



KI-Bedarfe der Wirtschaft am Standort Deutschland

Eine Analyse von Stellenanzeigen für KI-Berufe

Eine Studie im Rahmen des Projekts „Entwicklung und Messung der Digitalisierung der Wirtschaft am Standort Deutschland“ im Auftrag des Bundesministeriums für Wirtschaft und Energie

Stand März 2021

Jan Büchel und Dr. Armin Mertens (Institut der deutschen Wirtschaft)

Inhalt

1 Einleitung.....	3
2 Daten und Methode.....	4
2.1 Daten	4
2.2 Annotation der Daten	5
2.3 Identifikation der Klassen mittels Named Entity Recognition	7
3 Ergebnisse	8
3.1 Abschlüsse.....	9
3.2 Fachrichtungen	11
3.3 Berufserfahrung.....	12
3.4 Fähigkeiten	14
4 Fazit.....	17
Technischer Anhang.....	18
Technischer Anhang 1: Beschreibung des NER-Modells.....	18
Technischer Anhang 2: Evaluation der Modellergebnisse	19
Literaturverzeichnis.....	23
Abbildungsverzeichnis.....	26
Tabellenverzeichnis	26

1 Einleitung

Der Einsatz von Künstlicher Intelligenz (im Folgenden „KI“) im Arbeitsumfeld kann die Bedeutung und Anforderungsprofile bestehender Berufe verändern. KI kann dabei Aufgabenbereiche automatisiert und autonom übernehmen, die zuvor Arbeitskräfte manuell erbrachten. Diesem eher negativen Beschäftigungseffekt stehen gleichzeitig positive Effekte gegenüber: es kommt aus Sicht der Unternehmen zu Einsparungen aufgrund des höheren Automatisierungsgrades und demnach zu Produktivitätszuwächsen. Demnach werden Kapazitäten in den Unternehmen frei, sodass beispielsweise eine höhere Nachfrage der Unternehmen nach Arbeitskräften in nicht-automatisierten Aufgabenbereichen entstehen kann (Acemoglu/Restrepo, 2018). Zudem kann KI bislang nicht-existierende Berufsbilder neu entstehen lassen. Dazu zählen beispielsweise Trainer für Tonalität und Bedeutung der Kundensprache, die KI-Systemen beibringen, über die wörtliche Bedeutung einer Kommunikation hinauszublicken und Sarkasmus zu erkennen. Ein anderes Beispiel sind Daten-Analysten, die als wichtiges Bindeglied zwischen technisch komplexen KI-Systemen und Entscheidungsträgern in Unternehmen stehen und nicht-technischen Fachleuten beispielsweise erklären können, warum ein KI-gesteuerter Fertigungsprozess gestoppt wird oder warum Werbekampagnen nur eine bestimmte Teilmenge an Konsumenten adressieren (Wilson, et al., 2017). Die folgende Analyse fokussiert sich dabei nicht auf eine Einschätzung, ob der Einsatz von KI eher positive oder negative Beschäftigungseffekte mit sich bringt, sondern auf eine Bestandaufnahme der Kompetenzen, die in den durch KI veränderten oder neu entstehenden Berufsbildern derzeit nachgefragt werden (im Folgenden „KI-Kompetenzen“).

Dabei vereinen die durch KI veränderten Anforderungsprofile der Berufe einen Trend: Der Weggang von Routine-Jobs hin zu Berufen, die einen hohen Grad an kognitiven Fähigkeiten voraussetzen sowie manuell durchgeführt werden müssen und deshalb nicht von KI übernommen werden können. Die Veränderung der Berufsbilder erfolgt dabei sehr dynamisch, wie eine Schätzung des World Economic Forum (2016) zeigt: etwa 65 Prozent der heute eingeschulten Kinder werden wahrscheinlich später in Berufen arbeiten, die heute noch gar nicht existieren. Definieren sich heutige Jobs meist über ihre Berufsbezeichnung, sind zukünftige Berufe eher losgelöst von einer festen Bezeichnung und definieren sich eher über ihre Kompetenzen, die sie flexibel in verschiedenen Berufsfeldern oder Branchen einsetzen können. Dazu gehören unter anderem Berufsfelder, in denen KI zunehmend zum Einsatz kommt und eine entscheidende Rolle spielt. Neben der Fähigkeit zur Problembewältigung haben die dafür benötigten Kompetenzen eine hohe Digitalaffinität, Kreativität sowie eine lebenslange Anpassungsfähigkeit und Bereitschaft zur Umschulung gemein (Deakin University/Ford Australia/Griffith University, 2019).

Treiber der Entwicklung sind nicht zuletzt die Unternehmensleitungen. Sie gestalten nicht nur unternehmensinterne Prozesse, sondern ebenfalls ihre Geschäftsmodelle zunehmend digital und datengesteuert (Fritsch/Krotova, 2020, 7). Im Zuge dessen entsteht eine wachsende Nachfrage nach geschultem Personal, welches technische Konzepte, wie Machine Learning, beherrschen, anwenden und weiterentwickeln kann. Dabei gilt eine verstärkte Nachfrage der Unternehmen nach KI-Kompetenzen auf dem Arbeitsmarkt als Zeichen für eine positive Entwicklung von KI in Deutschland (Demary et al., 2020). Doch welche Fähigkeiten und Kompetenzen sind auf Arbeitnehmerseite besonders wichtig, um effektiv in Berufen unter Einsatz von KI zu arbeiten? Um diese Frage fundiert beantworten zu können, wird in diesem Papier der Status quo der KI-Bedarfe der Wirtschaft – also der nachgefragten KI-Kompetenzen im Jahr 2020 - abgebildet, indem aktuelle Stellenanzeigen in für KI relevanten Berufen genauer analysiert werden (im Folgenden „KI-Stellenanzeigen“). Vor die-

sem Hintergrund kann die Analyse die bereits heute in Deutschland von Unternehmen nachgefragten KI-Kompetenzen identifizieren.¹ Um die Fülle an KI-Stellenanzeigen effektiv untersuchen zu können, werden für die Analyse selbst wissenschaftliche Techniken aus dem KI-Bereich verwendet.² Unter Zuhilfenahme dieser Techniken kann ein Großteil der KI-Stellenanzeigen vom vorher trainierten Algorithmus automatisiert analysiert werden. Somit können KI-Kompetenzen in Anforderungsprofilen von KI-Stellenanzeigen identifiziert werden, auf welche die ausschreibenden Unternehmen der deutschen Wirtschaft besonders viel Wert legen.

Dabei wird in den folgenden Schritten vorgegangen: Kapitel 2 beschreibt die zugrundeliegenden Daten und erläutert die verwendete Methodik zur Erfassung der KI-Kompetenzen. Darauf folgend befasst sich Kapitel 3 mit den Ergebnissen der Analyse.

2 Daten und Methode

Die folgende Analyse basiert auf einem Datensatz, der deutschlandweit die Volltexte von Stellenanzeigen aus dem ersten Quartal 2020 beinhaltet (siehe Abschnitt 2.1). Für die Identifikation der KIFähigkeiten sowie relevanter Abschlüsse (Studium oder Ausbildung), der zugehörigen Fachrichtungen und der geforderten Berufserfahrung werden die Volltexte von Stellenanzeigen mittels computer-linguistischer Modelle, sogenannter „Named Entity Recognition“ (NER), analysiert. Dafür wird zunächst eine zufällige Stichprobe des gesamten Datensatzes als Trainingsdatensatz deklariert und in einem weiteren Schritt die Volltexte der Stellenausschreibungen in dieser Stichprobe von menschlichen Codierern annotiert. Dieser Annotationsprozess umfasst die Markierung relevanter Wörter oder Wortkombinationen und die Zuweisung zu einer der vordefinierten Klassen (siehe Abschnitt 2.2).

2.1 Daten

In der Analyse wird ein Datensatz über 21.607 KI-Stellenanzeigen verwendet (Demary et al., 2020). Dazu wurde über ein iteratives und mehrfach validiertes Verfahren ein Wörterbuch mit Stellenbeschreibungen erstellt, mittels dessen KI-Berufe anhand der Bezeichnung der Stelle wie beispielsweise „Machine Learning Engineer“ oder „Softwareentwickler – Robotic Process Automation“ aus einem Korpus mit rund drei Millionen Stellenanzeigen identifiziert wurden (ebenda, 21).³ Alle Stellenanzeigen wurden im ersten Quartal 2020 veröffentlicht. Die Daten wurden kommerziell vom Anbieter Textkernel erworben, die umfassend die Stellenanzeigen des deutschen Online-Stellenmarktes scrapen – also relevante Daten aus den Stellenanzeigen extrahieren und speichern, um sie

¹ Die Ergebnisse der Analyse sollten auch für die Suche nach Arbeitskräften über andere Wege, wie beispielsweise Headhunter gelten, da nicht zu erwarten ist, dass sich die KI-Kompetenzanforderungen der Unternehmen bei verschiedenen Suchwegen systematisch unterscheiden.

² Auch wenn es generell möglich ist, nachgefragte KI-Kompetenzen in den Stellenanzeigen aller Berufe zu identifizieren, ist die nachfolgende Analyse auf die Kompetenzen in KI-Berufen begrenzt. Neben dem inhaltlichen Schwerpunkt des Papiers begründet sich diese Begrenzung vor allem mit Blick auf die verwendete Methodik: Da die Grundgesamtheit an KI-Berufen in den Textkernel-Daten bei deutlich unter einem Prozent liegt (siehe Demary et al. 2020, S. 21), würde die Erstellung eines Trainingsdatensatzes von 1.000 Stellenanzeigen aus dem Datensatz aller Berufe (siehe Kapitel 2.2) dazu führen, dass eine zu geringe Menge an Stellenanzeigen mit KI-Bezug identifiziert und damit die Modellschätzung nicht adäquat durchgeführt werden könnte.

³ Die Autoren erstellen hierfür zunächst unter Bezugnahme auf einschlägige Fachliteratur ein initiales Wörterbuch mit Berufsbezeichnungen von KI-Berufen. Darauf aufbauend wurden in den Volltexten der mit diesem Wörterbuch identifizierten Stellenanzeigen relevante Stichwörter identifiziert, mit Hilfe derer in einer weiteren Volltextsuche neue KI-Berufsbezeichnungen gefunden werden konnten. Das finale Wörterbuch mit KI-Berufsbezeichnungen „ist das Ergebnis mehrfacher Wiederholungen dieses Prozesses und wiederholter manueller Validierung“ (Demary et al. 2020, S. 21).

anschließend analysieren zu können. Dabei sammelt Textkernel die Daten nicht nur von Online-Stellenanzeigenportalen wie Indeed oder Stepstone, sondern auch direkt von über 60.000 Firmenkariereseiten und Zeitungen. Da viele Stellenanzeigen auf mehreren Portalen oder an verschiedenen Zeitpunkten geteilt werden können, dedupliziert Textkernel die Daten. Die 21.607 KI-Stellenanzeigen wurden in einem weiteren Schritt um diejenigen Stellenanzeigen bereinigt, die sich auf Studentenjobs oder Praktika beziehen.⁴ Der so entstandene finale Datensatz enthält 11.841 KI-Stellenanzeigen, die im ersten Quartal 2020 in Deutschland online geschaltet wurden.

