

KI-Transfer BW: Praxispilot

BTB Betriebswirtschaftliche und

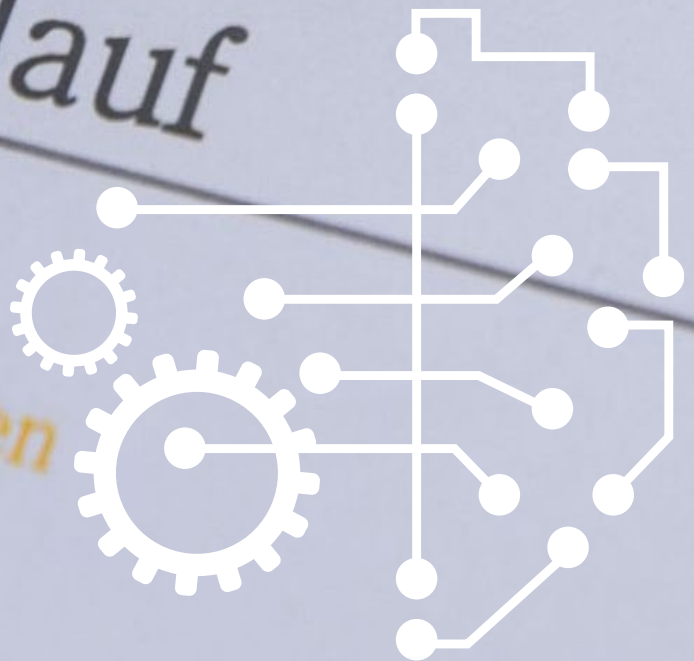
-technische Beratungs-GmbH

KI im HR: CV-Analyzer

Lebenslauf

Persönliche Daten

Geburtsdatum / -ort:
Staatsangehörigkeit:



Transferdokumentation für den Praxispiloten: KI im HR: CV-Analyzer

Beteiligte Partner

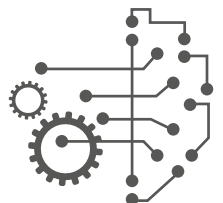
- BTB Betriebswirtschaftliche und -technische Beratungs-GmbH
- FZI Forschungszentrum Informatik

Autoren

- Alexander Dregger, FZI Forschungszentrum Informatik
- Oliver Denninger, FZI Forschungszentrum Informatik
- Ljubo Barac, BTB GmbH
- Kathrin Mayer, BTB GmbH
- Lukas Pesch, BTB GmbH
- Metin Dereli, BTB GmbH
- Oliver Steck, BTB GmbH
- Nico Krause, BTB GmbH

Ansprechpartner

- Ljubo Barac, Managing Consultant bei BTB
- Alexander Dregger, wissenschaftlicher Mitarbeiter am FZI Forschungszentrum Informatik
- Oliver Denninger, Bereichsleiter Software Engineering am FZI Forschungszentrum Informatik



Inhalt

1.	Management Summary	4
1.1.	Was ist ein Praxispilot?	4
1.2.	Kurzzusammenfassung des Praxispiloten mit BTB	4
2.	Einführung	6
2.1.	Ausgangssituation und Motivation	6
2.2.	Ziele und Nutzen der Zielgruppen	6
2.3.	Projektrahmen	7
3.	Inhaltliches	9
3.1.	Funktionale Anforderungen	9
3.2.	IT-Architektur	9
3.3.	Daten und Algorithmen	11
3.3.1.	Daten	11
3.3.2.	Datenaufbereitung	11
3.3.3.	Erstellen der Zielvariable	11
3.3.4.	Auswahl und Training des KI-Algorithmus	12
3.3.5.	Warum wird KI benötigt	12
3.3.6.	Protoypische Umsetzung des CV-Analyzers	13
3.3.7.	Evaluation	13
3.4.	Herausforderungen	14
4.	Resümee	15
4.1.	Wissensgewinn	15
4.2.	Ausblick	16
5.	Impressum	17
5.1.	Kontaktdaten	17
5.1.1.	Unternehmen	17
5.1.2.	KI-Transfer BW	17
5.1.3.	Fördergeber	17
5.2.	Bildnachweis	17
5.3.	Gestaltung und Umsetzung	17

1. Management Summary

1.1. Was ist ein Praxispilot?

Praxispiloten innerhalb des Projekts KI-Transfer BW sind kleine Projekte zwischen einem Unternehmen und KI-Transfer BW Experten, um einen konkreten Aspekt der Anwendung von KI-Lösungen zur Weiterentwicklung des Unternehmens zu untersuchen und erste Lösungsvorschläge zu erarbeiten. Im Rahmen der Praxispiloten werden, die in dem Projekt KI-Transfer BW entwickelten, vielfältigen Methoden, Werkzeuge und Vorgehensweisen in der betrieblichen Praxis getestet und so auch Weiterentwicklungspotenziale identifiziert.

