

# Optimiertes Risk & Fraud Management im E-Commerce

Wie Machine Learning die manuelle Prüfung revolutioniert – ein Praxisbeispiel eines internationalen Online-Händlers

# Abstract

Seit 2020 erlebte der E-Commerce **eine beeindruckende Entwicklung**, die durch die Corona-Pandemie noch beschleunigt wurde.

Durch die Schließung physischer Geschäfte fokussierten sich viele Verbraucher auf das Online-Shopping, was zu einem **exponentiellen Anstieg** des **Bestellaufkommens** und damit auch zu einer Zunahme potenzieller Betrugsversuche führte. Die Bedeutung der **Betrugsprävention** stieg erheblich. Um verdächtige Aktivitäten zu identifizieren und zu verhindern, nutzen viele Unternehmen neben der manuellen Prüfung bereits automatisierte Verfahren wie Machine-Learning-Modelle. Die manuelle Prüfung ist zeit- und kostenaufwendig und konfrontiert die Unternehmen mit einem hohen Bedarf an qualifizierten Prüfern.

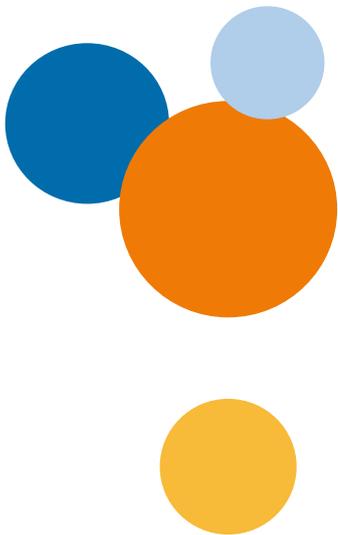
Durch den Einsatz von **Machine-Learning-Modellen** können Routineaufgaben automatisiert werden, was den manuellen Prüfaufwand erheblich reduziert und die Effizienz steigert. Ein Hybridansatz, bei dem automatisierte Systeme Prüfungen durchführen und nur komplexe Fälle an menschliche Prüfer weiterleiten, kann besonders effektiv sein. Dies ermöglicht es, Ressourcen gezielt einzusetzen und sich auf komplexe Fälle zu konzentrieren, die menschliches Urteilsvermögen erfordern.

Der deutsche Informationsdienstleister CRIF, der zur globalen CRIF-Gruppe mit Kunden in über 50 Ländern gehört, verfügt über 130 Jahre Erfahrung mit **datengetriebenen Risikolösungen** im B2B- bzw. B2C-Bereich. Das CRIF-Datenuniversum beinhaltet Informationen zu allen relevanten Aktionen im Lebenszyklus von Kunden,

nutzt **umfangreiche Datenquellen** zu Device-Informationen, Geo-Locations, Zahlungsdaten und Identitätsdaten und verarbeitet Profilverläufe und Muster, um Abweichungen identifizieren zu können. Mithilfe dieser Datenbasis hat CRIF bereits erfolgreich Machine-Learning-Modelle im Einsatz, die mit hoher Trennschärfe betrügerisches Verhalten und Fälle mit einem hohen Risiko für einen Zahlungsausfall vorhersagen. Durch eine modellgestützte Ablehnung dieser Fälle unterstützt CRIF seine Kunden dabei, den manuellen Prüfaufwand zu reduzieren.

Mit dem Ziel, die Aussteuerung in die manuelle Prüfung und die daraus resultierenden Kosten weiter zu optimieren, ist es CRIF nun gelungen, für ein großes E-Commerce-Unternehmen ein innovatives Machine-Learning-Modell zu entwickeln, dessen Fokus nicht auf dem Verhalten der Besteller liegt, sondern auf dem Verhalten des Prüfers. Im Gegensatz zu Modellen zur Betrugsprävention, die mit Zahlungsausfällen trainiert werden, nutzt dieses Modell die **qualifizierte Entscheidung des Prüfers als Zielgröße** und imitiert diese. Dies ermöglicht Online-Händlern, das breite Spektrum an Kriterien, welche die manuelle Prüfung beinhaltet, zu nutzen und eine modellgestützte Entscheidung auf Basis der vom Prüfer real beschlossenen Annahme oder Ablehnung zu treffen.