Für die Anwendung von NER sind hochqualitative Trainingsdaten auf Grund des hohen Aufwandes bei der Generierung selten aber für ein gutes Ergebnis essenziell (Lample et al., 2016). Deshalb wird für die folgende Analyse ein Trainingsdatensatz erstellt und manuell annotiert. Es wird für das Trainieren von Machine-Learning-Modellen ein solcher Datensatz verwendet, der aus einer zufälligen Stichprobe des gesamten Datensatzes besteht und in drei Partitionen unterteilt wird: Trainingsdaten, Validierungsdaten und Testdaten.⁵ Der gesamte Trainingsdatensatz basiert auf einer Stichprobe von 1.000 aus dem gesamten Datensatz zufällig gezogenen KI-Stellenausschreibungen, die manuell annotiert wurde. Für die weitere Analyse wird dieser annotierte Trainingsdatensatz in 80 Prozent Trainingsdaten, 10 Prozent Validierungsdaten und 10 Prozent Testdaten unterteilt. Die Testdaten werden beim Training der NER-Modelle zunächst außen vorgelassen. Die finalen Modelle werden dann auf diese Testdaten angewendet, um eine valide Aussage darüber treffen zu können, wie gut die Modellvorhersagen mit fremden beziehungsweise neuen Daten sind.

2.2 Annotation der Daten

Der Annotationsprozess der Trainingsdaten umfasst die maschinenlesbare Markierung und kategoriale Zuordnung relevanter Wörter oder Wortkombinationen durch menschliche Codierer. Die Klassen, die in der nachfolgenden Analyse in den Trainingsdaten markiert und aus dem gesamten Datensatz extrahiert werden, sind in Tabelle 1 dargestellt.

⁴ Informelle Stellenbesetzungen über ehemalige Studierende und Praktikantinnen und Praktikanten werden damit nicht einbezogen. Es wird davon ausgegangen, dass sich die für diese Stellen benötigten KI-Kompetenzen nicht von den KI-Kompetenzen der in der Analyse berücksichtigten Stellen unterscheiden.

⁵ Je nach Machine-Learning-Modell kann ein Datensatz auch nur in zwei Partitionen, nämlich in Trainings- und Testdaten, unterteilt werden.

Tabelle 1: Beispiele und Anzahl der verwendeten Klassen im Trainingsdatensatz

Die Häufigkeit bezieht sich auf die absolute Häufigkeit der im Trainingsdatensatz (1.000 KI-Stellenanzeigen) verwendeten Begriffe, die der jeweiligen Klassen zugeordnet wurden

Klasse	Beispiele der Ausprägungen	Häufigkeit
Fähigkeiten	„R“, „Python“, „Datenbanken“, „Data Warehouse“, „Google Cloud“, „Azure“, „Machine Learning“, „SAP“	11.908
Abschluss	„Hochschulstudium“, „Bachelor“, „Master“, „Promotion“, „Ausbildung“	962
Fachrichtung	„Mathematik“, „Wirtschaftswissenschaften“, „Informatik“, „Wirtschaftsinformatik“	1.924
Berufserfahrung	„erste Berufserfahrung“, „fundierte Berufserfahrung“, „mehrjährige Berufserfahrung“	639

Quelle: Eigene Berechnung auf Basis von Textkernel-Daten, 2020

Die Klassen wurden vor Beginn der Analyse definiert. Unter „Fähigkeiten“ sollen alle Fähigkeiten oder Skills extrahiert werden, die von den ausschreibenden Unternehmen für die KI-Berufe nachgefragt werden. Der Fokus liegt dabei jedoch auf technischen Fähigkeiten. Soft Skills wie „Teamfähigkeit“, „Kommunikationsfähigkeit“, „Belastbarkeit“, etc. werden hier nicht mitefassen. Die Klasse „Abschluss“ umfasst alle angeforderten beruflichen Abschlüsse wie etwa ein absolviertes Hochschulstudium oder eine Ausbildung. Die zugehörigen Studien- bzw. Ausbildungsfächer sollen in der Klasse „Fachrichtung“ zusammengefasst werden. Zuletzt bildet „Berufserfahrung“ eine Klasse für die gewünschte Berufserfahrung des Bewerbers oder der Bewerberin.

Während der Annotation wurden zwei Codierer eingesetzt. Orientiert an Benikova et al. (2014) wurden für die Annotation entsprechende Guidelines erstellt, die iterativ überarbeitet wurden. Zudem wurden alle Annotationen durch eine weitere Person überprüft und eventuelle Konflikte beseitigt, um eine hohe Qualität der annotierten Daten zu gewährleisten. Darüber hinaus wurden für die Annotation der „Fähigkeiten“-Klasse mehrere Studien und Umfragen konsultiert, um die notwendigen Kenntnisse über die geforderten Fähigkeiten in KI-Berufen aufzubauen. Dazu zählen der Report über künstliche Intelligenz der Universität Stanford (Raymond et al., 2019, 238) und die jährliche Entwicklerumfrage von Stack Overflow (Stack Overflow, 2020), in welcher Entwickler Angaben dazu machen, welche Tools und Fähigkeiten sie zurzeit benutzen und welche sie gerne in Zukunft benutzen würden.

Für die Annotation der KI-Stellenanzeigen wurde das frei zugängliche Annotations-Tool „Doccano“ verwendet (Nakayama, 2018). Da mehrere Codierer beteiligt waren, wurde Doccano auf der Cloud-Plattform Heroku aufgesetzt. Doccano ermöglicht die zeitgleiche, kollaborative Annotation von Freitexten und bietet die Möglichkeit, die annotierten Daten in gängigen Formaten zu exportieren. Abbildung 1 zeigt einen manuell annotierten Beispieltext.

Abbildung 1: Beispiel einer annotierten KI-Stellenbeschreibung

Als Data Scientist beim Institut der deutschen Wirtschaft sollten Sie **R SKILL** oder **Python SKILL** beherrschen und Erfahrung in **Microsoft Azure SKILL** mitbringen. Darüber hinaus sollten Sie einen **Master Abschluss ABSCHLUSS** in **(Wirtschafts-)Informatik FÄCHER** oder **Mathematik FÄCHER** und **zwei Jahre Berufserfahrung BERUFSERFABUNG** in relevanten Berufen vorweisen können.

Quelle: Eigene Darstellung

Hier werden die Programmiersprachen „R“, „Python“ und „Microsoft Azure“ als Fähigkeit (Skill) markiert, der absolvierte „Master Abschluss“ als Abschluss und die Studienfächer „(Wirtschafts-) Informatik“ und „Mathematik“ als entsprechende „Fachrichtung“. Die „zwei Jahre Berufserfahrung“ werden zusammengenommen als „Berufserfahrung“ klassifiziert, um gegebenenfalls auch Aussagen über die zeitliche Ausdehnung der gewünschten Erfahrung machen zu können.

Die Verteilung der Klassen auf die Stellenanzeigen im Trainingsdatensatz ist in Tabelle 1 dargestellt. Es zeigt sich, dass „Fähigkeiten“ mit 11.908 annotierten Wörtern oder Wortkombinationen mit Abstand die häufigste vergebene Klasse darstellt, wobei „Fachrichtung“ (1.924), „Abschluss“ (962) und vor allem „Berufserfahrung“ (639) deutlich seltener vorkommen. Dieses Ungleichgewicht zwischen den Klassen ist damit zu erklären, dass in den meisten KI-Stellenausschreibungen eine Reihe von gewünschten Fähigkeiten der Bewerberin oder des Bewerbers aufgezählt werden, während meist nur ein Abschluss gefordert wird. Darüber hinaus werden Informationen über die gewünschte Berufserfahrung in vielen KI-Stellenanzeigen nicht explizit formuliert oder sind nicht eindeutig zu annotieren.

2.3 Identifikation der Klassen mittels Named Entity Recognition

NER ist ein Teilgebiet der Informationsextraktion, welches das Ziel hat, vordefinierte Entitäten in unstrukturierten und unbekanntem Texten zu identifizieren und in Klassen zu gruppieren (Awashty et al. 2020). Dabei können benannte Entitäten etwa die Namen von Personen, Orten, Organisationen oder Fähigkeiten wie im obigen Beispiel die Programmiersprachen R oder Python sein. In diesem Papier werden die Entitäten, im Folgenden Ausprägungen genannt, in die angegebenen Klassen „Fähigkeiten“, „Abschluss“, „Fachrichtung“ und „Berufserfahrung“ eingruppiert.

Für die Analyse wird die in Python implementierte Bibliothek spaCy benutzt. spaCy umfasst verschiedene Module für die Anwendung von natürlichen Sprachverarbeitungsmodellen, wie die Tokenisierung von Texten (die Unterteilung von Volltexten in einzelne Wörter oder Zeichenketten), Part-of-Speech-Tagging (Wortarterkennung) und auch Named Entity Recognition. Für die Bestimmung benannter Entitäten in unbekanntem Texten mittels NER nutzt spaCy ein neuronales Netz.⁶

Abbildung 2 zeigt die Vorhersage des NER-Modells bezüglich der Klassen für das in Abbildung 1 dargestellte Textbeispiel. Um die Vorteile von NER-Modellen gegenüber regelbasierten Ansätzen (z. B. die Stichwortsuche mit Wörterbüchern) zu verdeutlichen, wurde die vorher verwendete beispielhafte KI-Stellenbeschreibung leicht angepasst, bevor das Modell auf sie angewendet wurde und die Einteilung der Textteile in Klassen vornahm. Explizit ist die angepasste beispielhafte KI-

⁶ Für eine detaillierte Beschreibung der in der Analyse benutzten NER-Modelle und deren Konfiguration siehe Technischer Anhang 1. Alle NER-Modelle wurden unter der Python Version 3.7.9 und spaCy Version 2.3.2 erstellt.

