

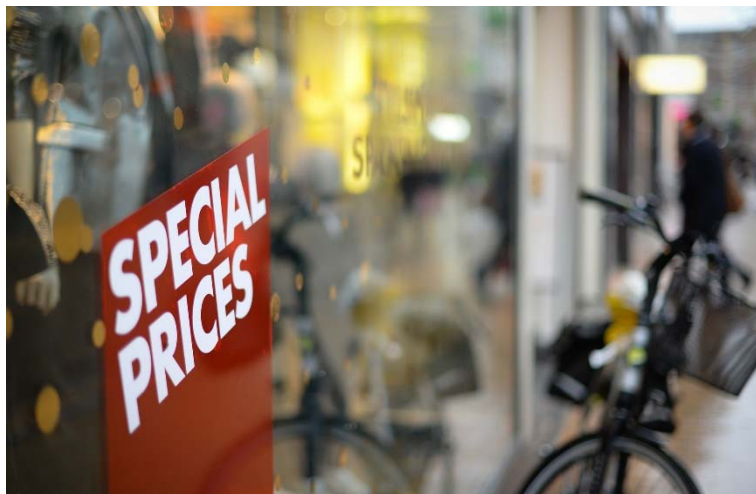
Mit dem SDSC-BW finden Produkte ihren geeigneten Platz.

Wie das SDSC-BW die richtigen Algorithmen für erfolgreiche Multilabel-Produktklassifizierung aufspürt.

Der Kunde

Die Price Intelligence GmbH (priceintelligence), Stuttgart, ist ein Anbieter für Marktüberwachungssoftware – mit Schwerpunkt auf der Optimierung von E-Commerce-Daten für Händler und Hersteller. Die Lösung erfasst und bewertet internationale Daten von Marktplätzen und Onlineshops. Preis- und Wettbewerbsbeobachtung, Dynamic Pricing sowie Sortimentsoptimierungen sind Features, von denen die Nutzer profitieren. www.priceintelligence.net

„Das Projekt mit dem SDSC-BW hat uns geholfen, unsere Ideen mit den richtigen Modellen in die Tat umzusetzen und den KI-Gedanken in der gesamten Gruppe voranzutreiben.“ - Sebastian Klumpp, Geschäftsführer der Price Intelligence GmbH.



priceintelligence

Die Daten

In modernen E-Commerce-Systemen fallen eine Vielzahl von Daten an – insbesondere, wenn Warenhäuser ihre großen Lager füllen und Produkte aus Lieferantenlisten in das hauseigene Produktinformationssystem integriert werden sollen. Auf dem Weg von der Lieferantenliste in das Product Information Management (PIM)-System werden Daten teilautomatisiert verarbeitet. Im Rahmen einer Potenzialanalyse untersuchte das SDSC-BW, wie der Prozess der Kategorisierung von Produkten in hierarchische Taxonomien innerhalb von PIM-Systemen vollautomatisiert werden könnte – und das auf Basis der Beschreibung nebst Produkttiteln. Als Datengrundlage stellte priceintelligence eine komplette PIM-Struktur sowie eine breite Masse an Lieferantenlisten zur Verfügung.

Die Herausforderung

Der Prozess der Multilabel-Produktklassifizierung unter Zuhilfenahme von aktuellen Algorithmen der natürlichen Sprachverarbeitung (NLP) benötigt Datensätze mit einer signifikanten Informationsgüte. Wichtig ist auch eine ausgewogene Verteilung über die zur Verfügung stehenden Kategorien, um Modelle für eine Vorhersage im Rahmen der Klassifizierung effizient zu trainieren. Die Herausforderung bei

priceintelligence bestand darin, einen vollautomatisierten Prozess für die Produktklassifizierung zu erarbeiten, der auch bei weniger ausgewogenen Verteilungen funktioniert. Parallel um priceintelligence-Entwicklerteam mussten die SDSC-BW Experten verschiedene Strategien verfolgen, wie Daten und Modelle bestmöglich auf- und verarbeitet werden können, um die Rate der korrekten Klassifizierungen zu erhöhen.

Die Potentialanalyse

Im Rahmen der kostenlos durchgeführten Potenzialanalyse untersuchten die Experten des SDSC-BW detailliert das PIM-System sowie die Lieferantendaten. Der Fokus der Analyse lag auf der Identifikation von textuellen Informationen mit signifikantem Informationsgehalt und deren Auftreten nebst Gewichtung. So konnte das Team Rückschlüsse auf Modellgenauigkeiten ableiten und eine geeignete Modellauswahl treffen.

Das Resultat

Auf Grundlage, der in der Potenzialanalyse untersuchten, Algorithmen bereitete das Expertenteam eine ideale Datenbasis mit maximal möglichem Informationsgehalt auf und trainierte drei Modelle für den Einsatz in der sogenannten flachen Produktklassifizierung. Durch die enge Zusammenarbeit der Experten und des Entwicklerteams und die damit verbundene Kopplung des Know-hows konnte priceintelligence zeitnah einen erfolgreichen und soliden Prozess für die automatische Produktklassifizierung implementieren.

Wie es weiter geht

Bereits während der Potenzialanalyse des SDSC-BW zeigte sich, wie interessant beispielsweise der Einsatz neuronaler Netze unter Verwendung von Gewichten ist, gerade wenn unbalancierte Daten in großen Mengen vorliegen. Nach Abschluss der Potenzialanalyse übermittelten die Spezialisten des SDSC-BW die Ergebnisse nebst weiterführenden Empfehlungen an die Datenexperten von priceintelligence. Aus letzteren gipfelte ein gemeinsamer Projektantrag, um eine Produktentwicklung anzustoßen, die den automatischen Klassifizierungsprozess um die Erkennung von Bildern erweitert und die Verarbeitung von unbalancierten sowie unvollständigen Daten optimiert. Darüber hinaus sollen gemeinsam ermittelte wissenschaftliche Leistungskennzahlen zukünftig im Rahmen von Folgeprojekten genutzt werden.

Das Smart Data Solution Center Baden-Württemberg

„Zusammenhänge erkennen. Potentiale nutzen.“ Unter diesem Motto ermöglicht das SDSC-BW kleinen und mittelständischen Unternehmen die Nutzung von Smart Data- Technologien. Und das als neutrale, unabhängige und vom Land Baden-Württemberg geförderte Institution.

In unseren kostenlosen Potentialanalysen zeigen wir Vorteile und Nutzungsmöglichkeiten von Daten – die in jedem Unternehmen jeder Branche anfallen- und helfen, diese zu bewerten.

Wie wertvoll sind Ihre Daten?

Finden Sie es heraus unter www.sdsc-bw.de