



Gluba, Alexander et al.

Lagedaten, Datenlage, Prognoseansätze. Perspektiven des „Predictive Policing“ in Niedersachsen. Teil 2

SIAK-Journal – Zeitschrift für Polizeiwissenschaft und polizeiliche Praxis (3/2020), 20-27.

doi: 10.7396/2020_3_B

Um auf diesen Artikel als Quelle zu verweisen, verwenden Sie bitte folgende Angaben:

Gluba, Alexander et al. (2020). Lagedaten, Datenlage, Prognoseansätze. Perspektiven des „Predictive Policing“ in Niedersachsen. Teil 2, SIAK-Journal – Zeitschrift für Polizeiwissenschaft und polizeiliche Praxis (3), 20-27, Online: http://dx.doi.org/10.7396/2020_3_B

©Bundesministerium für Inneres – Sicherheitsakademie / Verlag NWV, 2020

Hinweis: Die gedruckte Ausgabe des Artikels ist in der Print-Version des SIAK-Journals im Verlag NWV (<http://nwv.at>) erschienen.

Online publiziert: 01/2021

Lagedaten, Datenlage, Prognoseansätze

Perspektiven des „Predictive Policing“ in Niedersachsen Teil 2



ALEXANDER GLUBA,
*Landeskriminalamt Niedersachsen,
Kriminologische Forschung und
Statistik.*

Seit 2016 testet die Polizei Niedersachsen im Rahmen des Projekts PreMAP (Predictive Mobile Analytics for Police) Grenzen und Möglichkeiten der Prognose zukünftiger Wohnungseinbrüche. Während im Projekt zunächst die inhaltlichen, technischen und organisatorischen Voraussetzungen geschaffen wurden, konnte dann der entwickelte Ansatz pilotiert und evaluiert werden. Die Ergebnisse waren auf Seiten des Vorhersagealgorithmus und der Reaktionsebene teilweise nicht befriedigend und optimierungsbedürftig, dennoch sprachen sich Nutzerinnen und Nutzer dafür aus, die Potentiale des Predictive Policing weiter zu erforschen (siehe Gluba et al. 2020). Der vorliegende Beitrag zeigt in einer Art Werkstattbericht, welche grundsätzlichen Überlegungen in dieser Sache angestellt, welche Methoden genutzt und welche ersten Ergebnisse erzielt wurden und werden.



ALEXANDER PETT,
*Landeskriminalamt Niedersachsen,
Kriminologische Forschung und
Statistik.*

1. KONSEQUENZEN

Die Erfahrungen mit PreMAP (Predictive Mobile Analytics for Police) waren, dies zeigten die gewonnenen Erkenntnisse, heterogen. Zwar konnten einige Erfolge im Projekt erzielt werden – hier ist vor allem die mobile Bereitstellung von retrograden Lageinformationen zu nennen –, es trat aber auch einiger Optimierungsbedarf zu Tage: Der Prognoseansatz reagierte zu sensibel und wies zu viele Risikogebiete aus, die polizeilichen Reaktionen waren insbesondere vom Umfang her nicht geeignet, tatbereite Personen von ihrem Tun abzuhalten, und auch die interne Kommunikation bezogen auf PreMAP war an einigen Stellen suboptimal. Trotzdem wünschten die Nutzerinnen und Nutzer, dass PreMAP weiterhin genutzt und weiterentwickelt werden sollte.

Dies führte zu der Entscheidung, eine so genannte erweiterte Pilotierung durch-

zuführen, in deren Rahmen PreMAP in seiner bisherigen Form in insgesamt 15 von 33 Polizeiinspektionen in Niedersachsen seit November 2018 zum Einsatz kommt (Niedersächsisches Ministerium für Inneres und Sport 2018). Hierbei werden seither wichtige Rückmeldungen aus der Praxis gewonnen, was den dortigen Bedarf angeht. Der Bedarf wird durch eine eingesetzte Projektgruppe unter Federführung der Polizeidirektion Osnabrück zunächst geprüft und im Anschluss ggf. vorangetrieben. Der Bereich Kriminologische Forschung und Statistik des Landeskriminalamtes Niedersachsen hat in dieser Struktur die Aufgabe, den Prognoseansatz dergestalt zu überarbeiten, dass die Vorhersagen treffgenauer sind. Auch sollen Prognoseansätze für unterschiedliche Delikte, ergänzend zum Wohnungseinbruch, entwickelt werden, beispielsweise weitere Einbruchsdelikte in Keller/Dachböden oder

Geschäfte, Kfz-Aufbrüche oder Kfz-Totalentwendungen.

