



Präzise Datenabgleichslösung zur Beschleunigung der Vertriebsanalyse für Darnytsia_

Datenabgleichslösung für Darnytsia gleicht Daten aus verschiedenen Quellen mit über 96 % Genauigkeit ab

Branche: Pharmazeutika **Standort: Ukraine**

Mitarbeitende: mehr als 1.000



Über den Kunden

Website:

<https://darnytsia.ua/en>

Das Pharmaunternehmen „Darnytsia“ ist der größte Hersteller von Arzneimitteln in der Ukraine und kann auf eine lange pharmazeutische Tradition zurückblicken. In seiner über 90-jährigen Geschichte hat Darnytsia seit mehr als 20 Jahren die führende Position auf dem heimischen Pharmamarkt inne, was durch die jährlichen Indikatoren der Industrie- und Marktaktivität des Unternehmens sowie die maßgeblichen Branchenbewertungen bestätigt wird. Bei der digitalen Transformation macht Darnytsia erhebliche Fortschritte und setzt erfolgreich marktführende Technologien ein.

Zusammenfassung



Ziele

Konzipieren Sie eine Lösung für den Abgleich von Datentabellen aus verschiedenen Quellen und die Anpassung der Daten an ein einziges Format für weitere Vertriebs- und Marktanalysen, um die Datenverarbeitung zu beschleunigen und das Risiko menschlicher Fehler zu verringern.



Lösung

Die Datenwissenschaftler von Infopulse haben einen Hybridalgorithmus für den Datenabgleich implementiert, der eine Treffergenauigkeit von über 96 % für Medikamentennamen und über 85 % für Apothekennamen und -adressen aufweist.



Vorteile

Automatisierter Datenabgleich ohne das Risiko von menschlichen Fehlern, schnellere Datenanalyse und Entscheidungsfindung, optimierte Datenverarbeitung mit weniger Arbeitsaufwand und Kosten.



Erbrachte Dienstleistungen

Innovationsdienste, Intelligent Business, Intelligente Automatisierung, Smart Insights.

Geschäftliche Herausforderung

Die Zusammenarbeit zwischen Darnytsia und Infopulse hat eine lange Tradition, da sie bereits zahlreiche Projekte zur digitalen Transformation gemeinsam durchgeführt haben. Nach einer erfolgreichen Implementierung von [einem GenAI-Datenanalyse-Bot](#) und einer Verkaufsprognoselösung in Zusammenarbeit mit Infopulse, wandte sich das Unternehmen mit der Bitte um Unterstützung an uns, um eine weitere Herausforderung im Zusammenhang mit den für die Vertriebs- und

Marktanalyse verwendeten Daten anzugehen:

- Abgleich der Arzneimittelinformationen zwischen der internen Datenbank des Unternehmens und externen Tabellen, die von den mit Darnytsia zusammenarbeitenden Apotheken stammen
- Abgleich der Namen und Adressen der Apotheken zwischen einer Tabelle aus der internen Datenbank und externen Quellen

Zunächst wurde der Datenabgleich manuell durchgeführt. Diese Vorgehensweise ist jedoch zeitaufwändig und birgt das Risiko menschlicher Fehler, die sich negativ auf die Ergebnisse auswirken können. Der Pharmagigant gab sich nicht mit einer vorgefertigten Lösung zufrieden, die er bereits getestet hatte, und wandte sich an Infopulse, um einen maßgeschneiderten Abgleichalgorithmus zu entwickeln.

Lösung und Unternehmenswert

Gemeinsam mit Darnytsia hat Infopulse einen **hybriden Abgleichsalgorithmus** entwickelt, der Datensätze zwischen Datentabellen (intern und extern) mit einem **Gesamtfehler von weniger als 5 % abgleicht**.