# Inhalt



<b>1</b>	<b>Die manuelle Prüfung im E-Commerce</b>	<b>4</b>
<b>2</b>	<b>Die initiale Ausgangssituation des Kunden</b>	<b>6</b>
<b>3</b>	<b>Das Modell</b>	<b>8</b>
3.1	Modellentwicklung	9
3.2	Ergebnis	11
<b>4</b>	<b>Fazit</b>	<b>12</b>

# 1 Die manuelle Prüfung im E-Commerce

In Ländern, in denen der Kauf auf Rechnung eine große Rolle spielt, beschäftigen viele Online-Shop-Betreiber **große Teams von Spezialisten** zur manuellen Bearbeitung von Bestellanfragen, sowie zur initialen Anlage eines Kundenaccounts.

Die manuelle Prüfung zur Betrugsprävention im E-Commerce zielt darauf ab, verdächtige Aktivitäten zu identifizieren und betrügerische Transaktionen zu verhindern. Im Bestellprozess soll die Identitäts- und Bonitätsprüfung **möglichst modellgestützt** zur Annahme- oder Ablehnung des Rechnungskaufes führen. Trotz der zunehmenden Automatisierung bleibt die **manuelle Überprüfung unerlässlich**, um komplexe und nuancierte Betrugsmuster zu erkennen, die automatisierte Systeme möglicherweise übersehen. Die Mitarbeiter überprüfen unter anderem Unstimmigkeiten zwischen Rechnungs- und Lieferadresse, die Nutzung temporärer oder ungewöhnlicher E-Mail-Adressen, auffällige Wiederholungen in Bestellvorgängen, verdächtiges Bestellverhalten wie z. B. sehr hohe Bestellwerte oder plötzliche Änderungen im Kaufverhalten. Des Weiteren werden Bestellungen, in denen bekannte Paketweiterleitungsdienste oder Expressversand genutzt werden, eingehender untersucht.

Im nächsten Schritt kann eine **direkte Kontaktaufnahme** mit dem Kunden notwendig sein, um die Identität des Bestellers zu verifizieren und fragliche Details in der Bestellung zu klären. Im Falle einer Ablehnung nach erfolgter manueller Prüfung wird der Besteller

informiert und die Bestellung wird storniert. Weiterhin kann eine Sperrung des Kundenkontos oder die Aufnahme des Kunden in eine Blacklist erfolgen, um weitere Betrugsversuche zu verhindern.

Je nach der Komplexität des Falles können die Dauer und damit einhergehend die Kosten einer manuellen Bestellprüfung stark variieren. In einfachen Fällen, in denen lediglich die Kundendaten und Bestellinformationen auf Unstimmigkeiten überprüft werden müssen, liegt die Bearbeitungszeit bei wenigen Minuten.

Muss jedoch eine **eingehendere Untersuchung** der Bestell- und Zahlungshistorie erfolgen, um beispielsweise Identitätsdiebstähle oder koordinierte Betrugsversuche von Betrugsnetzwerken zu erkennen, kann die Bearbeitungszeit **mehrere Stunden** oder gar **einen gesamten Arbeitstag** in Anspruch nehmen. Dies kann insbesondere in Phasen von höherem Bestellaufkommen, beispielsweise während der Black Week oder in der Weihnachtszeit, zu **Zeit- und Kostendruck** führen.

## Ziel des Machine-Learning-Modells zur Optimierung der manuellen Prüfung



---

Ziel des von CRIF entwickelten Machine-Learning-Modells zur **Optimierung der manuellen Prüfung** ist es, die Fälle, die nach Ausschluss der modellgestützt festgestellten betrugsverdächtigen Bestellungen zur manuellen Prüfung übrig bleiben, weiter zu minimieren. Dies soll dem Prüfer die Möglichkeit geben, sich auf komplexe Fälle zu konzentrieren und kann gerade in bestellintensiven Phasen helfen, den Zeit- und Kostendruck zu senken.

---

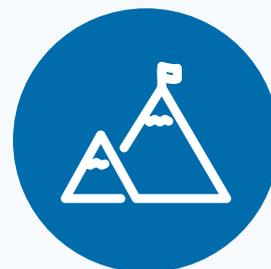
# 2 Die initiale Ausgangssituation des Kunden

Das Machine-Learning-Modell wurde für einen großen deutschen E-Commerce-Kunden entwickelt, der monatlich durchschnittlich **ca. 1 Mio. Online-Bestellungen** bearbeitet.

