

Bestärkendes Lernen

Unter einem sequentiellen Entscheidungsprozess (Sequential Decision Process, SDP) versteht man jedwede Art von dynamischem Prozess, der eine zeitgeordnete Entscheidungssequenz beinhaltet, die sich auf im Verlauf der Zeit wachsende Kosten oder Gewinne auswirkt. Es soll damit diejenige Entscheidungssequenz ermittelt werden, die zu einer Optimierung der definierten Kosten oder Gewinne führt.

Ein drittes Teilgebiet des maschinellen Lernens stellt das bestärkende Lernen, auch verstärkendes Lernen oder Reinforcement Learning genannt, dar. Die grundlegenden Merkmale decken sich teilweise mit denen des überwachten und unüberwachten Lernens, der Anwendungsbereich des bestärkenden Lernens bezieht sich allerdings ausdrücklich auf den Problembereich der sequentiellen Entscheidungsprozesse, weswegen hierfür auch auf eine andere Art von Daten zurückgegriffen wird als je beim überwachten bzw. unüberwachten Lernen. Die Theorie zum bestärkenden Lernen stützt sich größtenteils auf das in den 1950er Jahren von Richard Bellman eingeführte Konzept der dynamischen Programmierung (Bellman, 1957). Hieraus kann – wie weiter unten dargelegt – zwar ein Großteil der mathematischen Grundlage für das bestärkende Lernen gewonnen werden, allerdings ist die Rechenkapazität bei der dynamischen Programmierung begrenzt. Für die Praxis bedeutet dies, dass die dynamische Programmierung jedenfalls in ihrer konzeptgetreuen Ausführung keinen tauglichen Lösungsansatz für die Entwicklung von Anwendungen zur Nutzbarmachung von Hochdurchsatzdaten, wie sie heutzutage oftmals verfügbar sind, darstellt.

Eine besonders aufsehenerregende frühe Anwendungsform von bestärkendem Lernen bestand in der Entwicklung eines SDP, der als TD-Gammon bekannt wurde und Backgammon auf Wettbewerbsniveau spielen konnte. Erst vor Kurzem setzte auch DeepMind Technologies bestärkendes Lernen für die Entwicklung des Schachprogramms AlphaZero ein. Für solche Spielanwendungen eignet sich bestärkendes Lernen besonders gut, vor allem in Verbindung mit Modellen tiefer neuronaler Netze (tiefes bestärkendes Lernen oder Deep Reinforcement Learning). So wird es derzeit beispielsweise für die Entwicklung bemerkenswert leistungsstarker Videospieldanwendungen eingesetzt.