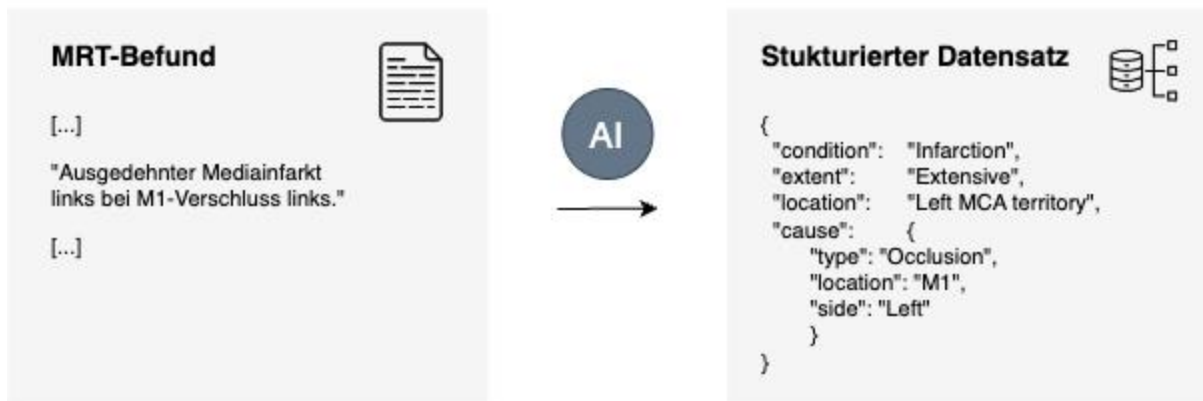
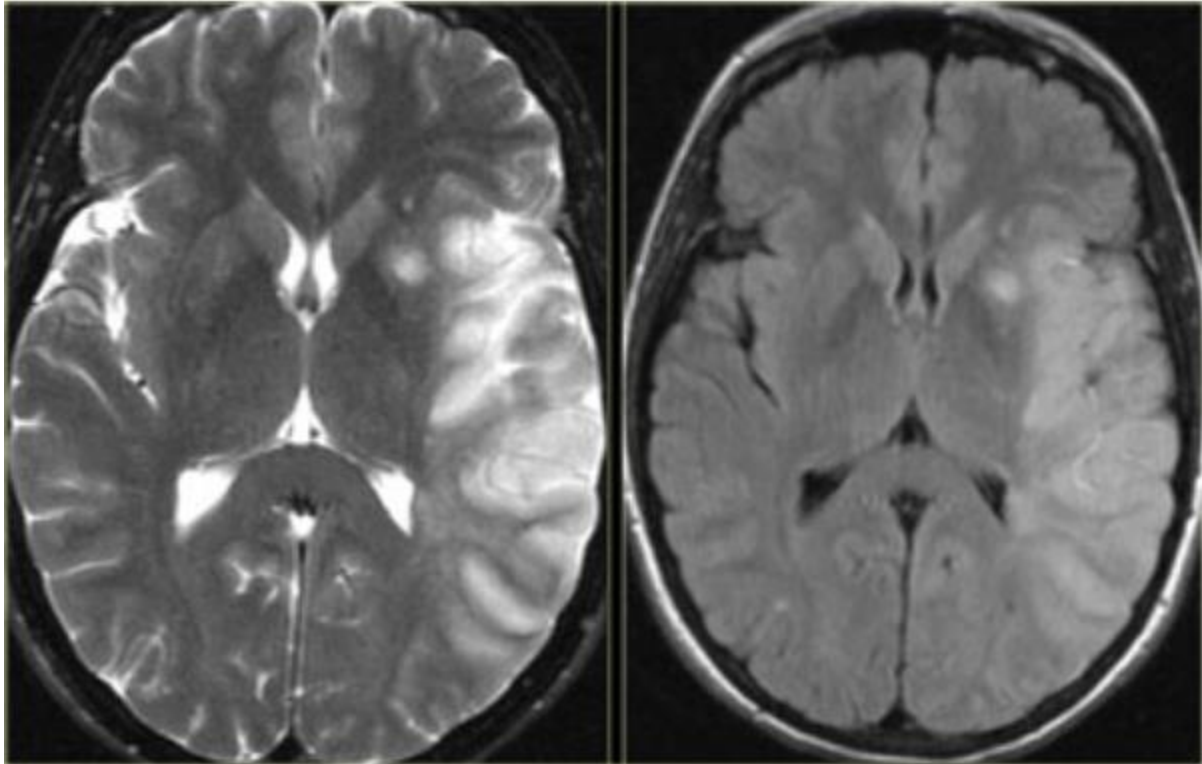