In saisonalen Phasen hohen Bestellaufkommens nimmt die Anzahl der Bestellungen im Vergleich zum Monatsdurchschnitt um ca. 30 % zu. Insbesondere im Neukundensegment fiel hier auf, dass nur ca. 7 % der Bestellungen, die in die manuelle Prüfung aus-gesteuert werden, abgelehnt wurden. Dies birgt ein **Optimierungspotential**, da 93 % des Prüfaufwands in Bestellungen investiert wird, die für den Prüfer offensichtlich nicht betrugsauffällig oder risikobehaftet zu sein scheinen.

Unter Annahme eines monatlichen Prüfaufkommens von 100.000 Bestellungen und durchschnittlichen Prüfkosten von 5 € pro Bestellung fallen **monatlich** bereits **465.000 € Prüfkosten** im Segment der geprüften und angenommenen Bestellungen an.

Die Aufgabenstellung des Kunden an uns war, ein Modell zu entwickeln, das den Prüfaufwand in diesem Segment reduziert, indem Fälle modellgestützt angenommen werden, die der Mitarbeiter in der manuellen Prüfung mit hoher Wahrscheinlichkeit positiv bewerten würde. Gleichzeitig muss bei Einsatz des Modells gewährleistet sein, dass der aktuelle Zahlungsausfall von 3 % noch unterschritten wird.





## Hohe Bestellaufkommen in saisonalen Phasen



In saisonalen Phasen hohen Bestellaufkommens nimmt die Anzahl der Bestellungen im Vergleich zum Monatsdurchschnitt um **ca. 30 %** zu.

# 3 Das Modell

CRIF nutzt für die Risikostrategie unseres Kunden bereits Machine-Learning-Modelle, die unter **Berücksichtigung der Ausfallwahrscheinlichkeit** zur Entscheidung auf Annahme oder Ablehnung einer Bestellung auf Rechnung beitragen.

So werden Bestellungen mit hoher Ausfallwahrscheinlichkeit oder bereits etablierten Betrugsmustern abgelehnt und müssen nicht mehr manuell bearbeitet werden. Unsere Kunden haben jedoch gute Gründe, **bestimmte Bestellungen** manuell zu prüfen, weil sie beispielsweise zusätzliche Bestellinformationen für die Entscheidung nutzen können und weitere Kriterien als die Ausfallwahrscheinlichkeit und gängige Betrugsmuster in Betracht ziehen wollen. Zudem ist es wirtschaftlich interessant, im Bereich hoher Bestellwerte manuelle Prüfungen durchzuführen, um **potenzielle Umsätze** nicht zu verlieren. Daher haben wir für einen unserer größten E-Commerce-Kunden einen neuen Ansatz verfolgt.

Um dem Kundenwunsch bestimmte Situationen manuell prüfen zu können, gerecht zu werden und gleichzeitig den manuellen Prüfaufwand und die Kosten bei unserem Kunden gering zu halten, hat CRIF ein Machine-Learning-Modell entwickelt, das als **Zielgröße** nicht den Zahlungsausfall berücksichtigt, sondern die Entscheidung des manuellen Prüfers. Als Zielgröße des Modells wird somit die **Ablehnungswahrscheinlichkeit des Prüfers** definiert. Mit diesem Ansatz werden

Bestellungen nicht vorab anhand der Ausfallwahrscheinlichkeit auf Annahme oder Ablehnung entschieden, sondern anhand des Prüferverhaltens. Das Modell berechnet eine Ablehnungswahrscheinlichkeit, auf deren Basis mit einem **individuellen Cut-Off** die Entscheidung auf Ablehnung oder Annahme der Bestellung getroffen werden kann. Dies ermöglicht es, Bestellungen, die anhand der Risikobewertung zur manuellen Prüfung ausgesteuert würden, aber mit einer sehr geringen Wahrscheinlichkeit vom Prüfer abgelehnt werden, modellgestützt anzunehmen und den Prüfaufwand so zu reduzieren.



## 3.1 Modellentwicklung

Für das Training des Modells nutzte CRIF neben eigenen Daten auch die Daten aus der manuellen Prüfung des Kunden, die neben Informationen wie Bestellzeitpunkt, Höhe des Warenkorbs, gewählte Produktgruppen, Kundenstatus, Mahnstatus, auch die Entscheidung des manuellen Prüfers beinhalteten, welche als **Zielgröße** definiert wurde. Diese wurden unter anderem

mit Identifikationsmerkmalen, Bonitätsdaten und geografischen Daten aus dem CRIF-Datenuniversum angereichert.

