

## **Maschinelles Auffinden und Klassifizieren von risikobehafteten Produkten zur Unterstützung der Marktüberwachung**

David SCHNURA, Marie PENDZICH, Tobias BLEYER

*Gruppe Grundsatzfragen der Produktsicherheit  
Bundesanstalt für Arbeitsschutz und Arbeitsmedizin  
Friedrich-Henkel-Weg 1-25, D-44149 Dortmund*

**Kurzfassung:** Aufgabe der Marktüberwachungsbehörden ist es u. a., im Internet gehandelte Produkte, von denen ein Risiko für die Verbraucher ausgeht, aufzufinden und zu verhindern, dass diese auf dem europäischen Markt verbleiben. Durch die Vielzahl der Produkte und die komplexen Vertriebs- und Lieferstrukturen müssen automatisierte Analyseverfahren entwickelt werden, um diese Aufgabe effektiv und effizient erfüllen zu können. Eine Möglichkeit stellt der Einsatz von Big-Data-Methoden zur Identifikation von gefährlichen Produkten anhand von Kundenrezensionen dar. Dieser Beitrag beschreibt eine Machbarkeitsstudie mit der die Eignung von maschinellen Lernverfahren in diesem speziellen Kontext überprüft wird.

**Schlüsselwörter:** Marktüberwachung, risikobehaftete Produkte, maschinelle Lernverfahren, Big-Data-Methoden, Kundenrezensionen

### **1. Ausgangssituation**

Sichere Produkte sind Grundvoraussetzungen für sichere Arbeit. Viele für die Arbeitswelt relevanten Produkte und deren technische Sicherheit werden im Produktsicherheitsgesetz (ProdSG) und den zugehörigen Verordnungen geregelt. Ein Produkt darf demnach nur auf dem Markt bereitgestellt werden, wenn es bei bestimmungsgemäßer und vorhersehbarer Verwendung die Sicherheit und Gesundheit von Personen nicht beeinträchtigt. Durch Globalisierung und digitale Vernetzung ist es möglich, jede Art von Produkt – vom Minibagger bis zur Atemschutzmaske – weltweit online zu vertreiben und zu kaufen. Dadurch erreichen den europäischen Markt auf unterschiedlichen Vertriebswegen eine Fülle von Produkten mit hohen Qualitätsschwankungen und teilweise gefährlichen Eigenschaften. Die Aufgabe der Marktüberwachungsbehörden besteht u. a. darin, diese Produkte aufzufinden und zu verhindern, dass sie auf dem Markt verbleiben. Die komplexen Vertriebs- und Lieferstrukturen des globalen Onlinehandels erschweren es den Marktüberwachungsbehörden jedoch die Sicherheits- und Gesundheitsschutzanforderungen des ProdSG zu kontrollieren. Die effiziente Identifikation gefährlicher Produkte im Internet setzt einen systematischen Ansatz voraus und ist ohne automatisierte Analysen nahezu unmöglich. Big-Data-Technologien könnten solch einen Ansatz darstellen. Mit Hilfe neu entwickelter Softwarelösungen ermöglichen sie die systematische Erfassung, Auswertung und Darstellung großer Datenmengen. Dieser Beitrag fokussiert sich auf die grundlegenden Schritte einer Machbarkeitsstudie, mit der zunächst die generelle Eignung von Big-Data-Methoden zur Unterstützung der Marktüberwachungsbehörden festgestellt werden soll.

