

## Mit Advanced Machine Learning Zeit und Geld sparen

Künstliche Intelligenz beschleunigt die Analyse juristischer Dokumente

Franz Kögl | Vorstand | IntraFind Software AG

16. März 2020

LR 2020, Seiten 37 bis 40 (insgesamt 4 Seiten)

---

Verträge und andere juristische Dokumente liegen meist in großen Mengen und in unstrukturierter Form vor. Machine-Learning-Verfahren helfen dabei, diese zu klassifizieren sowie relevante Informationen schnell zu sichten und effizienter zu analysieren. Damit wird ein klar messbarer Return of Investment erzielt. 1

ROI-Betrachtungen sind meistens schwierig. Denn für einen aussagekräftigen Vergleich muss die Prozessdauer und Prozessqualität erst ohne und dann mit Einsatz einer neuartigen Verbesserungskomponente gemessen werden, um die Amortisationsgeschwindigkeit einer neuen Software zu belegen. Ein Paradebeispiel für schnell erzielten ROI ist der Einsatz von Advanced-Machine-Learning-Verfahren wie sie bei der Vertragsanalyse verwendet werden. Diese bieten konkret messbares Zeiteinsparungspotential. Das Gute dabei ist, dass die komplexe KI im Hintergrund läuft und die Anwendung für den User einfach ist. 2

### I. Machine Learning reduziert Bearbeitungszeiten

Machine Learning hilft bei zwei unterschiedlichen Use-Cases im Rahmen der Vertragsanalyse. 3

#### 1. Sortierung großer unbekannter Dokumentenmengen

Die Sortierung großer unbekannter Dokumentenmengen in eine entsprechende vordefinierte Ablagestruktur, wie z.B. die automatische Einsortierung von unterschiedlichsten Dokumententypen einer Gewerbeimmobilie beim Kauf oder Verkauf. 4

Software für Textklassifikation ist in der Lage, Dokumente automatisch dem richtigen Thema oder einer bestimmten Kategorie zuzuordnen. Dafür wird ein Metadatum, zum

Beispiel ein bestimmtes Thema, zu einem beliebigen Textfragment erzeugt und das Dokument wird anschließend auf Grundlage seines Inhaltes diesem Thema zugewiesen. Das Machine-Learning-Verfahren übernimmt hierbei die Einsortierungsaufgaben. Egal, ob bei Due-Diligence-Verfahren sehr heterogene Datenbestände zu sortieren sind oder, ob es um eine Vielzahl von Dokumenten bei Massenverfahren geht – in all diesen Situation helfen Machine-Learning-Verfahren dabei, eine händisch nur schwer zu bewältigende Datenmenge in Sekundenschnelle zu klassifizieren und zu verschlagworten oder in eine Taxonomie einzuordnen. Automatische statt manueller Einsortierung spart Zeit und Kosten und ist eindeutig messbar!

## **2. Das automatische „Aufbrechen“ eines Dokuments in für die Bearbeitung relevante Klauseln, Findings oder Datenpunkte**

Sobald der Vertrag durch die Software analysiert wurde, können einzelnen Klauseln - oder letztlich allgemeiner - für den Benutzer relevante Textpassagen als Einzelinformationen schnell gefunden, bewertet oder miteinander verglichen werden. Der Benutzer kann sich so z.B. alle 50 Haftungsklauseln einer neuen Dokumentenmenge oder Klauseln mit Sonderkündigungsrechten mit wenigen Klicks anzeigen lassen und diese nach kritischen Aspekten durchforsten, kommentieren und Red Flags erkennen. Auch ist es möglich, sich für bestimmte Verträge oder ganze Vertragstypen, Exzerpte der für einen bestimmte Analyse relevanten Inhalte automatisch generieren zu lassen. Vertragspartner, Kündigungsfristen, Change-of-Control-Klauseln und Vertragsstrafen sind somit mit wenigen Handgriffen aus 10, 100 oder auch aus 10.000 Verträgen extrahiert und z.B. in Excel-Form zusammengefasst. Diese „Documents 2 Data“-Verarbeitung, dieses Machine Learning unterstützte Aggregieren von Kerninhalten bietet klar messbare Zeiteinsparungen und hilft dem Anwalt, seine Dienstleistungen wettbewerbsfähig anzubieten.

