

Predictive Policing

Künstliche Intelligenz prognostiziert Verbrechen eine Woche im Voraus

Eine KI kann anhand von zeitlichen Mustern und geografischen Standorten zukünftige Verbrechen eine Woche im Voraus mit einer Genauigkeit von 90 Prozent prognostizieren.

Chicago (U.S.A.). In vielen Ländern wird seit Langem an einer Kriminalitätsvorhersage, gearbeitet, um kriminelle Aktivitäten zu prognostizieren und Polizeiresourcen entsprechend zu verteilen. Es gibt für das sogenannte Predictive Policing unterschiedliche Ansätze, die entweder auf geographische oder individuelle Faktoren ausgerichtet sind und auf verschiedenen statistischen Methoden und Sozialforschungstechniken basieren. In der Forschung war die Kriminalitätsvorhersage bisher jedoch umstritten, da sie systematische Vorurteile in der Polizeiarbeit und deren komplexe Verbindung mit Kriminalität und Gesellschaft nicht berücksichtigte.

Wissenschaftler der University of Chicago um Ishanu Chattopadhyay haben nun eine neue Künstliche Intelligenz (KI) entwickelt, die Kriminalität vorhersagt, indem sie zeitliche Muster und geografische Standorte aus öffentlichen Daten über Gewalt- und Eigentumsdelikte analysiert. Laut der Publikation im Fachmagazin Nature Human Behaviour kann die KI zukünftige Verbrechen eine Woche im Voraus mit einer Genauigkeit von 90 Prozent prognostizieren.

Quelle: Online-Artikel der Website Forschung und Wissen – Fachmedien und Mittelstand digital

Autor: Klatt, R.; Erscheinungsdatum: 28.06.2023; <https://www.forschung-und-wissen.de/nachrichten/technik/kuenstliche-intelligenz-prognostiziert-verbrechen-eine-woche-im-voraus-13377722>

1. Lesen Sie den oben abgedruckten Teil einer Online-Publikation vom 28.06.2023 aufmerksam durch.
2. Erklären Sie den Begriff Predictive Policing und beschreiben Sie die im Text genannten Neuerungen in eigenen Worten.
3. Stellen Sie kurz zwei ethisch relevante Einwände dar, die gegen einen Einsatz von Predictive Policing in der Polizeiarbeit sprechen.

Aufgabenteil: Daten veranschaulichen

4. Laden Sie den Datensatz „philadelphiacrimebearbeitet.csv“ in Orange mit dem Widget „File“ ein und verschaffen sich mithilfe des Widgets „Data Table“ einen Überblick über den Datensatz.

Ermitteln Sie die Anzahl der Instanzen, die Zahl der Merkmale (features) und die festgelegte Zielvariable.

Formulieren Sie eine interessante Fragestellung, die mit dem Datensatz eventuell beantwortet werden könnte.

5. Wählen Sie das Widget „Distributions“, um zu ermitteln, zu welcher **Uhrzeit** die **wenigsten Gesamtdelikte** und zu welcher **Uhrzeit die meisten Tötungsdelikte (homicide)** begangen werden.

Ermitteln Sie zudem, über welchen Zeitraum die Daten im Datensatz erhoben wurden. Nutzen Sie hierfür das Widget „Data Table“.

Notieren Sie anschließend in absteigender Reihenfolge sowohl die Anzahl als auch die relative Häufigkeit der jeweiligen Typen an Gesetzesverstößen.

Tipp: Nutzen Sie hierfür die Einstellungsmöglichkeiten des Widgets „Distributions“.

6. Setzen Sie das Widget „Scatter Plot“ (Streudiagramm) ein, um sich die Verteilung der Gesetzesverstöße im Stadtgebiet (Ortsangabe aus Längen- und Breitengrad) anzeigen zu lassen. Nehmen Sie dafür die unten abgebildeten Einstellungen des Widgets vor und interpretieren Sie das Streudiagramm.

The image shows the configuration panel for a Scatter Plot widget. It is divided into several sections:






- Axes:** Axis x is set to 'Lon' and Axis y is set to 'Lat'. There is a button labeled 'Find Informative Projections' below the axes.
- Attributes:** Color is set to 'Type', Shape is '(Same shape)', Size is '(Same size)', and Label is '(No labels)'. There is a checked checkbox for 'Label only selection and subset'.
- Visuals:** Sliders for Symbol size, Opacity, and Jittering. A checkbox for 'Jitter numeric values' is present and unchecked.
- Options:** A list of checkboxes: 'Show color regions' (unchecked), 'Show legend' (checked), 'Show gridlines' (unchecked), 'Show all data on mouse hover' (checked), 'Show regression line' (unchecked), 'Treat variables as independent' (unchecked), and 'Show confidence ellipse' (unchecked).