Um zu prüfen, ob die Prognosegüte durch eine breitere Datenbasis verbessert werden kann, wurde durch Literaturrecherche und in einem Workshop mit Kriminologinnen und Kriminologen sowie Kriminalistinnen und Kriminalisten erarbeitet, welche polizeixternen sozioökonomischen und -demografischen Merkmale theoretisch Einfluss auf das Auftreten bestimmter Delikte haben können. In einem nächsten Schritt wurde recherchiert, woher entsprechende Daten bezogen werden können. Schnell zeigte sich, dass offizielle Stellen, wie Statistische Landesämter, nicht über Daten in der notwendigen Kleinräumigkeit verfügen und Kommunen nicht flächendeckend alle benötigten Informationen bereitstellen können. Daher musste auf kommerzielle Anbieter solcher kleinräumiger Daten zurückgegriffen werden. Nach einer intensiven Marktsichtung wurde entschieden, Daten der Firma Nexiga zu erwerben, da diese die meisten Anforderungen auf Ebene von Gebäuden oder zumindest Wohnquartieren abbilden konnte. Die Daten wurden zunächst nur für Hannover und Braunschweig erworben, um anhand dieser beiden Städte beispielhaft die Möglichkeiten der Weiterentwicklung des Prognoseansatzes zu testen.

2. AKTUELLER STAND DER WEITERENTWICKLUNG

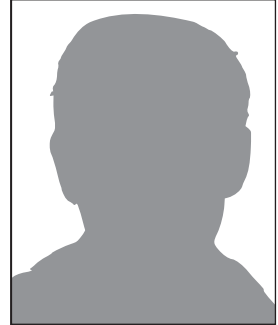
Alternative Ansätze zur tatbezogenen Prognose von Near Repeat-Delikten

Im methodischen Teil der zweiten Pilotierungsphase von PreMAP sollen unter anderem alternative Prognoseansätze erprobt werden – mit dem Ziel, die Prognosequalität des derzeit eingesetzten Ansatzes zu übertreffen. Bei dem in der ersten PreMAP-Pilotierungsphase entwickelten Ansatz mit merkmalspezifischen

Score-Beiträgen, der auf dem Near Repeat-Phänomen aufbaut und sich jetzt in der erweiterten Pilotierung im Einsatz befindet, wird für jedes einzelne erfasste Delikt das Ja-Nein-Merkmal „Delikt wird ein Near Repeat-Delikt haben: Ja/Nein“ prognostiziert. Als alternativer Ansatz für diese Prognose kommt grundsätzlich jeder Ansatz in Frage, der eine datengestützte Ja-Nein-Prognose bzw. Ja-Nein-Klassifizierung ermöglicht.

Ein in den empirischen Sozialwissenschaften verbreiteter Ansatz für diese Aufgabenstellung, die Logistische Regression, liefert zusätzlich zur reinen Ja-Nein-Prognose eine Wahrscheinlichkeitsprognose für das Auftreten des Wertes „Ja“. Neben der einfachen logistischen Regression und einer für die vorliegenden Daten mit zahlreichen Ja-Nein-Prädiktoren¹ besonders geeigneten Variante dieses Ansatzes – „Bayesian Logistic Regression“ (Genkin et al. 2007) – wurden der Decision Tree-Ansatz in der einfachen und in zwei durch Zufallsziehungen angereicherten Varianten – „Tree Ensemble“ (Breiman et al. 1984) und „Random Forest“ (Breiman 2001) – als mögliche Prognoseansätze darauf geprüft, ob sie eine gute Prognosequalität aufweisen (viele „True Positives“ und wenige „False Positives“²).

Zur vergleichenden Erprobung der Prognoseansätze wurden polizeiliche Einbruchdeliktdaten für die Jahre 2015 bis 2017 verwendet, und zwar aggregiert zu den gleichen Modus Operandi-, Tatmittel- und Stehlgut-Merkmalen, wie beim bereits im Einsatz befindlichen Prognoseansatz. Um die Prognosequalität mit und ohne Einbeziehung der kleinräumigen sozioökonomischen Daten des Datenanbieters Nexiga vergleichen zu können, die für die Stadtgebiete von Hannover und Braunschweig erworben wurden, wurden aus den gut 41.000 Delikten für ganz Niedersachsen diejenigen knapp 5.000 ausgewählt, die