5

### **II. Nur wenige Trainingsdaten erforderlich**

Der Nachteil vieler Lernverfahren, dass große Mengen an Trainingsdokumenten benötigt werden, sollte bei einer Software mit intelligenter Machine-Learning-Komponente eliminiert sein. Die Software muss auch mit wenigen Trainingsdaten zufrieden sein und aus dem Stand eine sehr gute Erkennungs- und Zuordnungsqualität liefern. Zudem sollte die Software beim Training des Lernverfahrens auch Feedback geben, was an den Trainingsdaten verändert werden sollte, damit das System selbst besser lernt. Damit können auch Nutzer, die keine Informationswissenschaftler sind, ein technisches Lernverfahren bedienen.

6

Das System muss aber auch dafür sorgen, dass sich seine Zuordnungsgüte in der produktiven Nutzung weiter verbessert: Einzelne Dokumente, bei denen sich das KI-Verfahren unsicher ist, müssen händisch eingruppiert werden können. Damit wird dann

7

eine Lernschleife angestoßen und das Verfahren wird durch dieses Relevanz-Feedback weiter trainiert und verbessert.

### **III. Effizienter im Job und mit mehr Spaß bei der Sache**

Der Nutzen von Machine Learning ist bei der tiefen Vertrags- und Dokumentenanalyse besonders schnell nachweisbar. Ein Test auf den zu analysierenden Daten kann ohne großen Aufwand umgesetzt werden und es ist belastbar zu sehen, was die automatische Sortierung von Daten und das Erkennen von beliebigen Findings für ein konkretes Einsparungspotential mit sich bringt. Die Rentabilitätsgrenze liegt bei einer relativ kleinen Dokumentenmenge. Das Machine-Learning-Verfahren spart nicht nur Zeit- und Personalkosten, sondern liefert aufgrund der objektiven Sortierungsarbeit auch bessere Ergebnisse als die menschlichen und damit subjektiven und tagesformabhängigen Einsortierungsentscheidungen. 8

Nicht zu vergessen ist die Motivationskomponente beim Einsatz von intelligenter Legal-Tech-Software: Sie entlastet Juristen von weniger spannenden Routinearbeiten und gibt ihnen mehr Freiraum für Tätigkeiten, die ihre menschliche Intelligenz und juristische Kompetenz erfordern. 9

### **IV. Diese Anforderungen sollte eine Vertragsanalyse-Software erfüllen**

Damit eine KI-basierte Software für Vertragsanalyse ihre User auch wirklich effizient unterstützt und einen schnellen ROI bringt, sollte man bei der Auswahl auf einige Features achten: 10

#### **1. Vortrainierte KI-Komponenten**

Der produktive Einsatz der Software lässt sich erheblich beschleunigen, wenn ihre KI-Komponenten bereits umfassend vortrainiert sind. Sie sollte deshalb in der Lage sein, die gängigsten Vertragsklauseln bereits von Haus aus automatisch zu erkennen. 11

#### **2. Unkompliziertes Weitertrainieren**

Um weitere relevante Inhalte selbstständig zu erkennen, kann die Software idealerweise von den Juristen unkompliziert für beliebige Rechtsgebiete und Textinhalte weitertrainiert werden. Advanced-Machine-Learning-Verfahren ermöglichen es, beliebige Klauseln, Datenpunkte oder Textpassagen von Null an neu mit sehr wenigen Trainingsdaten weiter anzulernen. 12

### **3. Unterstützung vielfältiger Formate**

Die Verarbeitung vielfältiger Dateiformate wie Word, PDF oder Excel sollte genauso wie die OCR-Texterkennung von gescannten Dokumenten standardmäßig in die Software integriert sein. 13

### **4. Flexibler Export**

Auch in umgekehrter Richtung ist die Software idealerweise möglichst flexibel. Damit Juristen die Ergebnisse ihrer Analysen weiterverarbeiten oder Kollegen zur Verfügung stellen können, sollten sie sich unkompliziert in Dateiformate wie Microsoft Word oder Excel exportieren lassen. Außerdem sollten die Daten über eine Programmierschnittstelle an eine nachfolgende Applikation wie ein Kanzleiverwaltungssystem oder eine andere Datenbank-Applikation weitergegeben werden können. 14

### **5. Kein Vendor-Lock-in**

Damit die Nutzer beim Betriebsmodell flexibel sind, sollte sich die Lösung sowohl on-premise, d.h. auf unternehmenseigenen Rechnern, als auch als Software-as-a-Service nutzen lassen. Bei der zweiten Variante muss aber sichergestellt sein, dass sich die eigenen Trainingsdaten exportieren und damit im Fall eines Anbieterwechsels weiterverwenden lassen. 15