KI-gestützte Extraktion strukturierter Datenparameter aus neuroradiologischen Befundberichten



Bildquelle: RadiologyAssistant © 2024.

Institut

Abteilung für diagnostische und interventionelle Neuroradiologie, Klinikum rechts der Isar (TUM)

Hintergrund

Radiologische Befundberichte enthalten detaillierte diagnostische Informationen und stellen somit eine wertvolle Datenquelle für wissenschaftliche und klinische Auswertungen dar. Allerdings erfordert die angemessene Verwertung dieser Daten eine intensive manuelle Datenkuration.

In der Vergangenheit haben Arbeiten im Bereich des Natural Language Processings (NLP) versucht, den Prozess der Datenextraktion zu automatisieren. Jedoch haben sich diese Methoden nur sehr begrenzt durchgesetzt, unter anderem aufgrund der eingeschränkten Verfügbarkeit der notwendigen annotierten Trainingsdaten.

Jüngste Entwicklungen auf dem Gebiet der Large Language Models (LLMs) eröffnen neue Möglichkeiten für die automatisierte Datenextraktion. Diese Modelle künstlicher Intelligenz (KI) sind in der Lage, große Mengen an Textdaten zu verarbeiten und erzielen in vielen Bereichen auch ohne ein domänenspezifisches Training eine hohe Genauigkeit. Der Einsatz dieser KI-Modelle für die Datenextraktion aus radiologischen Befunden ist bislang jedoch noch nicht ausreichend erforscht.

Projektziel

Ziel des Vorhabens ist es daher, die Genauigkeit unterschiedlicher LLMs bei der Extraktion strukturierter Datenparameter aus neuroradiologischen Befundberichten zu untersuchen. Zu den Krankheitsbildern, die im Fokus stehen, gehören zerebrale Aneurysmata, der ischämische Schlaganfall und Wirbelkörperfrakturen. Darüber hinaus soll das Potenzial der extrahierten Daten für sekundäre Zwecke (z.B. krankheitsspezifische Register) exploriert werden. Interne Vorarbeiten zur LLM-gestützten Datenextraktion aus Schlaganfall-Befundberichten liegen vor.

Durchführung

- Auswahl geeigneter neuroradiologischer Befundberichte aus der lokalen Datenbank anhand prädefinierter Kriterien
- Deidentifikation der Befundberichte

- Erstellung von Datenmodellen zur Abbildung der zu erhebenden Datenparameter
- Datenextraktion mittels verschiedener LLMs (mithilfe eines Python-Skriptes)
- Datenauswertung und -visualisierung

Voraussetzungen

- Interesse an KI-Anwendungen in der Medizin
- Dauer: 6 - 9 Monate
- Strukturierte und eigenständige Arbeitsweise
- Teamfähigkeit, Zuverlässigkeit, Lernfähigkeit
- Sehr gute Englischkenntnisse
- Programmierkenntnisse (z.B. Python) sind von Vorteil, aber nicht zwingend notwendig

Angebot

- Intensive Betreuung mit regelmäßigen Meetings
- Einarbeitung und Anleitung durch einen Clinician Scientist mit geschützter Forschungszeit
- Publikationen in peer-reviewed Journalen sind geplant

Ein Beginn ist ab sofort möglich. Bei Interesse bitte Bewerbung (inkl. Motivationsschreiben, Lebenslauf, Physikumszeugnis und Transcripts) an: suhwan.kim@tum.de

Projektbetreuer: Dr. med. Su Hwan Kim, MSc

Doktorvater: PD Dr. med. Dennis Hedderich, MHBA