in simulated classrooms. 16th annual International Conference on Education and New Learning Technologies.

[V9] Schönle, D., Reich, C., Abdeslam, D., Fiedler, D., Harms, U, & Poser, J. (2024). Automatic assessment of student answers using large language models: Decoding didactic concepts. The 2024 IARIA Annual Congress on Frontiers in Science, Technology, Services, and Applications.

Wichtigste Positionen des zahlenmäßigen Nachweises

Die Positionen des zahlenmäßigen Nachweises entsprechen den im Antrag formulierten Vorstellungen und werden im Schlussverwendungsnachweis dargelegt.

In erster Linie wurden die Mittel für Personalausgaben verwendet. Des Weiteren entstanden Kosten für Software-Entwicklung und für die Nutzung von Cloud-Diensten. Die übrigen Ausgaben ergeben sich durch sachliche Verwaltungsmittel in Form von Reisemitteln für nationale und internationale Tagungen.

Notwendigkeit und Angemessenheit der geleisteten Arbeit

Die durchgeführten Arbeiten sowie die dafür aufgewandten Ressourcen waren notwendig und angemessen, da sie (1) der im Projektantrag dargelegten Planung entsprachen und (2) alle im Arbeitsplan formulierten Aufgaben bearbeitet wurden. Im Teilprojekt HFU konnten die Ergebnisse des Projekts FiSK in einer erfolgreichen Dissertation, angefertigt von Daniel Schönle, verwertet werden.

Literaturverzeichnis

[1] M. W. Ashfaque, S. Tharewal, S. Iqhbal, and C. N. Kayte. A Review on Techniques, Characteristics and approaches of an intelligent tutoring Chatbot system. In Proceedings of the Conference on Smart Innovations in Design, Environment and Computing, 2020.

[2] Moschitti and Basili. Complex linguistic features for text classification: A comprehensive study. In European conference on information retrieval, 2004.

[3] S. Qin, J. Song, P. Zhang, and Y. Tan. Feature selection for text classification based on part of speech filter and synonym merge. In 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2015.

[4] K. Jasleen and R. Saini Jatinderkumar. POS word class based categorization of Gurmukhi language stemmed stop words. In Proceedings of First International Conference on Information and Communication Technology for Intelligent Systems: Volume 2, 2016.

[5] R. W. Numrich, L. Hochstein, and V. R. Basili. A metric space for productivity measurement in software development. In Proceedings of the second workshop on Software engineering for high performance computing system applications, 2005.

[6] R. Fischer, M. Jakobs, S. Mücke, and K. Morik. A Unified Framework for Assessing Energy Efficiency of Machine Learning. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pages 39–54, 2022.

Abbildungen

Abbildung 1 Architektur Adaptive Lernumgebung Gen 1 4
Abbildung 2: Userinterface Simulations-System Generation 1 5
Abbildung 3 Simulations-System Generation 2 7
Abbildung 4 Userinterface Simulations-System Generation 2 7
Abbildung 5 Motivation und Methoden von Machine Learning in E-Learning Systemen 10
Abbildung 6 Herausforderungen und Chancen bei Verwendung von Machine Learning in E-Learning Systemen 10
Abbildung 7 Beispiel für WUP-Filterung 11
Abbildung 8 Beispiel für LinPair Tokenisation 12
Abbildung 9 ML Implementation 14
Abbildung 10 Verteilungsmuster der Antworten (Legende siehe Tabelle 1) 16

Tabellen

Tabelle 1 Q-Codes 15

Anlagen

Inhaltlicher Kurzbericht

Erfolgskontrollbericht

Berichtsblatt

HOCHSCHULE FURTWANGEN UNIVERSITY HFU

Projekt FiSK

Effekte adaptiver Feedbackbots im Simulierten Klassenraum auf prozedurales Professionswissen

Teilprojekt HFU

Adaptive Feedbackbots durch Machine Learning Methoden

Kurzfassung Abschlussbericht

Zuwendungsempfänger:	Hochschule Furtwangen
Förderkennzeichen:	16DHB4005
Vorhabensbezeichnung:	FiSK – Effekte adaptiver Feedbackbots im Simulierten Klassenraum auf prozedurales Professionswissen –
	Teilprojekt Adaptive Feedbackbots durch Machine Learning Methoden
Projektleitung:	Prof. Dr. Christoph Reich
Laufzeit des Vorhabens:	01.03.2021 bis 29.02.2024, Verlängerung bis 31.07.2024

FiSK – Kurzer inhaltlicher Bericht

Das Projekt **FiSK** zur Erforschung von adaptiven Lernumgebungen wurde erfolgreich abgeschlossen. Die entwickelte exemplarische Lernumgebung erfüllt alle Anforderungen und unterstützt sowohl die universitäre Ausbildung von Biologielehrkräften als auch die Datenerhebung im Rahmen der Forschung. Sie kombiniert innovative Simulationstechnologien mit fortschrittlicher künstlicher Intelligenz und bietet ein umfassendes System, das adaptive, individuelle und formative Rückmeldungen liefert, um die Lernprozesse der Studierenden optimal zu fördern.

