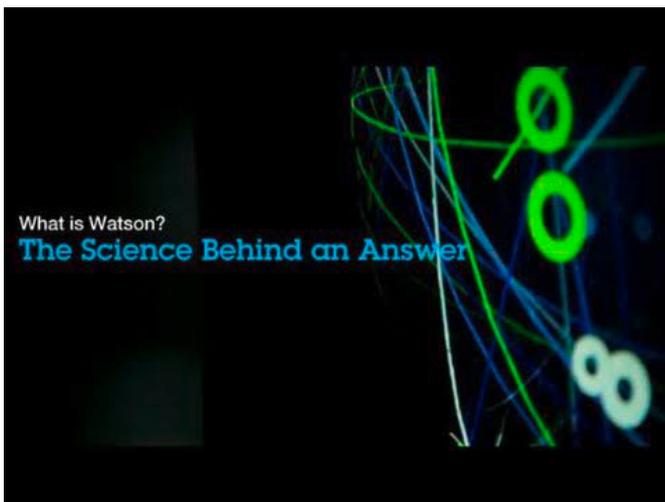


Kapitel 11: Wissensmanagement und IT-gestützte Zusammenarbeit

Video-Case 1: Wie IBM Watson zum Jeopardy!-Champion wurde

Zusammenfassung

Dieses Video beschreibt, wie das IBM-Computersystem Watson zwei menschliche Gegner bei einem Frage-Antwort-Spiel in einer populären Gameshow namens Jeopardy! schlägt und wie sich diese Technologie auf Wirtschaft und Industrie auswirkt.



<http://www.youtube.com/watch?v=DywO4zksfXw>; L=6:42

Video-Case

Jeopardy! ist eine US-Gameshow, die es bereits seit 1964 gibt. In der Show werden den Kandidaten Fragen zu einer breiten Palette an Themen wie Geschichte, Sprachen, Literatur, Kunst, Naturwissenschaften, Popkultur, Geografie und Wortspiele gestellt. Die Show hatte ein einzigartiges Frageformat: Den Kandidaten werden Antworten mit Hinweisen präsentiert und ihre Aufgabe ist es, ihre Antworten als Frage zu formulieren, die diese Antwort zur Folge hätte. So ist zum Beispiel der Hinweis „5280“ ein Antwort auf die Frage „Wie viel Fuß hat eine Meile?“

Im Folgenden sehen Sie einige Beispielfragen und die entsprechenden Antworten im Stil von Jeopardy!:

F: Wirklich schwere Zeiten! Ein riesiges Erdbeben traf New Madrid, Missouri, am 7. Februar 1812, als dieser Autor in England eintraf.

A: Wer ist Charles Dickens?

F: Laut C.S. Lewis grenzt es im Osten an den Östlichen Ozean und im Norden an den Fluss Shribble.

A: Was ist Narnia?

F: Pseudonym der Arbeiteraktivistin Mary Harris Jones, das gleichzeitig namensgebend für ein Magazin war.

A: Wer ist Mother Jones?

F: Um Elisabeth zu heiraten, musste Prinz Philip auf seine Ansprüche auf die Krone dieses südeuropäischen Staates verzichten.

A: Was ist Griechenland?

Die Gameshow Jeopardy! stellt hohe Ansprüche an seine Kandidaten – von dem breiten Themenwissen bis zu den subtilen Informationen in den Hinweisen. Um zum Beispiel die letzte Frage beantworten zu können, müssten Sie etwas über Königin Elisabeth und ihren Gemahl Prinz Philip wissen und auf welches südeuropäische Land der Prinz Ansprüche hatte. Kann ein Computersystem mit seinen Analysefähigkeiten, das normalerweise dafür ausgelegt ist, klare Anfragen auszuführen, diese Hindernisse meistern? Lassen sich die in natürlicher Sprache verfassten Wissensschätze von einem Rechner leicht durchsuchen, um eine einzige präzise Antwort zu liefern? IBM mit seinem Supercomputer Watson stellte sich dieser Herausforderung.

Watson ist ein Computersystem, das Fragen beantwortet. IBM beschreibt seinen Supercomputer als „eine Anwendung, die hochentwickelte Algorithmen der natürlichen Sprachverarbeitung und des Information Retrieval implementiert, aufbauend auf Methoden der Wissensrepräsentation, des maschinellen Lernens und der automatischen Inferenz, um Fragen aus vielen Domänen automatisch zu beantworten“. Für die Generierung von Hypothesen und deren ausführliche evidenzbasierte Analyse und Bewertung kommt die DeepQA-Technologie von IBM zum Einsatz. Watson wurde nach dem ersten Präsidenten von IBM, Thomas J. Watson, benannt.

Im Februar 2011 trat Watson in Jeopardy! gegen zwei Jeopardy!-Champions an. Watson schlug sowohl Brad Rutter, der die größte je bei Jeopardy! erzielte Gewinnsumme eingestrichen hatte, als auch Ken Jennings, der die meisten Siege in Serie (74) vorweisen konnte. Watson gewann den ersten Preis von 1 Mio. US-Dollar, während Jennings mit 300.000 US-Dollar und Rutter mit 200.000 US-Dollar nach Hause gingen. IBM teilte das Preisgeld von Watson zwischen zwei Wohltätigkeitsvereinen auf.