Stellenanzeige nur um die erfundene Programmiersprache „Fiktiv-X“ erweitert. Das Modell identifiziert diese fiktive Programmiersprache korrekt als solche und ordnet sie der Klasse Fähigkeiten (Skills) zu, obwohl keine sonstige Stellenausschreibung einen äquivalenten Begriff enthält. Da NER-Modelle lernen, in welchem Kontext ein Wort benutzt wird, konnte darauf geschlossen werden, dass es sich beim Begriff „Fiktiv-X“ um eine Programmiersprache handeln muss. Das NER-Modell kann also gegenüber regelbasierten Ansätzen generalisieren, also auch gänzlich neue Begriffe identifizieren und der korrekten Klasse zuordnen.

Abbildung 2: Modellvorhersage der angepassten beispielhaften KI-Stellenbeschreibung mit fiktiver Programmiersprache



Quelle: Eigene Darstellung

Da die in spaCy implementierten NER-Modelle verschiedene Konfigurationsmöglichkeiten zulassen, wurden mehrere Modelle getestet und mit üblichen Metriken evaluiert. Die Modellergebnisse zeigen insgesamt eine gute Performance der Modelle mit Blick auf die Prognosegüte und die Qualitätskennzahlen entsprechen denen in anderen Kontexten angewandeter NER-Modelle.⁷ Die im folgenden Kapitel dargestellten Ergebnisse wurden unter Anwendung des NER-Modells mit den präzisesten Vorhersagen auf den gesamten Datensatz an Online-KI-Stellenanzeigen ermittelt.

Die Analyse erlaubt nun, für jede Klasse („Fähigkeiten“, „Abschluss“, „Fachrichtung“ und „Berufserfahrung“) einzeln die am häufigsten auftretenden Ausprägungen zu identifizieren. Für die deskriptive Analyse wurde der Datensatz über 11.841 KI-Stellenanzeigen weiter bereinigt. Da die Volltexte einiger Stellenausschreibungen nicht aus dem gesamten Text, sondern nur aus einer Zusammenfassung der eigentlichen KI-Stellenanzeige bestehen, sind hier teilweise keine der oben genannten Klassen enthalten und können somit auch nicht identifiziert und extrahiert werden. Deshalb wurden solche KI-Stellenausschreibungen für die Berechnung der relativen Häufigkeiten entfernt. Der finale Datensatz besteht dann aus 10.958 KI-Stellenausschreibungen. Aus verschiedenen Blickwinkeln kann somit ein Bild des Anforderungsprofils in KI-Berufen gezeichnet werden.

3 Ergebnisse

Im Folgenden werden nun die Ergebnisse der in Kapitel 2 erläuterten Analyse vorgestellt. Dabei fokussiert sich die Analyse zunächst auf die benötigten Abschlüsse, geht dann auf Fachrichtungen und Berufserfahrungen ein und untersucht abschließend die geforderten Kompetenzen der Bewerberinnen und Bewerber.

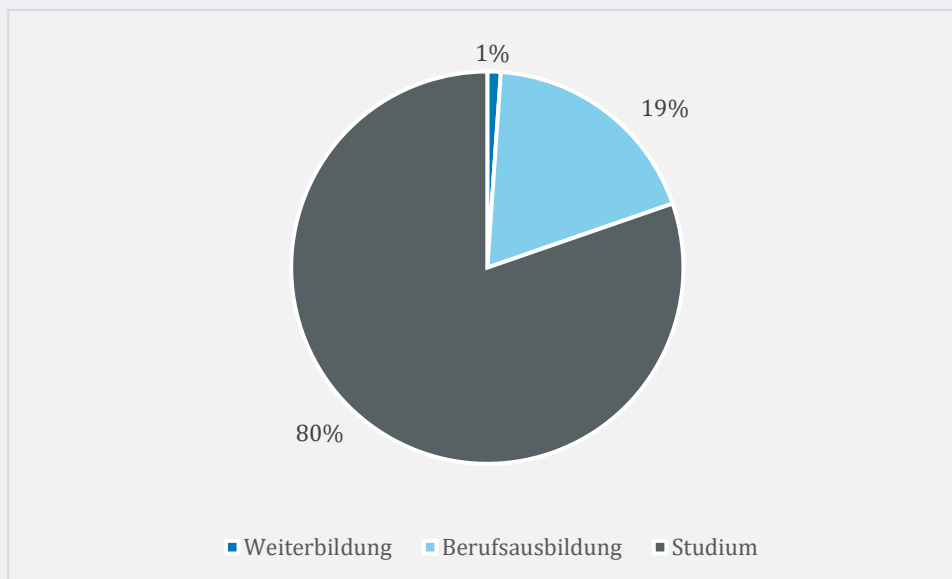
⁷ Siehe Technischer Anhang 2 für eine detaillierte Beschreibung der angewandten Evaluationsmetriken und eine ausführlichere Diskussion der Modellergebnisse.

3.1 Abschlüsse

Zunächst wird die Gesamtzahl aller in den KI-Stellenanzeigen genannten Abschlüsse genauer betrachtet. Dabei werden die einzelnen gesuchten Abschlüsse zunächst pro Stellenanzeige einer der Oberkategorien Studium, Berufsausbildung oder Weiterbildung zugeordnet.⁸ Existieren pro Stellenanzeige mehrere gesuchte Abschlüsse wie beispielsweise Bachelor oder Hochschulstudium, die inhaltlich der gleichen Oberkategorie Studium zuzuordnen sind, wird um diese Mehrfachnennungen bereinigt, sodass in der betrachteten KI-Stellenanzeige nur einmal die Abschlussvariante Studium gewertet wird. Jedoch kommen teilweise auch Studien- und Berufsabschlüsse in der gleichen Anzeige vor. In diesen Fällen werden in den betrachteten KI-Stellenanzeigen jeweils zwei verschiedene Abschlussvarianten erkannt und entsprechend gewertet. Insgesamt werden somit in den 7.001 KI-Stellenanzeigen, in denen auch tatsächlich mindestens eine Abschlussvariante gefordert wird, 8.334 gesuchte Abschlüsse der drei Varianten Studium, Berufsausbildung und Weiterbildung identifiziert (Abbildung 3). Demnach liegt der Fokus in den KI-Stellenanzeigen, die explizit mindestens einen Abschluss voraussetzen, nicht immer nur auf Abschlüssen einer Oberkategorie: im Durchschnitt werden gleichzeitig etwa 1,2 Abschlüsse der Varianten Studium, Berufsausbildung und Weiterbildung gefordert. Hauptsächlich betrifft dies die Kategorien Studium und Berufsausbildung, denn in 1.256 KI-Stellenanzeigen kommen sowohl ein Studien- als auch ein Berufsabschluss vor.

Abbildung 3: In KI-Stellenanzeigen gesuchte Abschlüsse nach Art des Abschlusses

Anteil der Abschlüsse an allen 8.334 identifizierten Abschlüssen in den 7.001 KI-Stellenanzeigen, in denen mindestens ein Abschluss erkannt wird, im ersten Quartal 2020 in Prozent



Quelle: Eigene Berechnung auf Basis von Textkernel-Daten, 2020

Bei 80 Prozent aller geforderter Abschlüsse wird ein Studienabschluss vorausgesetzt, gefolgt von 19 Prozent, in denen eine Berufsausbildung verlangt wird. Weiterbildungen bilden nur ein Prozent aller genannten Abschlüsse.

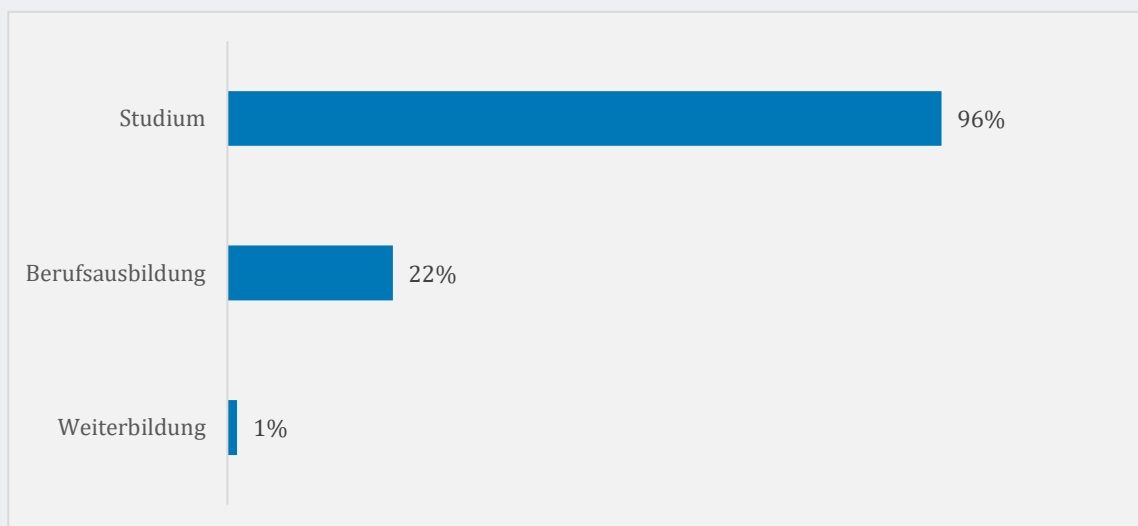
In einem zweiten Schritt wird anstelle der Gesamtzahl aller gesuchten Abschlüsse die KI-Stellenanzeigen betrachtet, in denen mindestens ein Abschluss erkannt wird. Somit wird genauer unter-

⁸ Im Zuge von Weiterbildungen können Arbeitnehmerinnen und Arbeitnehmer beispielsweise spezielle Zertifikate, Lizenzen oder Zertifizierungen erlangen (Burning Glass Technologies, 2017, 4).

sucht, welchen Stellenwert geforderte Abschlussniveaus im Gesamtkontext einnehmen. Es wird daher jeweils einzeln geprüft, in wie vielen der insgesamt 7.001 KI-Stellenanzeigen die Abschlussniveaus Studium, Berufsausbildung oder Weiterbildung gefordert werden (Abbildung 4). Infolgedessen zeigt sich die Bedeutung, die aus Arbeitgebersicht den Abschlussniveaus zugeschrieben wird.

