



Künstliche Intelligenz bei IP7

Inhaltsverzeichnis

Einsatz von KI	2
Allgemeine Informationen	2
Ordnerzuordnung	3
Classifier	3
Trainingsstatistik	4
Limitierungen	5
Automatische Klassifikation	5
Limitierungen	6

Einsatz der KI

Allgemeine Informationen

Zu viele Patente und zu wenig Zeit

In Zusammenarbeit mit Averbis, einem führenden Anbieter im Bereich der künstlichen Intelligenz, haben wir ein Konzept entwickelt, das spezifische Abläufe innerhalb unserer Software vereinfacht und die Effizienz erhöht.

Durch die Nutzung natürlicher Sprachmodelle, analysiert Averbis große Mengen von Patentdokumenten um sie mittels maschinellem Lernen gemäß Ihren individuellen Anforderungen und Richtlinien zu klassifizieren.

(Averbis: <https://averbis.com/de/patent-kategorisierung-und-monitoring/>)

Die KI kann in verschiedenen Bereichen innerhalb des IP7 Compass eingesetzt werden:

- ❖ Ordnerzuordnung bzw. automatische Klassifizierung von Patenten
- ❖ Automatische Aussortierung von "nicht relevanten" Treffern (demnächst verfügbar!)
- ❖ Sortierung von Ergebnislisten nach Relevanz (auch in Kürze verfügbar!)

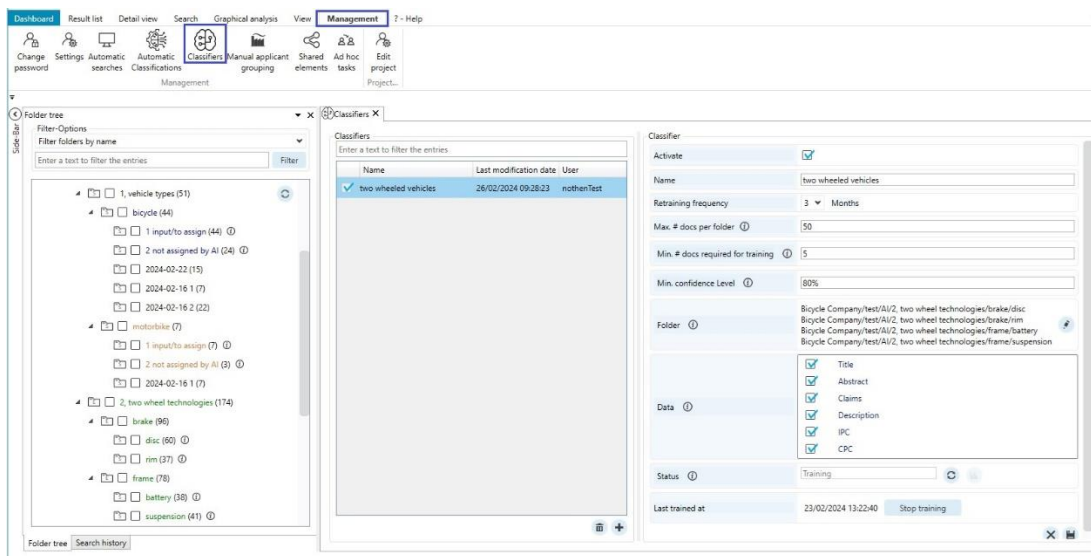
Das wesentliche Kennzeichen der KI besteht darin, dass sie durch vorgegebene Daten trainiert wird und dabei „lernt“. Das Ziel ist es, die KI durch stetige menschliche Anpassungen und wiederholtes Training kontinuierlich zu verbessern. Daher ist die Qualität der künstlichen Intelligenz immer abhängig von den bereitgestellten Daten und der menschlichen Unterstützung bzw. Korrektur.

Momentan befinden wir uns in der Pilotphase. Es ist möglich, dass sich einzelne Funktionen, Oberflächen und der Leistungsumfang in Zukunft ändern werden. Die KI ist standardmäßig nicht im Compass eingebunden, doch es gibt ein zusätzliches Modul, das nach Anforderung durch IP7 aktiviert werden kann. Für die Verwendung dieser Technologie entstehen extra Kosten. Gerne erstellen wir Ihnen hierfür ein individuelles Angebot.

Ordnerzuordnung

In dem IP7 Compass werden Ordnerstrukturen angelegt, um Patente gemäß individuellen Kriterien zu klassifizieren. Beispielsweise lässt sich ein Technologiebaum erstellen, der es ermöglicht, Patente entsprechenden Technologien zuzuweisen, auch Geschäftsbereichen oder Akten. Durch die Funktion "Automatische Klassifizierung" kann die künstliche Intelligenz diese Zuweisungsaufgabe selbstständig übernehmen.