MARKUS PULLEN,
*Landeskriminalamt Niedersachsen,
Kriminologische Forschung und
Statistik.*



MAXIMILIAN QUERBACH,
*Landeskriminalamt Niedersachsen,
Kriminologische Forschung und
Statistik.*

in einer der beiden Städte erfasst worden waren. Die Prognoseansätze wurden mit der Analysesoftware KNIME (Berthold et al. 2007) in der Version 4.0.2 implementiert. Als Prädiktoren wurden in alle geprüften Prognoseansätze in einer ersten Variante zunächst die in Verwendung befindlichen gut 90 Merkmale aus den polizeilichen Daten einbezogen, in einer zweiten Variante dann plausibel erscheinende Merkmale aus den kleinräumigen sozioökonomischen Nexiga-Daten (Datenstand: 01.01.2018) auf Wohnquartierebene (z.B. Anteil der Einpersonenhaushalte, durchschnittliche Angebotskaltmiete) und Straßenzugebene (z.B. mittlere Entfernung zum nächsten Autobahnanschluss).

Bei der Bewertung von Prognoseansätzen ist es üblich, einen Teil der Daten zum Trainieren des Prognosemodells zu verwenden („Trainingsmenge“) und dann die Prognosequalität bei den restlichen Daten zu untersuchen („Testmenge“). Um bei der Prüfung eines Ansatzes nicht nur ein einzelnes trainiertes Modell zu untersuchen, dessen Prognosequalität auf Grund der Auswahl der Trainingsmenge zufällig besonders hoch oder niedrig sein könnte, wurden für jeden Ansatz 100 verschiedene Trainingsmengen von je 90 % der Daten zufällig gezogen. Für jede davon wurde dann das Prognosemodell trainiert und auf die verbliebene 10 %-Testmenge angewendet. Delikteinzelfälle mit fehlenden Werten bei einzelnen Merkmalen wurden beim Trainieren des Modells weggelassen, in der Testmenge wurden fehlende Werte durch den Durchschnittswert des entsprechenden Merkmals in der gesamten Testmenge ersetzt („imputiert“). Zur Bewertung der Prognosequalität wurden anschließend für alle 100 Durchläufe die True Positive-Rate (TPR) und die False Positive-Rate (FPR) in der Trainingsmenge berechnet.³

Prognosequalität der alternativen Ansätze

Als ein vielversprechender Prognoseansatz hat sich eine Bayesianische Version der Logistischen Regression (im Folgenden kurz: BLR-Ansatz) nach Genkin u.a. (Genkin et al. 2007) erwiesen. Auf der linken Seite der Abbildung 1 (Seite 23) ist ein FPR-TPR-Streudiagramm für die Variante dargestellt, bei der nur die polizeilichen Deliktdaten in das Prognosemodell einbezogen wurden. Für jedes einzelne der 100 trainierten Prognosemodelle ist es wünschenswert, dass sich der Punkt, der die zugehörige FPR-TPR-Kombination darstellt, möglichst weit links oben befindet (hohe TPR, geringe FPR), zumindest jedoch oberhalb der eingezeichneten Diagonalen (TPR > FPR). Besonders ungünstig ist es, wenn FPR und TPR gleich hoch sind oder die FPR sogar höher ist (FPR ≥ TPR). Es finden sich 43 solcher Fälle, die genau auf der bzw. unterhalb der Diagonalen liegen. Nur in 57 der 100 Durchläufe wurde also für die Teilmenge der Delikte mit einem Near Repeat-Delikt ein höherer Anteil von Near Repeat-Delikten prognostiziert als für die Teilmenge ohne tatsächliches Near Repeat-Delikt. Die Delikte ohne tatsächliches Near Repeat-Delikt machen ohnehin den weit überwiegenden Anteil der Einbruchsdelikte aus (im untersuchten Tatbestand 88,8 % für ganz Niedersachsen, 83,0 % für die beiden untersuchten Stadtgebiete zusammen). Damit bedeutet eine hohe FPR als Anteil dieses großen Anteils eine besonders hohe absolute Zahl von False Positives, also von „falschen Alarmen“.

Im rechten Teil der Abbildung 1 (Seite 23) sind die TPR- und die FPR-Werte als Boxplots dargestellt. Die dort angegebenen Werte für den TPR- und den FPR-Median zeigen, dass die TPR zumindest „im Mittel“ (d.h. in der Median-Betrachtung) um 0,5 Prozentpunkte höher liegt als die FPR.