Die in den Praxispiloten gewonnenen Ergebnisse und Erkenntnisse werden vom KI-Transfer BW Team aufgenommen, dokumentiert und im Rahmen des Wissenstransfers verbreitet. Dies soll bei der Implementierung von KI in anderen KMU unterstützen. Dazu dient auch die vorliegende Dokumentation des Praxispiloten, der beim Unternehmen BTB Betriebswirtschaftliche und -technische Beratungs-GmbH in Zusammenarbeit mit dem FZI Forschungszentrum Informatik umgesetzt worden ist.

1.2. Kurzzusammenfassung des Praxispiloten mit BTB

Das Unternehmen BTB Betriebswirtschaftliche und -technische Beratungs-GmbH (im Folgenden BTB genannt) ist ein Anbieter von individueller Business-Software und von Sprachassistenten. Kernkompetenz von BTB ist die Entwicklung von individuellen Softwarelösungen für Unternehmen. Daneben bietet BTB auch eigene Software an und arbeitet im Bereich Personaldienstleistungen mit dem Unternehmen conneXion Network GmbH zusammen.

Im Sommer 2021 nach einem Vortrag von KI-Transfer BW auf dem Open Innovation Kongress kontaktierte das Unternehmen das das Projektteam von KI-Transfer BW, um sich über eine mögliche Zusammenarbeit im Bereich KI auszutauschen. Im gemeinsamen Austausch zeigte sich, dass sich der Themenbereich Human Resources (HR) für einen Praxispiloten eignen könnte. Das Unternehmen plant nämlich hierzu zeitnah ein Ausbildungsprojekt, das von Auszubildenden im Bereich der Fachinformatik betreut wird. In diesem Projekt sollte eine Lösung entwickelt werden, die mithilfe eines Online-Tools die Analyse von Lebenslauf- bzw. CV-Daten (Abkürzung für Curricu-

öffentliche Förderung durch das Ministerium für Wirtschaft, Arbeit und Tourismus Baden-Württemberg

Anbieter für individuelle Business-Software und Sprachassistenten

KI im HR

lum Vitae) erlaubt, um die Arbeit von Personalern zu erleichtern. Hierbei wünschte sich das Unternehmen, dass KI zum Einsatz kommen soll, um das strategische Know-how zu erweitern. In diesem Kontext wurde das FZI Forschungszentrum Informatik in die Entwicklung eingebunden, um Expertise im Themenfeld KI einzubringen. Das Ziel des Praxispiloten war es, den Use Case zu spezifizieren und diesen im Rahmen eines ersten Proof-of-Concept prototypisch umzusetzen.

U-NET
Machine Learning
KI basierte Auswertung
Lebensläufe
Bildsegmentierung
GPT2 Ausbildungsprojekt
Datenqualität
Natural Language Processing

Unternehmen	BTB Betriebswirtschaftliche und -technische Beratungs-GmbH
Branche	Softwareentwicklung
Unternehmensbereich	HR
Unternehmensgröße	50 Mitarbeitende
Technologie und Methoden	Machine Learning, Semantische Segmentierung, Natural Language Processing
KI-Lösung	CV-Analyzer
Ziel	Erhöhung des Komforts Zeitersparnis und evtl. bessere Qualität bei der Dokumentation von Prüfergebnissen
Dauer	3 Monate
Personenzahlen (beteiligte Personengruppen)	9 Mitarbeitende

2. Einführung

2.1. Ausgangssituation und Motivation

Ausgangspunkt des Praxispiloten war ein Kontakt, der zwischen dem KI-Transfer BW Team und BTB im Nachgang des Open Innovation Kongresses 2021 entstanden ist. Das Unternehmen war dabei insbesondere interessiert mithilfe von konkreten Projekten strategisch das eigene KI-Know-how zu erweitern. Hierzu wurden auch schon erste Initiativen bei BTB angestoßen, jedoch wünschte sich das Unternehmen dies mit weiteren Projekten zu intensivieren.

Im Rahmen eines geplanten Ausbildungsprojektes sollte ein CV-Analyzer gebaut werden, der zunächst intern bei BTB eingesetzt wird. Der CV-Analyzer hatte dabei folgende Ziele:

- Entwicklung einer web-basierten Oberfläche, um CVs einzulesen
- Prüfung des CVs auf wesentliche Informationen
- Rückmeldung der Informationen mittels web-basierter Oberfläche und per Email
- Nutzung von KI-basierter Bildsegmentierung, um unterschiedliche gestaltete CVs in eine standardisierte Form zu bringen
- Nutzung von KI-Algorithmen zur Auswertung von Textinhalten des CVs

Die beiden letzten Zielsetzungen wurden im Rahmen des Praxispiloten ergänzt.

2.2. Ziele und Nutzen der Zielgruppen

Der Praxispilot sollte einen Proof-of-Concept entwickeln, also einen ersten Versuch eines Lösungsansatzes für den Anwendungsfall. Ziel ist es hierbei, die Machbarkeit des Anwendungsfalls zu prüfen, um am Ende das Potential von KI besser evaluieren zu können und zu entscheiden, ob der Anwendungsfall weiterverfolgt wird. Im Rahmen dieses Proof-of-Concept war es zentral, den Auszubildenden neben Techniken zur Konkretisierung eines Use Cases aus dem UX- und Change-Management-Bereich Methoden und Ansätze der künstlichen Intelligenz zu vermitteln. Diese KI-Methoden sollten dann unter Supervision des KI-Experten Oliver Denninger vom FZI selbstständig eingesetzt werden, um eigene Erfahrungen hiermit sammeln zu können.