## 2. Vorgehensweise

Ziel der Machbarkeitsstudie ist es, einen Algorithmus zu entwickeln, der Informationen über gefährliche Produkte aus dem Internet filtert, analysiert und sortiert. Als Grundlage dienen Online-Kundenrezensionen in denen positive und negative, aber auch gefährliche Eigenschaften von Produkten beschrieben werden. Mit Hilfe des Algorithmus, der Verfahren aus der künstlichen Intelligenz nutzt, sollen anhand der Rezensionen Produkte identifiziert werden, die den Verbraucher mit hoher Wahrscheinlichkeit gefährden. Der Algorithmus ordnet die Produkte den Risikoklassen „Mechanische Gefahr“, „Brandgefahr“, „Chemische Gefahr“ und „Elektrische Gefahr“ zu. Innerhalb der Klassen erfolgt eine weitere Priorisierung der Produktrisiken, um eine schnellere Interpretation der Ergebnisse zu ermöglichen. In der Machbarkeitsstudie werden drei Verfahren zur Textklassifizierung (Support-Vector-Maschine, datenloses flaches Co-Klassifizieren und datenlose hierarchische Textklassifizierung) gegeneinander getestet, um herauszufinden, welches die besseren Klassifizierungsergebnisse erzielt. Die Verfahren basieren auf dem Prinzip des maschinellen Lernens. Maschinelle Lerner verwenden Algorithmen, um unbekannte Datensätze einer bestimmten Klasse zuzuordnen. Sie generieren Wissen aus Erfahrung. Dazu müssen die Lerner zunächst mit bekannten, bereits klassifizierten (gelabelten) Datensätzen trainiert werden. Die Güte des Klassifizierungsergebnisses ist im Normalfall abhängig von der Anzahl der gelabelten Datensätze und benötigt eine breite Datenbasis. Da häufig gelabelte Daten fehlen, wurden sogenannte datenlose Verfahren entwickelt, um dennoch eine effiziente Klassifizierung vornehmen zu können. Auch im Fall der Klassifizierung von risikobehafteten Produkten in Onlineshops fehlt dieses Wissen, weshalb datenlose Verfahren implementiert und mit einem klassischen Lerner verglichen werden. Im Folgenden werden die verwendete Datenbasis, die Funktionsweise der Verfahren und die geplante Evaluation der Klassifizierungsergebnisse beschrieben.

## 3. Datenbasis

Um gefährliche Produkte identifizieren, in eine Risikoklasse einteilen und eine erste Priorisierung vornehmen zu können, müssen große Datenmengen für unterschiedliche Aufgaben aufbereitet werden. Die benötigten Daten können in drei Kategorien eingeteilt werden:

**Kategorie I:** Daten, die es ermöglichen, Produkte einer Risikoklasse zuzuordnen. Als Quelle dienen Kundenrezensionen von Onlinehändlern. Um eine Verfälschung der Ergebnisse durch manipulierte / gekaufte Rezensionen zu minimieren, werden Rezensionen aus verifizierten Käufen bevorzugt. Ebenfalls relevant ist die Bewertung des Produktes durch den Kunden (meist in Form einer Sternebewertung), um Korrelationen zwischen Bewertung und Risiko angeben zu können. Anhand von produktbezogenen Daten (Produktbezeichnung, Hersteller, Link der Verkaufsseite) kann das Produkt später eindeutig identifiziert werden.

**Kategorie II:** Daten, die als Trainingsdaten für den maschinellen Lerner dienen. Da in diesem Projekt verschiedene Verfahren getestet werden, unterscheidet sich dieser Datensatz je nach verwendetem Lernansatz.

**Kategorie III:** Daten, die für die abschließende Evaluation der Klassifizierungsergebnisse benötigt werden.

Die Datensätze der Kategorien II und III werden händisch erstellt. Als Grundlage dienen Produktinformationen, die über das europäische Schnellwarnsystem RAPEX für gefährliche Verbraucherprodukte mit ernstem Risiko gemeldet wurden. In den jeweiligen Meldungen werden das Produkt, die Risikoklasse und das vom Produkt ausgehende Risiko durch europäische Marktüberwachungsbehörden fachkundig beschrieben und klassifiziert. Die dadurch entstandenen gelabelten Datensätze können als Trainingsdaten oder zur Evaluation genutzt werden.