Anschließend wurde ein **Gradient-Boosting-Modell** trainiert und konnte nach Identifikation der relevanten Features einen Gini-Koeffizienten von 0.71 erreichen.

### Gradient Boosting



---

Gradient Boosting ist ein maschinelles Verfahren zur Lösung von Klassifikations- und Regressionsproblemen. Es basiert auf der Idee, schrittweise **schwache Modelle** zu **starken Modellen** zu **kombinieren**, um die Genauigkeit zu erhöhen.

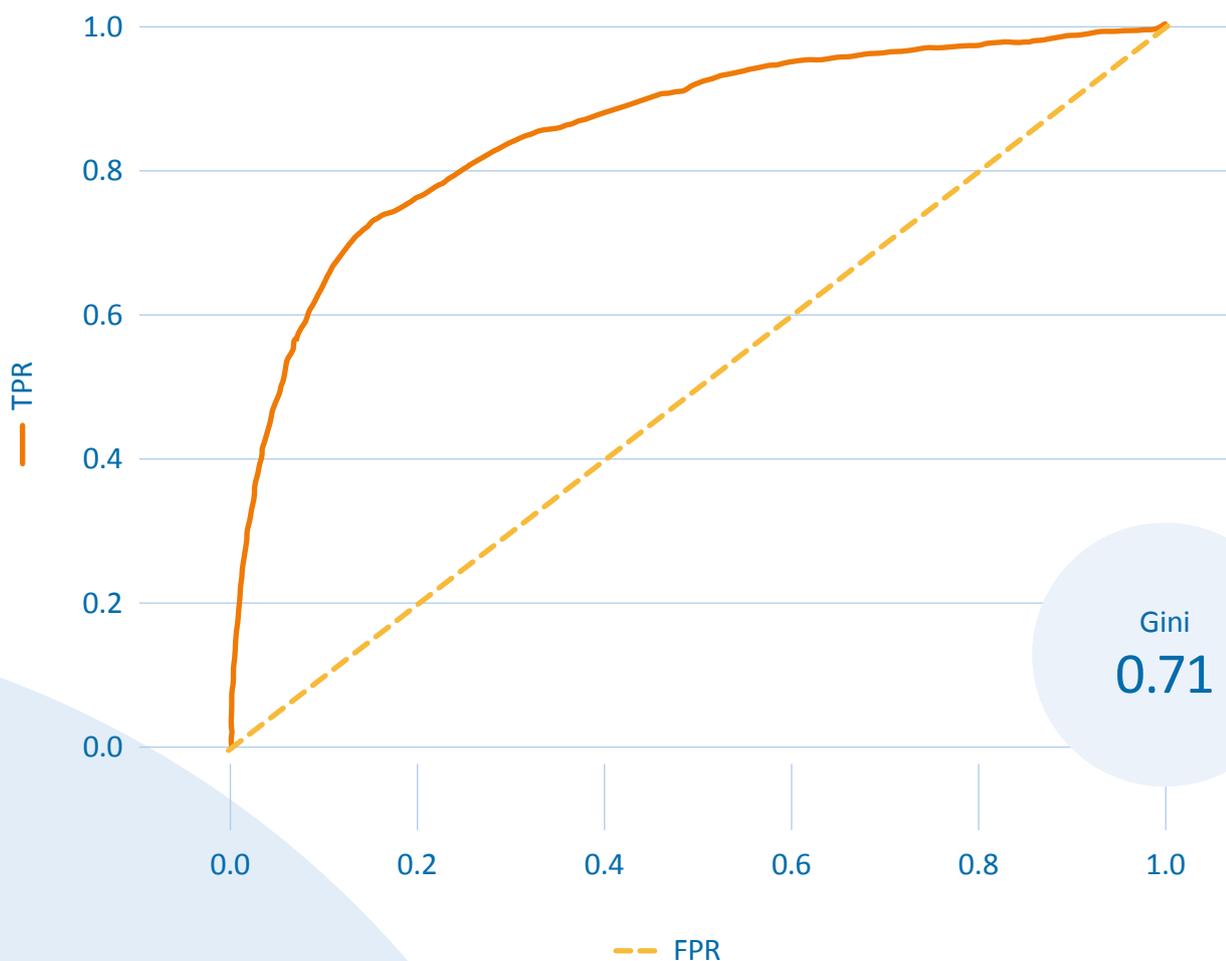
Zunächst wird ein **Entscheidungsbaum** erstellt, welcher eine grobe Schätzung der Zielgröße liefert. Im nächsten Schritt wird die Abweichung zwischen den vorhergesagten Werten und den tatsächlichen Werten und basierend darauf ein **Gradient** berechnet. Dieser gibt an, wie das Modell angepasst werden kann, um die Genauigkeit zu erhöhen. Mit dieser Information wird ein neues Modell berechnet. Dieser Prozess wird iterativ wiederholt und führt zu einem kombinierten Modell, welches mit jedem neuen Modell die Abweichung zwischen vorhergesagten und tatsächlichen Werten weiter minimiert.