Quelle: Gluba/Pett/Pullen/Querbach

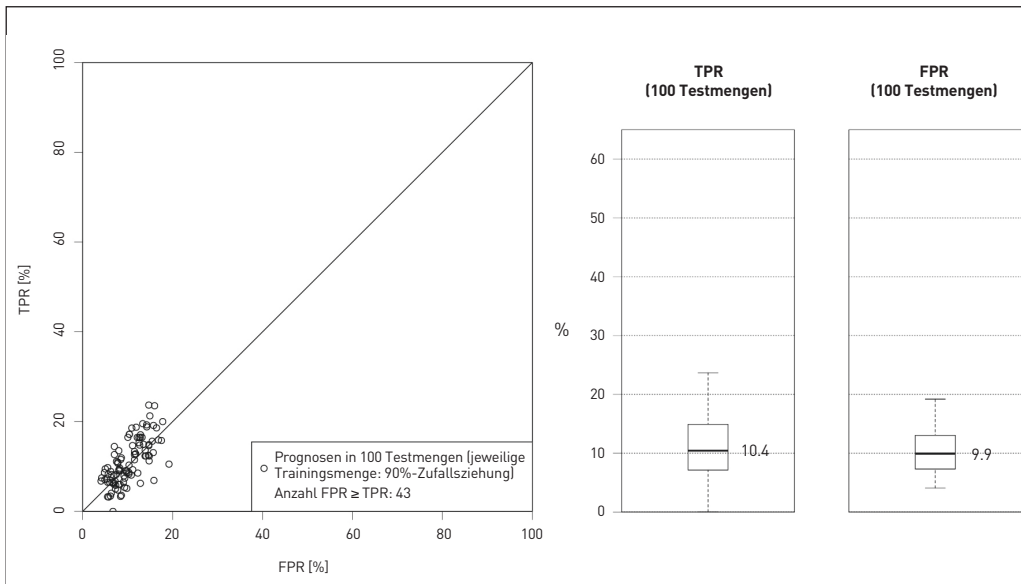


Abb. 1: Prognose durch Bayesianische Logistische Regression mit polizeilichen Deliktdaten⁴

Der linke Boxplot weist allerdings noch deutlicher als das Streudiagramm darauf hin, dass in mindestens einem der 100 Durchläufe eine TPR von Null aufgetreten ist. Wünschenswert erscheint es auf jeden Fall, einen Prognoseansatz zu finden, bei dem die Punktwolke im Streudiagramm weiter oben und damit auch der ganze TPR-Boxplot weiter oben liegt. Dabei sollte sich im günstigsten Fall die Punktwolke nicht zugleich weiter nach rechts und damit der FPR-Boxplot auch nicht weiter nach oben verschieben. Im Hinblick auf möglichst einheitliche Qualität der unterschiedlichen Prognosemodelle, die sich aus zufällig variierenden Trainingsdaten ergeben können, ist zudem eine kompakte Punktwolke wünschenswert, bei der die Punkte möglichst nahe beieinanderliegen. Eine solche Punktwolke würde in beiden Boxplots zu Boxen von geringer Höhe sowie zu Maximal- und Minimalwerten nahe bei den Boxrändern führen.

Ein Teil dieser Qualitätsziele kann durch die Einbeziehung der Merkmale aus den Nexiga-Daten in der zweiten BLR-Variante erreicht werden, wie Abbildung 2 (Seite 24) zeigt. Die Punktwolke liegt tatsächlich

weiter oben als in Abbildung 1, aber auch weiter rechts. Immerhin ist der TPR-Median auf 46,8 % angestiegen, der FPR-Median nur auf 38,7 %. Der Abstand der beiden Mediane hat sich also auf 8,1 Prozentpunkte erhöht, die TPR-Box im Boxplot liegt vollständig oberhalb der FPR-Box, und im linken Streudiagramm liegen nur noch neun der 100 FPR-TPR-Kombinationen unterhalb (oder genau auf) der Diagonalen (FPR \geq TPR). Festzuhalten ist, dass die Einbeziehung der zusätzlichen Merkmale die Prognosequalität des BLR-Ansatzes im Hinblick auf die TPR („möglichst weit oben“) und den Abstand zwischen TPR und FPR („möglichst weit oder zumindest möglichst häufig oberhalb der Diagonalen“) sichtbar erhöht. Diese Verbesserungen sind allerdings damit verbunden, dass der FPR-Boxplot in seiner Gesamtheit nach oben verschoben ist, was im Widerspruch zum Qualitätsziel „möglichst weit links im Streudiagramm“ steht.