Classifier



Bevor die künstliche Intelligenz in der Lage ist, Patente in Ordner einzuordnen, ist es notwendig, dass sie von bereits bestehenden Zuordnungen lernt. Hierfür ist eine vorhandene Ordnerstruktur mit zuvor korrekt klassifizierten Patenten erforderlich. Anschließend kann ein "Classifier" angelegt werden. Bei diesem Vorgang werden verschiedene **Einstellungen** festgelegt, von denen viele einen **direkten Einfluss** auf den Trainingsprozess haben:

- **Aktivieren:** Hier wird entschieden, ob ein automatisches Training regelmäßig erfolgt bzw. ob der Classifier verwendet werden kann.
- **Name**
- **Trainingsfrequenz:** Bestimmt die Häufigkeit des automatischen Trainings des Classifiers. Da sich die Inhalte der Ordner über die Zeit ändern, ist es sinnvoll den Classifier regelmäßig trainieren zu lassen.
- **Maximale Anzahl:** Legt fest, wie viele Dokumente die KI höchstens pro Ordner für Trainingszwecke nutzen darf. Eine zu geringe Anzahl an Dokumenten je Ordner kann die Trainingsqualität mindern, während eine zu hohe Anzahl die Trainingsdauer erheblich verlängert. Ideal ist eine konstante Dokumentenanzahl pro Ordner.

Beispiel: Angenommen Sie möchten eine große Menge an Ordnern automatisch klassifizieren. In den meisten Ordnern befinden sich ca. 50 Dokumente. In einigen wenigen

Ordnern befinden sich jedoch mehr als 500 Dokumente. Dann sollte der Wert für die Maximale Anzahl auf 50 Dokumente gesetzt werden, um ein ausgewogenes Training und somit bessere Ergebnisse zu erzielen. Die Obergrenze liegt bei 1.000 Dokumenten pro Ordner. Sind mehr Dokumente in einem Ordner vorhanden, werden über eine zufällige Auswahl 1.000 Dokumente aus diesem Ordner ausgewählt.

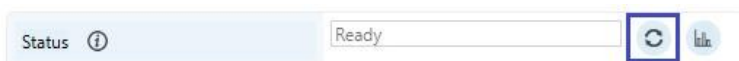
- **Minimale Anzahl:** Gibt an, wie viele Dokumente mindestens in einem Ordner vorhanden sein müssen, damit sie für das Training berücksichtigt werden.

Beispiel: die KI kann aus einem Ordner mit nur 3 Dokumenten keine guten Erkenntnisse gewinnen. Der Mindestwert liegt bei 10 Dokumenten. Ordner mit einer geringeren Anzahl werden für das Training ausgeschlossen.

- **Minimales Vertrauensniveau:** Das "Confidence Level" gibt in Prozent an, wie sicher sich die KI bei einer Zuordnung ist. Alles unterhalb des minimalen Wertes wird in den Ordner "nicht klassifizierbar" verschoben. Dieser wird später in der automatischen Klassifizierung festgelegt.
- **Ordner:** Bestimmt, welche Ordner für das Training genutzt werden. In diese Ordner wird später die "automatische Klassifizierung" die Zuordnungen vornehmen. Es können bis zu 500 Ordner ausgewählt werden.
- **Daten:** Legt fest, welche Daten die KI für das Training erhält (Titel, Zusammenfassung, Ansprüche, Beschreibung, IPC, CPC). Die Auswahl von "Beschreibung", liefert der KI wesentlich mehr Trainingsdaten, was sich entsprechend auf die Dauer des Trainings auswirken wird.
- **Status:** Zeigt den aktuellen Zustand des Classifiers an:

"Training notwendig": Nach der Erstellung des Classifiers ist ein Training erforderlich.
"Training": Der Classifier befindet sich im Trainingsprozess. Währenddessen können keine automatischen Klassifizierungen, die diesen Classifier nutzen, durchgeführt werden.
"Bereit": Der Classifier ist für die automatische Klassifizierung einsatzbereit, was auch bedeutet, dass das letzte Training erfolgreich abgeschlossen wurde.
"Fehler": Es ist ein Fehler aufgetreten. Ein erfolgreiches Training ist Voraussetzung für den Einsatz in einer automatischen Klassifizierung.

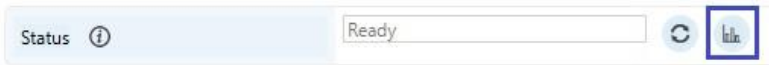
- **Letztes Training:** Hier wird angezeigt, wann das letzte erfolgreiche Training durchgeführt wurde. Der aktuelle Status kann über den „Aktualisieren“-Button abgerufen werden:



Sobald der "Classifier" mit dem Training fertig ist (Status "Bereit"), kann dieser in einer automatischen Klassifizierung eingesetzt werden.