7. **Für Schnelle:** Sie können Streudiagramme auch für eine einzelne Verbrechenart erstellen, indem Sie im Widget „Data Table“ eine einzelne Verbrechenart auswählen und danach mit einem zweiten „Scatter Plot“-Widget verbinden.

Versuchen Sie es beispielsweise bei „Public Drunkenness“ und „Prostitution“ und interpretieren Sie nun die einzelnen Streudiagramme.

Aufgabenteil: Klassifikatoren einsetzen, vergleichen und überprüfen

8. Lesen Sie den Datensatz „Philadelphia_crime_bearbeitet“ in ein zweites „File“-Widget auf der gleichen Benutzeroberfläche ein. Öffnen sie das „File“-Widget per Doppelklick und entfernen Sie im Datensatz das Merkmal „Date“. Dazu wählen Sie für das Merkmal 1 (name: Date) bei „Role“ **skip** aus.

Columns (Double click to edit)			
Name	Type	Role	Values
1 Date	 datetime	skip	
2 Time	 datetime	feature	
3 Type	 categorical	target	Family Abuse, Gam...
4 Lat	 numeric	feature	
5 Lon	 numeric	feature	

8.1 Erstellen Sie nun zwei Klassifikatoren zur Vorhersage. Nutzen Sie hierfür die Widgets „Test and score“ und „Confusion Matrix“.

Tipp: Begrenzen Sie im Widget „Tree“ die Baumtiefe (bei: limit max. tree depth) auf 9, um die Rechenzeit zu begrenzen. Es ergibt sich ein ziemlich unübersichtlicher Baum, der aber später gebraucht wird.

Legen Sie beim kNN die Anzahl nächsten Nachbarn auf 5 fest.

Verwenden Sie zum Testen der Klassifikatoren im Widget „Test & Score“ die Methode „Leave one out“.

Alternative:

Begrenzen Sie im Widget „Tree“ die Baumtiefe (bei: limit max. tree depth) auf 5, um Rechenzeit zu begrenzen und den Baum überschaubar zu halten.

Legen Sie beim kNN die Anzahl nächsten Nachbarn auf 5 fest.

Verwenden Sie zum Testen der Klassifikatoren im Widget „Test & Score“ **nicht** die Methode „leave one out“, sondern testen Sie mit Cross Validation.

8.2 Vergleichen Sie die Korrektklassifikationsrate (KKR od. CA) bei den beiden Klassifikatoren und markieren Sie den „besseren“ Klassifikator!

8.3 Erläutern Sie einen möglichen Grund für das schlechtere Abschneiden eines der beiden Klassifikatoren.

8.4 Bei der Polizeidienststelle in Philadelphia geht um 23:40 Uhr ein Notruf einer völlig verzweifelten Person ein. Nach wenigen Sekunden wird der Anruf jäh unterbrochen. Zum Glück bestand die Verbindung ausreichend lange, so dass die folgenden GPS-Daten erfasst werden konnten: Längengrad: -75,0731 Breitengrad: 40,0322

Police cadet Jo will sofort losfahren, um den Sachverhalt alleine aufzuklären. „Sicher, wieder ein Betrunkener!“ Doch der Chief bremst ihn aus. „Vorsicht, Junge“, meint er, „alleine is nicht!“

Ist die Vorsicht des Chief's berechtigt?

Überlegen Sie, welche beiden Möglichkeiten Sie zur Verfügung haben, um vorherzusagen, welche Verbrechenart vorliegt.

Viele weitere Aufgaben dazu denkbar:

Konfusionsmatrix auswerten

Klassifikationsergebnisse verbessern (Parameter ausprobieren)

Ethische Diskussion zu Beginn aufgreifen und Datenanalyse aufgreifen!

etc.

Show probabilities for		Classes known to the model	<input checked="" type="checkbox"/> Show classification errors				
	Tree		error	Type	Time	Lat	Lon
1	0.00 : 0.00 : 1.00 : 0.00 : 0.00 : 0.00 → Homicide			?	23:40:00	40.0322	75.0731