Erfolge der Lernumgebung

- **Simulierter Klassenraum:**
  Ein realitätsnaher virtueller Klassenraum wurde erfolgreich implementiert. Lehramtsstudierende können in interaktiven Szenarien Schülerfragen analysieren, um verschiedene Stufen des Lernfortschritts zu diagnostizieren.
  - **Themenschwerpunkte:** Drei Themenbereiche, wie z. B. Evolution, wurden vollständig entwickelt. Die Inhalte decken Konzepte wie Evolutionsprinzipien, Schwellenkonzepte und Fehlvorstellungen der Schüler ab.
  - **Lernprozess:** Die Studierenden stellen vordefinierte Fragen, analysieren die Antworten der virtuellen Schüler und diagnostizieren das Vorhandensein oder Fehlen wichtiger Lernindikatoren. Dieser Prozess wird erfolgreich durch eine summative Bewertung ergänzt.
- **Interaktive Feedbacksysteme:**
  Das System bietet Lernenden eine Multiple-Choice-Umgebung, in der sie gezielt diagnostische Entscheidungen treffen können. Dies ermöglicht ein fundiertes Training im Umgang mit komplexen Lehrsituationen.

Chatbot-System: Reaktive und Proaktive Funktionen

Das Chatbot-System wurde vollständig umgesetzt und kombiniert reaktive und proaktive Fähigkeiten:

- **Reaktive Funktionen:** Der Chatbot beantwortet zuverlässig Fragen zu Lehrmaterialien und bietet kontextbezogenes Feedback.
- **Proaktive Funktionen:** Der Chatbot greift gezielt in den Lernprozess ein, um den Lernweg zu lenken. Dies verbessert die Lernerfahrung und sorgt für eine individuell angepasste Unterstützung.

Systemarchitektur

Eine hochmoderne Lernsystemarchitektur wurde erfolgreich entwickelt:

- **Weboberfläche:** Benutzerfreundliche und intuitive Gestaltung zur einfachen Bedienung durch Studierende.
- **Klassenzimmersimulations-Backend:** Zuverlässige Simulation von Unterrichtsszenarien.
- **Cloud-basiertes Chatbot-Backend:** Ermöglicht Skalierbarkeit und nahtlose Funktionalität.
- **Middleware:** Die proaktiven Funktionen des Chatbots werden durch eine leistungsfähige Middleware nahtlos in das Gesamtsystem integriert.

Forschungserfolge im Bereich KI und Textklassifikation

Der Fokus auf die automatisierte Klassifikation von Freitextinhalten führte zu bahnbrechenden Erkenntnissen und Innovationen:

- **Weighted-Unimportant-Part-of-Speech (WUP):**
  Dieses Verfahren reduziert erfolgreich gering-relevante Wortarten und steigert so den Informationsgehalt der Texte. Die Klassifikationsgenauigkeit konnte deutlich verbessert werden.
- **LinPair:**
  Die linguistischen Informationen aus Texten wurden gezielt genutzt, was die maschinelle Verarbeitung und Klassifikation signifikant optimierte.
- **Effizienz-Metrik:**
  Eine neu entwickelte Theorie zur Effizienz maschinellen Lernens wurde in die **Quality Focused Compact Efficiency Metric** umgesetzt. Die Metrik ermöglichte die präzise Messung der Effekte der entwickelten Verfahren und bestätigte deren Effizienz.

Ergebnisse und Auswirkungen

- **Verbesserte Fehlererkennung:**
  Die Diagnostik von Schülerfehlern wurde dank der entwickelten Verfahren signifikant verbessert.
- **Höhere Effizienz:**
  Die Systeme verarbeiten Informationen schneller und zielgerichteter, was zu einer erhöhten Effektivität in der Textklassifikation führte.
- **Breite Anwendbarkeit:**
  Die entwickelten Methoden zeigen vielversprechende Potenziale für weitere Anwendungen in der automatisierten Textverarbeitung und im maschinellen Lernen.

Das Projekt FiSK zeigt, wie adaptive Technologien und innovative KI-Verfahren die universitäre Lehre revolutionieren können. Alle gesetzten Ziele wurden erreicht, und die erzielten Ergebnisse bieten sowohl für die Ausbildung von Lehrkräften als auch für die Weiterentwicklung von KI-Systemen bedeutende Vorteile.