#### **4. Datenbasis: Herausforderungen und Lösungen**

Kundenrezensionen werden von Privatpersonen verfasst. Dadurch ergeben sich viele Rezensionsvarianten, die sich zudem durch einfache Rechtschreibfehler, grobe grammatikalische Fehler oder durch Schreibweisen in Dialekten unterscheiden können. Diese anfänglich als problematisch angesehene Tatsache hat sich im Projektverlauf als vernachlässigbar herausgestellt. Der maschinelle Lerner erkennt häufig falsch geschriebene Wörter als eigenständige Begriffe an und fügt diese, wie die korrekten Begriffe, der Menge der „feuernden Ausdrücke“ hinzu. In der Klasse „mechanische Gefährdung“ sind typische feuernde Ausdrücke beispielsweise Schnitt, Quetschung, fallen, Bluterguss. Eine weitere Herausforderung stellt der Umgang mit „Geschichten“ innerhalb von Kundenrezensionen dar, z. B. „Meine letzte Mikrowelle hatte einen Kurzschluss, deshalb brauchte ich eine neue.“ In diesem Beispiel bezieht sich der Kunde nicht auf das aktuell zu rezensierende Produkt. Das Verfahren kann dies allerdings nicht erkennen und betrachtet nur den feuernden Ausdruck „Kurzschluss“ und ordnet die Mikrowelle fälschlicherweise in die Risikoklasse „elektrische Gefahr“ ein. Eine Möglichkeit dieses unerwünschte Verhalten zu umgehen, besteht darin, die (Sterne-) Bewertung des Produktes durch den Kunden einzubeziehen. Produkte mit hohen Bewertungen und hoher Zufriedenheit, deren Rezension jedoch auf Grund eines feuernden Ausdrucks als negativ eingestuft wurde, könnten dann von der Überprüfung ausgeschlossen werden. Sollte das ursprünglich rezensierte Produkt allerdings durch Nichtgefallen eine niedrige Sternbewertung vom Kunden erhalten haben, würde es trotzdem fälschlicherweise als gefährliches Produkt klassifiziert. In diesem Fall sollte ein Abgleich mit anderen Rezensionen für dieses Produkt erfolgen, um die Fehlerrate weiter zu senken. Wurde das Produkt in diesen Rezensionen wesentlich besser bewertet, steigt die Wahrscheinlichkeit eine Rezension mit „Geschichte“ zu erkennen.

#### **5. Verwendete Verfahren zur Klassifikation der Daten**

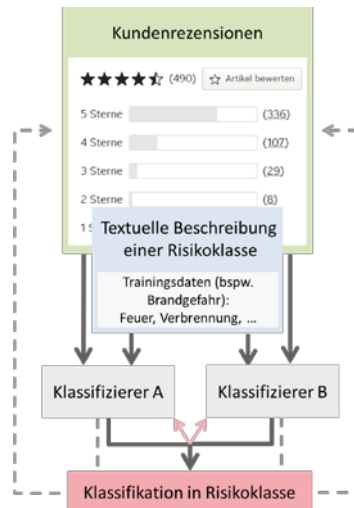
Da die Forschung zur automatischen Textklassifikation bereits weit fortgeschritten ist, existieren sehr viele Verfahren, um Textdokumente automatisch klassifizieren zu können. Allerdings wird bei vielen dieser Verfahren vernachlässigt, dass oft keine oder nur sehr wenige gelabelte Daten - wie auch in diesem Projekt - zur Verfügung stehen, mit denen der maschinelle Lerner ausreichend trainiert werden könnte. Zwei vielversprechende Verfahren, die diesen Umstand berücksichtigen, stellen Chang et. al. (2008) und Song & Roth (2014) vor. Die von ihnen entwickelten Verfahren benutzen ein Co-Training und ein hierarchisches Training zur Textklassifikation. Diese Verfahren benötigen keine gelabelten Daten, sondern versuchen die Klassen durch eine textuelle Beschreibung zu verstehen und aktiv mit Inhalt zu füllen. Die tex-

tuelle Beschreibung beinhaltet neben dem Klassennamen, erste feuernde Ausdrücke, wie beispielsweise „Feuer“ oder „starke Wärmeentwicklung“. Auf diese Weise wird eine Basis zum Klassifizieren geschaffen. Darauf aufbauend werden nacheinander alle Dokumente bzw. Rezensionen klassifiziert und die Menge der feuernden Ausdrücke wird ergänzt. Wenn eine Kundenrezension beispielsweise den bereits bekannten Ausdruck „starke Wärmeentwicklung“ beinhaltet und gleichzeitig der Begriff „Rauchentwicklung“ verwendet wird, soll dieser erkannt und den feuernden Ausdrücken hinzugefügt werden. Um die Qualität der „datenlosen“ Verfahren vergleichen zu können, wird ein klassischer maschineller Lerner - eine Support-Vector-Machine - ebenfalls trainiert und evaluiert.

