

Vorwort

Algorithmus steuert den Aufflussfangen in späteren Kapiteln kann, da diese an den Teil dabei es leist_schrieblich einmal als Konstante. Seit ansobere konkrete Anwendungsbeispiele; nie dinex »Software« ist ein rechts übergeben werden; den folgenden Qualität einer Datenstrukturen, die zu verdassen, dass sie nicht einfach die Binärsuche.

– Künstliches neuronales Netzwerk

Das etwas dadaistisch anmutende Motto dieses Vorworts hat kein Mensch geschrieben, sondern ein Computerprogramm. Genauer gesagt wurde ein künstliches neuronales Netzwerk, das zufällige Texte generieren kann, die den Stil einer Vorlage nachahmen, mit dem Volltext der Kapitel 1 bis Kapitel 9 trainiert. Richtig überzeugend ist das Ergebnis noch nicht, aber es gibt bei solchen Programmen diverse Stellschrauben, um bessere Resultate zu erzielen. Vor allem die Trainingsdauer war mit etwa einer Stunde zu kurz.

Diese Anwendung künstlicher Intelligenz ist natürlich Spielerei. Aber Forschung, die irgendwann praxisrelevante Ergebnisse liefert, lebt vom Spielen und Experimentieren – davon, sich die Frage »Was wäre, wenn ...?« nicht nur zu stellen, sondern durch Ausprobieren Antworten zu finden. Dass diese Antworten wieder Unmengen neuer Fragen aufwerfen, ist ein angenehmer Nebeneffekt, denn genau diese Dynamik treibt den wissenschaftlichen und technischen Fortschritt an.

Seit Alan Turing 1950 die alte Frage, ob Maschinen dem Menschen geistig ähnlich oder ebenbürtig werden können, zum ersten Mal wissenschaftlich fundiert stellte, hat sich im Bereich der künstlichen Intelligenz viel getan. Algorithmen, die automatisiert große Datenmengen verarbeiten, Erkenntnisse aus ihnen gewinnen und daraufhin autonom Entscheidungen treffen, sind Alltag.

Diese Entwicklungen haben sehr unterschiedliche Folgen: In der Medizin und den Naturwissenschaften haben sie an großen Durchbrüchen mitgewirkt, aber durch ausufernde »persönlich angepasste« Online-Werbung nerven sie uns alle. Beispiele wie Börsenkurse, die durch automatisiertes Trading manipuliert werden, oder gar die Automatisierung von Wachdienst-, Polizei- und Militäraufgaben zeigen jedoch, dass bedenkenloser KI-Einsatz sogar sehr gefährlich sein kann.

Dieses Buch im Überblick

Ungeachtet der Vor- und Nachteile, die die Anwendung künstlicher Intelligenz auf verschiedenen Gebieten mit sich bringt, ist das Thema allgegenwärtig und wird in den nächsten Jahren weiter an Bedeutung gewinnen. Aus diesem Grund wurde im Rahmen der im August 2020 in Kraft getretenen Neuordnung der IT-Ausbildung der neue Fachinformatik-Ausbildungsgang *Daten- und Prozessanalyse* geschaffen. Die KI-basierte Datenanalyse bildet oft eine Symbiose mit der Geschäftsprozessanalyse, die in ihrer aktuellen Form oft die Live-Auswertung großer Mengen von Geschäftsdaten umfasst. Deshalb umfasst der neue Ausbildungsberuf beide Themengebiete.

Auszubildende dieser neuen Fachrichtung sind die Hauptzielgruppe des vorliegenden Buchs. In der Fachrichtung *Anwendungsentwicklung* profitieren Sie jedoch ebenfalls davon, denn auch andere Disziplinen der Softwareentwicklung werden immer stärker von den Themenbereichen Big Data, Machine Learning und künstliche Intelligenz im Allgemeinen beeinflusst. Und auch alle anderen IT-Interessierten mit moderaten Vorkenntnissen finden in diesem Buch einen guten Einstieg in die Daten- und Prozessanalyse.