---

## Receiver Operating Characteristic

Die dargestellte ROC-Kurve (siehe Grafik) zur Beurteilung von Klassifikationsmodellen zeigt zu Beginn einen steilen Anstieg, was darauf hindeutet, dass das Modell eine **hohe Rate an „True Positives“**, d. h. an richtig

erkannten Ablehnungen aufweist, während die „False Positive Rate“, d. h. die Rate an Annahmen, die fälschlicherweise als Ablehnung klassifiziert werden, gering ausfällt.



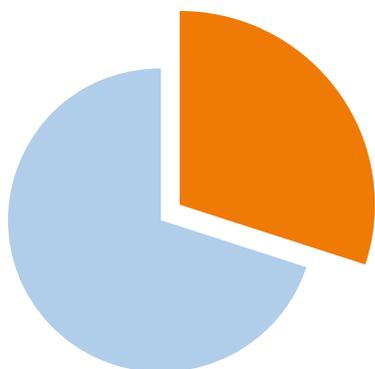
## 3.2 Ergebnis

Nach Erstellung des Modells wurde ein **Cut-Off** definiert, um zu bestimmen, bei welcher durch das Modell berechneten Ablehnungswahrscheinlichkeit die **Entscheidung auf Annahme** getroffen werden kann. Bestellungen, die eine höhere Ablehnungswahrscheinlichkeit zeigen als der definierte Cut-Off, werden weiterhin in die manuelle Prüfung ausgesteuert, während Bestellungen mit einer Ablehnungswahrscheinlichkeit unterhalb des Cut-Offs angenommen werden.

Ziel ist es, einen Cut-Off zu wählen, der gewährleistet, möglichst viele Bestellungen modellgestützt anzunehmen und gleichzeitig die **Ausfallquote** der angenommenen Bestellungen **minimal** zu halten. Insbesondere

bei ausoptimierten Risiko- und Betrugspräventionsstrategien ist diese Vorgehensweise sinnvoll, um sich weiter einer optimalen Steuerung zu nähern.

Seit Einführung des Modells konnten 30 % der Bestellungen, die bisher in der manuellen Prüfung bearbeitet wurden, freigegeben werden, was zu einer durchschnittlichen **jährlichen Kostenersparnis** von **87.000 €** führte. Innerhalb der im Modell angenommenen Aufträge wurde eine **Ausfallquote von 1,2 %** festgestellt, welche deutlich unter dem eingangs berechneten durchschnittlichen Zahlungsausfall von 3 % liegt. Die durch das Modell angenommenen Bestellungen sind also in Bezug auf den Ausfall statistisch unauffällig.



**30 %** der Bestellungen,  
die bisher in der manuellen Prüfung bearbeitet wurden,  
konnten **modellgestützt freigegeben** werden.



# 4 Fazit

CRIF ist es mit der Entwicklung des Machine-Learning-Modells zur manuellen Prüfung gelungen, den **Prüfaufwand** und die **Prüfkosten** unseres Kunden massiv **zu reduzieren** und Bestellungen, welche eine geringe Ablehnungswahrscheinlichkeit seitens des Prüfers aufweisen, anzunehmen.

Dies kann gerade in **bestellintensiven Phasen** helfen, den Aufwand und die Kosten zu senken, ohne dabei **Prüfgenauigkeit** einzubüßen. Die Verwendung der Prüferentscheidung als Zielgröße ermöglicht es, ein Modell zu entwickeln, das die Stärken des menschlichen Urteilsvermögens noch besser mit der **Effizienz und Skalierbarkeit** automatisierter Systeme kombiniert.

Die E-Commerce-Branche wird weiterhin von solchen **technologischen Innovationen** profitieren, die die Balance zwischen automatisierten Systemen und menschlicher Intervention optimieren. Dies ist ein entscheidender Schritt, um den Herausforderungen der modernen digitalen Wirtschaft gerecht zu werden und die Integrität des Online-Handels zu sichern.





