



FH Salzburg

Angebot

Blumatix: LLM Document Extraction with Confidence Score

der

Fachhochschule Salzburg GmbH

Urstein Süd 1, 5412 Puch/Salzburg

vertreten durch den Head of Research Department Creative Technologies, FH-Prof. Dr. Markus Tatzgern

Projektleitung: FH-Prof. Dr. Clemens Havas

UID: ATU 44554503

Firmenbuchnummer: FN 166054y

Firmenbuchgericht: Landesgericht Salzburg

(im Folgenden „FHS“ genannt)

an

Blumatix Intelligence GmbH

Schwarzstraße 48, 5020 Salzburg

vertreten durch Geschäftsführer, DI Loiperdinger Martin

Projektleitung: Martin Loiperdinger

UID: ATU70948479

Firmenbuchnummer/ZVR: FN 450980 g

Firmenbuchgericht: Landesgericht Salzburg

(im Folgenden „Blumatix“ genannt)

1. Projektbeschreibung

Blumatix arbeitet an der automatisierten Digitalisierung von Dokumenten mit dem Ziel, Informationen mithilfe von Machine-Learning-Methoden möglichst präzise aus Dokumenten zu extrahieren. Im Rahmen eines Mini-Projekts mit dem Ziel eines Proof of Concepts soll untersucht werden, ob ein Confidence Score zu Machine Learning Modellen im speziellen Large Language Models (LLMs) hinzugefügt werden kann, der Kunden eine Rückmeldung, ob die Sicherheit der Ergebnisse werden soll.

Das Proof-of-Concept untersucht, ob LLMs bei der Extraktion typischer Rechnungsinformationen wie Gesamtbetrag oder Rechnungsdatum verlässliche Confidence-Scores im Bereich von 0 bis 1 liefern können. Diese Scores sollen anzeigen, wie sicher sich das Modell bei der Erkennung ist und ob extrahierte Werte automatisch weiterverarbeitet werden dürfen.

Darüber hinaus wird erforscht, wie LLMs als „Judge“ eingesetzt werden können, um synthetisch generierte Frage-Antwort-Paare für wissensbasierte Sprachmodell-Systeme zu validieren. Beide Ansätze lassen sich kombinieren: Das LLM bewertet einerseits die Korrektheit (Judge) und liefert gleichzeitig einen Confidence-Score, der signalisiert, wie zuverlässig dieses Urteil ist.

Technisch können solche Konfidenzwerte aus den von der API bereitgestellten Token-Logits bzw. Wahrscheinlichkeitsverteilungen abgeleitet werden (z. B. über logprobs in der Azure OpenAI API). Dadurch lässt sich nachvollziehen, mit welcher Wahrscheinlichkeit das Modell bestimmte Tokens oder Antworten ausgewählt hat. Ziel ist es zu prüfen, ob diese Wahrscheinlichkeiten konsistent in einen praxisnahen Score zwischen 0 und 1 übersetzt werden können und ob sie als verlässliche Grundlage für automatische Entscheidungen dienen.

Der Auftraggeber stellt rund 500 annotierte Rechnungsdokumente (PDF) inklusive Ground-Truth-Daten im JSON-Format zur Verfügung. Zum Einsatz kommen die Modelle GPT-4o-mini und Mistral Document AI.

Für die Umsetzung des Projektes wird Blumatix Lizenzen für die Modelle, Cloud-Ressourcen (Microsoft Azure) und einen Arbeitsplatz unentgeltlich zur Verfügung stellen.

Mehr Details zu dem Beschreibung der Projektbeschreibung befindet sich im beigefügten Dokument „SteadyLearning_POC_v2.pdf“. Die angegebenen Werkzeuge von 10 Tagen sind ein Richtwert für die Umsetzung des Projektes.

Verwertung

Das Mini-Projekt ergänzt das Forschungsprojekt „GenText“ des Departments Creative Technologies (CT) der FHS, in dem Methoden zur synthetischen Datengenerierung für Dokumente erforscht werden. Während sich „GenText“ primär mit der Erzeugung realistischer Dokumentendaten beschäftigt, konzentriert sich das Mini-Projekt auf Erklärbarkeit der Analyse von Dokumenten.

Dokumentenanalyse ist ein bedeutendes und dynamisches Forschungsfeld mit international anerkannten wissenschaftlichen Konferenzen wie zum Beispiel ICDAR. Da die im Rahmen dieses Projekts gewonnenen Erkenntnisse auch für die wissenschaftliche Gemeinschaft von Relevanz sind, verpflichtet sich Blumatix, der FHS die aggregierten Projektergebnisse zur Veröffentlichung bereitzustellen und die Verwertung im Rahmen einer wissenschaftlichen Publikation zu ermöglichen. Dabei ist sicherzustellen, dass keine sensiblen oder unternehmensspezifischen Daten von Blumatix offengelegt oder Dritten zugänglich gemacht werden.

2. Projektziel

Ergebnisse des Projekts:

1. Implementierung eines exemplarischen Proof-of-Concept-Workflows (z.B. Python oder Jupyter-Notebook) der die Extraktion der definierten Rechnungsinformationen mit zugehörigen Confidence-Scores zeigt.
2. Quantitative Auswertung der Genauigkeit und Verlässlichkeit des Confidence-Score.
3. Zusammenfassung der Ergebnisse und Analysen in einem klar strukturierten Markdown-Dokument.

Das Projekt gilt mit Übergabe des Endberichts und des Sourcecodes als abgeschlossen.

3. Projektlaufzeit

Das Projekt beginnt mit 3.11.2025 und endet mit 15.12.2025.

4. Auftragsentgelt & Zahlungsmodalitäten

Die in Punkt (1) beschriebenen Leistungen bietet die FHS gemäß ihrer Kalkulation zu einem Gesamtbetrag von EUR 5.500 (inkl. 10% Umsatzsteuer) an.

Dieser Betrag ist wie folgt zu zahlen:

50 % bei Auftragserteilung

50 % nach Übergabe der Ergebnisse **mit Datum 15.12.2025**

Der Rechnungsbetrag ist jeweils binnen 14 Tagen nach Rechnungslegung zur Zahlung fällig.

5. Sonstige rechtliche Bestimmungen

Die dem Angebot beigefügten **Allgemeinen Bedingungen zur Beauftragung & Durchführung von Forschungs- & Entwicklungsleistungen der Fachhochschule Salzburg GmbH** (Stand: 27.07.2023) bilden integrale Bestandteile dieses Angebotes sowie des Beauftragungsvertrages.

6. Gültigkeit des Angebots

Dieses Angebot ist unter den genannten Bedingungen bis 15.10.2025 gültig.

Für die Fachhochschule Salzburg GmbH:

Puch/Salzburg, am

.....
Head of Research Department Creative Technologies
FH-Prof. Dr. Markus Tatzgern

.....
Projektleitung
FH-Prof. Dr. Clemens Havas