

Retrospektive Klassifikation von Sepsis-Patienten in intensivmedizinischen Sekundärdaten mit Machine Learning

(Bachelorarbeit)



CLEMENS KÖHLER

Motivation

Im Rahmen des BMBF-geförderten Projekts "SMITH – Smart Medical Information Technology for Healthcare", welches eine Schnittstelle zwischen verschiedensten Forschungsteams und Unikliniken bietet um die datenbasierte Medizin zu standardisieren, wird das Krankheitsbild der Sepsis untersucht.

Sepsis ist eine oft tödlich verlaufende Komplikation, bei der das körpereigene Immunsystem auf eine in die Blutbahn übergegangene Infektion überreagiert. Auch im Falle einer Genesung kann es häufig zu bleibenden Organschäden kommen, daher ist die frühe Erkennung dieser Erkrankung mit entsprechender Therapie von entscheidender Bedeutung.

Die Erkennung von Sepsis ist auch nach aktuellen medizinischen Standards sehr schwierig, darum ist es sinnvoll die Verwendung von einer durch Machine-Learning unterstützte Analyse zu untersuchen um eine Sepsis früher erkennen zu können.

Stand der Technik

In der Literatur finden sich einige Ansätze welche sich mit der Klassifikation von Sepsis in Patientendaten beschäftigen. Dabei sind einige Machine Learning Methoden vielversprechend. Hierbei muss jedoch beachtet werden, dass die Ansätze abweichende Sepsis Definitionen verwenden, die das Krankheitsbild verschieden stark eingrenzen.

Dabei wird zwischen verschiedenen Verfahren unterschieden um die Daten aus den Patientendatensätzen zu extrahieren, aber auch zwischen verschiedenen Machine Learning Verfahren. Oft gesehen werden hier Verfahren wie Naïve Bayes oder die Verwendung von Support Vector Machines. Die meisten Veröffentlichungen stammen aus den USA und verwenden den MIMIC Datensatz, weswegen eine Anwendung auf europäische Datensätze sinnvoll ist.

Zielsetzung

Ziel der Arbeit soll es sein, durch Literaturrecherche und eigenes Testen in vorhandenen Datensätzen, Sepsis zu klassifizieren und dabei herauszufinden welche Methoden zur Sepsisklassifikation am besten geeignet sind. Es wird sich hierbei ausschließlich auf die nicht neuronalen Machine Learning Verfahren beschränkt. Dabei wird ein aus der Literatur ausgewähltes Verfahren adaptiert und auf unseren Daten evaluiert.

Geplante Vorgehensweise

Die ersten Schritte beinhalten eine ausführliche Literaturrecherche um vielversprechende Methoden zur Sepsisklassifikation in der Literatur zu finden. Nach einer Einarbeitung in die entsprechenden Verfahren werden diese gegenübergestellt, miteinander verglichen und die besten bestimmt. Nach einer Einarbeitung in die Datensätze sollen die ausgewählten Methoden (re-)implementiert werden und die eigenen Ergebnisse mit den Ergebnissen aus der Literatur verglichen werden.