

Das Konzept des Machine Learnings existiert seit vielen Jahren. Bereits in den 1990 Jahren kamen erste Versuche auf, Maschinen intelligenter zu machen. Während zu diesen Zeiten die Anzahl der Datensätze klein war, änderte sich mit der Zeit die Nutzung des Internets und somit auch die Anzahl der Daten von wenigen Hunderten auf Millionen, wenn nicht Milliarden. Doch ohne das Wissen, wie diese zu verarbeiten sind, hilft selbst der Beste Algorithmus nicht.

In diesem Artikel wird anhand von Beispielen aus Bereichen des Gesundheitssystems, der Nationalen Sicherheit und unter anderem der Fertigungsindustrie gezeigt, wie Probleme des Machine Learning effizient behoben werden können. Konkret spricht man von drei großen Problemen wie, den Inhalt der Daten, die effiziente Verarbeitung großer Inhalte und die Wahrhaftigkeit der Aussagen.

Während frühere Ansätze aus der Mathematik mittels Linearen Algebra oder der Set Theorie sich als ineffizient herausstellen, musste man neue Methoden finden. Als man Systeme zu parallelisieren, bemerkte man eine deutliche Steigerung der Performance unter Verwendung von Graph Algorithmen.

Mit all dem neuen Wissen, kam man auch zu neuen Ereignissen. Der Versuch High Performance Computing mit Big Data Architekturen zu verknüpfen, stellte eine ineffiziente Lösung dar. Weiters machten wir die Erkenntnis, dass der leistungsstärkste Algorithmus das beste Verständnis über seine Daten hat. Dies war einer der Gründe, weshalb Datensätze einheitlich strukturiert sein müssen. Trotz all dem sind Daten vielfältiger. Nicht jeder Datensatz wird als Matrix abgebildet. Viele Daten wie, Fotos, Videos oder Texte sind in sich einheitlich, aber unterschiedlich zu verarbeiten.