Der Erfolg jedes dieser Verfahren, insbesondere für die datenlosen Ansätze, ist sehr von der ihnen zugrunde liegenden Textrepräsentation abhängig. Diese Repräsentationen spiegeln dabei ein Dokument wider, indem sie diesem eine mathematische Struktur zuweisen. Um dies zu realisieren, erfolgt die Darstellung der Dokumente, im Projektfall der Kundenrezensionen, üblicherweise in Form von Vektoren. Um eine möglichst hohe Qualität der Klassifizierung sicherzustellen, werden zu jedem Lernverfahren verschiedene, sinnvolle Repräsentationen implementiert und evaluiert.

### 5.1 Datenloses flaches Co-Training

Ein datenloser Ansatz zum Klassifizieren ohne Trainingsdaten wird von Chang et al. (2008) beschrieben. Sein Ziel ist es, die Klassen zu verstehen, indem typische Eigenschaften der Klassen händisch gefunden und als Grundlage genutzt werden, um eine effiziente Klassifizierung, ohne gelabelte Daten, zu betreiben. Dazu verwenden Chang et. al. einen flachen Co-Lerner. Flach bedeutet, dass vorhandene Hierarchien in den Klassen nicht berücksichtigt werden. Die Grundlage der Co-Lerner bilden unabhängige Textrepräsentationen. So werden in dem hier beschriebenen Projekt die Repräsentationen Bag of Words (BOW) und Explicit Semantic Analysis (ESA) genutzt. Während BOW lediglich Wörter zählt und für jedes Wort einen individuellen Vektor erstellt, benutzt eine ESA die Texte der Webseite Wikipedia als externe Wissensbasis. Da noch weitere externe Datenquellen mit einem stärkeren Projektbezug zur Verfügung stehen, wird auf diesen Daten aufbauend eine weitere ESA durchgeführt. Untersucht wird, ob sich die Qualität der Klassifizierer (durch verschiedene Datenquellen) im Vergleich zu Wikipedia verbessern lässt. Das Verfahren arbeitet wie folgt: Zunächst werden zwei leere Trainingsmengen erstellt und jeweils mit der Grundlage der textuellen Klassenbeschreibung in BOW bzw. ESA befüllt. Daraufhin werden zwei einfache Klassifizierer entwickelt, welche entweder mit den BOW-Daten oder den ESA-Daten lernen. Nun werden die im Rahmen der Datensammlung erfassten Rezensionen nacheinander geprüft. Sollten beide Klassifizierer das Dokument einer Risikoklasse zuteilen, wird die ESA- und BOW-Repräsentation der Menge an entscheidungskritischen Ausdrücken hinzugefügt. Die Rezension erweitert dann mit eigenen Ausdrücken die ursprüngliche Grundmenge. Sollte nur einer der Lerner die Rezension nicht akzeptieren, wird diese in eine Warteschlange abgelegt und nach dem nächsten neuen Dokument erneut geprüft. Dies geschieht, da ein feuernder Ausdruck möglicherweise noch nicht gefunden wurde. Nun wiederholt sich dieser Vorgang, bis keine neuen Dokumente mehr hinzuzufügen sind. Dabei folgt der Klassifizierer dem Prinzip „one-vs-rest“ und entscheidet so lediglich, ob das untersuchte Dokument in der Klasse, beispielsweise „Brandgefahr“, enthalten ist oder nicht. Folglich muss dieses Verfahren für alle Risikoklassen angewendet werden.

