

Proposal zur Bachelorarbeit

Entwicklung einer KI-gesteuerten Erweiterung zur Umwandlung von Audioaufnahmen in verarbeitbare Objekte für eine Verwaltungssoftware

Nathan Gebremichael

Betreuung

- Prof. Dr. Oliver Hofmann, Betreuender Dozent der Technischen Hochschule Nürnberg
- Christian Koch, Informatiker, Lead Senior Developer bei mediendesign AG

Ausgangssituation

Die Ausgangssituation dieses Projekts ist eine von mediendesign AG-entwickelten Verwaltungssoftware, die von verschiedenen Versicherungen zur Verwaltung von Schadensfällen verwendet wird. Sachbearbeiter aus diesen Versicherungen können dabei neue Schadensmeldungen aufnehmen und bearbeiten, die relevanten Informationen in Formulare eintragen und im System abspeichern.

Eine mögliche Erleichterung dieses Prozesses könnte die Verwendung von Sprachaufnahmen sein, die der Bearbeiter vor Ort bei der Besichtigung des Schadensobjekts erstellt. Die Idee ist, diese Audioaufnahmen durch Künstliche Intelligenz (KI) in ein Objekt umzuwandeln, das dann im Backend der Verwaltungssoftware weiterverarbeitet werden kann. Dies würde den Prozess der Dateneingabe erheblich vereinfachen und beschleunigen, da der Bearbeiter nur eine Aufnahme abschicken muss und der Eintrag automatisch erstellt und entsprechend abgespeichert wird.

Die Herausforderung besteht darin, eine Lösung zu entwickeln, die in der Lage ist, die Audioaufnahmen effizient und genau zu transkribieren und in ein geeignetes Format für die Backend-Verarbeitung umzuwandeln. Darüber hinaus muss diese Lösung lokal gehostet werden können und lokale Modelle verwenden, um Datenschutz- und Sicherheitsanforderungen zu erfüllen.

Insgesamt bietet diese Ausgangssituation eine spannende Gelegenheit, die Möglichkeiten der Künstlichen Intelligenz zur Verbesserung von Verwaltungsprozessen zu erforschen und praktische Lösungen für reale Herausforderungen zu entwickeln.

Zielsetzung

Das Hauptziel dieses Projekts ist die Entwicklung einer KI-gesteuerten Erweiterung für eine Verwaltungssoftware, die in der Lage ist, Audioaufnahmen eines Bearbeiters entgegenzunehmen und diese automatisch in ein Format umzuwandeln, das von der Software verarbeitet werden kann. Dies soll erreicht werden, indem die Audioaufnahmen in Text transkribiert und dann in verarbeitbare Objekte umgewandelt wird.

Insgesamt soll dieses Projekt dazu beitragen, den Einsatz von Künstlicher Intelligenz in der Verwaltungssoftware zu erweitern und gleichzeitig praktische Lösungen für bestehende Herausforderungen zu bieten.

Vorgehensweise

Die Entwicklung einer KI-gesteuerten Erweiterung zur Umwandlung von Audioaufnahmen in verarbeitbare Objekte für Verwaltungssoftware erfordert eine sorgfältige Planung und Durchführung. Der erste Schritt in diesem Prozess ist die Anforderungsanalyse und Planung. Hierbei werden die genauen Anforderungen und Ziele des Projekts definiert, wobei besonderes Augenmerk auf die Beschränkungen und Anforderungen, wie die lokale Hosting-Anforderung und die Verwendung lokaler Modelle, gelegt wird.

Nachdem die Anforderungen klar definiert sind, folgt die Auswahl des Modells. Die Huggingface-Modelbibliothek bietet eine Vielzahl von Modellen, die speziell für Sprachverarbeitungsaufgaben wie Transkription oder Textgenerierung entwickelt wurden. [1] Ein passendes Modell wird ausgewählt und für die Integration in die Anwendung vorbereitet.

Der nächste Schritt ist die Entwicklung der Anwendung selbst. Dies umfasst die Einrich-

tung eines Servers für das Hosting der Anwendung, die Implementierung der Audioaufnahme-Funktionalität und die Integration des ausgewählten Modells. Besondere Aufmerksamkeit wird auf die Integration von LangChain [2] oder AutoGen [3] gelegt, um die Audioaufnahmen in Text umzuwandeln.

Sobald die Anwendung entwickelt ist, wird sie umfangreichen Tests unterzogen. Diese Tests stellen sicher, dass alle Komponenten korrekt interagieren und dass die Anwendung wie erwartet funktioniert. Eventuell notwendige Optimierungen werden in diesem Schritt durchgeführt.

Schließlich, nachdem sichergestellt wurde, dass alles korrekt funktioniert, wird die Anwendung bereitgestellt. Dieser Schritt markiert den Abschluss des Projekts und den Beginn der Nutzung der Anwendung durch Endbenutzer.

Diese Vorgehensweise stellt sicher, dass das Projekt systematisch und effizient durchgeführt wird, wobei jeder Schritt auf dem vorherigen aufbaut und zum erfolgreichen Abschluss des Projekts beiträgt.

Literatur

- [1] HuggingFace. *Hugging Face Hub documentation*. URL: <https://huggingface.co/docs/hub/index> (besucht am 27.10.2023).
- [2] LangChain. *Agents — Langchain*. Agents — Langchain. URL: <https://python.langchain.com/docs/modules/agents/> (besucht am 27.10.2023).
- [3] Microsoft. *Getting Started — AutoGen*. URL: <https://microsoft.github.io/autogen/docs/Getting-Started> (besucht am 27.10.2023).