Der geschäftliche Nutzen eines solchen Outputs für Darnytsia umfasst Folgendes:

- Die **vollständige Automatisierung** des Datenabgleichs mit 20-mal weniger Zeitaufwand für die Mitarbeiter
- Einen **geringeren Zeit- und Kostenaufwand** für die entsprechenden Vorgänge: Der Abgleich von Zehntausenden von Datensätzen dauert **weniger als 10 Minuten**

- **85–96 % Treffergenauigkeit**
- **Kein Risiko für menschliche Fehler** aufgrund der Automatisierung und des Wegfalls manueller Tätigkeiten
- Eine **schnellere Datenanalyse**, Entscheidungsfindung und Erkenntnisgewinnung

Die Durchführung des Projekts dauerte zwei Wochen.

Technische Informationen

Von unserem Kunden erhielten wir zwei Excel-Datentabellen mit Arzneimittelinformationen (jeweils 300 bis 400 Datensätze) und drei Tabellen mit Apothekendaten (jeweils bis zu 10.000 Datensätze). Die Daten waren inkonsistent und in verschiedenen Formaten eingestellt, was die ursprünglich gestellte Aufgabe erschwerte. Es galt, eine Lösung zu entwickeln, die Übereinstimmungen zwischen Datensätzen findet und dabei unterschiedliche Formate und Tippfehler berücksichtigt.

Mithilfe dieser ersten Datensätze haben wir nach geeigneten Abgleichmetriken gesucht und verschiedene Ansätze und verfügbare Algorithmen getestet. Daraufhin hat Infopulse einen **hybriden Abgleichsalgorithmus** entwickelt, der nach Ähnlichkeiten zwischen zwei Datenstrings sucht und den gesamten Ähnlichkeitswert errechnet.

Dabei wurden die folgenden Metriken verwendet:

- Ratcliff-Ähnlichkeit
- Levenshtein-Abstand

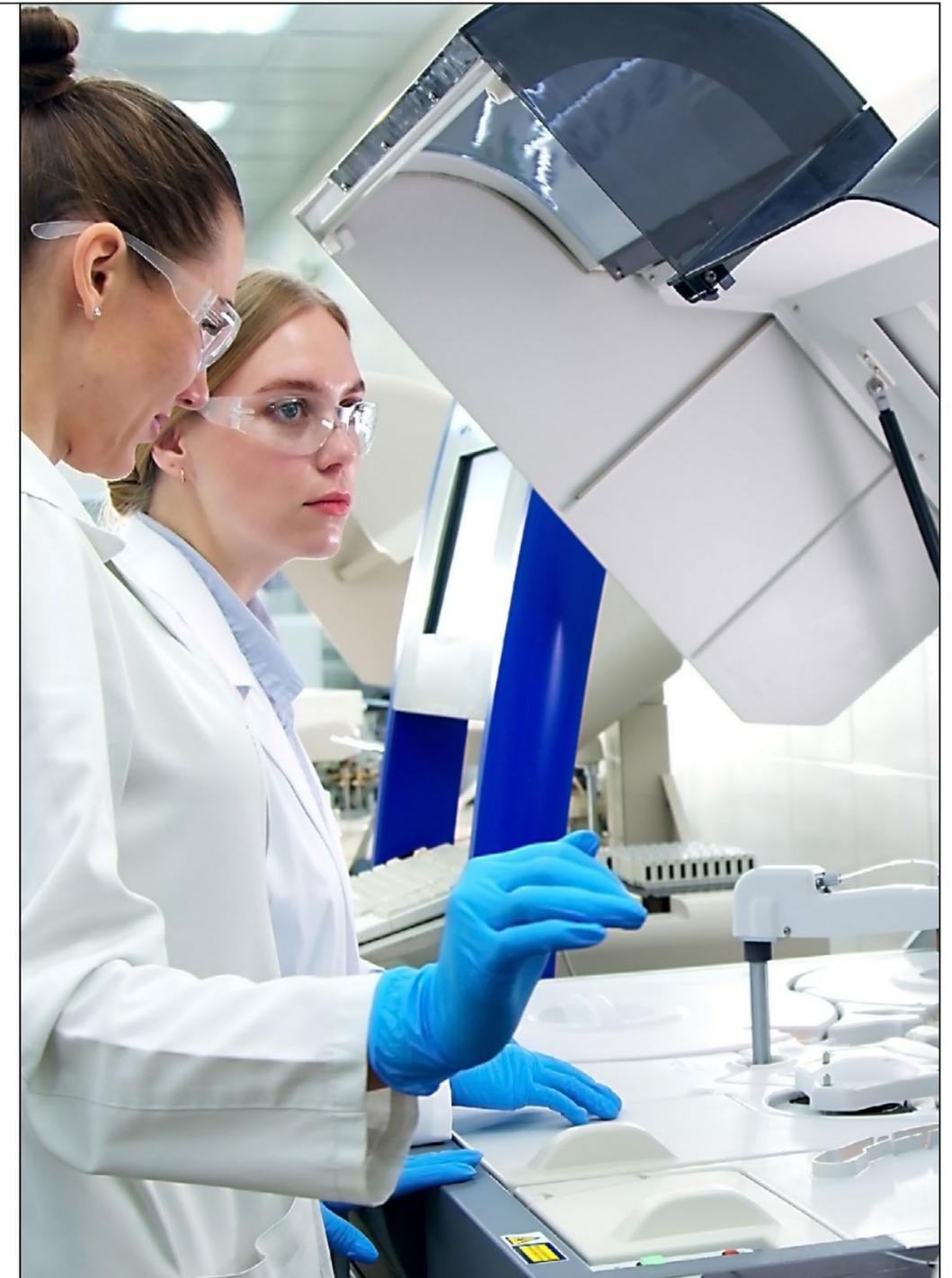
- TF-IDF (Begriffshäufigkeit-umgekehrte Dokumentenhäufigkeit)