Zielsetzungen CV-Analyzer

eigene Erfahrungen mit KI

Aus Sicht potentieller Endnutzender in der HR sollte die Lösung dazu dienen, diese in ihrer Arbeit zu entlasten. Personaler und Personalerinnen müssen tagtäglich viele CVs sowohl in firmeneigenen Personalabteilungen als auch bei externen Personaldienstleistern bearbeiten. Dabei sind CVs sehr unterschiedlich gestaltet und um möglichst effizient ohne Nutzung von Formularen Informationen aus dem CV zu evaluieren, sollte der CV-Analyzer den Personaler unterstützen. Neben dem Extrahieren von Informationen ist es hierbei wichtig die Darstellung des CVs in eine standardisierte Form zu bringen, sodass schneller eine Vergleichbarkeit zwischen verschiedenen Bewerbenden erzielt werden kann.

verbesserte Analyse von CVs mittels KI

2.3. Projektrahmen

Das Projekt wurde von September 2021 bis Dezember 2021 durchgeführt. Dabei wurden folgende Schritte im Rahmen des Projektes absolviert:

1. **Anforderungsanalyse:** Das BTB-Team bestehend aus den Auszubildenden und Mitarbeitenden Kathrin Mayer, Lukas Pesch, Marc Oberkirch, Metin Dereli, Oliver Steck und Nico Krause und dem Projektleiter Ljubo Barac hatte bereits eine erste Oberfläche für den CV-Analyzer entwickelt und war dabei, erste Funktionen zu entwickeln. Unklar war jedoch, wie genau KI im Rahmen des CV-Analyzers zum Einsatz kommen sollte. Um dies zu konkretisieren wurde eine Stakeholder- und Nutzerbefragung durchgeführt, die die Erwartungen an den Praxispiloten konkretisiert. Hierzu wurden die Auszubildenden und Mitarbeitenden durch das KI-Transfer BW-Team gecoacht. Anschließend führten Sie mit Stakeholdern aus der Personalabteilung und Firmenleitung vier Interviews durch. Die Ergebnisse wurden als Anforderungen und potentielle Anwendungsfälle aufbereitet.
2. **Auswahl des umzusetzenden Anwendungsfalls:** Im Rahmen der Anforderungsanalyse wurden zwei Anwendungsfälle identifiziert: Eine Möglichkeit bestand beim CV-Analyzer darin, ein KI-basiertes Scoring des CVs vorzunehmen. Anhand von bestimmten Parametern würde hierbei die KI eine erste Bewertung eines CVs vornehmen. Eine weitere Möglichkeit wurde darin gesehen, CVs, die sehr unterschiedlich aufgebaut sein können, in eine standardisierte Form mittels KI zu bringen, sodass Personaler diese schneller bearbeiten können. Am Ende wurde gemeinsam durch BTB und das

Interviews zur Klärung der Anforderungen

KI-Transfer BW-Team entschieden, den zweiten Anwendungsfall umzusetzen, da dieser eine gute Ausgangsbasis zur Weiterentwicklung schafft und es wurde vermutet, dass dieser weniger komplex in der Umsetzung ist.

3. **Aufbereitung der Daten:** Eine wesentliche Herausforderung lag für das Projekt in der Verfügbarkeit von einer ausreichenden Anzahl von qualitativ hochwertigen Lebensläufen. Es griff dabei auf frei verfügbare Musterlebensläufe im Internet zurück. Darüber hinaus wurden auch gezielt weitere CVs generiert und Data Augmentation Methoden zur Variation der Datenbasis eingesetzt.
4. **Auswahl von geeigneten KI-Algorithmen:** In Abstimmung mit dem KI-Experten Oliver Denninger wurde das weitere Vorgehen umgesetzt. Dieser half dabei die passende KI-Methodik zu identifizieren. In diesem Zusammenhang wurde für die Bildsegmentierung entschieden auf den U-Net-Algorithmus zurückzugreifen, einem Ansatz aus dem Bereich der Convolutional Neural Networks.
5. **Training der KI-Algorithmen:** Anhand von Beispieldatensätzen machten sich die Auszubildenden und Mitarbeitenden mit der generellen Funktionsweise des Algorithmus vertraut. In einem zweiten Schritt bereiteten sie die Daten für das Training auf, indem sie die Daten mithilfe von Masken labelten, um im Rahmen des Trainingsdatensatzes festzulegen, welche Bereiche durch den U-Net-Algorithmus im CV erkannt werden sollten. Anschließend wurde das Training der KI mit den aufbereiteten CVs durchgeführt. Unter Abstimmung mit Oliver Denninger wurden dann iterativ verschiedene Parameter angepasst, um ein möglichst optimales Ergebnis zu erzielen.
6. **Weiterentwicklung des CV-Analyzers:** Parallel zum Einsatz der KI-Algorithmen wurde der CV-Analyser hinsichtlich seiner Programmoberfläche und nicht-KI-bezogenen Funktionalitäten weiterentwickelt.
7. **Aufbereitung und Bewertung der Ergebnisse:** Der Einsatz des U-Net-Algorithmus zur KI-basierten Bildsegmentierung erwies sich als herausfordernd. Grund hierfür waren vor allem die Daten, denn die CVs wiesen eine höhere Varianz in der Gestaltung als erwartet auf. Vor diesem Hintergrund waren die Ergebnisse der Bildsegmentierung noch nicht