Abbildung 4: Stellenwert von gesuchten Abschlüssen in KI-Stellenanzeigen

Anteil der KI-Stellenanzeigen, die den Abschluss voraussetzen, an den 7.001 KI-Stellenanzeigen, in denen mindestens ein Abschluss erkannt wird, im ersten Quartal 2020 in Prozent, Mehrfachnennungen



Quelle: Eigene Berechnung auf Basis von Textkernel-Daten, 2020

In 96 Prozent der KI-Stellenanzeigen, in denen mindestens ein Abschluss gefordert wird, bildet ein Studium demnach eine Grundvoraussetzung und macht somit eine deutliche Charakteristik der KI-Kompetenzen aus. Bezogen auf die Gesamtzahl aller 10.958 KI-Stellenanzeigen – also inklusive der KI-Stellenanzeigen, in denen kein Abschluss gefordert wird – fordern mit 61 Prozent mehr als die Hälfte aller KI-Stellenanzeigen explizit einen Studienabschluss. Jedoch wird nicht in allen KI-Stellenausschreibungen, in denen mindestens ein Abschluss erkannt wird, ausschließlich ein Studium verlangt. 22 Prozent der KI-Stellenausschreibungen suchen Bewerberinnen und Bewerber, die einen Berufsabschluss aufweisen. Somit werden auch Kompetenzen, die in einer Berufsausbildung erworben werden, im Kontext von KI-Berufen benötigt und nachgefragt. Nur in ein Prozent der KI-Stellenanzeigen mit geforderten Abschlüssen gilt der Nachweis einer relevanten Weiterbildung als hinreichende Qualifikation.

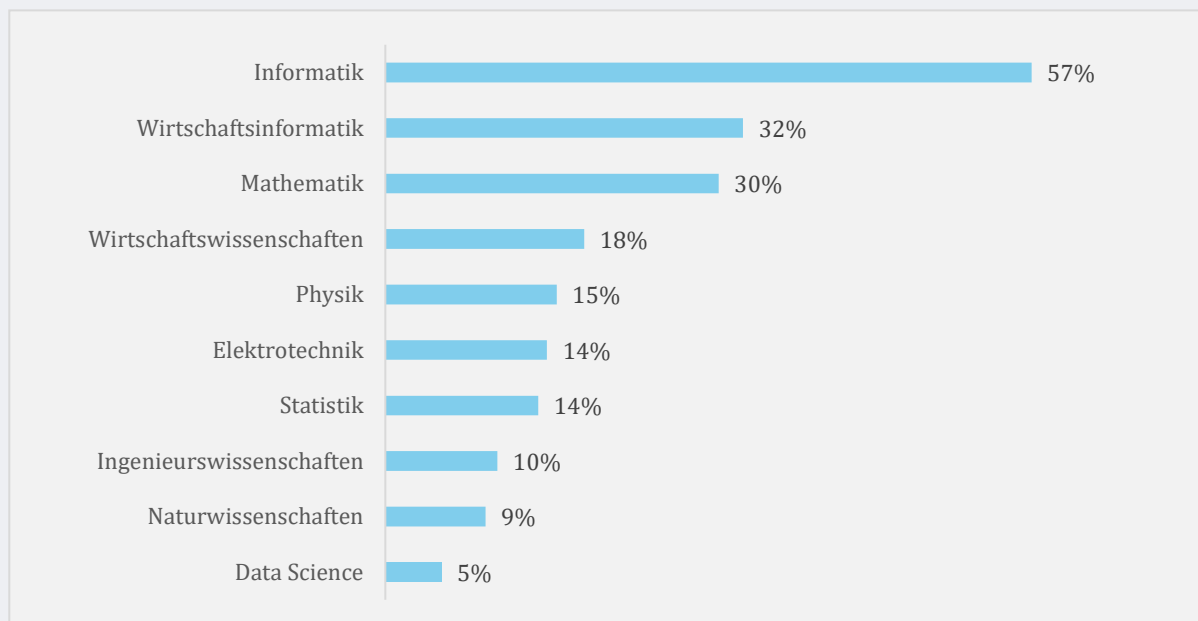
In rund einem Drittel aller KI-Stellenanzeigen wird keine Angabe zu einem expliziten Abschluss gemacht oder die Angabe lässt sich keiner der drei Abschlussvarianten zuordnen. Ein möglicher Erklärungsansatz hierfür könnte sein, dass aus Arbeitgebersicht der Stellenwert, der Abschlüssen zugeschrieben wird, relativ gering ist und vielmehr andere Qualifikationsmerkmale wie bereits gesammelte Berufserfahrung und demnach gesammeltes Wissen im KI-relevanten Bereich im Vordergrund stehen.

3.2 Fachrichtungen

Die Analyse der in den KI-Stellenanzeigen genannten Studienfächer zeigt, welche Fachrichtungen Unternehmen bei ihren gesuchten Bewerberinnen und Bewerbern präferieren (Abbildung 5).

Abbildung 5: In KI-Stellenanzeigen gesuchte Fachrichtungen

Anteil der KI-Stellenanzeigen, die die Fachrichtung voraussetzen, an den 6.784 KI-Stellenanzeigen, in denen mindestens ein Fach erkannt wird, im ersten Quartal 2020 in Prozent, Mehrfachnennungen



Quelle: Eigene Berechnung auf Basis von Textkernel-Daten, 2020

Mit einem Anteil von etwa 57 Prozent sind Absolvierende aus der Fachrichtung Informatik am gefragtesten. Der Informatikbereich schließt dabei auch die englischsprachige Bezeichnung „Computer Science“ mit ein. Rund 32 Prozent der KI-Stellenanzeigen, in denen mindestens eine Fachrichtung erkannt wird, sind für Bewerberinnen und Bewerber aus der eng verwandten und spezialisierteren Fachrichtung Wirtschaftsinformatik ausgeschrieben. Nur in etwa 62 Prozent aller KI-Stellenanzeigen werden Anforderungen an eine bestimmte Fachrichtung definiert. Bei Betrachtung der KI-Stellenanzeigen mit geforderter Berufsausbildung weichen die erforderlichen Fachrichtungen dabei nicht sonderlich ab. Fast ausschließlich treten die Abschlussalternativen Berufsausbildung und Studium innerhalb einer KI-Stellenanzeige in Kombination auf. Dabei wird eine vergleichbare Berufsausbildung meist als Alternative zu einem Studium aus einer der Fachrichtungen aus Abbildung 5 genannt. Inhaltlich kommt es daher zu keinen signifikanten Unterschieden. Zudem sind bestimmte Ausbildungstitel meist sehr spezifisch. Demnach kann es aus Unternehmenssicht sinnvoll sein, einen zu speziellen Fokus auf bestimmte Ausbildungsberufe zu vermeiden. Dies ist gerade vor dem Hintergrund bedeutend, dass KI-Kompetenzen relevant sind, die Bewerberinnen und Bewerber aus verschiedenen Ausbildungen im MINT-Bereich mitbringen können.

Allgemein zeigt sich, dass vorwiegend Bewerberinnen und Bewerber aus den MINT-Fachrichtungen, vor allem aus dem Informatik-Bereich, gesucht werden. Bezogen auf alle MINT-Fachrichtungen (Mathematik, Informatik, Naturwissenschaften, Technik) ist im Zeitraum von 2012 bis 2020 mit 87 Prozent der größten Beschäftigungszuwachs bei IT-Expertenberufen zu verzeichnen, die in

der Regel von Akademikern besetzt sind (Anger et al., 2020, 65). Das bislang gezeichnete Anforderungsprofil der KI-Kompetenzen von vorwiegend Studienabsolvierenden aus der Fachrichtung Informatik lässt sich in der Kategorie IT-Expertenberufe wiederfinden. Auch unter den fachlich ausgerichteten MINT-Berufen, zu denen in der Regel Ausbildungsberufe zählen, gab es im IT-Bereich im gleichen Zeitraum mit 52 Prozent das höchste Wachstum (ebenda, 66). Damit fragen Unternehmen mit KI-Stellenanzeigen Fachrichtungen nach, die in einem derzeit stark wachsenden Berufssegment angesiedelt sind. Es liegt demnach ein intensiver Wettbewerb um die entsprechenden Fachkräfte vor.

Jedoch zeigt die Analyse auch, dass in den KI-Stellenanzeigen oftmals nicht nur eine, sondern mehrere alternative Fachrichtungen angegeben werden. So führen die insgesamt 6.784 KI-Stellenanzeigen, die mindestens eine Fachrichtung angeben, im Durchschnitt etwa zwei der in Abbildung 5 genannten Fachrichtungen auf. Es besteht demnach kein Fokus auf eine Kernfachrichtung, sondern eine gewisse Flexibilität auf Arbeitgeberseite, Bewerberinnen und Bewerber aus mehreren alternativen Fachrichtungen zuzulassen. Für KI-Kompetenzen stehen also Fähigkeiten im Vordergrund, die vorwiegend in den MINT-Fachrichtungen erlernt werden, aber nicht zwangsläufig an eine der Fachrichtungen gebunden sind.⁹ Gerade diese Fähigkeiten ändern sich im Zeitverlauf jedoch auch schnell. Dies zeigt beispielsweise die generell höhere Dynamik in den Anforderungsprofilen der Berufe aus den MINT-Fachrichtungen: es werden schnell neue Kompetenzen verlangt und ältere werden obsolet (Deming/Noray, 2020). Infolgedessen werden eine gewisse Flexibilität und Anpassungsfähigkeit vorausgesetzt. Zudem haben alle MINT-Fachrichtungen eine gewisse Datenaffinität gemein.

Nur in fünf Prozent der KI-Stellenanzeigen, in denen mindestens eine Fachrichtung erkannt wird, suchen Unternehmen Bewerberinnen und Bewerber aus der Fachrichtung Data Science, der oftmals in engem inhaltlichen Zusammenhang zu KI gebracht wird oder teilweise als Voraussetzung gilt, um weitergehende Studiengänge mit Bezug zu KI zu studieren (Das Ingenieurstudium, 2020). Der niedrige Anteil kann unter anderem damit zusammenhängen, dass Data Science meist noch ein relativ junger und spezialisierter Studiengang ist, der erst wenige Absolvierende hervorgebracht hat (Studieren.at, 2020). Eine erneute Analyse in den kommenden Jahren wird zeigen, ob sich ein positiver Trend zur erhöhten Nachfrage nach Bewerberinnen und Bewerbern aus der Fachrichtung Data Science ergibt.