Wie ursprünglich mit Darnytsia vereinbart, wird der Output in Form von Python-Code präsentiert, den unser Kunde bei Bedarf ausführen kann.

Die Tests des Abgleichsalgorithmus ergaben folgende Ergebnisse:

- **96 % Genauigkeit** beim Abgleich zweier Arzneimitteltabellen mit jeweils **300–400 Datensätzen**
- **über 85 % Genauigkeit** beim Abgleich von drei Apothekentabellen (Namen und Adressen abgeglichen) mit jeweils **10.000 Datensätzen**

Der Zeitaufwand für die Ausführung des Abgleichsalgorithmus liegt **unter einer Minute** für Arzneimitteltabellen und **unter 10 Minuten** für Apothekentabellen. Die Lösung ist nicht an einen bestimmten Datenumfang gebunden und kann für jedes Datenvolumen eingesetzt werden.



Verwendete Technologien



Python



Difflib



SequenceMatcher



RapidFuzz



Sklearn



Über Infopulse

Infopulse, ein Teil des führenden nordischen Unternehmens für digitale Dienstleistungen Tietoevry, ist ein internationaler Anbieter von Dienstleistungen in den Bereichen Solution Engineering, KI/ML, Cloud & IT Operations und intelligente Automatisierung für KMUs und Fortune-100-Unternehmen weltweit.

Bei Infopulse finden Sie intelligente Lösungen zur Automatisierung von Geschäftsprozessen, einschließlich Service Desk Support, für das Bauwesen, die Fertigung, die Logistik, die Landwirtschaft, den Einzelhandel und andere Branchen.

Viele namhafte Unternehmen vertrauen auf Infopulse, darunter Allianz Bank, BICS, Bosch, Credit Agricole, Delta Wilmar, ING Bank, LMT, Microsoft, Metinvest, Offshore Norge, OLX, OTP Bank, Santander, SAP, UkrSibbank BNP Paribas Group, Vodafone, Zeppelin und andere.

Für weitere Informationen besuchen Sie bitte www.infopulse.com/de

Kontaktieren sie uns

DE +49 (69) 505-060-4719

PL +48 (221) 032-442

US +1 (888) 339-75-56

UA +38 (044) 585-25-00

BG +359 (876) 92-30-90

NL +31 (70) 89-10-8534

 info@infopulse.com

