

Predictive Policing - Kriminalanalysen im Zeitalter maschinellen Lernens

Masterarbeit

zur Erlangung des akademischen Grades „Master of Science (M.Sc.)“ im Studiengang
Wirtschaftswissenschaften der Wirtschaftswissenschaftlichen Fakultät der Leibniz
Universität Hannover

vorgelegt von:

Name: Tappert
[REDACTED]

Vorname: Chris Robin
[REDACTED] [REDACTED]

Prüfer: Prof. Dr. Michael H. Breitner

Hannover, den 28.09.2018

Inhaltsverzeichnis

| | |
|---|-----------|
| Abbildungsverzeichnis | IV |
| Tabellenverzeichnis | V |
| 1 Einleitung | 1 |
| 2 Grundlagen und Notwendigkeit | 4 |
| 2.1 Zahlen aus Kriminalstatistiken | 4 |
| 2.2 Mögliche Problemfelder innerhalb der Kriminologie | 9 |
| 2.3 Google Frameworks | 19 |
| 2.3.1 Tensorflow | 19 |
| 2.3.2 Tensorboard | 20 |
| 2.3.3 Integration in Python | 23 |
| 2.4 Forschungslücke | 25 |
| 3 Maschinelle Lernmethoden zur Bildung von Prognosen | 26 |
| 3.1 Von Altbewährtem zu Neuem | 26 |
| 3.2 Vorhersage potentieller Täter | 26 |
| 3.2.1 Random Forest | 30 |
| 3.2.2 Neuronales Netzwerk | 32 |
| 3.2.3 Support Vector Machines | 36 |
| 3.3 Ort und Zeit des Verbrechens | 40 |
| 3.3.1 Temporale und geospatiale Analysen | 40 |
| 3.3.2 Data-Mining | 46 |
| 3.4 Vergleichbarkeit der Methoden | 47 |
| 4 Empirische Auswertungen | 50 |
| 4.1 Datensets | 50 |
| 4.2 Auswertungen | 51 |
| 4.2.1 Vorhersage potentieller Täter | 51 |
| 4.2.2 Vorhersage potentieller Tatzeiten und Tatorten | 64 |
| 4.3 Verzerrungen | 69 |
| 5 Diskussion, Limitationen und Empfehlungen | 73 |
| 5.1 Evaluation und Relevanz | 73 |
| 5.2 Umsetzung in die Praxis | 75 |
| 5.3 Politische und rechtliche Aspekte | 81 |
| 5.3.1 Datenbeschaffung | 81 |
| 5.4 Limitationen | 84 |

| | |
|---|------------|
| 5.5 Ausblick | 85 |
| 6 Fazit | 90 |
| 7 Anhang | 91 |
| 8 Ehrenwörtliche Erklärung der Abschlussarbeit | 165 |

1 Einleitung

In unserer heutigen, digitalen Welt stehen wir alle häufig im Fokus von Vorhersagen, ohne es wirklich zu bemerken. Sei es bei unserer nächsten Kaufentscheidung oder bei der Auswahl unserer Freizeitdestination: vieles wird heute mittels Maschinen versucht so gut wie möglich vorherzusagen, um platziert Werbungen zu schalten, die uns zum Kaufen bzw. Erleben von Dingen anregen sollen, die wir sonst unter Umständen gar nicht gekauft oder ausprobiert hätten. Vorhersagen werden allerdings nicht nur zu Werbezwecken entwickelt. Auch die Bewältigung gesellschaftlicher Probleme durch frühzeitige Erkennung ist immer häufiger das Hauptanliegen bei der Entwicklung von Vorhersagen. Idealerweise sollte der Erkennungszeitpunkt schon dann erfolgen, wenn das Problem noch gar nicht vorhanden ist, um die Schäden, die dieses mit sich bringen würde, gänzlich zu vermeiden. Man spricht dann von der sogenannten prädikativen Analyse ('Predictive Analysis'). Im Zuge der prädikativen Analyse stellen sich Forscher immer wieder die Frage, ob auch menschliches Verhalten überhaupt vorhersagbar ist oder ob dort die Grenzen des Vorhersagbaren überschritten werden.⁴ Menschliches Verhalten könnte im Gegensatz zu maschinellem Verhalten zu zufällig und irrational sein.⁵ Auch Maschinen werden hinsichtlich ihrer Arbeitsweise analysiert, um herauszubekommen, wann die nächste Wartung am kosteneffizientesten vollzogen werden sollte.⁶ So besteht eine theoretische Rechtfertigung für diese Arbeit, da Menschen genauso wie Maschinen in Gewohnheiten und immer gleichen Zyklen agieren könnten. Gerade instinktives Verhalten hinsichtlich krimineller Vergehen könnten so berechenbar werden.