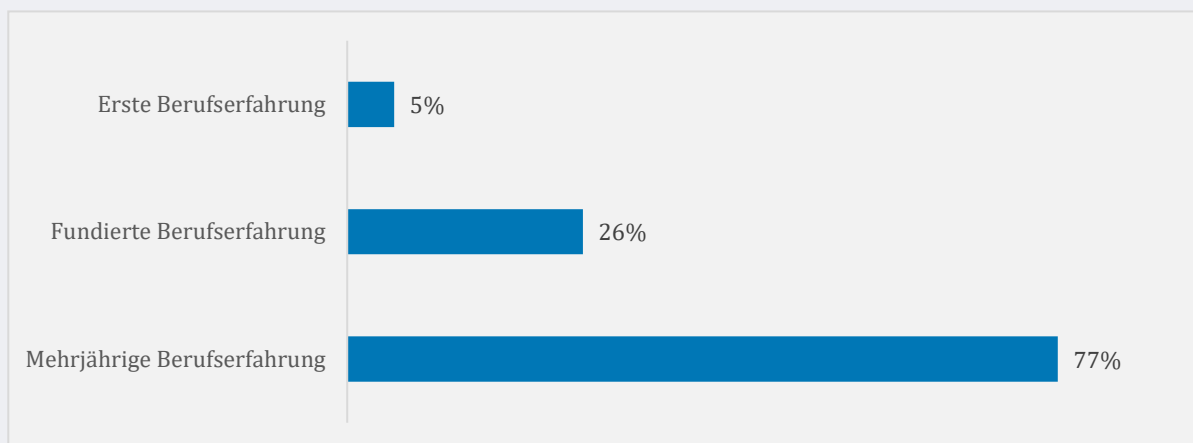
3.3 Berufserfahrung

Die Untersuchung der in den KI-Stellenanzeigen gewünschten Berufserfahrung zeigt, dass besonders berufserfahrene Bewerberinnen und Bewerber gesucht werden (Abbildung 6).

⁹ Die in den KI-Stellenanzeigen geforderten Fähigkeiten werden in Abschnitt 3.4 detaillierter thematisiert.

Abbildung 6: In KI-Stellenanzeigen geforderte Berufserfahrung

Anteil der KI-Stellenanzeigen, die diese Berufserfahrung voraussetzen, an den 5.730 KI-Stellenanzeigen, in denen mindestens eine Angabe zur Berufserfahrung erkannt wird, im ersten Quartal 2020 in Prozent, Mehrfachnennungen



Quelle: Eigene Berechnung auf Basis von Textkernel-Daten, 2020

Konkret fordern Arbeitgeber in 52 Prozent aller KI-Stellenanzeigen explizit Berufserfahrung. In den übrigen Fällen werden hierzu keine Angaben gemacht oder die Berufserfahrung kann nicht explizit zugeordnet werden. Berufserfahrung wird üblicherweise in drei Stufen unterteilt: erste Berufserfahrung, fundierte Berufserfahrung und mehrjährige Berufserfahrung (academics, 2017). An den drei Stufen orientiert sich die Gruppierung der Angaben zur Berufserfahrung in den KI-Stellenanzeigen in Abbildung 6.

Weisen die Bewerberinnen und Bewerber erste Berufserfahrungen auf, haben sie bereits einen strukturierten Arbeitsalltag kennengelernt, beispielweise durch Praktika oder als studentische Hilfskraft während des Studiums. Es handelt sich aber um Berufseinsteigerinnen und -einsteiger, die in der Regel noch nicht fest und mehrjährig in ihrem erlernten Berufsfeld nach Erlangung ihres Abschlusses beschäftigt waren. Explizit werden Berufseinsteiger nur in einem geringen Anteil von fünf Prozent der KI-Stellenanzeigen, in denen konkret Angaben zur Berufserfahrung gemacht werden, gesucht.

Fundierte Berufserfahrung besitzen Bewerberinnen und Bewerber, wenn sie nach ihrem Berufseinstieg bereits ein bis zwei Jahre Erfahrung in ihrem Berufsfeld gesammelt haben, wozu Praktika und Tätigkeiten in einem fachfremden Bereich nicht zählen (academics, 2017). Insgesamt werden in 26 Prozent der KI-Stellenanzeigen mit expliziten Angaben zur Berufserfahrung Arbeitnehmerinnen und Arbeitnehmer mit fundierter Berufserfahrung gesucht.

Gesammelte Berufserfahrung, die zeitlich hierüber hinausgeht – also ab drei Jahren Erfahrung im Bereich ihres Berufsfelds –, wird als mehrjährige Berufserfahrung definiert. In den KI-Stellenanzeigen, die auch tatsächlich Berufserfahrung fordern, werden mit einem Anteil von 77 Prozent vorwiegend Fachkräfte mit mehrjährigem Erfahrungslevel gesucht. Der hohe Anteil zeigt deutlich, dass derzeit zu den gewünschten KI-relevanten Qualifikationen meist nicht nur ein Studienabschluss, sondern auch gesammelte Erfahrungen zählen. Bewerberinnen und Bewerber sollten ihre Fähigkeiten also auch bereits angewendet und im Berufsleben unter Beweis gestellt haben. Ein möglicher Erklärungsansatz hierfür ist, dass Unternehmen neben den ausgeschriebenen KI-Stellen möglicherweise über unzureichend erfahrenes und geschultes Personal mit notwendigem Know-how

im KI-Bereich verfügen. Deshalb sehen sie sich nicht in der Lage, Berufseinsteiger adäquat bei ersten praktischen Anwendungen ihrer KI-Kompetenzen begleiten und unterstützen zu können. Stattdessen werden bereits erfahrende Fachkräfte gesucht, die ihre Expertise im Umgang mit KI unmittelbar und reibungsloser im Unternehmen einbringen können. Beispielweise fehlen rund 62 Prozent deutscher Unternehmen aus Branchen der Industrie und industrienahen Dienstleistungen das nötige Know-how für Big-Data-Analysen und demnach besitzen sie keine geeigneten personellen Ressourcen (Engels/Goecke, 2019, 46). Diese Tendenz ist auch bei kleinen und mittelständigen Unternehmen ausgeprägt, bei denen etwa 64 Prozent angeben, dass fehlendes Know-how die Erzielung von Mehrwert aus Daten verhindert (BMW, 2019). Unternehmen fragen jedoch hochqualifiziertes IT-Personal durchaus nach. Sogar in dem Maße, dass nicht nur ein intensiver Wettbewerb um diese Fachkräfte entsteht, sondern ebenfalls ein Nachfrageüberhang. Eine genauere Betrachtung der Fachkräftelücke in MINT-Berufen und vor allem in IT-Berufen, wie zum Beispiel Informatikern, zeigt: Fachkräfteengpässe in IT-Berufen sind seit 2014 deutlich angestiegen und erreichten im Oktober 2019 einen Wert von 52.100 offenen Stellen ohne passend qualifizierte Arbeitslose (Anger et al., 2020). Coronabedingt sank die IT-Lücke im Oktober 2020 auf 26.000, im Vergleich zu anderen MINT-Berufen fiel der Rückgang jedoch relativ gering aus. Vor diesem Hintergrund ist es ebenfalls nicht verwunderlich, dass der Anteil der IT-Lücke an der gesamten MINT-Fachkräftelücke von etwa 14 Prozent im Jahr 2014 kontinuierlich auf rund 24 Prozent im Jahr 2020 zunahm. Zudem ist der IT-Fachkräfteengpass am deutlichsten in IT-Expertenberufen ausgeprägt, also Beschäftigten mit höherem Bildungsabschluss, die auch in den KI-Stellenanzeigen hauptsächlich nachgefragt werden (vgl. Abbildung 4).

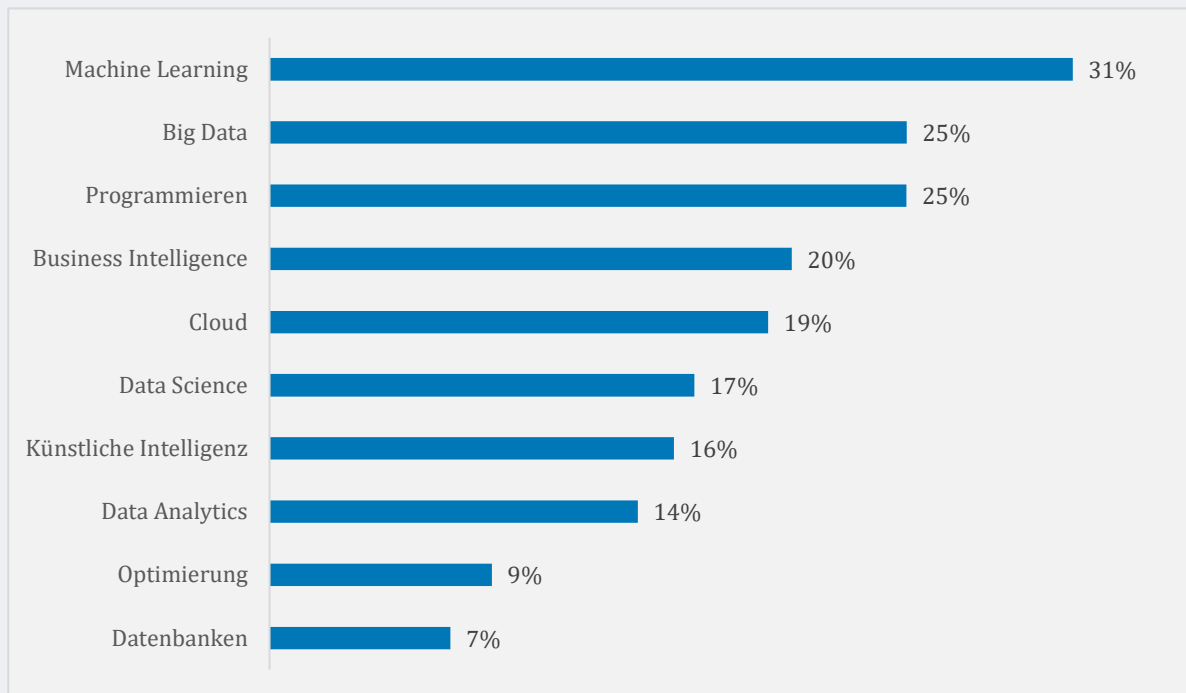
Interessant ist auch die Frage, ob in den KI-Stellenanzeigen eher allgemeine oder einschlägige Berufserfahrung – also spezifische praktische Kenntnisse aus genau dem Bereich oder der Branche der Stellenanzeige – nachgefragt werden. Nur in drei Prozent der 5.730 KI-Stellenanzeigen, die auch tatsächlich Berufserfahrung fordern, wird einschlägige Berufserfahrung verlangt. Dies zeigt deutlich den vielseitig ausgerichteten Charakter von KI-Kompetenzen: sie sind nicht an eine gewisse Branche gebunden, sondern definieren sich vielmehr über ihre Skills, die universell in verschiedenen Branchen oder Berufsfeldern eingesetzt werden können. Die Skills sollten sie allerdings bereits praktisch unter Beweis gestellt haben, wenngleich auch nicht zwangsläufig im gleichen Berufsumfeld. Gleichzeitig könnte der niedrige Anteil der geforderten einschlägigen Berufserfahrung auch eine Folge der angespannten Angebotssituation bei Fachkräften mit KI-Kompetenzen sein: Da diese knapp sind, vergrößern die nachfragenden Unternehmen unter Umständen auf diese Weise auch ihren Suchradius.