Zum Vergleich: In dem im Einsatz befindlichen und nur auf polizeilichen Delikt-daten beruhenden Prognoseansatz ergaben sich für den Zweimonatszeitraum

Quelle: Gluba/Pett/Pullen/Querbach

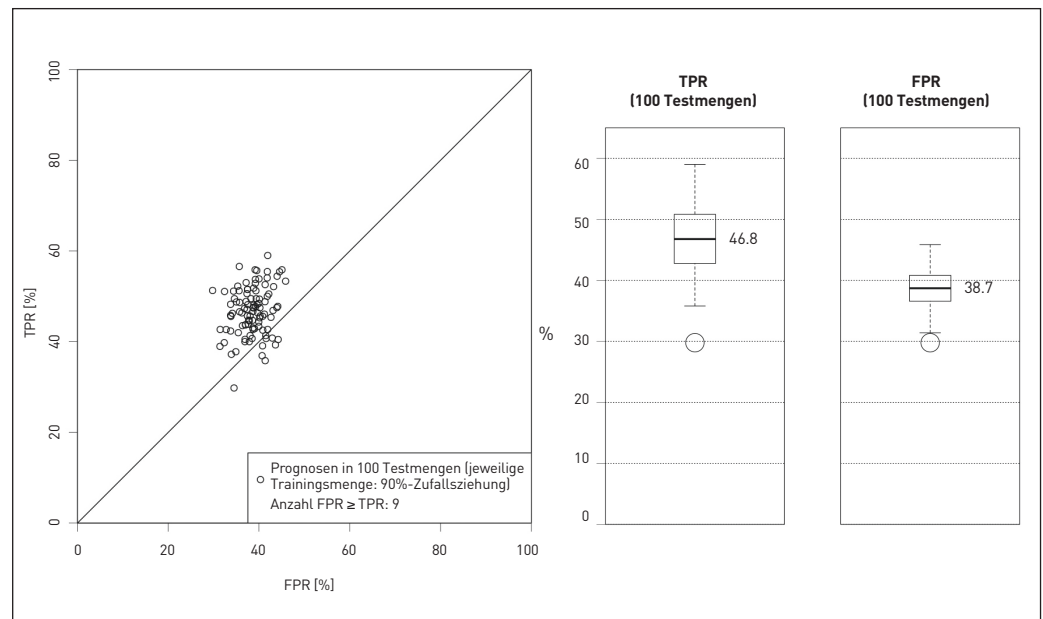


Abb. 2: Prognose durch Bayesianische Logistische Regression mit polizeilichen Delikt- und kleinräumigen sozioökonomischen Daten⁴

November/Dezember 2018 für knapp 1.300 Einbruchsdelikte eine TPR von 26,3 % und eine FPR von 18,3 % für die beteiligten 15 Polizeiinspektionen zusammen.⁵ Für den Monat Januar 2019 ergaben sich für gut 400 Einbruchsdelikte eine TPR von 21,8 % und eine FPR von 14,3 %. Für den Zweimonatszeitraum September/Oktober 2019 ergaben sich für knapp 700 Einbruchsdelikte eine TPR von 16,2 % und eine FPR von 10,5 %. Auch hier zeigen sich eine gewisse Schwankungsbreite sowie ein stets deutlicher Abstand zwischen TPR und FPR. Die TPR-Werte liegen sichtbar über dem TPR-Median der BLR-Variante ohne die Nexiga-Daten, aber deutlich unter dem TPR-Median der BLR-Variante einschließlich der Nexiga-Daten. Die FPR-Werte bewegen sich im Bereich des FPR-Boxplots für die BLR-Variante ohne Nexiga-Daten. Im Hinblick auf die TPR liefert die BLR-Variante, die zusätzlich die sozioökonomischen Daten einbezieht, also eine Verbesserung der Prognosequalität gegenüber dem aktuellen Ansatz, aber um den Preis von höheren FPR-Werten.