Bildsegmentierung mittels KI

zufriedenstellend. Ursächlich hierfür könnten die Qualität und Menge der Daten sein. Auch hinsichtlich der Textanalyse zeigten sich nicht die erhofften Ergebnisse.

3. Inhaltliches

3.1. Funktionale Anforderungen

Der Praxispilot „KI im HR“ hat im Wesentlichen vier Ziele:

- die Präzisierung des Anwendungsfalls für den KI-Einsatz
- die Entwicklung einer Oberfläche und wesentlicher damit verbundener Features des CV-Analyzers
- die KI-basierte Segmentierung von CVs, um diese in standardisierter Form an Personaler weiterzuleiten
- die Auswertung von Textinhalten des CVs unter Nutzung von KI-Algorithmen

Das erste Ziel wurde im Rahmen einer Anforderungsanalyse umgesetzt. Auf Basis von vier Interviews im Bereich HR, Geschäftsführung und IT wurden Anforderungen für den KI-Einsatz im Praxispiloten definiert. Parallel hierzu fand die Entwicklung der Oberfläche und zahlreicher Features statt. Diese stehen aber nicht im Fokus des Praxispilotenberichts. Im Folgenden wird nun auf die Umsetzung der KI-bezogenen Ziele näher eingegangen.

**# Bildsegmentierung und
Textanalyse mittels KI**

3.2. IT-Architektur

Für den Praxispilot wurde eine Microservice-Architektur gewählt. Diese ist prinzipiell eine Weiterentwicklung von serviceorientierter Architektur (SOA). Die Anwendung wurde in mehrere Anwendungen (sog. Microservices) aufgeteilt:

- **Daten-Service:** die notwendigen Daten werden in eine Microsoft Azure Datenbank gespeichert (Microsoft Azure ist eine Cloud-Computing-Plattform von Microsoft).
- **Frontend-Service:** verantwortlich für die Benutzeroberfläche. Dieser wurde in AngularJS umgesetzt.
- **Backend-Service:** verantwortlich für die Logik. Der Backend-Service wurde in Java umgesetzt.

Für die Datenreinigung und -vorbereitung sowie für das Trainieren der Algorithmen wurde ausschließlich auf Python 3 basierende Open-Source-Software verwendet. Innerhalb des Python-

Ökosystems wurden die folgenden Open-Source-Bibliotheken genutzt:

- **Numpy, Pandas, Pandas-Profiling und Matplotlib:** Software-Bibliotheken zum Einlesen der Daten, zur Datenexploration, zur Datenaufbereitung und für die visuelle Darstellung
- **Detectron2, Google Colab, openCV und TensorFlow:** Softwarebibliotheken zum Trainieren der Algorithmen aus dem Bereich des maschinellen Lernens

Datenaufbereitung mit Open-Source-Bibliotheken

Die Microservices laufen unabhängig voneinander und sind mit einem Mausklick skalierbar. Die Administration findet in Azure statt.

