

Das FIF will seine Möglichkeiten nutzen, den Fehlentwicklungen der Informatiknutzung entgegenzuwirken und ihren Einsatz zum Nutzen der Gesellschaft zu fördern. Wir tun dies durch unsere inhaltliche Arbeit, durch unsere Kampagnen, Publikationen, Konferenzen und durch unsere Stellungnahmen. Und wir verleihen nun zum sechsten Mal den FIF-Studienpreis.

Mit unserem Studienpreis wollen wir

- hervorragende Abschlussarbeiten auszeichnen und die dahinterstehende
- Aufmerksamkeit für Themengebiete schaffen, die wir für die Entwicklung einer „digitalen“, „digitalisierten“ Gesellschaft für essentiell wichtig halten.

Technik ist von der historischen Entwicklung her auf militärische, industrielle und wirtschaftliche Zwecke ausgerichtet. Funktionierende Technik wird stets auch angewendet – unabhängig von ih-

rer Rechtmäßigkeit und ihrer Sozialverträglichkeit. Doch es ist die gesellschaftliche Aufgabe der Informatikerinnen und Informatiker, technische Systeme auch von ihren ethischen, sozialen und rechtsstaatlichen Anforderungen her zu denken, um eine Technik zu verhindern, die zum Selbstzweck wird und schädliche Nutzung als „Sachzwang“ etabliert. Mit unserem Studienpreis wollen wir

reife von Arbeiten bei uns einbedanken. Eine Jury, in diesem

- Professorin Britta Schinzel aus Freiburg,
- Professor Jochen Koubek aus Bayreuth,
- Rainer Rehak aus Berlin,
- und mir selbst, Stefan Hügel aus Frankfurt am Main,

hat aus den Einreichungen für den Studienpreis 2017 eine Arbeit ausgewählt, die wir heute hier prämiieren werden.

erschienen in der FIF-Kommunikation,  
herausgegeben von FIF e.V. - ISSN 0938-3476  
[www.fif.de](http://www.fif.de)

FIF e.V. – Stefan Hügel: Laudatio für den 1. Preis

## Tobias Krafft: Qualitätsmaße binärer Klassifikatoren im Bereich kriminalprognostischer Instrumente der vierten Generation

### Masterarbeit an der Technischen Universität Kaiserslautern

Die Anwendung der künstlichen Intelligenz – beispielsweise in Form maschinellen Lernens und statistischer Klassifikation – breitet sich aus und nährt bei manchen große Hoffnungen, zur Lösung vieler Probleme beitragen zu können. Ihre Instrumente und Methoden, wie statistische Klassifikatoren, die in der Öffentlichkeit gelegentlich, modisch schick und populär vereinfachend, kurzerhand als „Algorithmen“ bezeichnet werden – sollen für unterschiedliche Zwecke genutzt werden, die nicht zuletzt neue ethische Fragen aufwerfen.

Besonders schwierig werden diese Fragen, wenn Entscheidungen davon abhängen, die für Betroffene weitreichende Folgen haben oder haben können.

In der Kriminalprognostik ist dies zweifellos der Fall. Es geht um die Frage, wie hart Straftäterinnen und Straftäter bestraft werden sollen, ob sie in Sicherungsverwahrung genommen werden oder ob sie vorzeitig auf Bewährung entlassen werden – kurz: wie ihr Leben nach der Straftat verlaufen wird. Wie Tobias Krafft in seiner Arbeit, die wir heute auszeichnen werden, gleich in der Einleitung feststellt:

*„Der Bürger eines jeden demokratischen Rechtsstaates erwartet von der Justiz Gerechtigkeit und einen höchstmöglichen Schutz vor jeglicher Art von Angriff auf seine Person und Rechte. Jedoch sollte ein Angeklagter im Falle der nachgewiesenen Schuld ein gerechtes Urteil im Hinblick auf seine Straftat und eine dementsprechend angemessene Strafe erwarten können. In früheren Zeiten oblag es allein dem Richter und etwaigen Beratern, Persönlichkeit und zukünftiges Legalverhalten eines*



*Angeklagten einzuschätzen und beispielsweise zu entscheiden, ob dieser eine Gefahr für die Öffentlichkeit darstellt und verwahrt werden muss, oder ob die Strafe zur Bewährung ausgesetzt werden kann.“*

Da die Entscheidung einer Richterin oder eines Richters von persönlichen Einschätzungen und Bewertungen, unsicheren Vermutungen, individuellen Wertvorstellungen und kulturellen Rahmenbedingungen abhängt, ist die Versuchung groß, sie einem „objektiven“ Rechner aufzuerlegen und damit zu „versachlichen“.

Durch unterschiedliche Kategorien von Systemen der künstlichen Intelligenz erhofft man sich Hilfe in der Forensik, Kriminologie und bei richterlichen Entscheidungen: Da gibt es einmal die Systeme des maschinellen Lernens mittels der symbolischen

bzw. subsymbolischen KI, etwa konnektionistische Systeme bzw. künstliche Neuronale Netze. Hier werden auch Support-Vector-Maschinen für überwachtes und unüberwachtes Lernen als binäre Klassifikatoren, etwa zur Lügendetektion, benutzt. Zum anderen gibt es die auf statistischen bzw. stochastischen Methoden beruhenden Kriminalprognoseinstrumente wie Systeme der algorithmischen Entscheidungsfindung, welche in der von uns ausgezeichneten Arbeit ausschließlich betrachtet werden.



