

Data Driven Diagnosis - Diagnostic and Documentation Quality

Studiengang: BSc in Informatik | Vertiefung: Data Engineering
Betreuer: Prof. Dr. Jürgen Vogel

In dieser Arbeit wurden medizinische Diagnosen mit mehreren Machine Learning Verfahren nach Niereninsuffizienzen klassifiziert. Dabei wurde einerseits mit klassischen Modellen gearbeitet, jedoch auch mit eher neueren neuronalen Netzen. Die Arbeit wurde in Zusammenarbeit mit der Insel Gruppe durchgeführt.

Ausgangslage

Ein Forschungsteam der Insel Gruppe hat eine Studie zur genauen Bestimmung von Graden der Niereninsuffizienz gemacht. In dieser Studie wurde ein Algorithmus, mit dem Niereninsuffizienzen mittels Blutwerten automatisch erkannt und genau definiert werden können, entwickelt.

Durch diesen Algorithmus wurde festgestellt, dass viele Niereninsuffizienzen nicht in den schlussendlich kodierten Diagnosen auftauchen. Es soll nun ein Modell erarbeitet werden um herauszufinden, ob die Ärztinnen und Ärzte die Insuffizienzen nicht erkennen oder ob diese nicht relevant für die Abrechnung sind und daher nicht kodiert werden. Dazu sollen Abschlussberichte mit einem Modell für maschinelles Lernen klassifiziert werden. Bisher wurde das mit einer einfachen Schlagwortsuche gemacht.

Ergebnisse

Während der Literaturrecherche wurde festgestellt, dass dies keine neue Aufgabenstellung ist und es insbesondere für den Englischen und Chinesischen Sprachraum bereits erste Versuche gegeben hat. Die besten Ergebnisse wurden mit neuronalen Netzwerken erzielt. Aus diesem Grund sollte in dieser Arbeit einerseits die generelle Machbarkeit der Fragestellung an sich erforscht, andererseits der direkte Vergleich zwischen einem neuronalen Netzwerk und klassischen Modellen für maschinelles Lernen gezogen werden.

Klassische Modelle für maschinelles Lernen

Es wurden vier klassische Modelle für maschinelles Lernen implementiert: K-Nearest Neighbor, Logistische Regression, Naive Bayes und Random Forest. Zudem wurden die beiden Vektorisierer Bag of Words und TF-IDF umgesetzt. Die Präzision der verschiedenen Kombinationsmöglichkeiten konnte durch verschiedene Anpassungen an den Modellen (z.B. Zwei-Stufen Klassifizierung) und am Inhalt der Daten (mit Part of Speech Tagging) von 22% auf 58% ange-

hoben werden. Niereninsuffizienzen generell konnten mit einer Präzision von bis zu 97% auf den Testdaten bestimmt werden.

Neuronales Netzwerk

Als neuronales Netz wurde ein BERT Modell auf das Dataset der Insel Gruppe angewendet. Aus Zeitgründen wurde ein vortrainiertes Modell der Firma deepset.ai verwendet. Dieses wurde mit Daten aus Wikipedia, Rechtssprechungen und Nachrichtenartikeln trainiert und erzielt für verschiedene NLP-Tasks sehr gute Werte. Während dem Training, welches sich über mehrere Tage hinzog, traten verschiedene Probleme mit den Hyperparametern auf, welche zu keinem akkuraten Ergebnis führten. Das Potenzial eines BERT Modells ist jedoch sehr gross, wie in verwandten Werken dokumentiert wurde.

Fazit

Es konnten mit beiden Methoden nur ungenügende Ergebnisse erzielt werden. Dies liegt einerseits am sehr unausgeglichene Datenset mit relativ wenigen Datensätzen pro Kategorie (teilweise nicht einmal zweistellig). Deshalb war ein Zusammenfassen zu generelleren Klassen unumgänglich. Dadurch ging jedoch ebenfalls die Möglichkeit verloren, akut auf chronische (acute-on-chronic) Niereninsuffizienzen zu erkennen.

Auch bei den zusammengefassten Datensets konnten nur ungenügende Ergebnisse erzielt werden, da es trotzdem noch immer sehr wenig Datensets waren. Die Performance der verschiedenen Modelle konnten jedoch kontinuierlich verbessert werden und erreichten eine Verbesserung, verglichen mit dem ursprünglichen Modell, von fast 20% (39% auf 58%).

Ausblick

Diese Arbeit hilft der Insel Gruppe und sie können auf dieser Basis weiter forschen. Die Modelle könnten mit Blutwerten oder Angaben von Medikamentengaben ergänzt und so verbessert werden.



Steve Blaser



Michelle Lüscher