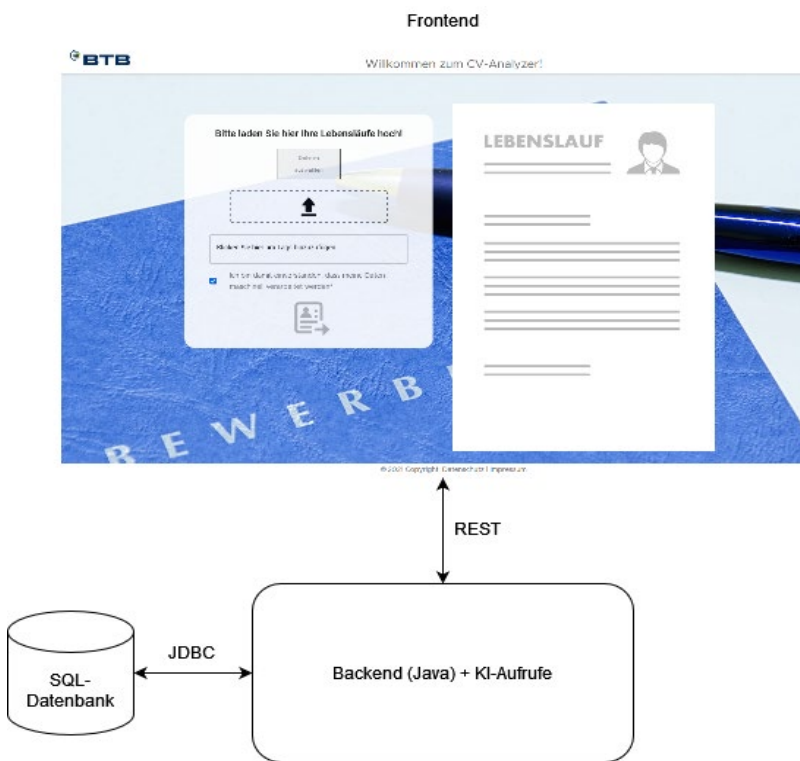


Abbildung 1, Übersicht CV-Analyzer: Dieser besteht aus drei Anwendungen, die parallel im Einsatz sind; Kommunikation zwischen Weboberfläche (Frontend) und Backend erfolgt über REST (Programmierschnittstelle, die sich an WWW-Paradigmen hält). Eine SQL-Datenbank kommuniziert mit dem Backend über einer Datenbankschnittstelle für Java (JDBC). Die Microservice-Architektur sorgt für hohe Skalierbarkeit und eine klare Trennung div. Schichten.

3.3. Daten und Algorithmen

3.3.1. Daten

Die Trainingsdaten wurden vom CV-Analyzer-Team auf Basis von Musterlebensläufen selbst erstellt, da keine echten Daten von der Personalabteilung bereitgestellt wurden. Das Team erstellte viele Lebensläufe in verschiedensten Layouts, um die hohe Bandbreite an Gestaltungsmöglichkeiten bei den Lebensläufen abzudecken. Die unterschiedlichen Layoutgestaltungen der Lebensläufe (z.B. Einbindung von Bildern, Icons, Textpositionierung) waren dabei weniger die Herausforderung, sondern die Generierung von passenden Inhalte hierfür. Mit rund zweihundert Lebensläufen wurden dann erste Trainingsversuche eines Modells gestartet.

3.3.2. Datenaufbereitung

Um aus den Daten, die zur Verfügung standen, mehr Trainingsdaten zu erstellen, wurde die Technik der Datenaugmentierung genutzt und das Layout der Lebensläufe geändert: durch Maßnahmen wie das Verschieben von Bewerbungsfotos oder die Vergrößerung von Textblöcken. Nachdem jeder Lebenslauf in ein Bilddateiformat umgewandelt worden ist, musste man diese maskieren, damit die künstliche Intelligenz später beim Training wusste, welche Bildbereiche sie erkennen sollte.

Für das Maskieren der Bereiche in den Bilddateien wurde das Programm Labelme benutzt, in dem man entweder grobe Rechtecke um die Textblöcke in den Lebensläufen zog oder eine genauere Umrandung um die einzelnen Zeilen erstellte. Das Resultat der Maskierung von zwei- bis dreihundert Lebensläufen waren ebenso viele JSON-Dateien, in denen sich die Maske in schwarz-rot, das normale Bild und beides übereinander befanden, jeweils als JPG-Dateien. Um diese zu extrahieren wurde eine Batch-Datei geschrieben, die aus den einzelnen JSON-Dateien jeweils Maske, Bild und Kombination der zwei entnahm und gesammelt in eigene Ordner speicherte. Diese konnten dann sofort im Python-Skript aufgerufen und trainiert werden.

Datenaugmentierung bei Lebensläufen

Labelling der Daten mithilfe von Maskierung

3.3.3. Erstellen der Zielvariable

Im Laufe des Entwicklungsprozesses veränderte sich die Zielsetzung für die KI, da verschiedene Umsetzungsmöglichkeiten bestanden. Manche dieser Möglichkeiten stellten sich allerdings insbesondere aufgrund des engen zeitlichen Rahmens als nicht

umsetzbar heraus. Ein erster Entwicklungsansatz bestand darin, das Design, das Layout und die Qualität eines Lebenslaufes zu bewerten. Dieser Ansatz wurde jedoch verworfen, da eine Grundlage für eine solche Bewertung fehlte.

Nachdem der Bewertungsansatz verworfen worden ist, wurde als neues Ziel definiert, die einzelnen Textblöcke, also Stationen in einem Lebenslauf, zu erkennen und zu extrahieren. Ergänzend gab es die Überlegung aufgrund dieser Blöcke eine kurze Zusammenfassung zu dem jeweiligen Bewerber zu schreiben. Um dieses Ziel zu erreichen, wurden Modelle zur Bildsegmentierung und Textgenerierung verwendet. Aufgrund der verschiedenen Zielsetzungen konnte das Team verschiedene KI-Methoden ausprobieren. Dies war für den Lerneffekt im Rahmen des Ausbildungsprojektes förderlich, weil das Team damit mehrere Formen von künstlicher Intelligenz kennenlernte.

3. 3. 4. Auswahl und Training des KI-Algorithmus

Für die Bildsegmentierung wurde ein U-Net-Algorithmus benutzt, der zur Segmentierung einzelner Textblöcke im komplexen Layout der CVs eingesetzt wurde. Diese Segmentierung ist die Basis, um z. B. CVs in eine standardisierte Form umzustrukturieren oder Informationen zur weiteren Verarbeitung zu extrahieren. Das Modell wurde über die Keras-API mit Tensorflow trainiert und die Masken wurden in Labelme erstellt.

