

Einsatz von Machine Learning Modellen in der Authentizitätskontrolle von Fruchtsäften

Wirtschaftszweig: Lebensmittel- und Getränkeindustrie

Anwendungsfall und Projektziel

Dieser Anwendungsfall konzentriert sich auf die Lebensmittelauthentizität und -sicherheit, genauer gesagt auf die Erkennung von Verfälschungen in der Fruchtsaftindustrie. Derzeit wird diese Analyse größtenteils manuell von qualifizierten Experten durchgeführt. Die Experten stützen sich stark auf ihre Erfahrung, und ihre Zahl ist begrenzt. Sie sind in erster Linie für die Erstellung einer Referenzdatenbank zuständig, die später für die Authentifizierung von Fruchtsäften verwendet wird. Während einfache univariate Analysen bereits durchgeführt werden können, ist die Ermittlung komplexer Zusammenhänge zwischen mehreren Variablen noch ein unerforschtes Gebiet. Der Einsatz von maschinellem Lernen zur Ermittlung von Mustern und Zusammenhängen zwischen mehreren Variablen kann neue Perspektiven für die Bestimmung der Echtheit von Fruchtsäften eröffnen, da dieser Ansatz schneller und genauer ist und sich an die im Laufe der Zeit auftretenden Schwankungen anpassen kann.

Betrüger passen ihre Methoden ständig an, um ihre Gewinne zu steigern und gleichzeitig einer Aufdeckung zu entgehen. Viele Unternehmen, die Geräte und Verfahren zur Erkennung von Verfälschungen in Lebensmitteln entwickeln, verlagern ihren Fokus häufig auf andere profitable Branchen wie die Weinindustrie, wodurch die Saftindustrie noch anfälliger für solche betrügerischen Handlungen wird. Vor diesem Hintergrund möchte die SGF Service Plus GmbH künstliche Intelligenz (KI) einführen, die anhand umfangreichen Referenzdaten aus zuvor authentifizierten Verfahren trainiert werden kann. Ein gut trainiertes Modell kann Muster und Zusammenhänge zwischen verschiedenen Parametern erkennen, die von menschlichen Prüfern zuvor möglicherweise übersehen wurden. In nachfolgenden Phasen können die Modelle mithilfe zeitreihenbasierter Analysen selbst sehr subtile Veränderungen in den Mustern erlernen, die durch Faktoren wie Klimaveränderungen oder Änderungen der Anbaumethoden entstehen können. Diese winzigen Schwankungen sind selbst für den erfahrensten Fachmann schwer zu erkennen und zu dokumentieren.

Lösungsansatz und Herausforderungen

Die größte Herausforderung bei der Entwicklung der Lösung ist die enorme Produktvielfalt in der Saftindustrie. Säfte können aus zahlreichen Obstsorten gewonnen werden, und jede Obstsorte weist unterschiedliche Sorten und Herkunftsorte mit jeweils eigenen Profilen auf. Hinzu kommen viele externe Faktoren wie die Erntemethode, Höhenunterschiede der Anbauflächen, die Unvorhersehbarkeit der einzelnen Erntesaisonen (trockener Sommer vs. feuchter Sommer) und unterschiedliche Verarbeitungstechniken, die das Profil eines Saftes verändern können und selbst für erfahrene Fachleute manchmal schwer zu erfassen sind.

Angesichts der vielfältigen Herkunft der Fruchtsäfte stellt auch die Variabilität der Parameter eine Herausforderung dar. In den Referenzdaten des Unternehmens ist nicht für jedes Muster die gleiche Anzahl an Parametern vorhanden. Manchmal ist der Datensatz zudem sehr lückenhaft. Daher muss man sich auf jene Parameter konzentrieren, für die ausreichend Daten vorliegen. Techniken wie die Datenimputation können hier hilfreich sein.

gefördert vom

Die Datenaufbereitung wird eine große Rolle für das Ergebnis spielen. Bei ausreichender Datenqualität sind Modelle des maschinellen Lernens sehr gut darin, Muster zu klassifizieren und verborgene Muster zu erkennen. Die entwickelten Modelle können das NMR-Spektrogramm von Säften analysieren und bestimmte Aspekte einer Produktdeklaration (z. B. Herkunft, Produktkategorie, Obstsorte) überprüfen, sowie die Zugabe nicht deklarerter Obstsorten (z. B. Äpfel oder Pfirsiche in Aprikosen, Trauben in Granatäpfeln) erkennen. Durch die präzise Berechnung des Fruchtanteils mithilfe der Modelle können Experten indirekt den Zuckerzusatz erkennen. Die automatische Aktualisierung der Modelle auf Basis der ständig wachsenden Datenbank, einschließlich der automatischen Validierung und Bewertung der Modellqualität, ist ebenfalls in diesem Projekt geplant.

Potenziale

Die SGF Service Plus GmbH verfügt über eine der größten Datenbanken der Branche (rund 5 Millionen Datenpunkte), die seit 1992 aufgebaut wurde. Das Unternehmen kooperiert mit SGF International e.V., der über ein Budget von ca. 800.000 Euro jährlich verfügt, um die Datenbank mithilfe von Analysedaten akkreditierter Labore zu erweitern und zu aktualisieren. Es wird bereits in den Aufbau einer Infrastruktur investiert, die die Erstellung und Aktualisierung von Modellen für maschinelles Lernen in einer Azure MLOps-Umgebung unterstützt. Dazu gehören ergänzende Programmierungen der Datenbank, um Machine-Learning-Modelle in den Routinebetrieb zu integrieren. Für dieses Projekt wurde die Unterstützung durch erfahrene Fachleute zugesagt. Insgesamt sind die Chancen für eine erfolgreiche Umsetzung der Lösung hier sehr hoch.

Nach der Implementierung kann diese Lösung über eine webbasierte Plattform Dritten angeboten werden. Die entwickelten Modelle können auch auf andere Bereiche der Lebensmittelindustrie (z. B. Honig, Gewürze, Fisch) übertragen werden, in denen ein hohes Risiko der Verfälschung besteht. Das Unternehmen kann diese Modelle für eine Zusammenarbeit mit Universitäten und Forschungsinstituten für Forschungszwecke zur Verfügung stellen.

Umsetzendes KMU	Forschungspartner
SGF Service Plus GmbH Marie-Curie-Ring 10a 55291 Saulheim	Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI) Smart Data&Knowledge Services Dept. https://www.dfki.de/web

Weitere Informationen zu KI4KMU-RLP finden sie unter: www.ki4kmu-rlp.de

gefördert vom