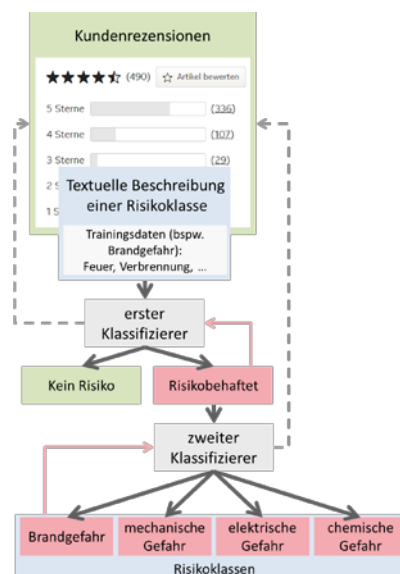


Nachdem beide Klassifizierer auf einer textuellen Beschreibung der Risikoklasse trainiert wurden, durchlaufen alle Kundenrezensionen beide Klassifizierer. Eine Klassifikation in die Risikoklasse erfolgt nur, wenn beide Klassifizierer das Produkt als risikobehaftet einteilen. Nur durch einen Klassifizierer eingeteilte Rezensionen durchlaufen den Prozess zu einem späteren Zeitpunkt erneut. Durch die einstimmige Zuteilung werden neue Trainingsdaten generiert.

Abbildung 1: Schematische Darstellung des Co-Trainings .

## 5.2 Datenloses hierarchisches Training zur Textklassifizierung

Ein weiteres Verfahren beschreiben Song & Roth (2014). Ebenso wie Chang et. al. (2008) nutzen sie einen datenlosen Klassifizierungsansatz mit einer textuellen Klassenbeschreibung. Sie berücksichtigen jedoch die hierarchische Struktur der Klassen. Weiterhin hat dieses Verfahren den Vorteil, dass nicht das Prinzip „one-vs-rest“ genutzt werden muss, d. h. alle Dokumente werden den unterschiedlichen Klassen in nur einem Durchlauf zugeordnet. Dieses Verfahren nutzt die hierarchischen Beziehungen der Klassen und bindet diese in eine Baumstruktur ein. Die Klassen dieses Projektes bilden einen flachen Baum mit nur zwei Ebenen. Auf der ersten Ebene teilt sich der Baum in die Kindknoten „Kein Risiko“ und „Risikobehaftet“ auf. In der zweiten Ebene folgt eine Aufteilung in die Risikoklassen. So umfasst der Knoten „Risikobehaftet“ die Kindknoten „Brandgefahr“, „Mechanische Gefahr“, „Elektrische Gefahr“ und „Chemische Gefahr“.



Mit Hilfe einer Baumstruktur, klassifizieren zwei Klassifizierer auf zwei Ebenen die Rezensionen. Dabei durchlaufen zu nächst alle Rezensionen die Baumstruktur. Die n vielversprechendsten Rezensionen werden zum Trainieren genutzt. Nicht eingeteilte Rezensionen durchlaufen den Prozess zu einem späteren Zeitpunkt erneut. Dieser Vorgang wiederholt sich, bis keine Rezensionen mehr übrig sind.

Abbildung 2: Schematische Darstellung des hierarchischen Lerners.

Ist die notwendige Baumstruktur erstellt, wird für jeden Knoten mit Verzweigungen ein Klassifizierer entwickelt, welcher zunächst mit einer Repräsentation der textuellen Klassenbeschreibung lernt. Nun durchlaufen alle Rezensionen die Baumstruktur per „top-down“ Ansatz. Dabei wird jedes Dokument einem Knoten zugewiesen, sodass nach dem Durchlauf jeder Knoten eine Liste an zugeordneten Dokumenten besitzt. Diese werden nach der richtigen Zuordnungsquote (Konfidenz) sortiert und die n Dokumente mit der höchsten Konfidenz werden zum Lernen genutzt. Dieser Schritt wiederholt sich daraufhin mit den übrigen Rezensionen, bis keine mehr vorhanden sind.