Trainingsstatistik

Eine Analyse über den Trainings-Lauf kann durch den folgenden Botton ausgelöst werden:

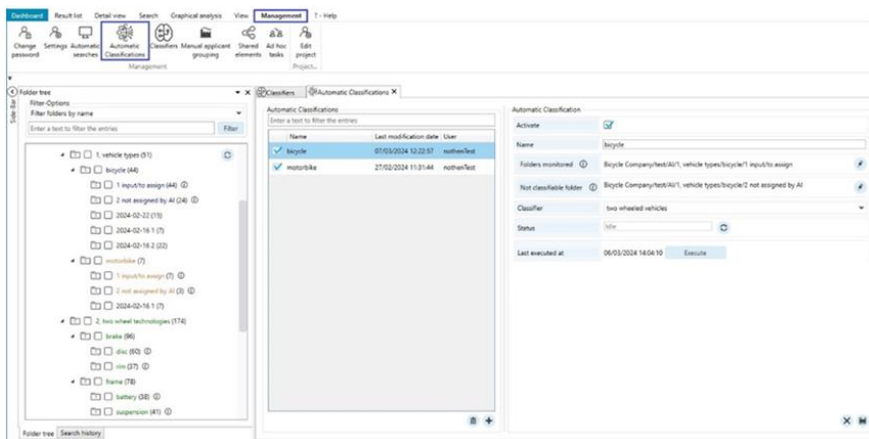


Weitere Informationen zu den Begriffen Precision, F1Score und Recall:
https://en.wikipedia.org/wiki/Precision_and_recall

Limitierungen

Maximal werden 10.000 Patente aus allen im Classifier einbezogenen Ordnern für das Training herangezogen. Sollte die Gesamtzahl der Patente in den Ordnern diese Zahl überschreiten, werden sie proportional reduziert. Falls dadurch die Patentanzahl in einem Ordner unter die festgelegte Mindestmenge fällt, wird die Zahl der ausgewählten Patente in diesem Ordner wieder bis zur Mindestgrenze erhöht. Das führt dazu, dass die tatsächliche Obergrenze leicht über 10.000 liegen kann.

Automatische Klassifikation



Bei der Einrichtung einer automatischen Klassifizierung stehen folgende Einstellungsmöglichkeiten zur Verfügung:

- **Aktivieren:** Diese Option bestimmt, ob die Klassifizierung automatisch durchgeführt wird. Mithilfe dieser Option kann eine aktive automatische Klassifizierung bei Problemen schnell gestoppt werden.
- **Name**
- **Überwachte Ordner:** Bestimmt die "Eingangsordner", deren Inhalte von der KI klassifiziert werden. Das umfasst alle derzeitigen und zukünftigen Patente in diesen Ordnern. Die Eingangsordner sollten nicht mit den Ordnern des Classifiers verwechselt werden: Die Ordner, in die die KI die Patente zuordnet, werden im Classifier definiert.
- **Ordner für "nicht klassifizierbar":** Hier werden alle Patente eingeordnet, die von der KI nicht klassifiziert werden können.
- **Classifier:** Auswahl des zuvor erstellten Classifiers.

➤ **Status:**

"idle" (inaktiv) – Die automatische Klassifizierung läuft zu dem Zeitpunkt nicht.
"running" (aktiv) – Die automatische Klassifizierung wird im Moment ausgeführt.

➤ **Letzte Ausführung:** Die letzte Durchführung der automatischen Klassifizierung.

Mehrere autom. Klassifizierungen:

Ein Classifier kann theoretisch für mehrere automatische Klassifizierungen verwendet werden.

Beispiel:

Es existieren mehrere Fahrzeugtypen, die überwacht werden und dann in einen Technologiebaum zugeordnet werden sollen:

1 „Fahrzeugarten“ -> "Fahrrad" und "Motorrad"

Es existiert eine Ordnerstruktur bzw. ein Technologiebaum, der alle Arten von Zweiradfahrzeugen aufnehmen soll:

2 „Technologien für Zweiräder“

- ➔ Die Ergebnisse, die von der KI nicht eindeutig zugeordnet werden können, müssen jedoch separat gespeichert werden. Aus diesem Grund wird für "Fahrrad" und "Motorrad" jeweils eine eigene automatische Klassifizierung erstellt. Allerdings muss der Classifier für diesen Fall nur einmal angelegt werden.

Limitierungen

In einem Durchgang können bis zu 5.000 Treffer bzw. Patente klassifiziert werden.

Überschreitet die Anzahl der Patente diese Grenze, wird die automatische Klassifizierung nicht durchgeführt und automatisch deaktiviert. Sobald der Lauf manuell ausgelöst wird, erscheint eine entsprechende Warnmeldung. Wird der Durchgang dennoch gestartet, beschränkt sich die Klassifizierung auf maximal 5.000 Patente.

Bereits in einem Lauf klassifizierte Patente müssen nicht aus dem Eingangsordner entfernt werden. Sie werden in nachfolgenden Durchgängen als schon klassifiziert identifiziert und nicht erneut bearbeitet. Um einzelne oder mehrere Patente nochmals klassifizieren zu lassen, ist es erforderlich, diese zunächst aus dem Ordner zu entfernen und danach erneut hinzuzufügen. Mit dieser erneuten Zuordnung gelten sie nicht länger als bereits klassifiziert.