

# Aktivitätsklassifizierung aus Daten eines Multi-Sensor-Gerätes mit verschiedenen Maschinen-Lern-Algorithmen

Hornberger A, Marscholke M, Wolf KH  
 Institut für Medizinische Informatik Technische Universität Braunschweig, Deutschland  
 a.hornberger@tu-bs.de

**Einleitung und Fragestellung** Die Zunahme der Lebenserwartung in Verbindung mit sinkenden Geburtenraten kehrt das Verhältnis zwischen jungen und alten Menschen in Deutschland um. Die Minderung der Lebensqualität im höheren Alter beruht zum Teil auf einer Abnahme an Eigenständigkeit, die wiederum auf Einschränkungen der physischen Aktivität zurückzuführen ist [1]. Besonders bei alten Menschen ist die Gefahr hoch, dass sich Bewegungsradius und körperliche Aktivitäten mit der Zeit immer weiter verringern. Ziel assistierender Gesundheitstechnologien ist es, die Aktivität von alten Menschen zu fördern und dadurch ihre Eigenständigkeit möglichst lange zu erhalten. Das Ziel dieser Arbeit ist es, Algorithmen zu vergleichen, die körperliche Aktivitäten aus Sensordaten klassifizieren und es erlauben, automatisch ein Aktivitätsprofil zu erstellen

**Material und Methode** Basis für die Erstellung von Aktivitätsprofilen sind Messwerte wie z.B. Beschleunigungen, Hautwiderstand, Wärmefluss, Haut- und Umgebungstemperatur. Die Messdaten wurden mit einem Multi-Sensor-Gerät erfasst (SenseWear Pro 2, Bodymedia, Pittsburgh, USA) [2]. Dieses Gerät wird am hinteren Oberarm getragen und stellt insgesamt 20 Datenkanäle zur Verfügung. Im Rahmen einer Übung zur Vorlesung „Assistierende Gesundheitstechnologien“ zeichneten fünf Studenten, die alle den gleichen Parcours an Aktivitäten durchliefen, Daten auf und annotierten diese. Der Parcours bestand aus den Aktivitäten Stehen, Sitzen, Essen, Bildschirmarbeit, Fernsehen, Fahrrad-Ergometrie, Gehen, Treppensteigen und Liegen. Zusätzlich absolvierten die Studenten ein Training in einem 3-Achsen Trainingsgerät („FPZ 3D: The Spacecurl“). Aus den aufgezeichneten 20 Datenkanälen wurden für das Testen der Algorithmen 13 Kanäle ausgewählt (vgl. Tab. 1). Die übrigen Kanäle beinhalten aus anderen Kanälen berechnete Informationen und die Zeit. Die Daten wurden zur weiteren Verarbeitung in Blöcke von 15s und 60s zusammengefasst. Alle Blöcke, die keine Annotationen oder einen Wechsel der Aktivitäten enthielten, wurden entfernt. Die Aggregation umfasst die Summen bzw. Mittelwerte der einzelnen Parameter. Für die Analyse der Daten wurde das WEKA Toolkit (Waikato Environment for Knowledge Analysis, ver 3.4.7) verwendet, das viele Maschinen-Lern-Algorithmen zur Verfügung stellt [3]. Erste Tests mit den Algorithmen zeigten, vor allem bei den sitzenden Aktivitäten wie Fernsehen, Essen und Bildschirmarbeit, fehlerhafte Klassifikationen. Daher wurden diese drei Aktivitäten in „Sitzen“, sowie Treppensteigen in „Gehen“ zusammengefasst. Eine Literaturrecherche ergab, dass frühere Arbeiten vor allem die Algorithmen BayesNet, NaiveBayes, C4.5 und K-nearest-neighbor erfolgreich verwendeten. NaiveBayes arbeitet mit Standardwahrscheinlichkeiten, BayesNet erstellt einen gerichteten, nicht-zyklischen Graphen mit Wahrscheinlichkeiten, C4.5 erzeugt einen Entscheidungsbaum und K-nearest-neighbor wählt den Wert mit der geringsten euklidischen Distanz. Die Algorithmen wurden jeweils mit den Daten von vier Probanden trainiert und dann mit denen des fünften Probanden getestet. Der anschließend berechnete Mittelwert der korrekten Klassifikationen dient zum Vergleich der Algorithmen.

**Ergebnisse** Der Anteil an richtig klassifizierten Tätigkeiten liegt bei allen vier Algorithmen zwischen 36,28 % und 78,86 % bei der Aggregation auf 60s und zwischen 41,14 % und 72,08 % bei der Aggregation auf 15s. Der Algorithmus BayesNet schneidet im Mittel über alle Kombinationen von Datensätzen mit 60,81 % (60s) und 59,29 % (15s) am besten ab. K-nearest-neighbor liefert die besten Einzelergebnisse, im Schnitt klassifiziert er allerdings nur mit 54,65 % (60s) und 54,85 % (15s) Genauigkeit. In der typischen Confusion Matrix (vgl. Tab. 2), des BayesNet Algorithmus, trainiert mit den Daten (15s) des Autors AH, ist zu erkennen, dass Sitzen und Stehen am häufigsten falsch klassifiziert sind. Die Confusion Matrizen der anderen Algorithmen sind vergleichbar. Ebenso klassifizieren sie die Fahrrad-Ergometrie häufig falsch als Sitzen oder Spacecurl-Training.