Die Juroren bei der Preisverleihung, Fotos: Dagmar Boedicker

Nicht zu vergessen die Frage der Verantwortung: Eine Maschine – so unsere Überzeugung – kann niemals so etwas wie „Verantwortung“ übernehmen. Doch wer kann es dann? Die Wissenschaftlerinnen und Wissenschaftler, die einen Klassifikationsalgorithmus entwerfen, auf dessen Basis lange Zeit später eine falsche Rückfallprognose gestellt wird? Die Programmierinnen und Programmierer, die irgendwann einmal die dabei verwendete Software implementiert haben? Die Richterinnen und Richter, die sie anwenden, ohne das Ergebnis anhand von anderen Kriterien zu überprüfen?

Kann es unser Ziel sein, die Urteile von in diesen Fragen erfahrenen Richterinnen und Richtern durch die Einschätzung von – vielleicht in rechtlichen Dingen völlig unerfahrenen – Informatikerinnen und Informatikern zu ersetzen? Statt einer erfahrenen Beurteilung im konkreten Einzelfall eine verallgemeinerte, in einen Klassifikationsalgorithmus gegossene Entscheidung über das Leben des betroffenen Menschen zu fällen?

Es liegt nahe, die Ergebnisse der algorithmischen Klassifikation an der Wirklichkeit zu erproben: Wie viele Fehleinschätzungen treffen solche Klassifikationen, die durch Klassifikationsalgorithmen vorgenommen werden? Wie viele *false negatives* gibt es – ein Straftäter wird gegen die Erwartung rückfällig –, und wie viele *false positives* – eine Straftäterin verbleibt in Sicherungsverwahrung, obwohl sie niemals wieder eine Straftat verüben würde?

Dies zu untersuchen, unternimmt die Arbeit von Tobias Krafft, die wir heute mit dem FIF-Studienpreis 2017 auszeichnen wollen.

Der Autor beginnt damit, solche (binären) Klassifikatoren statistisch zu überprüfen und Evaluationsstrategien vorzustellen. Nach einem historischen Abriss beschreibt er den aktuellen Stand der Beurteilung von *Algorithmic-Decision-Making-Prozessen* (ADM-Prozessen). Betrachtet werden zwei Qualitätsmaße und deren Diskrepanz wird aufgezeigt:

- *Area under the Receiver Operating Characteristic* (AUC),
- *Positive Predictive Value* (PPV<sub>k</sub>).

Nach einem Abriss der Entwicklung standardisierter Prognoseinstrumente zur Rückfälligkeitanalyse und einer mathematischen Analyse der Abweichung zwischen AUC und PPV<sub>k</sub> wird diese Diskrepanz anhand eines realen Datensatzes aus den USA nachgewiesen, für das Werkzeug COMPAS, einem kriminalprognostischen Instrument der vierten Generation.

Der Autor beurteilt die in der Praxis häufige Nutzung des AUC anhand seiner Untersuchungen als kritisch: die Abweichung vom PPV<sub>k</sub>, der näher am richterlichen Entscheidungsprozess evaluiert, bleibe deutlich hinter den Erwartungen zurück.

*„Sollte im deutschen Justizwesen die Einführung ADM-gesteuerter Prozesse zur Debatte stehen, wäre Deutschland in der glücklichen Lage, von den Erfahrungen und Fehlern anderer Länder, wie den USA zu profitieren. Daher ist zu hoffen, dass dies ohne überstürzten politischen Aktionismus, sondern mit Bedacht nach einer ausführlichen, gesellschaftlichen Debatte erfolgt“,*

so das Fazit des Autors. Es bedürfe dieser Debatte, die sich aber ohne präzisere Einsicht in die algorithmische Entscheidungsfindung nicht führen lasse. Dem ist unbedingt zuzustimmen – ob der Appell angesichts der sich verstärkenden Kriminalitäts- und Terrorangst, ob sie nun berechtigt ist oder nicht, dem Vormarsch von Law-and-Order in den öffentlichen und parlamentarischen Debatten und einer manchmal erstaunlichen Technikgläubigkeit gehört wird, bleibt abzuwarten.

Aus unserer Sicht bestätigen die Ergebnisse die Einschätzung, dass eine vollständig automatisierte Entscheidungsfindung im kriminologischen Bereich kaum verantwortet werden kann und davon Abstand genommen werden sollte. Auf jeden Fall ist dem Preisträger, dem Institut, an dem er die Arbeit verfasst hat, und der damit verbundenen Initiative AlgorithmWatch in unser aller Interesse Erfolg bei ihrer weiteren, kritischen Arbeit zu wünschen.

Als kleine Nebenbemerkung müssen wir jedoch, bei allem großen Lob, die Verwendung des Algorithmusbegriffs kritisieren: auch wenn er auf statistische Programme angewandt wird, bleibt für algorithmische (warum nicht einfach nur Software-) Entscheidungssysteme die Bezeichnung Algorithmus fragwürdig, denn es handelt sich um komplex zusammengesetzte Software-Systeme, nicht um einen einzelnen definierten Algorithmus, der universell operieren würde. Wir halten dies aus theoretischen wie Verantwortungsgründen für problematisch.

Das tut der Arbeit aber keinen Abbruch: Mit ihrer Untersuchung der Konsequenzen der computergestützten Entscheidungsfindung behandelt die Arbeit ohne jeden Zweifel ein aktuelles und gesellschaftlich bedeutsames Thema. Sie ist interdisziplinär aufgebaut und kommt zu wichtigen Ergebnissen für die Debatte um unsere Rechte und unsere Sicherheit. Aus diesem Grund hat sich die Jury des FIF-Studienpreises einhellig für die Auszeichnung der Arbeit entschieden.

**Herzlichen Glückwunsch, Tobias Krafft, zum FIF-Studienpreis 2017.**