Ein weiterer vielversprechender Prognoseansatz ist der Tree Ensemble-Ansatz (im Folgenden: TE-Ansatz) nach Breiman u.a. (Breiman et al. 1984). Bei der Variante mit den kombinierten Polizei- und Nexiga-Daten liegt der TPR-Median der 100 Durchläufe mit 53,0 % sogar 6,2 Prozentpunkte über dem der entsprechenden BLR-Variante, dafür liegt aber auch der FPR-Median um 8,4 Prozentpunkte höher bei 47,1 %. Damit beträgt der Abstand zwischen dem TPR- und dem FPR-Median in dieser TE-Variante nur 5,9 Prozentpunkte. Die Punktwolke ist deutlich weniger kompakt als in der zugehörigen BLR-Variante, und für zwölf der 100 Durchläufe gilt die ungünstige Eigenschaft $FPR \geq TPR$.⁶

Die übrigen geprüften Prognoseansätze (einfache Logistische Regression, einfacher Decision Tree, Random Forest) lieferten deutlich ungünstigere Ergebnisse im Hinblick auf die FPR-TPR-Kombinationen in 100 probeweisen Durchläufen und können deshalb als Kandidaten für eine praktische Umsetzung verworfen werden.

Zusammenfassende Bewertung der alternativen Prognoseansätze und nächste Schritte

Der Bayesianische Logistische Regressionsansatz unter Einbeziehung der sozio-ökonomischen Daten weist also unter den geprüften Ansätzen für eine tatbezogene Prognose des Merkmals „Delikt wird ein Near Repeat-Delikt haben: Ja/Nein“ die günstigsten kombinierten TPR- und FPR-Ergebnisse auf. Im Hinblick auf die TPR-Werte ist der Tree Ensemble-Ansatz noch positiver zu bewerten, dabei müssen aber höhere FPR-Werte und eine größere Streuung der FPR-TPR-Kombinationen in Kauf genommen werden. Ob einer dieser Ansätze dem in Betrieb befindlichen Ansatz mit merkmalspezifischen Score-Beiträgen vorzuziehen ist, kann auf Grund der Ergebnisse nicht eindeutig entschieden werden. Beide Alternativansätze liefern zwar eine deutlich höhere True Positive-Rate – selbst die TPR-Minimalwerte bei den 100 geprüften Trainingsmengen liegen in beiden Ansätzen deutlich über dem höchsten der hier exemplarisch genannten TPR-Werte aus der Praxis –, allerdings liegt aber bei beiden Ansätzen auch eine durchwegs höhere False Positive-Rate vor als beim derzeitigen Ansatz.

Ein klarer Vorteil des BLR-Ansatzes gegenüber dem TE-Ansatz besteht in seiner Skalierbarkeit mit Hilfe der in der Prognose enthaltenen Wahrscheinlichkeit für den Prognosewert „Ja“. Wenn den Praxisanwendern die FPR zu hoch erscheint und sie bereit sind, eine geringere TPR in Kauf zu nehmen, kann man beim BLR-Ansatz die Wahrscheinlichkeitsschwelle für eine „Ja“-Prognose vom Standardwert 0,5 auf einen höheren Wert anheben. Dann werden insgesamt tendenziell weniger „Positives“ prognostiziert. Im günstigen Fall sinkt dadurch die Anzahl der False Positives, im besonders günstigen Fall sogar stärker als die Zahl der True Positives.

Die Möglichkeit, den Schwellenwert für eine „Ja“-Prognose anzuheben und damit die FPR und die TPR zu beeinflussen, besteht auch beim derzeitigen Ansatz.

Im weiteren Verlauf der zweiten Pilotierungsphase sollen neben weiteren Verbesserungsideen zu den hier beleuchteten tatbezogenen Ja-Nein-Prognoseansätzen noch raumbezogene Prognoseansätze und die Übertragbarkeit beider Ansatztypen auf weitere Deliktbereiche untersucht werden.

3. FAZIT

Die niedersächsische Predictive Policing-Lösung PreMAP ist einer kritischen Bewertung unterzogen worden. Durch unterschiedliche quantitative und qualitative Methoden konnte im Rahmen der Evaluation gezeigt werden, dass das System noch optimierungsbedürftig ist. Dies gilt für den Prognoseansatz an sich, der zu sensibel reagierte und zu viele Risikogebiete identifizierte, sowie Ausmaß und Art der getroffenen polizeilichen Maßnahmen als Reaktion. Insbesondere die Kommunikation über grundsätzliche Aspekte von PreMAP oder Prognosen für den Streifen- oder Ermittlungsdienst zeigten sich als verbesserungswürdig.

