

Predictive Analytics

Was ist Predictive Analytics?

Bei der prädiktiven Analytik geht darum eine Statistik z.B. über das Verhalten einer Person zu erstellen. Die einzelnen Informationen werden dann aus den Datenmengen extrahiert und verglichen, um Zusammenhänge in diesen Datenmengen zu entdecken und eine Vorhersage des Verhaltensmusters der ausgewählten Person zu treffen. Meist sind es wiederkehrende Entscheidungen, die epochal sind bei der Prognose der Handlungsweise eines Nutzers.

Prädiktive Analytik ist eine statistische und analytische Technik um ein Modell zu entwickeln, dass die zukünftigen Ereignisse und Handlungsweisen darstellt. Das Modell unterscheidet sich stark nach welchen Vorhersagen sich gerichtet wird.

Charles Nyce, „Predictive Analytics White Paper“, unter: <https://www.the-digital-insurer.com/wp-content/uploads/2013/12/78-Predictive-Modeling-White-Paper.pdf>, S.1 (abgerufen am 01.11.2017).

Das wichtigste bei dieser und wie bei vielen anderen analytischen Techniken ist es zuverlässige Daten mit hoher Qualität zu haben, um eine relative saubere Analyse durchzuführen. „Predictive Analytics“ ist in vielen Branchen gegenwärtig sehr ausgeprägt vorhanden. Von dem Vorhersehen von Flugpreisen, Kundenkaufverhaltens oder das Aufstellen von Prognosen hinsichtlich der Behandlungserfolge an Patienten gibt es ein breites Spektrum der Einsatzmöglichkeit dieses Verfahrens. Jedoch gibt es einige Bedenken bei der Ausführungsgewalt der „Predictive-Analytics-Software“. Falls die Software so entworfen ist, dass der Endbenutzer falsche Parameter eingeben kann, so wird es zur einer falschen Vorhersage kommen und einen gravierenden Schaden beim Unternehmen verursachen. Um dem entgegenzuwirken ist es unabdinglich die Software so zu gestalten, dass es keine falschen Parameter geben kann

Heather McIlvaine (21.03.2013), „Was ist eigentlich Predictive Analytics“, unter: <https://news.sap.com/germany/was-ist-eigentlich-predictive-analytics/> (abgerufen am 01.11.2017).

„Predictive Analytics Process“

Bei dem Ablauf der prädiktiven Analytik gibt es eine feste Reihenfolge, die jedes Mal abgearbeitet werden muss um eine vollwertige Analyse zu bekommen.

Der **erste** Schritt bezieht sich auf das „Bestimmen des Projekts“. Hier wird, wie der Name schon erahnen lässt das Projekt definiert und es werden Datensätze mit denen man arbeiten wird identifiziert und bereitgelegt.

Der **zweite** Schritt ist dafür gedacht die Daten zu sammeln. Ein sehr wichtiger Begriff ist das „Data Mining“ was sich auf die große Daten Menge bezieht und durch Methoden diese verarbeitet. Bei diesem Schritt ist die Benutzerinteraktion unabdingbar mit einbezogen, wodurch sich ein Einblick in diese Interaktion ermöglicht wird.

Im **dritten** Schritt wird die eigentliche Datenanalyse durchgeführt. Durch die Inspektion von Daten werden wichtig eingestuft Informationen aufgefunden.

Beim **vierten** Schritt wird eine Statistik anhand der vorgefundenen Daten erstellt.

Die Statistik dient als **fünfter** Schritt dann als ausschlaggebendes Instrument zur Modellierung der Vorhersagungsmodelle.

Der **vorletzte** Schritt ist die Entwicklung. Dieser Abschnitt widmet sich völlig der Entwicklung der Vorhersagungsmodelle. Durch Entscheidungen und Verhaltensmuster die im Alltag getroffen werden, können diese als großes Ganzes dafür genutzt werden, um die Modellentwicklung zu automatisieren.

Die Modellüberwachung ist der **siebte** Schritt und bildet das Schlusslicht dieses Prozesses. Am Ende wird das Modell wie so gut wie jedes verwaltet und überwacht. Diese gilt der Überprüfung und beugt mögliche Fehler vor.

„What is Predictive Analytics“, unter: <https://www.predictiveanalyticstoday.com/what-is-predictive-analytics/> (abgerufen am 01.11.17)

API (Application Programming Interface)

Google selbst bietet die Google Cloud Prediction API an, die vor allem Funktionen der „Predictive Analytics“ ermöglicht. Diese RESTful-API ermöglicht es die Erstellung eines Modells für maschinelles Lernen und lässt damit zu Daten zu analysieren und verschiedene Funktionen zu Ihren Anwendungen hinzuzufügen. Leider ist die API veraltet und wird ab dem 30.04.2018 außer Dienst gestellt.

Welche Aufgaben hat die Prediction API?

Die API führt laut dem offiziellen Entwicklerleitfaden der „Cloud Prediction API“ zwei Aufgaben durch:

1. *„Bei einem neuen Element wird ein numerischer Wert vorhergesagt, der für dieses Element gilt. Dieser Wert basiert auf ähnlich bewerteten Beispielen in den Trainingsdaten.“*
2. *„Bei einem neuen Element wird eine Kategorie gewählt, die es am besten beschreibt. Voraussetzung dafür ist ein Satz von ähnlich kategorisierten Elementen in den Trainingsdaten.“*

„Entwicklerleitfaden“, unter: https://cloud.google.com/prediction/docs/developer-guide#what_is_the_prediction_api (abgerufen am 01.11.17)

Der Nachfolge von Google Cloud Prediction API ist „Cloud Machine Learning Engine“. Wie Cloud Prediction erfüllt es alle gewünschten Funktionen und verfügt über das leistungsstarke TensorFlow-Framework, das hinter vielen Produkten von Google steht. TensorFlow selbst, ist eine Open Source Software-Bibliothek für numerische Berechnungen mit Hilfe von Datenflussgraphen.

Besonderheiten von „Cloud Machine Learning Engine“:

1. Einfache Modelle für maschinelles Lernen anfertigen, die mit allen Datentypen und -größen arbeiten können.
2. TensorFlow-Framework (Deep Learning Funktionen)
3. Skalierbare prognostische Analysen

„Cloud Machine Learning Engine“, unter: <https://cloud.google.com/ml-engine/> (abgerufen am 01.11.17)