

Zusammenfassung

Der *septische Schock* stellt trotz jahrelanger medizinischer Forschung den Intensivmediziner vor ein großes Rätsel. Warum führt der septische Schock zum Tod bei einer hohen Anzahl der Patienten und warum überleben wiederum andere? Diese Fragestellung ist zentral im DFG-Projekt *MEDAN* (Abk. für: Medizinische Datenanalyse mit Neuronalen Netzen), das den Anstoß für diese Dissertation gab. Es werden aber allgemeiner aktuelle Methoden aus den Bereichen *Data Mining*, *Knowledge Discovery*, *Machine Learning* und *Intelligent Data Analysis* vorgestellt und zusammen mit Methoden der *Statistik* unter dem Begriff *Adaptive Datenanalyse* subsumiert.

In dieser Arbeit werden also Methoden der adaptiven Datenanalyse anhand der Originalliteratur vorgestellt, weiterentwickelt und angewendet. Um neben allgemeineren Betrachtungen eine Vertiefung zu erreichen, wird insbesondere die *Regelgenerierung* mit integrierter *Klassifikation* als vielversprechende Analysemethode genauer untersucht. Neue Algorithmen werden formuliert und bekannte werden weiterentwickelt, sowohl für metrische (*Neuro-Fuzzy-System*) als auch für symbolische Daten (*Generalisierungsregeln*). Im Hinblick auf die Anwendung wird auch eine Kombination der betrachteten Systeme untersucht.

Aufbauend auf dem Neuro-Fuzzy-System von K.-P. Huber und M. Berthold (*Fuzzy-RecBF-DDA*) werden entscheidende Verbesserungen und Erweiterungen für die Analyse metrischer Daten vorgenommen, die insbesondere das Überlappungsverhalten der Aktivierungsbereiche der Neuronen betreffen, und die eine robustere Regelgenerierung mit integrierter Klassifikation ermöglichen. Die Ansätze zur Generierung von *durchschnittsbasierten Generalisierungsregeln* (*GenDurchschnitt*) aus symbolischen Daten werden komplett neu entwickelt und auf ein theoretisches Fundament gestellt.

Die Leistungsfähigkeit dieser Entwicklungen wird anhand von *Benchmark-Daten* evaluiert. Um vertiefte Erkenntnisse und wissenschaftlich verwertbare Ergebnisse über die Problematik des septischen Schocks zu gewinnen, werden anhand der Daten dreier großer *Patienten-Datenbanken* umfangreiche Analysen vorgenommen. Dabei werden auch die Eigenschaften und die Grenzen der verwendeten Verfahren aufgezeigt, da die untersuchten Realweltdaten im Vergleich zu den Benchmark-Daten schwieriger zu handhaben sind, z.B. aufgrund fehlender Werte oder aufgrund der Individualität der intensivmedizinischen Patienten. Da insbesondere in der Medizin ein sorgfältiges Vorgehen wichtig ist, um sinnvolle, verwertbare Ergebnisse zu erhalten, wird immer (nach adäquater Vorverarbeitung der Daten) sorgfältig mit *Testdaten* gearbeitet und es werden immer *Versuchswiederholungen* durchgeführt.

Neben den Anwendungsanalysen wird die *Extraktion von scharfen Regeln* aus unscharfen Neuro-Fuzzy-Regeln untersucht, u.a. um optimale Regeln zu erhalten. Einheitlich für die verschiedenen Regelverfahren wird ein neues Paradigma der Merkmalswahl anhand einer neu eingeführten *Wichtigkeit* diskutiert. Ein neu entwickeltes *Ähnlichkeitsmaß* für Regeln und Regelmengen ermöglicht eine redundanzfreie Präsentation von Regeln und eine automatische *Modellwahl*. Überlegungen zur *Frühzeitigkeit* von Regeln werden ebenfalls angestellt. Die Erkenntnisse werden in die Anwendung umgesetzt. Dem Intensivmediziner können statistisch plausible, verständliche, exakte Regeln präsentiert werden, die manuell erzeugten Regeln deutlich überlegen sind.

Anhand der Daten des septischen Schocks wird aufbauend auf der Regelgenerierung ein *Alarmsystem* erstellt, das den Arzt warnt, wenn sich ein septischer Schock-Patient in einem kritischen bzw. sehr kritischen Zustand befindet. Die Leistung des Alarmsystems wird mit den in der Medizin oft verwendeten *Scores* verglichen. Dieses Alarmsystem könnte in Zukunft den z.Z. leistungsfähigsten Score (SOFA-Score) in der klinischen Praxis ablösen, da es für eine sehr gute Diagnoseleistung nur drei anstelle von zehn Variablen zur Bewertung eines Patienten benötigt.

Schlüsselworte: Medizinische Datenanalyse, Septischer Schock, Adaptive Datenanalyse, Neuronale Netze, Klassifikation, Score, Regel, Neuro-Fuzzy, Maschinelles Lernen, Generalisierungsregel, Wichtigkeit, Ähnlichkeit, Frühzeitigkeit, Alarmsystem