Die im Zuge von PreMAP erkannten Befunde stehen dabei weitgehend im Einklang mit denen anderer Evaluationen. So finden sich für die Ansätze P4 in Baden-Württemberg oder SKALA in Nordrhein-Westfalen durchaus ähnliche Probleme und Handlungsfelder (Gerstner 2018; Landeskriminalamt Nordrhein-Westfalen 2018). Und auch international können die wenigen existierenden Studien den Mehrwert des Predictive Policing nicht eindeutig nachweisen (Mohler et al. 2015; Hunt et al. 2014; Ratcliffe et al. 2020).

Trotz der benannten Probleme haben die Nutzerinnen und Nutzer von PreMAP sich auch in Ermangelung erfolgreicher,

bewährter oder etablierter Konzepte und Ansätze dafür ausgesprochen, dass weiter an PreMAP gearbeitet werden soll. In der Folge wurde die nächste Phase, die so genannte erweiterte Pilotierung, in insgesamt 15 Polizeiinspektionen begonnen.

In dieser Phase wird ergebnisoffen geprüft, ob und wie sich in den erkannten Bereichen Optimierungen realisieren lassen. Eine wichtige Fragestellung besteht darin, ob durch einen alternativen Prognoseansatz mit oder ohne Einbeziehung von zusätzlichen sozioökonomischen Daten eine Verbesserung der Prognosequalität erreicht werden kann. Umfangreiche Probeläufe mit fünf Alternativansätzen deuten darauf hin, dass eine Verbesserung möglich ist. Ob eine Erhöhung der Zahl korrekter Prognosen realisierbar ist, ohne dass die Zahl der „Fehlalarme“ unerwünscht stark steigt, müssen weitere Analysen zeigen. Schon jetzt konnte aber bereits eine Strategie vorgeschlagen werden, wie einer der geprüften Prognoseansätze im Hinblick auf eine erwünschte mittlere Fehlalarm-Quote (False Positive-Rate) kalibriert werden kann. Daraus ergibt sich dann die im Mittel erreichbare Quote von zutreffenden Near Repeat-Prognosen (True Positive-Rate).

Ferner soll zukünftig auch die Kommunikation im Zusammenhang mit PreMAP verändert werden, beispielsweise soll im Zuge des länderübergreifenden EU-Projektes „Cutting Crime Impact“ (CCI)⁷ eine neue inhaltliche Struktur der Briefings im Streifen dienst entwickelt und getestet werden.

Risikobewertungen sind ein Kernelement von PreMAP. Doch schon immer existierten daneben Lageinformationen, die es den Polizeikräften erleichtern, das Kriminalitätsgeschehen zu erschließen und zu analysieren. Derartige Lagekomponenten – zum Beispiel durchgeführte Kontrollen, Informationen zu Spuren oder Lagebildern – werden bei der Weiterentwicklung von PreMAP künftig verstärkt einbezogen werden.

Mit dem Ende der erweiterten Pilotierung Ende Juli 2020 ist ein weiterer wichtiger Schritt getan, in dem Bestreben, die vorhandenen polizeilichen Informationen retro- und prospektiv in geeigneter Weise für die Ermittlungs- und Einsatzkräfte aufzubereiten. Neue Kriminalitätserscheinungsformen, polizeilich-strategische Schwerpunkte sowie technische oder wissenschaftliche Weiterentwicklungen werden allerdings ein ständiges Reflektieren und gegebenenfalls Anpassen notwendig machen.

¹ Als Prädiktoren werden die in das Prognosemodell einbezogenen „erklärenden“ Merkmale bezeichnet. Die Prädiktoren aus den polizeilichen Daten sind meist Ja-Nein-Prädiktoren (z.B. „Stehlgut Schmuck: Ja/Nein“).

² Als „True Positives“ bezeichnet man diejenigen Fälle, in denen ein (später oder in Testdaten) tatsächlich beobachteter Wert „Ja“ korrekt als „Ja“ prognostiziert wurde, als „False Positives“ die Fälle, in denen ein tatsächlich beobachtetes „Nein“ unzutreffend als „Ja“ prognostiziert wurde.

³ Die TPR ist der Anteil der „Ja“-Prognosen an jenen Delikten, die tatsächlich ein Near Repeat-Delikt hatten (nicht: [...] waren). Die FPR ist der Anteil der „Ja“-Prognosen an den Delikten ohne tatsächliches Near Repeat-Delikt. Da für alle Delikte in der Testmenge bekannt ist, ob sie ein tatsächliches (polizeilich erfasstes) Near Repeat-Delikt hatten, kann man für diese Daten die benötigten Zahlen einfach ermitteln.