Für die automatische Textgenerierung – Erstellung kurzer textueller Zusammenfassungen der Stationen im Lebenslauf eines Bewerbers – wurde ein deutsches GPT-2 Modell verwendet. Mittels der Transformers-Bibliothek wurde das Modell mit Texten aus Lebensläufen weitertrainiert.

3. 3. 5. Warum wird KI benötigt

Eine künstliche Intelligenz zum Analysieren von Lebensläufen wird benötigt, um alle verschiedenen Formen von Textblöcken zu erkennen. Natürlich könnte man ein Programm mithilfe von regulären Ausdrücken schreiben, welches einzelne Textblöcke erkennt, aber dann müsste man alle Möglichkeiten einen solchen Block aufzubauen in Betracht ziehen. Dieser reguläre Ausdruck wäre ziemlich lang, unverständlich, schwer zu warten und unzuverlässig. Aufgrund der Vielfalt der Strukturen und Layouts von Lebensläufen ist es wichtig einen einzelnen Textblock zu erkennen, unabhängig vom Inhalt, da dieser später zugeordnet

werden kann. Genau dafür wird eine künstliche Intelligenz verwendet, die unabhängig vom Textinhalt nur die Form des Textes zuordnet und erkennt. Dadurch kann die KI besser mit der Vielfalt an CVs umgehen.

3. 3. 6. Protoypische Umsetzung des CV-Analyzers

Als herausfordernd erwies sich die Bildsegmentierung mittels KI, da diese nicht in der Lage war, die relevanten Textblöcke in den CVs mit akzeptabler Genauigkeit zu erkennen. Ein solcher Textblock wäre z. B. die Beschreibung mit dem Datum einer beruflichen Station im Lebenslauf gewesen. Hätte die Erkennung eines solchen Textblocks zuverlässig funktioniert, dann hätte man die Koordinaten des Textblocks nutzen können, um diese an das Backend weiterzuleiten und mithilfe von Texterkennungsverfahren die Inhalte auszulesen. Dadurch hätten die Informationen des Lebenslaufs beispielsweise in eine Datenbank überführt werden können.

3. 3. 7. Evaluation

Mithilfe der unterschiedlichen KI-Ansätze, die im Rahmen des Praxispiloten ausprobiert wurden, konnte BTB viel über die Funktionsweise und die Herausforderung von KI lernen. Zwar konnten die Ergebnisse noch nicht in einen fertigen Prototypen umgesetzt werden, aber die Erkenntnisse waren umfangreich.

Zunächst kann man sagen, dass die Anzahl an Trainingsdaten für die kaum zu überblickenden verschiedenen Möglichkeiten an Formen, Designs und Layouts bei CVs schlicht zu gering waren, um eine vernünftige Erkennung zu bewerkstelligen. Formen wurden zwar in den CVs vom Bildsegmentierungsalgorithmus erkannt, doch oft hatten diese nichts mit den eigentlichen Textblöcken zu tun, oder der gesamte Text wurde unabhängig von Form oder Position erkannt. Um Blöcke zuverlässig zu erkennen, hätte man mehr Masken mit derselben Form gebraucht. Die Anzahl der Trainingsdaten hätte um ein Vielfaches größer sein müssen, was für das Projekt, vor allem auch in Hinsicht auf die Teamgröße und den zeitlichen Rahmen, schlichtweg nicht möglich war. In der wissenschaftlichen Literatur gibt es zwar vereinzelt Arbeiten, die Segmentierung von Seitenlayouts für die komplexen Layouts von Lebensläufen erwähnen, es gibt aber wenig Information dazu, auf welchen Daten diese Arbeiten stattgefunden haben. Insbesondere gibt es keine verfügbaren

Datensätze mit Lebensläufen, die für das Training eines Modells hätten genutzt werden können.

Neben der Segmentierung und Neustrukturierung der Lebensläufe wurde im Rahmen des Ausbildungsprojekts zum Schluss versucht, auch textuelle Zusammenfassungen aus den Stationen des Lebenslauf zu generieren. Das dafür eingesetzte deutsche GPT-2 Modell kam jedoch kaum mit Begrifflichkeiten aus dem Kontext IT und entsprechenden Berufsbildern zurecht. Dass Begriffe wie Administration oder Front-End-Entwicklung nicht vom Modell verstanden werden, war erwartet worden. Dass allerdings auch Begriffe wie Praktikum oder Projektleitung nicht durch das Modell erfasst werden, war nicht erwartet worden. Insgesamt deuteten die generierten Sätze darauf hin, dass das verwendete Modell zu einem erheblichen Anteil mit älteren literarischen Texten trainiert wurde (vermutlich da diese keinen Schutzrechten mehr unterliegen). Aufgrund der großen Diskrepanz zwischen dem Vokabular des Modells und dem Vokabular der zu generierenden Texte, war auch das punktuelle weitertrainieren des Modells mit Texten auf Basis des Zielvokabulars nicht erfolgreich. Analog zu Bildsegmentierung hätten hier wesentlich umfangreichere Trainingsdaten gesammelt und aufbereitet werden müssen, was im Rahmen des Piloten nicht möglich war.