3.4 Fähigkeiten

Es schließt sich die Frage an, welche Fähigkeiten explizit von den Unternehmen nachgefragt werden. Im Rahmen der Analyse werden diese Kompetenzen zunächst inhaltlich in vier Segmente gruppiert (Konzepte, Programmiersprachen, Plattformen und Frameworks/Bibliotheken). In diesen werden dann jeweils einzeln die häufigsten Ausprägungen ausgewiesen. Dabei schreiben Unternehmen, die KI-Kompetenzen nachfragen, diesen Fähigkeiten generell eine enorme Bedeutung zu. In etwa 89 Prozent aller KI-Stellenanzeigen wird mindestens eine Fähigkeit gewünscht. In Abbildung 7 liegt der Fokus zunächst auf den Konzepten, die in den KI-Stellenanzeigen gewünscht sind.

Abbildung 7: In KI-Stellenanzeigen geforderte Kompetenz im Umgang mit Konzepten

Anteil der KI-Stellenanzeigen, die die Fähigkeit voraussetzen, an den 9.795 KI-Stellenanzeigen, in denen mindestens eine Fähigkeit erkannt wird, im ersten Quartal 2020 in Prozent, Mehrfachnennungen



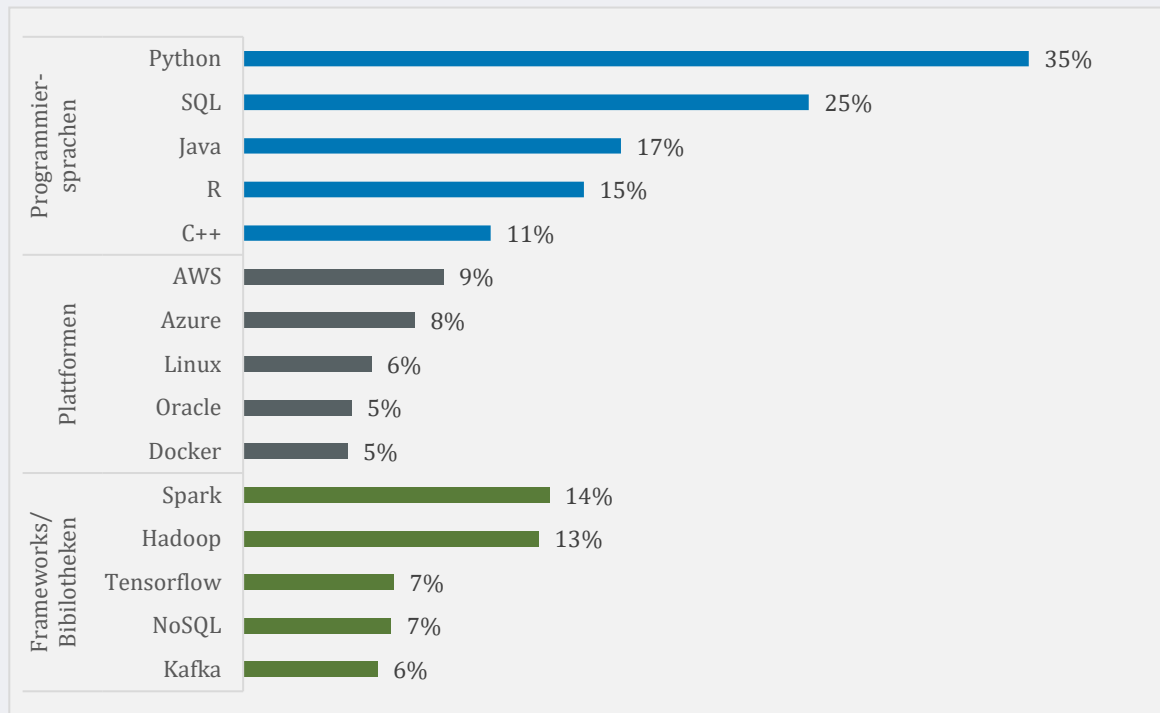
Quelle: Eigene Berechnung auf Basis von Textkernel-Daten, 2020

Dabei verlangen rund 31 Prozent der KI-Stellenanzeigen, in denen jeweils Angaben zu mindestens einer Fähigkeit erkannt werden, ein tiefgreifendes Verständnis von Machine-Learning, gefolgt von Big Data und einem generellen Programmierverständnis mit einem Anteil von jeweils 25 Prozent. Neben der Analyse der Fachrichtungen (vgl. Abbildung 5) tritt Data Science auch bei den Konzepten auf. Dabei erkennt das Modell, dass Data Science in den KI-Stellenanzeigen häufiger im Satzkontext von Fähigkeiten beziehungsweise Konzepten (17 Prozent) als im Kontext von Fachrichtungen (fünf Prozent) genannt wird und klassifiziert es entsprechend. Diese Erkenntnis bekräftigt die Vermutung, dass Data Science aus Arbeitgebersicht derzeit eher als Konzept denn als Fachrichtung wahrgenommen und explizit in KI-Stellenbeschreibungen gefordert wird.

Neben technischen Konzepten wird in den KI-Stellenanzeigen auch der Umgang mit expliziten Programmiersprachen, Plattformen sowie Frameworks – also dem Rahmen, innerhalb dessen der Programmierer eine Anwendung erstellt - und Bibliotheken verlangt (Abbildung 8).

Abbildung 8: In KI-Stellenanzeigen gefragte Kompetenzen im Umgang mit Programmiersprachen, Plattformen sowie Frameworks und Bibliotheken

Anteil der KI-Stellenanzeigen, die die Fähigkeit voraussetzen, an den 9.795 KI-Stellenanzeigen, in denen mindestens eine Fähigkeit erkannt wird, im ersten Quartal 2020 in Prozent, Mehrfachnennungen



Quelle: Eigene Berechnung auf Basis von Textkernel-Daten, 2020

Gemessen an den KI-Stellenanzeigen, in denen mindestens eine Fähigkeit erkannt wird, fallen Programmiersprachen wie Python (35 Prozent), SQL (25 Prozent) oder Java (17 Prozent) am bedeutendsten ins Gewicht. Diese Erkenntnis ist konsistent zur Analyse der Konzepte, denn auch hier nimmt die Programmierfähigkeit (25 Prozent) eine bedeutende Rolle ein. Konkrete Nennungen von Plattformen oder Frameworks und Bibliotheken sind seltener: Ein fundiertes Verständnis von und ein sicherer Umgang mit Plattformen wie AWS wird in neun Prozent der KI-Stellenanzeigen mit expliziten Angaben zu Fähigkeiten aufgeführt. Frameworks und Bibliotheken wie Spark werden in 14 Prozent der KI-Stellenanzeigen nachgefragt. Bedeutende Programmiersprachen zu beherrschen ist demzufolge ähnlich wichtig wie der Umgang mit bedeutenden KI-Konzepten. Dieses Ergebnis ist analog zu dem oben erläuterten Resultat, dass spezifische Berufserfahrung selten gefragt wird. Konkrete Anwendungen der Kenntnisse in einzelnen Plattformen, Frameworks und Bibliotheken stehen hinter eher allgemeinen Programmierkenntnissen in der Häufigkeit der Nennungen zurück.

4 Fazit

Zusammenfassend lässt sich ein typisches Anforderungsprofil der KI-Stellenanzeigen in Deutschland als Status quo der KI-Kompetenzen im ersten Quartal 2020 wie folgt charakterisieren:

- In vielen Fällen wird ein Studium und vor allem ein universitäres Studium gefordert. Ein Ausbildungsabschluss als Qualifikation wird weniger häufig angegeben, ist aber auch nicht unüblich.
- Studienabsolventinnen und -absolventen sollen vorwiegend aus dem Informatik-Bereich oder angrenzenden MINT-Fachrichtungen kommen. Dennoch ist auf Nachfragerseite eine Flexibilität festzustellen, denn meist wird nicht ein Abschluss einer bestimmten Fachrichtung gewünscht, sondern mehrere alternative Fachrichtungen aufgeführt.
- Es werden vorwiegend Bewerberinnen und Bewerber mit mehrjähriger Berufserfahrung gesucht. Demnach sollten sie ihre theoretisch angeeigneten Fähigkeiten auch bereits angewendet und im Berufsleben unter Beweis gestellt haben.
- Der Wunsch nach allgemeiner anstelle von spezifischer Berufserfahrung veranschaulicht die Vielseitigkeit von KI-Kompetenzen: die Fähigkeiten stehen im Vordergrund und können über verschiedene Branchen und Bereiche hinweg vielseitig eingesetzt werden.
- Dabei werden nicht nur Programmiersprachen wie Python abverlangt, sondern optimale Bewerbende müssen neben technischen Fähigkeiten auch ein Verständnis von KI-Konzepten vorweisen.

Auf Basis dieses derzeitigen Status quo der in KI-Stellenanzeigen nachgefragten KI-Kompetenzen werden zukünftig vergleichbare Analysen durchgeführt, um die Entwicklung der Anforderungsprofile in Deutschland zu beobachten.

Technischer Anhang

Technischer Anhang 1: Beschreibung des NER-Modells

In der Analyse von Online-KI-Stellenanzeigen wird die in Python implementierte Bibliothek für natürliche Sprachverarbeitung (NLP) spaCy benutzt. Die Bibliothek parst zunächst unstrukturierten Text in ein spaCy-Objekt. Ein Parser analysiert die eingegebenen Zeichenketten und bereitet diese so auf, dass diese maschinell weiterverarbeitet werden können. Das resultierende spaCy-Objekt kann dann in einer NLP-Pipeline weiterbearbeitet werden. Eine solche Pipeline umschreibt eine Reihe von Verarbeitungsschritten wie etwa die Tokenisierung von Texten (die Unterteilung von Textdaten in einzelne Wörter oder Zeichenketten), Part-of-Speech-Tagging (Wortarterkennung) und auch Named Entity Recognition. Hier wird zunächst ein Parser durchlaufen, der die Textdaten einliest und die Abhängigkeitsstruktur (die grammatikalische Abhängigkeit eines Wortes von anderen Wörtern) bestimmt. Dieser Parser orientiert sich an sogenannten „shift-reduce Parsern“ (siehe Lample et al. 2016). Darauf aufbauend nutzt spaCy ein vierstufiges Named Entity Recognition Framework, das auf neuronalen Netzen basiert und hier in seinen Grundzügen beschrieben werden soll:

(1) Zunächst werden sogenannte „word embeddings“ erstellt, die Wörter mathematisch in ihrem jeweiligen Kontext darstellen. Word embeddings werden meist dazu genutzt, um Wörter in numerische, multidimensionale Vektor-Repräsentation zu übertragen, um diese dann in verschiedenen Modellen weiter bearbeiten zu können. Dabei stellen die Dimensionen die Semantik eines jeden Wortes dar. Mit den resultierenden Wortvektoren können so z. B. arithmetische Berechnungen durchgeführt werden. Ein konkretes Beispiel wäre etwa, dass das Ergebnis der arithmetischen Operation des Wortvektors für „König“ minus des Wortvektors für „Mann“ plus des Vektors für „Frau“ den Wortvektor „Königin“ ergibt ($\text{König} - \text{Mann} + \text{Frau} = \text{Königin}$). Gegenüber traditionellen vor-trainierten word embeddings wie Word2Vec oder GloVe nutzt spaCy sogenannte „Bloom embeddings“ (Serrà/Karazoglou, 2017). Diese sind durch Komprimierungstechniken deutlich effizienter und schneller als herkömmliche Modelle.