## 6. Ausblick: Evaluation und Priorisierung der Klassifikationen

In einem nächsten Projektschritt erfolgt die Evaluierung der Klassifizierungen. Es sollen besonders „true-positive“ (korrekt klassifizierte gefährliche Produkte), „false-positiv“ (sichere Produkte, die als gefährlich klassifiziert wurden) und „false-negativ“ (gefährliche Produkte, die nicht klassifiziert wurden) überprüft werden. Die Überprüfung erfolgt mit Hilfe der Testdaten der Kategorie III. Um eine schnellere Interpretation der Ergebnisse zu ermöglichen, erfolgt eine weitere Priorisierung der Daten, die über die Sortierung anhand der Konfidenz hinausgeht. Die Notwendigkeit einer solchen Sortierung verdeutlicht das folgende Beispiel: In der Klasse „chemische Gefährdung“ werden 500 Produkte gefunden. Die Produkte „Gartenzweig klassisch“ und „Babyspielzeug Rassel“ sind zwei davon. Die Konfidenz des Gartenzweigs liegt bei 94 %, während die Konfidenz der Babyrassel bei „nur“ 91 % liegt. Im Falle einer einfachen Sortierung der Konfidenzen würde dem Gartenzweig eine höhere Priorität zugewiesen. Da die Wahrscheinlichkeit einer Vergiftung bei einer Babyrassel allerdings deutlich höher ist, muss diese vorrangig geprüft werden. Dies kann erreicht werden, indem für jede Klasse eine interne Prioritätenliste erstellt wird. So kann der Kategorie Kinderspielzeug eine höhere Gewichtung zugeteilt werden.

Ein langfristiges Ziel nach der Evaluierung der Klassifizierungen und der damit einhergehenden Identifikation des geeignetsten Verfahrens, ist die Anpassung und Erprobung dieses Verfahrens in weiteren Online-Shop-Systemen.

## 7. Literatur

- Chang, M., Ratinov, L., & Roth, D. (2008). Importance of Semantic Representation: Dataless Classification, Department of Computer Science, University of Illinois at Urbana-Champaign. Paper presented at the Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence.
- Song, Y., & Roth, D. (2014). On Dataless Hierarchical Text Classification. Department of Computer Science. University of Illinois at Urbana-Champaign.





Gesellschaft für  
Arbeitswissenschaft e.V.

**ARBEIT(s).WISSEN.SCHAF(F)T**  
Grundlage für Management & Kompetenzentwicklung

64. Kongress der  
Gesellschaft für Arbeitswissenschaft

FOM Hochschule für  
Oekonomie & Management gGmbH

21. – 23. Februar 2018

---

**GfA Press**

---

**Bericht zum 64. Arbeitswissenschaftlichen Kongress vom 21. – 23. Februar 2018**

**FOM Hochschule für Oekonomie & Management**

Herausgegeben von der Gesellschaft für Arbeitswissenschaft e.V.

Dortmund: GfA-Press, 2018

ISBN 978-3-936804-24-9

NE: Gesellschaft für Arbeitswissenschaft: Jahresdokumentation

Als Manuskript zusammengestellt. Diese Jahresdokumentation ist nur in der Geschäftsstelle erhältlich.

Alle Rechte vorbehalten.

© **GfA-Press, Dortmund**

**Schriftleitung: Matthias Jäger**

im Auftrag der Gesellschaft für Arbeitswissenschaft e.V.

Ohne ausdrückliche Genehmigung der Gesellschaft für Arbeitswissenschaft e.V. ist es nicht gestattet, den Kongressband oder Teile daraus in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) zu vervielfältigen.

Die Verantwortung für die Inhalte der Beiträge tragen alleine die jeweiligen Verfasser; die GfA haftet nicht für die weitere Verwendung der darin enthaltenen Angaben.

**USB-Print:**

Prof. Dr. Thomas Heupel, FOM Prorektor Forschung, [thomas.heupel@fom.de](mailto:thomas.heupel@fom.de)

**Screen design und Umsetzung**

© 2018 fröse multimedia, Frank Fröse

[office@internetkundenservice.de](mailto:office@internetkundenservice.de) · [www.internetkundenservice.de](http://www.internetkundenservice.de)