

Einsendeaufgabe 8 - ML Introduction Exercise

Welche Bedeutung hat die Verallgemeinerung in dem Papier?

Die Verallgemeinerung ist eine kostengünstige und oft durchführbare Methode im Bereich des maschinellen Lernens. Sie wird genutzt, um Algorithmen anhand von verallgemeinerten Beispielen lernen zu lassen, wie wichtige Aufgaben durchgeführt werden können. Die Verallgemeinerung hilft besonders dort, wo manuelle Programmierung nicht möglich ist. Diese Bedeutung ist als wichtige Erkenntnis gekennzeichnet. Die Verallgemeinerung sollte über die Beispiele im Trainingsset hinaus gehen, weil es nicht wahrscheinlich ist, den gleichen Beispielen in der Testzeit noch einmal zu sehen.

Was sind die vielen Gesichter der Überanpassung?

Die vielen Gesichter der Überanpassung werden in dem Papier als „Schreckgespenst“ des maschinellen Lernens angesehen. Es wird von vielen Gesichtern gesprochen, weil die Überanpassung in vielen Formen auftreten kann. Die Überanpassung ist vergleichbar mit dem menschlichen Halluzinieren. Der Algorithmus erkennt zum Beispiel nicht vorhandene Muster in einem Datensatz.

Warum haben Menschen in höheren Dimensionen Probleme?

Ich würde sagen, wir Menschen sind einfach nicht dafür gemacht, in höheren Dimensionen zu denken. Unsere Vorstellungskraft reicht dafür nicht aus, sie ist schlicht begrenzt, um in hohen Dimensionen zu verstehen was dort passiert.

Was ist Feature Engineering?

Feature Engineering steht für den Prozess, in welchem Features aus den Rohdaten des jeweiligen Projekts abgeleitet bzw. erzeugt werden. Des Weiteren sollen dabei auch die besten Features für die spezifische Aufgabe gefunden werden.

Warum schlagen mehr Daten kluge Algorithmen?

Um so mehr Daten vorliegen, um so mehr kann man Trainieren und Testen. Dadurch werden immer mehr Abweichungen aussortiert und somit werden die Ergebnisse immer exakter.

Was ist Ensemble-Lernen?

Ensemble-Lernen ist das Verwenden verschiedener Werkzeuge / Algorithmen in Kombination um bessere Ergebnisse zu erzielen. Bei einem Wettbewerb von Netflix war der Gewinner einer Kombination von Ensemble aus über 100 Lernenden.

Was ist Genauigkeit in der Datenwissenschaft?

Die Genauigkeit in der Data Science ist eine errechenbare Zahl, die uns Auskunft darüber gibt, wie genau zum Beispiel Vorhersagen beim ML sind. Sie kann je nach Bedarf verschieden definiert werden.