(2) Der zweite Schritt umfasst die Kodierung von Wörtern in eine Satzmatrix. Dadurch kann auch der Kontext jedes Wortes im jeweiligen Satz berücksichtigt werden. Hierfür wird ein künstliches neuronales Netz – ein convolutional neural network (CNN) – verwendet.

(3) Der nächste Schritt bzw. die nächste Ebene des Modells umfasst die Entscheidung, welche Informationen wichtig für die Identifikation der vorgegebenen Klassen sind. Hier wird ein sogenannter „attention mechanism“ benutzt, der wie neuronale Netze generell versucht, das menschliche Gehirn zu imitieren, indem der Fokus selektiv auf wichtige Informationen gelenkt werden soll, während unwichtige Informationen ignoriert werden.

(4) Die oben erstellten Vektoren gehen in einem letzten Schritt in ein mehrlagiges Perzeptron (Neuronales Netz) ein. Das Ergebnis dieses Prozesses ist die Vorhersage bzw. die Bestimmung der relevanten Klassen. Als Ergebnis werden die Wörter bzw. Wortkombination im Volltext und die jeweilige Klassenzuordnung ausgegeben.

Innerhalb dieses NER-Modells gibt es die Möglichkeit, verschiedene Hyperparameter so zu setzen, dass die Ergebnisse bzw. die Vorhersagegenauigkeit optimiert wird. Hyperparameter sind Parameter, die im Trainingsalgorithmus verändert werden können und vor dem Training des Modells festgesetzt werden müssen. Ein wichtiger Hyperparameter ist dabei die sogenannte „dropout rate“. Diese gibt an, mit welcher Wahrscheinlichkeit (von 0 bis 100 Prozent) individuelle Wörter oder Repräsentationen während der verschiedenen Iterationen, die das Modell durchläuft, vergessen werden sollen. Ein höherer Wert macht es dem Modell schwieriger, sich an die Trainingsdaten zu „erinnern“ und verhindert somit das Overfitting des Modells. Overfitting geschieht dann, wenn das Modell zu sehr auf die Eigenheiten des Trainingsdatensatzes zugeschnitten ist. Dadurch können generellere Trends übersehen und neue Daten schlechter vorhergesagt werden. Der zweite wichtige Hyperparameter im NER-Modell ist die Mini-Batch-Größe, die in jeder Iteration zufällig ausgewählt werden. Durch Nutzung von Mini-Batch wird in jeder Iteration nicht der gesamte Trainingsdatensatz, sondern nur eine Teilmenge davon benutzt. Beide Hyperparameter wurden über eine Raster-suche optimiert. Die so ermittelten Werte betragen 0,4 für die dropout rate und eine Spannweite von 2 bis 128 für den Mini-Batch Parameter.

Um die Ergebnisse weiter zu optimieren, werden zudem zusätzliche Named Entity Recognition Modelle trainiert, die zunächst durch Vektoren vortrainierter Sprachmodelle initialisiert werden (Pre-Training). Hierfür wird, ähnlich wie in Googles BERT System, transfer learning verwendet, um domänenspezifische Spracheigenheiten besser erfassen zu können. Konkret bedeutet das, dass die gesamten Textdaten der KI-Stellenanzeigen (also nicht nur die annotierten Trainingsdaten) in das Modell eingegeben werden können, damit dieses die Eigenheiten der Sprache in KI-Stellenanzeigen besser verstehen und modellieren kann.

Technischer Anhang 2: Evaluation der Modellergebnisse

Für die Evaluation der Ergebnisse werden verschiedenen Evaluationsmetriken berechnet, die die Qualität des Modells insgesamt und für jede der vier Klassen darstellen. Diese Metriken umfassen Precision, Recall und den F1-Wert, die üblicherweise für die Bewertung von überwachten Klassifikationsalgorithmen verwendet werden und sich auf die Konfusionsmatrix (siehe Abbildung A - 9) beziehen.

Abbildung A - 9: Konfusionsmatrix

		Ermittelte Klasse	
		positiv	negativ
Tatsächliche Klasse	positiv	TP (richtig positiv)	FN (falsch negativ)
	negativ	FP (falsch positiv)	TN (richtig negativ)

Quelle: Eigene Darstellung

Alle Wörter oder Wortkombinationen, die von einem NER-Modell die korrekte Klasse zugeordnet bekommen, werden als „richtig positiv“ deklariert (TP). Wörter, die einer der Klassen aus Tabelle 1 zugeordnet wurden, aber eigentlich keiner oder einer anderen Klasse zugeordnet werden müssten, werden als „falsch positiv“ (FP) bezeichnet. Demgegenüber werden solche Wörter, denen das Modell keine Klassenzugehörigkeit attestiert, aber eigentlich eine eindeutige Klassenzugehörigkeit aufweisen als „falsch negativ“ (FN) angesehen. Wörter, denen richtigerweise keine Klasse zugeordnet wurde, bezeichnet man als „richtig negativ“ (TN).

Gegenüber traditionellen Klassifikationsaufgaben wie etwa der Markierung von Spam-E-mails (ist die E-Mail eine Spam-Nachricht oder nicht) gibt es bei der Bewertung von NER-Modellen augenscheinliche Grauzonen. Denn Wortkombinationen können *teilweise* korrekt einer Klasse zugeordnet werden. So kann etwa bei der Klassifikation der Wortkombination „mehrere Jahre Berufserfahrung“ nur der Teilausschnitt „Berufserfahrung“ der jeweiligen Klasse zugeordnet werden. Da hier die Länge der Zeichenkette und damit auch die Start- und Endposition der Zeichenfolge im Volltext nicht genau übereinstimmen, geht die Identifikation der Klasse als „falsch“ in die Berechnung der Metriken mit ein. Zudem können zusammengesetzte Wörter in der deutschen Sprache, die theoretisch zwei Klassen zugeordnet werden können, wie etwa „Informatikstudium“, durch das Modell in die jeweils falsche Klasse eingruppiert werden. Solche Fälle werden in der Evaluation von NER-Modellen üblicherweise als gänzlich falsche Klassifikation angesehen. Geht man also davon aus, dass solche Fälle einer Klassifikation in Grauzonen häufiger vorkommen, geben die unten angegebenen Evaluationsmetriken lediglich die Untergrenze, bzw. die pessimistische Schätzung der Qualität des jeweiligen Modells an.

Die Evaluationsmetrik der Precision (Spezifizität) bezeichnet alle korrekt vorhergesagten Einheiten, die genau der Zeichenlänge der handkodierten Zuweisung entsprechen, gemittelt über alle vorhergesagten Einheiten. Somit wird die Precision basierend auf der Konfusionsmatrix berechnet durch:

$$Precision = \frac{TP}{TP + FP}$$

Wie der Name sagt, bezeichnet Precision wie präzise bzw. akkurat das Modell in den Vorhersagen ist. Demgegenüber wird die Evaluationsmetrik der Recall (Sensitivität) wie folgt berechnet:

$$Recall = \frac{TP}{TP + FN}$$

Recall gibt den Anteil aller korrekt klassifizierten Wörter bzw. Wortkombinationen an der Gesamtheit aller zu klassifizierenden Einheiten an und somit den Anteil der erkannten Klassen an allen Einheiten die theoretisch erkannt werden müssten.

Zuletzt beschreibt der F1-Wert das harmonische Mittel der beiden Werte von Precision und Recall:

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Auf einer Skala von 0 (wenn entweder Precision oder Recall einen Wert von 0 haben) bis +1 (wenn Precision und Recall beide einen Wert von 1 haben) gibt der Wert also den Genauigkeitsgrad des Modells insgesamt an.

Tabelle A - 2 zeigt die Metriken Precision, Recall und den F1-Wert für das jeweils durch Rasteroptimierung ermittelte beste Modell mit und ohne Pre-Training. Es zeigt sich, dass die Modelle mit Pre-Training besser performen als die Modelle ohne. Vor allem der Recall-Wert, also der Anteil der annotierten Wörter und Wortkombinationen, der korrekt erfasst wird, ist durch vorausgegangenes transfer learning deutlich höher. Für die Analyse wird deshalb das beste Modell mit Pre-Training genutzt.

Tabelle A - 2: Evaluationsmetriken der besten NER Modelle

	Ohne Pre-Training	Mit Pre-Training
<i>Precision</i>	73,57	74,46
<i>Recall</i>	71,33	75,16
<i>F1</i>	72,43	74,81

Quelle: Eigene Berechnung

Die Qualität der Modelle insgesamt einzuordnen fällt ohne weitere Vergleiche allerdings schwer. Eine erste Einordnung gelingt durch die Betrachtung von kontextspezifischen NER-Modellen aus anderen Bereichen. So erreichen Tarcar et al. (2020) für die Extraktion medizinischer Begriffe einen F1-Wert von 75,7. Hierfür wurde allerdings ein mit medizinischen Texten vortrainiertes Modell benutzt. Für das Modell ohne Pre-Training beträgt der F1-Wert lediglich 67,9. Für die Identifikation der typischen NER-Klassen „Person“ und „Ort“ in historischen deutschen Texten berichten Labusch et al. (2019) F1-Werte zwischen 50 und 60. Bei oft genutzten Datensätzen für die Evaluation verschiedener NER-Modelle hingegen werden teilweise F1-Werte von über 90 erreicht (siehe etwa Akbik et al., 2019). Es zeigt sich also, dass die Spannweite möglicher Evaluationswerte sehr groß ist und die Bewertung eines Modells immer vom jeweiligen Kontext abhängt. Insgesamt kann aber konstatiert werden, dass der in Tabelle A - 2 berichtete F1-Wert von 74,81 in eine sehr gute Richtung geht, zumal hiermit, wie oben diskutiert, lediglich die pessimistischste Schätzung der Modellqualität ausgedrückt wird.

