

Information » Transformer-Modell «

Ein Transformer-Modell ist eine spezielle KI-Architektur, die 2017 von Google entwickelt und vorgestellt wurde. Es hat das maschinelle Lernen, insbesondere im Bereich der natürlichen Sprachverarbeitung (NLP) revolutioniert. Das Besondere an Transformern ist der sogenannte Aufmerksamkeitsmechanismus, der es dem Modell ermöglicht, Beziehungen zwischen verschiedenen Teilen einer Eingabesequenz zu erkennen und zu gewichten.

Ein Transformer besteht aus zwei Hauptteilen: dem Encoder, der den Eingangstext verarbeitet, und dem Decoder, der den Ausgangstext erzeugt. Diese Architektur erlaubt es, Informationen effizient zu verarbeiten und mit einer Parallelverarbeitung viele Wörter gleichzeitig zu betrachten. Transformer-Modelle sind die Grundlage für viele Anwendungen wie maschinelle Übersetzung und Textgenerierung und haben sich als sehr leistungsfähig erwiesen.