⁴ „True Positive-Rate“ (TPR) und „False Positive-Rate“ (FPR) für 100 Durchläufe auf unterschiedlichen 10 %-Testmengen (verbleibende 90 % der Daten als Trainingsmenge jeweils zufällig gezogen) – links: FPR-TPR-Streudiagramm, rechts: Boxplots.

⁵ In diesen TPR- und FPR-Angaben sind nicht exakt die durch den aktuell im Einsatz befindlichen Ansatz prognostizierten „Ja“-Prognosen für Near Repeat-Delikte als „Positives“ enthalten, sondern die aktivierten Risikogebiete. Mögliche Abweichungen ergeben sich aus folgenden Gründen: Menschliche Bewerter können abweichend vom Modellergebnis Risikogebiete aktivieren oder deaktivieren (Erhöhung oder Absenkung der „Positives“

gegenüber dem Modell). Für Delikte, deren tatsächlicher oder unterstellter Tatzeitpunkt bei der Erfassung bereits mehr als 72 Stunden zurückliegt oder bei denen der Tatzeitraum nicht auf höchstens vier Tage eingegrenzt werden kann, wird kein Risikogebiet aktiviert (Absenkung der „Positives“). Im Zähler der hier betrachteten TPR und FPR werden alle nach diesen Korrekturen aktivierten Risikogebiete berücksichtigt. Ob TPR und FPR dadurch höher oder niedriger liegen, als es bei den reinen „Modell-Positives“ der Fall wäre, lässt sich nicht eindeutig sagen.

⁶ In der Variante, die nur die polizeilichen Daten einbezieht, ergibt sich im Streudiagramm interessanterweise keine Punktwolke, die weiter links unten liegt wie beim BLR-Ansatz, sondern eine Punktwolke, die sich entlang der gesamten Diagonale erstreckt, wobei für 61 Durchläufe $FPR \geq TPR$ gilt. Dass der TPR-Median dabei 81,4 % beträgt, ist demzufolge kein Argument für die Verwendung dieser Variante, zumal der FPR-Median sogar bei 82,8 % liegt.

⁷ Eine Darstellung des Projekts findet sich unter <https://www.cuttingcrimeimpact.eu/>.

Quellenangaben

Berthold, Michael R. et al. (2007). KNIME – The Konstanz Information Miner, in: Preisach, Christine et al. (Eds.) *Data Analysis, Machine Learning and Applications: Proceedings of the 31st Annual Conference of the Gesellschaft für Klassifikation e. V.*, New York, 319–326.

Breiman, Leo (2001). *Random Forests, Machine Learning* (1), 5–32.

Breiman, Leo et al. (1984). *CART – Classification and Regression Trees*, Belmont.

Gerstner, Domonik (2018). *Predictive Policing in the Context of Residential*

Burglary: An Empirical Illustration on the Basis of a Pilot Project in Baden-Württemberg, Germany, *European Journal for Security Research* (2), 115–138.

Genkin, Alexander et al. (2007). *Large-Scale Bayesian Logistic Regression for Text Categorization*, *Technometrics* (3), 291–304.

Gluba, Alexander et al. (2020). *Lagedaten, Datenlage, Prognoseansätze. Perspektiven des „Predictive Policing“ in Niedersachsen, Teil 1*, *SIAK-Journal – Zeitschrift für Polizeiwissenschaft und polizeiliche Praxis* (2), 31–41.

Hunt, Priscilla et al. (2014). *Evaluation of the Shreveport predictive policing experiment*, Santa Monica.

Landeskriminalamt Nordrhein-Westfalen (2018). *Kooperative Evaluation des Projektes SKALA*, Düsseldorf.

Mohler, George O. et al. (2015). *Randomized controlled field trials of predictive policing*, *Journal of the American Statistical Association* (512), 1399–1411.

Niedersächsisches Ministerium für Inneres und Sport (2018). *Zweite Pilotphase startet: Einsatz der Prognosesoftware PreMAP wird zum 1. November weiter intensiviert*, Online: https://www.mi.niedersachsen.de/aktuelles/presse_informationen/zweite-pilotphase-startet-einsatz-der-prognosesoftware-premap-wird-zum-1-november-weiter-intensiviert-170622.html (21.11.2019).

Ratcliffe, Jerry H. et al. (2020). *The Philadelphia predictive policing experiment*, in: *Journal for Experimental Criminology*, Online: <https://link.springer.com/article/10.1007/s11292-019-09400-2> (13.01.2020).