Da die Evaluationsmetriken teilweise stark zwischen verschiedenen Klassen variieren können, werden in Tabelle A - 3 alle Metriken zudem pro Klasse angegeben.

Tabelle A - 3: Evaluationsmetriken nach Klasse

	Abschluss	Fachrichtungen	Berufserfah- rung	Fähigkeiten
<i>Precision</i>	79,71	81,89	68,97	73,51
<i>Recall</i>	78,57	82,54	78,43	73,86
<i>F1</i>	79,14	82,21	73,40	73,68

Angegeben sind die Werte für das beste NER-Modell mit Pre-Training

Quelle: Eigene Berechnung

Die in Tabelle A - 3 dargestellten Metriken pro Klasse machen deutlich, dass es hier große Varianz zwischen der Identifikation und Zuweisung annotierter Begriffe gibt. Während Abschlüsse (F1 = 79,14) und dazugehörige Fachrichtungen (F1 = 82,21) generell sehr gut erkannt werden, fällt die Identifikation von Fähigkeiten (F1 = 73,68) und vor allem auch von Berufserfahrung (F1 = 73,40 mit einer Precision von 68,97) schwerer. Diese Diskrepanz ist zum einen durch die deutlich höhere Homogenität in den Klassen Abschlüsse und Fachrichtungen gegenüber der Fähigkeiten-Klasse zu erklären. So finden sich in den KI-Stellenanzeigen deutlich weniger verschiedene Begriffe für mögliche Abschlüsse und Fachrichtungen. Bei Fähigkeiten hingegen gibt es eine deutlich größere begriffliche Vielfalt. Einige Fähigkeiten kommen dabei in nur sehr wenigen, spezialisierten KI-Stellenausschreibungen vor und können so wesentlich schwerer korrekt identifiziert werden. Darüber hinaus ist ein Problem bei der Klasse Berufserfahrung, dass es generell recht wenig annotierte Entitäten gibt.

Literaturverzeichnis

Acemoglu, Daron / Restrepo, Pascual, 2018, Artificial intelligence, automation and work, Working Paper No. 24196, National Bureau of Economic Research

Anger, Christina / Kohlisch, Enno / Koppel, Oliver / Plünnecke, Axel, 2020, MINT-Herbstreport 2020. MINT-Engpässe und Corona-Pandemie: kurzfristige Effekte und langfristige Herausforderungen, Gutachten für BDA, BDI, MINT Zukunft schaffen und Gesamtmetall, Köln

academics, 2017, Was zählt als Berufserfahrung? Ausbildung, Praktikum, Promotion – Berufserfahrung? <https://www.academics.de/ratgeber/bewerbung-was-zaehlt-als-berufserfahrung> [23.11.2020]

Akbik, Alan / Bergmann, Tanja / Vollgraf, Roland, 2019, Pooled Contextualized Embeddings for Named Entity Recognition, in: Proceedings of the NAACL-HLT 2019, S. 724-728

Awasthy, Parul / Moon, Taesun / Ni, Jian / Florian, Radu, 2020, Cascaded Models for Better Fine-Grained Named Entity Recognition, arXiv preprint, <https://arxiv.org/abs/2009.07317> [9.11.2020]

Benikova, Darina / Biemann, Chris / Rezniek, Marc, 2014, NoSta-D Named Entity Annotation for German: Guidelines and Dataset, in: Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), S. 2524-2531

BMWI - Bundesministerium für Wirtschaft und Energie, 2019, Mehrwerte aus Daten – Potenziale und Handlungsoptionen für den Mittelstand, <https://www.mittelstand-digital.de/MD/Redaktion/DE/Publikationen/mehrwerte-aus-daten.html> [25.11.2020]

Bonin, Holger / Gregory, Terry / Zierahn, Ulrich, 2015, Übertragung der Studie von Frey/Osborne (2013) auf Deutschland, ZEW, Kurzexpertise Nr. 57, Mannheim

Burning Glass Technologies, 2017, The Narrow Ladder: The Value of Industry Certifications in the Job Market, <https://www.burning-glass.com/research-project/certifications> [12.01.2020]

Das Ingenieurstudium, 2020, Künstliche Intelligenz Studium, <https://www.ingenieurwesen-studieren.de/studiengaenge/kuenstliche-intelligenz/> [7.12.2020]

Demary, Vera / Engels, Barbara / Goecke, Henry / Koppel, Oliver / Armin Mertens / Christian Rusche / Marc Scheufen / Jan Wendt, 2020, KI-Monitor 2020. Status quo der Künstlichen Intelligenz in Deutschland, Gutachten im Auftrag des Bundesverbands Digitale Wirtschaft (BVDW) e.V., Köln

Deming, David J. / Noray, Kadeem, 2020, Earnings dynamics, changing job skills, and STEM careers. *The Quarterly Journal of Economics*, 135(4), 1965-2005

Engels, Barbara / Goecke, Henry, 2019, Big Data in Wirtschaft und Wissenschaft, IW-Analysen, Nr. 130, Köln

Frey, Carl Benedikt / Osborne, Michael A., 2017, The Future of Employment: How Susceptible are Jobs to Computerization?, in: *Technological Forecasting and Social Change*, Nr. 114, S. 254–280

Gregory, Terry / Salomons, Anna / Zierahn, Ulrich, 2019, Racing with or Against the Machine? Evidence from Europe, IZA Discussion Paper Nr. 12063, Bonn

Kaspersky, 2020, Schöne neue Jobwelt? Arbeit 5.0: Wie die Deutschen unter 31 den Einsatz Künstlicher Intelligenz in der Arbeitswelt von morgen beurteilen, Teil 5 der Kaspersky-Report-Serie über die Generation KI, https://media.kasperskydaily.com/wp-content/uploads/sites/96/2020/09/07085315/K_Kurzreport_GenerationKI_5_Neue_Jobwelt_final-1.pdf [24.11.2020]

Labusch, Kai / Neudecker, Clemens / Zellhöfer, David, 2019, BERT for Named Entity Recognition in Contemporary and Historical German, in: *Proceedings of the 15th Conference on Natural Language Processing (KONVENS 2019)*

Lample, Guillaume / Ballesteros, Miguel / Subramanian, Sandeep / Kawakmi, Kazuya / Dyer, Chris, 2016, Neural Architectures for Named Entity Recognition, in: *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, S. 260-270

Ford Australia, Deakin University, Griffith University, 2019, 100 Jobs of the future, <https://100jobsofthefuture.com/report/> [24.9.2020]

Fritsch, Manuel / Krotova, Alevtina, 2020, Wie datengetrieben sind Geschäftsmodelle in Deutschland. Analyse des Status quo, IW-Report, Nr. 9, Köln

Nakayama, Hiroki / Kubo, Takahiro / Kamura, Junya / Taniguchi, Yasufumi / Liang, Xu, 2018, doccano: Text Annotation Tool for Human, <https://github.com/doccano/doccano> [9.11.2020]

Raymond, Perrault / Shoham, Yoav / Bryniolfsson, Erik / Clark, Jack / Etchemendy, John / Grosz, Barbara / Lyons, Terah / Manyika, James / Niebles, Juan Carlos / Mishra, Saurabh, 2019, Artificial Intelligence Index Report 2019, https://hai.stanford.edu/sites/default/files/ai_index_2019_report.pdf [9.11.2020]

Serrà, Joan / Karatzoglou, Alexandros, 2017, Getting deep recommenders fit: Bloom embeddings for sparse binary input/output networks, in: Proceedings of the eleventh ACM conference on recommender systems, ACM, S. 279-287

Stack Overflow (2020) Stack Overflow Developer Survey <https://insights.stackoverflow.com/survey/2020> [9.11.2020]

Studieren.at, 2020, Data Science Studium, <https://www.studieren.at/studien/data-science/> [7.12.2020]

Tarcar, Amogh Kamat / Tiwari, Aashis / Rao, Dattaraj / Dhaimodker, Vineet Naique / Rebelo, Penjo / Desai, Rahul, 2020, Empirical Evaluation of Healthcare NER Model Performance with Limited Training Data, in: Proceedings of Health Search and Data Mining Workshop (HSDM 2020) in the 13th ACM International WSDM Conference (WSDM 2020). ACM, Houston, TX, USA

Textkernel, 2020, Textkernel Jobdatenbank, <https://www.textkernel.com/de/solution/jobfeed/> [30.11.2020]

Udemy, 2019, 2020 Workplace Learning Trends Report: The Skills of the Future, https://humanresources.report/Resources/Whitepapers/LP/1ae5558a-ed87-44cd-b64b-35c5d3bd8606_2020_Workplace_Learning_Trends_Report.pdf [24.9.2020]

Wilson, H. James / Daugherty, Paul R. / Morini-Bianzino, Nicola, 2017, The Jobs That Artificial Intelligence Will Create. MIT Sloan Management Review, 58(4): 14-16.

World Economic Forum, 2016, Global Challenge Insight Report: The Future of Jobs Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution, http://www3.weforum.org/docs/WEF_Future_of_Jobs.pdf [24.9.2020]

Abbildungsverzeichnis

<i>Abbildung 1: Beispiel einer annotierten KI-Stellenbeschreibung</i>	7
<i>Abbildung 2: Modellvorhersage der angepassten beispielhaften KI-Stellenbeschreibung mit fiktiver Programmiersprache</i>	8
<i>Abbildung 3: In KI-Stellenanzeigen gesuchte Abschlüsse nach Art des Abschlusses</i>	9
<i>Abbildung 4: Stellenwert von gesuchten Abschlüssen in KI-Stellenanzeigen</i>	10
<i>Abbildung 5: In KI-Stellenanzeigen gesuchte Fachrichtungen</i>	11
<i>Abbildung 6: In KI-Stellenanzeigen geforderte Berufserfahrung</i>	13
<i>Abbildung 7: In KI-Stellenanzeigen geforderte Kompetenz im Umgang mit Konzepten</i>	15
<i>Abbildung 8: In KI-Stellenanzeigen gefragte Kompetenzen im Umgang mit Programmiersprachen, Plattformen sowie Frameworks und Bibliotheken</i>	16
<i>Abbildung A - 9: Konfusionsmatrix</i>	19

Tabellenverzeichnis

<i>Tabelle 1: Beispiele und Anzahl der verwendeten Klassen im Trainingsdatensatz</i>	6
<i>Tabelle A - 2: Evaluationsmetriken der besten NER Modelle</i>	21
<i>Tabelle A - 3: Evaluationsmetriken nach Klasse</i>	22