**Diskussion** Die Algorithmen lassen sich anhand der erzielten Klassifikationsgenauigkeit deutlich in eine Reihenfolge bringen. Keiner der Algorithmen liefert wirklich gute Ergebnisse. Sie liegen im Rahmen der in vergleichbaren Arbeiten erzielten Genauigkeiten [4-5]. Ursächlich für die wenig exakte Klassifizierung durch alle Algorithmen dürfte die Art der verwendeten Daten sein. Die meisten Fehler entstehen bei der Unterscheidung von Stehen und Sitzen. Im zugrunde liegenden Laborversuch bemühten sich die Studenten so gut wie möglich zu sitzen und zu stehen. Möglicherweise verfälscht die – im Vergleich zum Alltag – extrem ruhige Haltung die Ergebnisse. Darüber hinaus wird eine inter-individuelle Klassifikation durchgeführt. Die Algorithmen lernen von einem Individuum und erzielen die schlechten Ergebnisse bei dem Versuch, die Daten einer anderen Person zu klassifizieren. Die Erstellung eines Klassifizierers mit den realen Daten der Person, deren weitere Aktivitäten zu klassifizieren sind, dürfte die Erkennungsgenauigkeit deutlich erhöhen. Bei einem solchen Klassifizierer könnte darüber hinaus die Tageszeit in die Bewertung eingehen, was weitere Anhaltspunkte für die Klassifikation liefern dürfte. Weitere Kontextinformationen, wie die Wahrscheinlichkeit des Überganges von einer Aktivität zur nächsten, grenzen die Entscheidungsmöglichkeiten der Algorithmen weiter ein und führen zu genaueren Ergebnissen. Um eine größere Trennschärfe zwischen Sitzen und Stehen zu erreichen, wäre die Verwendung weiterer Sensoren z.B. am Oberschenkel möglich. Das Gleiche gilt für eine bessere Klassifizierung der Fahrrad-Ergometrie. Die verschiedenen Aggregationen (15s und 60s) zeigen keine signifikanten Unterschiede in der Fehlerrate (vgl. Tab. 3). Das deutet darauf hin, dass der Spielraum bei der Wahl eines Aggregationsintervalls sehr groß ist. Weitere Untersuchungen sollten die für die Klassifikation optimale Länge des Intervalls ermitteln. Schlussfolgernd kann gesagt werden, dass alle Algorithmen geeignet sind, um ein passables Aktivitätsprofil zu erstellen. Weitere Arbeiten sind notwendig, um die Klassifikationsgenauigkeit zu erhöhen.

## Literatur

- [1] E. McAuley, J. F. Konopack, R. W. Motl, K. S. Morris, S. E. Doerken, and K. R. Rosengren, "Physical activity and quality of life in older adults: influence of health status and self-efficacy," *Ann Behav Med*, vol. 31, pp. 99-103, 2006.
- [2] D. Andre and A. Teller, "Health. Care. Anywhere. Today," *Stud Health Technol Inform*, vol. 118, pp. 89-110, 2005.
- [3] E. Frank, M. Hall, L. Trigg, G. Holmes, and I. H. Witten, "Data mining in bioinformatics using Weka," *Bioinformatics*, vol. 20, pp. 2479-81, 2004.
- [4] J. Parkka, M. Ermes, P. Korpipaa, J. Mantjarvi, J. Peltola, and I. Korhonen, "Activity classification using realistic data from wearable sensors," *IEEE Trans Inf Technol Biomed*, vol. 10, pp. 119-28, 2006.
- [5] L. Bao and S. S. Intille, "Activity Recognition from User-Annotated Acceleration Data," presented at Pervasive, 2004.

Anzahl der Extrema in der transversalen Beschleunigung
Anzahl der Extrema in der longitudinalen Beschleunigung
Mittlerer Wärmefluss
Hauttemperatur
Mittlere Beschleunigung transversal
Mittlere Beschleunigung longitudinal
Umgebungstemperatur
Mittlere Wertdistanz der transversalen Beschleunigung
Mittlere Wertdistanz der longitudinalen Beschleunigung
Schrittzahl
Mittlerer Hautwiderstand
Energieverbrauch pro Minute
Anmerkung

Tab. 1 Verwendete Parameter zur Klassifizierung durch die Algorithmen

a	b	c	d	e	f	<-- klassifiziert als
165	0	0	0	3	0	a = Sitzen
1	98	0	0	0	0	b = Gehen
14	1	39	0	0	0	c = Liegen
0	8	0	24	0	15	d = Spacecurl
66	1	0	0	0	0	e = Stehen
28	7	0	15	0	20	f = Ergometer

Tab. 2 Confusion Matrix des Algorithmus BayesNet mit 15s Blöcken eines Probanden

	BayesNet	NaiveBayes	K-nearest-neighbor	C4.5
15s	59,29	53,40	54,85	52,91
60s	60,81	55,74	54,65	52,73

Tab. 3 Klassifikationsgenauigkeiten der Algorithmen in Prozent