Im Bereich der Klassifikation von Straftätern in 'Gefährlich' oder 'Ungefährlich' wird schon seit geraumer Zeit auf (logistische) Regressionen gesetzt⁷, welche daher die Basismethode dieser Arbeit darstellen wird. Ausgehend von dieser Basismethode wird versucht, weg von heuristischen Ansätzen hin zu modernen Algorithmen zu gelangen, die aufgrund der Datenmenge schon einer sicheren Vorhersage sehr nahe kommen könnten. Verschiedenste Datensätze und deren Variablen sollten dabei idealerweise miteinander in Verbindung gebracht werden, um die Datenmenge dadurch so weit zu erhöhen, dass man dieser sicheren Vorhersage so nahe wie möglich kommt. Dabei müssen vor allem rechtliche und ethische Aspekte berücksichtigt werden, da es sich bei den Beobachtungsgegenständen eben nicht um einfache Gegenstände, sondern um Menschen handelt, deren Würde nach Artikel 1 des Grundgesetzes unantastbar ist. Der Fall um die Analysefirma Cambridge Analytica in jüngster Vergangenheit sollte aufgezeigt haben, welche Steine ins Rollen gebracht werden können, wenn bei Analysen personenbezoge-

⁴Song et al., 2010

⁵Perry, 2013, S. 3

⁶Huhtanen et Jung, 2018

⁷Thomas, 1929

ner Daten Schlupflöcher in Gesetzen ausgenutzt werden. Auch wird man sich von der Utopie befreien müssen, dass man jemals die perfekte Vorhersage treffen wird, da es so etwas wie perfekt in der Mathematik zu selten gibt und die Modelle allesamt lediglich auf relativen Wahrscheinlichkeiten beruhen und nicht auf absoluten Sicherheiten. In Deutschland gibt es jedoch umfangreiche Datenschutzgesetze, die vor rechtlichen Malheurs bewahren sollten, solange man sich gewissenhaft an diese hält. In den Vereinigten Staaten von Amerika ist die Lage diesbezüglich eine andere, wo doch dort mit Daten viel freizügiger umgegangen wird als hierzulande. Jedoch können in den USA bereits große Erfolge im Bereich 'Predictive Policing' verzeichnet werden, wie diverse Beispiele zeigen, die Perry⁸ ausführlichst schildert.

Auch in Deutschland existieren viele Pilot-Projekte, wie bspw. mit der Software 'Precobs' in Baden-Württemberg, Bayern sowie Niedersachsen oder mit dem SPSS Modeller der Firma IBM im bevölkerungsreichen Rhein/Ruhrgebiet. Entstanden ist das Projekt 'Predictive Policing' 2010 durch die UCLA (University of California, Los Angeles), indem ein Bezug der Kriminologie zur Seismologie untersucht wurde.⁹ So spielt dieses Projekt auch in der aktuellen Literatur eine Rolle. Der in Deutschland sehr bekannte Autor Sebastian Fitzek hat 2015 ein Buch über Predictive Policing mit dem Titel 'Das Joshua-Profil'¹⁰ verfasst, in welchem aufgezeigt wird, wie nützlich diese Technologie auf der einen Seite, aber auch wie gefährlich und anfällig sie auf anderen Seite für Verzerrungen sein kann. Gerade diese negativen Konsequenzen von einer zu mächtigen künstlichen Intelligenz werden in Orwells Meisterwerk '1984'¹¹ deutlich, in dem Menschen kaum individuelle Rechte mehr besitzen und nur noch als bewertbare Objekte angesehen werden, die bei Nichterbringung von Leistung bestraft bzw. aus der Gesellschaft ausgestoßen werden.

