

Anfertigen eines CAT bzw. PIKE

Autor: Andreas Schau

Kann ... // Algorithmus x // exact ... // (Kunden)-Problem ... berechnen / lösen?

Ja. Mit KNIME konnte ein Datensatz, der Persönlichkeitswerte junger Menschen und deren Konsumverhalten verschiedener legaler und illegaler Drogen enthält analysiert und so optimiert werden, dass ein Machine-Learning-Algorithmus bei bestimmten Drogen, mit mehr als 70% Genauigkeit vorhersagen kann, ob dieser Persönlichkeitstyp diese Droge schon einmal konsumiert hat oder nicht.

Big Data / Data Mining / Data Science Kernaussage:

(P) roblem):

[Welcher Frage hat für die Lösung des Kunden / Auftraggeber die größte Bedeutung?]:

Kann mithilfe von KNIME und Machine-Learning vorausgesagt werden, welche Persönlichkeits-Typen eher zum Konsum bestimmter Drogen neigen?

(I) ntervention:

(Bibliotheken- und Algorithmen-Auswahl, ... z.B. pandas für Finanzdaten ...)

[Welche Berechnung erwäge ich vornehmlich?]:

Es kam die Open-Source-Software KNIME zum Einsatz. Dieser wurden aus dem KNIME-Hub noch weitere Knoten aus der Javascript-Bibliothek hinzugefügt.

Via dieser Software wurden die Daten aufbereitet, Analysen und Machine-Learning angewendet, sowie Visualisierungen erstellt.

(K) ontrollintervention

(falls erforderlich: Bibliotheken- und Algorithmen-Auswahl ... z.B. scikit-learn für Finanzdaten ...)

[Was ist die andere Möglichkeit?]:

Die Visualisierungen hätte man auch nach der Daten Aufbereitung in Python in Tableau erreichen können. Python hätte ebenfalls beim Machine-Learning via scikit-learn Bibliothek und den Visualisierungen via Seaborn Bibliothek verwendet werden können.

(E) rgebnismaß (Zielgröße(n)) – Die Evidence

[Was möchte ich / der Kunde erreichen? Z.B. Prädiktor oder Klassifikator erstellen ...]:

Dem Autor fiel in seinem Umfeld auf, dass es manchen Menschen leichter und Anderen schwerer fällt mit verschiedenen Drogen umzugehen. Etwa wenn diese Menschen in Ihrer Jugend mit dem Rauchen angefangen haben und es nicht mehr so einfach los zu werden scheinen oder wenn durch Operationen eine Abhängigkeit zu Schmerzmitteln entstanden zu sein scheint.

Die hier verwendeten Daten ermöglichten eine Untersuchung, ob und wie stark die eigenen, psychologisch ermittelten Persönlichkeitswerte, Vorhersagen zum Konsumverhalten erlauben könnten.

Sind manche Menschen schon durch Ihren Charakter vorprogrammiert mit bestimmten Drogen Probleme zu bekommen, weil sie viel eher dazu neigen diese je auszuprobieren?

Gibt ihnen ihr Charakter eher vor nach solchen Erfahrungen zu suchen?

Anmerkungen:

Literaturhinweise:

Datensatz: [Daten von kaggle](#)

Die Suche nach der besten Evidenz

1. Problem

Kann eine Bestimmung der wichtigsten Charakterzüge helfen Drogen-Prävention effektiver zu gestalten? Zum Beispiel, indem man Persönlichkeiten, die zu bestimmten Konsummustern neigen eine speziell auf sie angepasste Aufklärung anbietet?

2. Definition einer wichtigen suchbaren Frage

Welche Zusammenhänge bestehen unter diesen Gesichtspunkten zwischen Charakter und Drogenkonsum?

3. Auswahl der wahrscheinlichsten Quelle für diese Evidenz

Via Korrelationsmatrix und Heatmap kann man erste Einsichten zwischen Charakter und Drogenkonsum gewinnen. Ein Machine-Learning-Algorithmus kann verwendet werden um die Zusammenhänge zu lernen und Vorhersagen zu machen. Je höher man die Genauigkeit der Vorhersagen bekommt, desto besser konnten die versteckten Abhängigkeiten gelernt werden.

Ein so trainiertes Modell könnte jungen Menschen zeigen, wofür sie möglicherweise eine Schwäche haben und sie könnten dann an Material verwiesen werden, welches speziell Sie am besten unterstützt.

4. Erstellung einer Suchstrategie

Erst kann man versuchen, alle Daten zum lernen zu verwenden. Wenn dies keine guten Ergebnisse liefert, kann man gezielt den Lernumfang verringern auf die relevantesten Features des Datensatzes. So kann der Konsum-Raum von verschiedenen Zeiträumen um seine temporale Dimension gekürzt werden und nur noch aussagen, ob eine spezielle Droge schon konsumiert wurde oder noch nicht.

5.0 Zusammenstellung der Evidenzausbeute

5.1 Falle Ausbeute schlecht

**Auswahl der zweit-wahrscheinlichsten
Quelle für diese Evidenz
Erstellung einer Suchstrategie
Zusammenfassung der Evidenz
Anwendung der Evidenz**

Dieser Ansatz kann dann für jede Droge einzeln durchgeführt werden und so kann pro Droge herausgefunden werden, welche Persönlichkeiten für diese am anfälligsten sind.

6. Anwendung der Evidenz

Ein Random-Forest wurde verwendet um den LSD Konsum aus den Persönlichkeitswerten heraus zu lernen und dieser konnte mit einer Genauigkeit von 78% voraussagen treffen, ob LSD anhand dieser schon von einem Menschen konsumiert wurde oder noch nicht.

Diese Genauigkeit wirkt nicht überragend groß, jedoch sollte man beachten, dass es ja keinen 100% Zusammenhang zwischen Charakter und Drogenkonsum geben muss, daher könnte dieser Wert auch schon sehr gut sein.

Anmerkung: Der Autor ist weder Psychologe noch Pharmakologe und trifft nur Aussagen nach bestem Gewissen und Überlegungen.