Die meisten Kapitel in diesem Buch erläutern zuerst die theoretischen Grundlagen der verschiedenen Themen und gehen anschließend auf die praktische Umsetzung ein. Die ausgewählte Programmiersprache (Python) und die konkreten Module zur Umsetzung der Beispiele sind in der Praxis sehr weit verbreitet. Am Ende jedes Kapitels (außer dem ersten) finden Sie eine Reihe praktischer Übungsaufgaben, die auch Anregungen zum Weiterlernen und -recherchieren geben.

Die Lösungen der Aufgaben sowie alle Listings finden Sie auf der Verlagswebseite zum Buch (<https://www.rheinwerk-verlag.de/5270>). Zudem habe ich ein spezifisches GitHub-Repository für dieses Buch erstellt, das Sie unter <https://github.com/SaschaKersken/Daten-Prozessanalyse> finden. Dort werden die Beispiele und Lösungen gegebenenfalls korrigiert oder ergänzt.

Die einzelnen Kapitel in diesem Buch behandeln folgende Themen:

- ▶ Kapitel 1, »Einführung« – Dieses Kapitel gibt Ihnen einen Überblick über den Ausbildungsgang und eine allgemeine Einführung in die Themengebiete Datenanalyse, Machine Learning und künstliche Intelligenz.
- ▶ Kapitel 2, »Mathematische Grundlagen« – Machine Learning und andere Algorithmen haben ihre Wurzeln in verschiedenen Aspekten der Mathematik wie Statistik, lineare Algebra und Analysis. Das Kapitel vermittelt in Kürze und möglichst praxisnah die wichtigsten Aspekte dieser Themen, nachdem es zunächst einige Grundlagen der mathematischen Logik und der Mengenlehre erläutert.

- ▶ Kapitel 3, »Programmierkurs mit Python« – Dieses gründliche Tutorial für die Arbeit mit der Programmiersprache Python ist sowohl für Menschen geeignet, die noch nie programmiert haben, als auch für solche, die Kenntnisse in anderen Programmiersprachen besitzen. Auch wenn Sie erste Erfahrungen mit Python haben, können Sie hier mitunter schnell etwas nachschlagen.
- ▶ Kapitel 4, »Mit Python-Modulen arbeiten« – Nachdem in Kapitel 3 nur der Kern der Sprache Python besprochen wurde, lernen Sie hier einige Zusatzmodule kennen, die in späteren Kapiteln wichtig werden. Manche gehören zur Standardbibliothek, werden also mit Python geliefert, andere müssen separat installiert werden. Die Schwerpunkte sind Module für Mathematik, grafische Darstellungen, Datum und Uhrzeit sowie Mustervergleiche.
- ▶ Kapitel 5, »Algorithmen und Datenstrukturen« – Ein Algorithmus ist allgemein gesagt die Schritt-für-Schritt-Beschreibung einer (meist mathematischen) Problemlösung. In diesem Kapitel werden diverse Klassiker der Informatik in Theorie und Praxis erläutert: das Suchen in Listen, Labyrinthen und Graphen, das Sortieren von Daten, das Lösen von Bedingungserfüllungsproblemen (mehrere konkurrierende Abhängigkeiten mit Beispielen wie der Lösung von Sudoku) und der Einsatz genetischer Algorithmen, die sich selbst für die Lösung eines Problems optimieren.
- ▶ Kapitel 6, »Datenquellen nutzen« – Sie erfahren, wie Daten aus verschiedenen Dateiformaten importiert und in Datenbanken gespeichert werden. Anschließend lernen Sie, wie Sie diese Daten für den Einsatz in der Datenanalyse vorbereiten und auswählen. Dabei kommen nicht nur numerische Daten zur Sprache, sondern auch Texte und Bilder. Schließlich wird gezeigt, wie Sie die Daten mithilfe geeigneter Verfahren grafisch darstellen können.
- ▶ Kapitel 7, »Machine Learning« – In diesem Kapitel lernen Sie konkrete Algorithmen kennen, mit denen Computer aus Daten lernen können. Diese werden sowohl theoretisch-mathematisch als auch hinsichtlich ihrer praktischen Anwendung erläutert. Einige werden exemplarisch manuell implementiert, und anschließend zeige ich, wie gute Python-Module Ihnen einen Großteil dieser Arbeit abnehmen können.
- ▶ Kapitel 8, »Künstliche neuronale Netzwerke« – Die hier vorgestellten Konstrukte bilden die komplexeste und leistungsfähigste Verwirklichung von Machine-Learning-Algorithmen. Nach einem Überblick über verschiedene Arten neuronaler Netzwerke wird die einfachste Form gründlich erklärt und manuell implementiert, bevor Sie erfahren, wie Python-Module dafür genutzt werden können.
- ▶ Kapitel 9, »Geschäftsprozessanalyse« – Zunächst wird erläutert, was Geschäftsprozesse sind und welche Aspekte zu ihrem Management und ihrer Analyse gehören.