So muss darauf geachtet werden, die negativen Aspekte vermeintlich großer Errungenschaften stets mit den positiven abzuwägen, um sich in der Gesellschaft nicht zu abhängig von maschineller Dominanz zu machen. Durch wissenschaftliche Errungenschaften ist es so in der Geschichte bereits häufig gelungen, die Ergreifung eines Täters herbeizuführen. Was wäre, wenn Sir Alec Jeffreys niemals den genetischen Fingerabdruck 1984 entdeckt hätte, durch welchen jedes Individuum zweifelsfrei identifizierbar ist und dadurch bereits unzählige Morde aufgeklärt werden konnten. Solche Ereignisse sind es, die die Forschung in jeglichen Bereichen der Kriminologie zu einer derart wichtigen Angelegenheit hervorhebt. Durch die Digitalisierung könnte nun ein neues

⁸Perry, 2013, S. 77

⁹Friend, 2013

¹⁰ISBN-13: 978-3785725450

¹¹ISBN-13: 978-3548234106

Zeitalter angebrochen sein, das uns diverse Möglichkeiten bietet, weitere großartige Errungenschaften zu erreichen.

6 Fazit

In der vorliegenden Arbeit ging es um die Fragestellung, inwieweit der Mensch durch den Einsatz künstlicher Intelligenzen in seiner Tätigkeit abgelöst werden kann und wie viel Entscheidungsgewalt einer künstlichen Intelligenz bei moralischen Fragen überlassen werden darf. Gerade im Bereich des Militärs, wo eine zu starke künstliche Intelligenz unmittelbar über Menschenleben entscheidet, ist diese Frage äußerst gewissenhaft zu diskutieren, da der Mensch sonst im schlimmsten Fall seine Kontrolle verliert. Ein Kontrollmechanismus könnte dabei durch eine Blockchain eingeführt werden, durch die sich Manipulationen im Evaluationsprozess weitestgehend verhindern lassen könnten. Dr. Mark Benecke, ein deutscher Kriminalbiologe, hat in einem mündlichen Vortrag über Spurenanalysen am Tatort am 04.09.2018 in Berlin mit dem Satz 'auch wenn noch so viel Logik und die Statistik eindeutig dafür sprechen, bedarf es immer noch Spuren für einen eindeutigen Beweis' eine grundlegende Aussage getroffen.

Die hier vorgestellten Algorithmen konnten mit ihrer Performance überzeugen. Es bleibt allerdings stets ein Restzweifel, den weder Mensch noch Maschine komplett tilgen können. Es ist stets darauf zu achten, Algorithmen nicht allzu viel Macht einzuräumen und sie gänzlich alleine entscheiden zu lassen, sondern diese stets nur als beratende, alarmierende oder hinweisende Instanz betrachten. Man sollte aufpassen, dass das Schlagwort der künstlichen Intelligenz nicht zu sehr 'gehyped' wird, da sonst sehr mächtige Werkzeuge irgendwann nur noch belächelt werden könnten. So würde der Diskussion um die maschinellen Lernmethoden in Verbindung mit der künstlichen Intelligenz mehr Sachlichkeit gut tun.

Auch die Gefahr des massenhaften Jobverlusts ist durchaus gegeben, wenn eine künstliche Intelligenz repetitive Aufgaben von Menschen übernehmen kann. In diesem Zug werden neue Gesellschaftsformen diskutiert werden müssen, wie die Einführung eines bedingungslosen Grundeinkommens, um die Menschheit vor Massenarmut zu schützen und den Wohlstand aufrecht zu erhalten. Wie bereits im alten Rom, in dem Menschen nicht arbeiten mussten, da Sklaven ihre Arbeit übernahmen, könnten in einem Gedankenexperiment die Sklaven der heutigen Zeit Roboter sein, sodass der Mensch mehr freie Zeit gewinnt, ohne dass dabei eine Minderheit entwürdigende Arbeit übernehmen muss. Der Mensch hätte nicht nur neu gewonnene Freizeit, in der er das Leben mehr genießen könnte, sondern auch weniger Frust, den er aus der entwürdigenden Arbeit erhält und den er unter Umständen bei anderen ablässt. Die Zeichen stehen auf einen starken, gesellschaftlichen Wandel durch die Digitalisierung.