Eine Form der künstlichen Intelligenz, die funktioniert und im Rahmen des Projektes genutzt wurde, ist eine Bild-zu-Text-Erkennung. Wenn aus einem Dokument der Text nicht gelesen werden kann, weil es sich um ein Bild im Dokument handelt, wird der Text mithilfe von künstlicher Intelligenz erkannt und ausgegeben. Dadurch werden viele Dokumente trotzdem ausgelesen, die zuvor mit anderen Methoden nur fehlerhaft auslesbar waren.

3.4. Herausforderungen

Bei der Entwicklung des CV-Analyzers ergaben sich einige Herausforderungen:

Zentral für das Gelingen der KI ist die Verfügbarkeit von Daten in ausreichender Menge und Qualität. In diesem Praxispiloten zeigt sich, dass insbesondere bei personenbezogenen Daten wie z. B. Lebensläufen diese Verfügbarkeit besonders komplex ist. Neben der Menge der CVs ist auch wichtig, dass diese möglichst gut verschiedene Gestaltungen von CVs repräsentieren. Hierbei stellte sich im Laufe des Entwicklungsprozesses heraus, dass

**# Datenverfügbarkeit und
-qualität**

Bewerbende mit größerer Varianz als gedacht ihre Lebensläufe gestalten. Dies erschwerte die Umsetzung deutlich, denn die unterschiedlichen Gestaltungsansätze müssen mit einer ausreichenden Menge an CVs im Datensatz repräsentiert werden. Dies gelang nur eingeschränkt.

Eine weitere Herausforderung bestand in der beschränkten Zeit der Umsetzung. Gerade die Generierung von qualitativ hochwertigen Daten ist mit einem bestimmten Zeitaufwand verbunden, da Schritte wie das Labelling durch Menschen erfolgen müssen. Zudem existiert mittlerweile eine Vielzahl von KI-Algorithmen, deren Testung Zeit beansprucht, um einen möglichst optimalen Ansatz zu identifizieren bzw. diesen zu optimieren. Mit einem umfangreicheren Zeitbudget hätten noch intensiver der KI-Algorithmus entwickelt und die Datenmenge und -qualität verbessert werden können.

zeitliche Begrenzung des Praxispiloten

4. Resümee

4.1. Wissensgewinn

Der Praxispilot hat wesentlich zu einem besseren Verständnis von KI bei dem Unternehmen BTB beigetragen. Hierbei war insbesondere hilfreich, dass im Rahmen des Ausbildungsprojektes die Mitarbeitenden des Unternehmens selbst die KI-Algorithmen unter Supervision eines KI-Experten entwickeln konnten. Dadurch bekommt das Unternehmen eine bessere Vorstellung von KI-Projekten und kann besser einschätzen, inwiefern es selbst in diesem Bereich in Zukunft Kundenprojekte durchführen möchte.

Aufbau von strategischem KI-Know-how im Unternehmen

Auch hat das Unternehmen im Rahmen des Praxispiloten ein verändertes Verständnis darüber bekommen, welche Bedeutung Daten für das Gelingen einer KI-Lösung hat. Im Generieren und Aufbereiten der Daten liegt einer der zeitaufwändigsten Faktoren für die KI-Entwicklung.

Zusätzlich hat der Praxispilot das Unternehmen darin bestärkt, weiter in den Aufbau von KI-Know-how zu investieren und Kooperationen hierzu mit Forschungspartnern anzustreben, die bei KI-Lösungen ihre Expertise einbringen können.

Kooperation mit Forschung erwünscht

4.2. Ausblick

Durch die Durchführung des Praxispilots sind BTB weitere Möglichkeiten der Ausbildung eigener Mitarbeitenden sowie Unterstützung deren Kunden bewusst geworden. Das Unternehmen konzentriert sich hierbei auf Daten bestehender Kunden und versucht gemeinsam mit den Kunden KI-Pilotprojekte umzusetzen. Im Rahmen des Praxispilots ist Entscheidungsträgern klar geworden, dass KI-Projekte mit deutlich mehr Risiko verbunden sind. Um dieses Problem zu lösen, sowie das Projekt für Klienten attraktiver zu machen wurde ein Vorgehensmodell entwickelt:

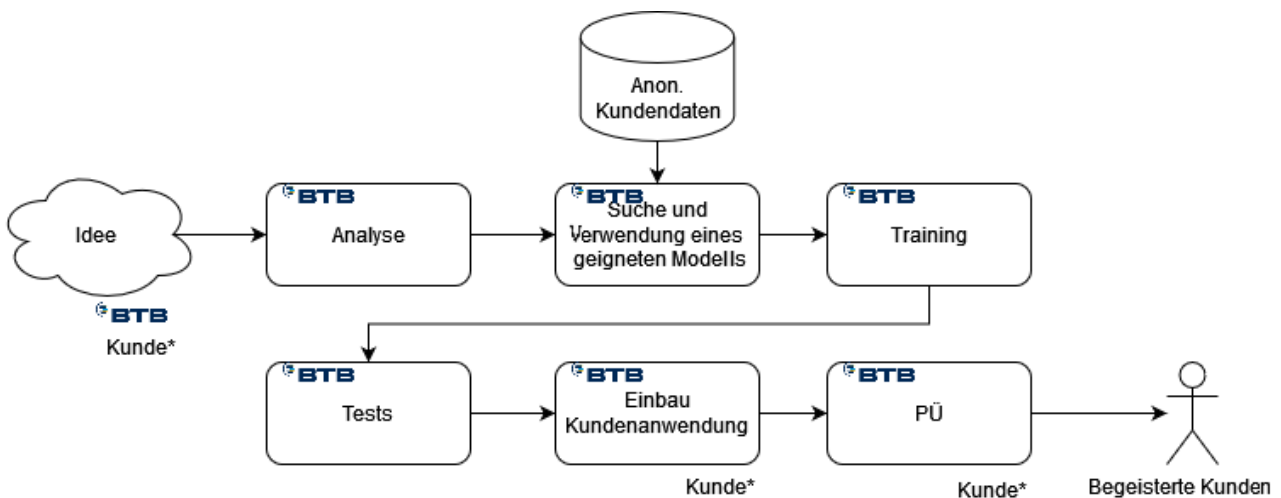


Abbildung 2: Vorgehensmodell für ein potenzielles KI-Pilotprojekt: BTB übernimmt den größten Teil der Aufgaben – von der ersten Idee bis zur Produktübernahme (PÜ). Der Kunde beteiligt sich erst nach der ersten kritischen Phase (Training bzw. Tests), wenn die gewünschte Genauigkeit des Modells (z. B. > 90%) erreicht wird.

Potenzielle Pilotprojekte könnten in dem Fall mit bestehenden Kunden entworfen werden. Entwicklung, Training und die Tests übernimmt die BTB. Die Entwicklung erfolgt im Rahmen eines Ausbildungsprojekts. Sind die Tests erfolgreich, so kann die Lösung dem Kunden vorgestellt werden oder sogar in bestehende Projekte integriert werden.

Der Praxispilot sowie das obige Modell haben bereits zum ersten Pilotprojekt geführt, d. h. die BTB wird im ersten Quartal 2022 ein KI-Pilotprojekt umsetzen. Darüber hinaus wurde das Interesse am Thema KI sowohl bei der Geschäftsführung als auch bei den Mitarbeitenden geweckt.

5. Impressum

5.1. Kontaktdaten

Sie haben eine Frage zum Praxispiloten? Gerne können die Vertreter der Praxispilotpartner bei Fragen und Anmerkungen direkt angesprochen werden:

5.1.1. Unternehmen

BTB Betriebswirtschaftliche und -technische
Beratungs-GmbH
Ljubo Barac, BTB GmbH
E-Mail: Ljubo.Barac@btb.gmbh
Wilhelm-Haas-Str. 6
Leinfelden-Echterdingen 70771

5.1.2. KI-Transfer BW

DIZ Digitales Innovationszentrum
Alexander Dregger
Wissenschaftlicher Mitarbeiter
E-Mail: dregger@diz-bw.de
Telefon: +49 721 9654-659
Haid-und-Neu-Straße 18
76131 Karlsruhe

FZI Forschungszentrum Informatik
Oliver Denninger
E-Mail: denninger@fzi.de
Haid-und-Neu-Straße 18
76131 Karlsruhe

5.1.3. Fördergeber

Ministerium für Wirtschaft, Arbeit und Tourismus
Baden-Württemberg

Postfach 10 01 41
Schlossplatz 4 (Neues Schloss) 70001 Stuttgart
Telefon +49 711 123-2869
Fax +49 711 123-2871
E-Mail: pressestelle@wm.bwl.de
www.wm.baden-wuerttemberg.de

Projektverantwortliche
Mascha Ananda Eckhardt
Referat 31 „Industrie- und Technologiepolitik,
Digitalisierung“
Ministerium für Wirtschaft, Arbeit und Tourismus
Baden-Württemberg
Postanschrift: Schlossplatz 4 (Neues Schloss)
70173 Stuttgart
Dienststz: Willi-Bleicher-Straße 19
70174 Stuttgart
Telefon: +49 711 123-2442
E-Mail: mascha.eckhardt@wm.bwl.de

Der Praxispilot ist im Rahmen des vom Ministerium für Wirtschaft, Arbeit und Tourismus geförderten Projektes KI-Transfer BW entstanden. Weitere Informationen hierzu finden Sie unter:
<https://www.wirtschaft-digital-bw.de/ki-labs>

5.2. Bildnachweis

S. 1: pixabay.com, S. 18: unsplash.com

5.3. Gestaltung und Umsetzung

DIZ Digitales Innovationszentrum
Sabine Schneider

„KI-Projekte im Mittelstand können auch wesentlich durch die jungen Kräfte im Unternehmen gefördert werden. So lernen diese und das KMU nicht nur die neuen technischen Möglichkeiten kennen, sondern es können auch erste KI-Erfolge entstehen.“

Ljubo Barac, Managing Consultant bei BTB Betriebswirtschaftliche und -technische Beratungs-GmbH