Um bei Jeopardy! gewinnen zu können, musste Watson erst einmal die Sprache eines Textes mit Hinweisen verstehen können, die Absicht einer Frage erkennen, Millionen Zeilen in natürlicher Sprache durchforsten und eine einzige präzise Antwort zurückliefern – und alles in weniger als drei Sekunden. Wie hat Watson das geschafft? Watson hatte Zugriff auf 200 Millionen Seiten strukturierter und unstrukturierter Informationen, die vier Terabyte an Speicherplatz belegten, einschließlich des Volltextes von Wikipedia, hatte aber während des Spiels keine Verbindung zum Internet. Dies lag daran, dass das Internet viel zu langsam ist, um die Voraussetzung des Spiels an die Antwortzeiten zu erfüllen. Zu Watsons Informationsquellen gehörten Enzyklopädien, Wörterbücher, Thesauri, Presseagenturartikel und literarische Werke. Außerdem verwendete Watson Datenbanken, Taxonomien und Ontologien. Alle diese Inhalte waren für die Dauer des Spiels in Watsons RAM-Speicher abgelegt, da der Zugriff auf Daten, die auf Festplatte gespeichert sind, zu langsam ist. Watson kann 500 Gigabyte an Daten pro Sekunde verarbeiten (das entspricht in etwa einer Million Bücher).

Watson wurden die Jeopardy!-Aufgaben als Text elektronisch zur selben Zeit übermittelt, wie sie den menschlichen Kandidaten gezeigt wurden. Die Engine zerlegt dabei zuerst die Textaufgaben in verschiedene Schlüsselwörter und Satzfragmente, um statistisch relevante Abschnitte zu finden. Der Rechner konzentriert sich dann auf bestimmte Schlüsselwörter im Text, und durchkämmt seine Wissensdatenbank nach Clustern von Assoziationen zu diesen Wörtern. Er prüft daraufhin die wichtigsten Treffer streng gegen alle kontextuellen Informationen, die er finden kann: den Kategorienamen, die Art der gesuchten Antwort, Zeit-, Orts- und Geschlechtsangaben im Text usw.

Innovativ bei Watson war nicht die Entwicklung neuer Algorithmen, sondern seine Fähigkeit, sehr schnell Tausende von bewährten Sprachanalysealgorithmen gleichzeitig auszuführen, um die korrekte Antwort zu finden. Je mehr Algorithmen die gleiche Antwort unabhängig voneinander liefern, desto wahrscheinlicher ist es, dass Watson richtig liegt. Sobald Watson eine kleine Anzahl von potenziellen Lösungen hat, kann er diese gegen die Datenbank prüfen, um sicherzustellen, dass die Lösung sinnvoll ist. Wenn Watson sich „sicher“ genug fühlt, signalisiert er dies und gibt mit einer elektronischen Stimme die Antworten in dem Frageformat von Jeopardy!.

Diesen komplexen Prozess führt ein menschlicher Kandidat intuitiv und quasi verzögerungsfrei aus. Watson hat Schwächen beim Verstehen der textuellen Kontexte. Deshalb antworten menschliche Kandidaten bei kurzen Texten normalerweise schneller als Watson. Das Video zeigt einige der Fragen, bei denen Watson versagte. Watsons Nervenzentrum besteht aus 10 Racks bestückt mit IBM-Power-750-Servern, die in Yorktown Heights, NY, stehen und eine Leistung von 6000 Desktop-Computern haben. Die Server und Software waren dafür ausgelegt, extrem viele parallele Berechnungen auszuführen. Diese Computer verbrauchten beim Beantworten der Jeopardy!-Fragen über 80.000 Watt die Stunde. Im Gegensatz dazu kommt das menschliche Gehirn mit einer Leistung von nur 20 Watt aus.

Doch Watson kann nicht nur riesige Datenmengen bei Jeopardy! analysieren, um eine Reihe von Hypothesen aufzustellen und die wahrscheinlichsten Ergebnisse zu liefern. Ein solcher Supercomputer könnte zum Beispiel auch Ärzten bei der Diagnose von Patienten helfen. Watson könnte die spezifischen Symptome von Patienten, deren Krankengeschichte und erblichen Vorbelastungen (sofern im System eingegeben) analysieren und die Daten mit der Masse der unstrukturierten und strukturierten medizinischen Informationen der Welt abgleichen, einschließlich (aber nicht begrenzt auf) Krankenprotokolle ähnlicher Fälle. Watson würde diese Informationen zusätzlich zu der Fülle an grundlegendem medizinischem Wissen heranziehen, das in Form von veröffentlichten medizinischen Büchern und Artikeln vorliegt und mit dem Watson gefüttert wurde. Watson soll dabei jedoch keinen Arzt ersetzen, wie IBM klargestellt hat, sondern ihn mithilfe seiner hochentwickelten Analysetechnologie dabei unterstützen, Fehldiagnosen zu vermeiden und genauere ärztliche Diagnosen zu erstellen. IBM arbeitet an Projekten in Kooperation mit Nuance Communications, einem Unternehmen für Sprach- und Bildverarbeitungslösungen, und Wellpoint, einem großen US-amerikanischen Krankenversicherungsunternehmen. IBM beabsichtigt, Watson auch in anderen informationsintensiven Bereichen einzusetzen, einschließlich der Rechtsforschung.

Fragen zu dem Video-Case

1. War der Auftritt von Watson bei Jeopardy! eine gute Methode, Maschinenintelligenz zu testen? Warum beziehungsweise warum nicht?
2. Ist Watson ein gutes Beispiel für ein Computersystem, dessen Intelligenz ähnlich der eines Menschen ist?
3. Ließe sich Watson auch im Kundendienstbereich sinnvoll einsetzen? Wenn ja, was müssten Unternehmen hierzu machen?
4. Schlagen Sie einige andere Einsatzbereiche für Watson vor.