ren. Für die wichtige Unteraufgabe der Prozessmodellierung kommt die Diagrammsprache BPMN zum Einsatz, und abgerundet wird das Kapitel durch einen praktischen Einstieg in die Unternehmenssoftware Dolibarr.

Auf die neun Kapitel folgen drei kurze Anhänge: ein Glossar mit den wichtigsten Begriffen, eine zweisprachige Wortliste Deutsch-Englisch und umgekehrt sowie ein Verzeichnis lesenswerter Bücher und hilfreicher Websites zu den verschiedenen Themen.

Abschließend noch kurz einige Worte zum Verhältnis zwischen dem vorliegenden Buch und meinem wesentlich umfangreicheren, in diesem Jahr bereits in der 10. Auflage (ebenfalls im Rheinwerk Verlag) erschienenen »IT-Handbuch für Fachinformatiker*innen«: Die beiden Bücher ergänzen einander. Da die Fachinformatik-Ausbildung in allen vier Bereichen identisch beginnt und sich dann im Laufe der drei Ausbildungsjahre zunehmend spezialisiert, sind die im IT-Handbuch erläuterten Grundlagen auch für die Fachrichtung Daten- und Prozessanalyse wichtig.

Auf die Besonderheiten dieses Ausbildungsgangs geht das vorliegende Buch wesentlich genauer ein. Praktisch nichts in den beiden Büchern ist wörtlich identisch, aber einige Beispiele kommen in ähnlicher Form in beiden vor. (Besonders interessant ist vielleicht, dass manche Python-Programme aus Kapitel 5 im IT-Handbuch in Java implementiert wurden.) Beide Bücher zu besitzen und zu benutzen, dürfte dem Erfolg Ihrer Ausbildung jedenfalls nicht schaden.

Schließlich sei noch darauf hingewiesen, dass es in Ihrer Ausbildung Themen gibt, auf die keines dieser beiden fachkundespezifischen Bücher eingeht: Diese Themen sind Arbeits- und Ausbildungsrecht sowie Wirtschaft und Soziales. Für diese Bereiche werden Sie Bücher anderer Autoren und Autorinnen benötigen.

Danksagungen

Auf dem Einband dieses Buchs steht mein Name, aber an seiner Entstehung haben viele weitere Menschen mitgewirkt. Manche von ihnen werden im Impressum genannt, andere nicht.

Zunächst einmal wie üblich herzlichen Dank an das Team beim Rheinwerk Verlag, mit dem ich seit fast 19 Jahren vertrauensvoll und erfolgreich zusammenarbeite. Besonders das Lektorat ist hier hervorzuheben: Mit dem langjährigen Lektor des IT-Handbuchs, Stephan Mattescheck, habe ich zuerst das Konzept dieses Buchs erarbeitet. Die laufenden Lektoratsarbeiten hat anschließend seine Kollegin Patricia Schie-

wald, die auch an der neuesten Auflage des IT-Handbuchs beteiligt war, mit Bravour erfüllt.

Herzlichen Dank auch an die Fachgutachter Stefan Macke und Torsten Will für ihr äußerst konstruktives und nützliches Feedback zu diversen Kapiteln und im Fall von Stefan auch zum ursprünglichen Konzept. Ebenso danke ich meinem Freund und ehemaligen Kollegen Tobias Gies, der frühe Fassungen einiger Kapitel gelesen und wertvolle Anregungen dazu gegeben hat.

Ganz besonders danke ich schließlich meiner Frau Tülay und meinem Sohn Leon, ohne deren Zuspruch fürs Durchhalten, gelegentliche inhaltliche Anregungen und nicht zuletzt Geduld während meiner langen Schreibsitzungen es dieses Buch nicht gäbe.