CELENT

OPERATIONALISIERUNG DER BETRUGSPRÄVENTION AUF IBM Z16

Verringerung von Verlusten im Banken-, Karten- und Zahlungssektor

Neil Katkov

5. April 2022

Dieser Bericht wurde von IBM in Auftrag gegeben. IBM beauftragte Celent mit der Planung und Durchführung einer Celent-Studie für IBM. Sämtliche Analysen und Schlussfolgerungen stammen ausschließlich von Celent. IBM hatte keine redaktionelle Kontrolle über den Berichtsinhalt.

INHALT

ZUSAMMENFASSUNG

Fortschritte bei der künstlichen Intelligenz (KI), wie beispielsweise Deep Learning, ermöglichen erhebliche Verbesserungen bei der Betrugserkennung. Große Banken und Zahlungsverarbeiter, die KI-Modelle einsetzen, lassen diese jedoch aufgrund von Durchsatz- und Latenzbeschränkungen bei ihren Betrugserkennungen oft nur für einen Bruchteil der Transaktionen laufen. Infolgedessen bleiben viele betrügerische Transaktionen unüberwacht und unentdeckt.

IBM Integrated Accelerator for AI, Teil des neuen IBM Telum Mainframe-Prozessors, wurde entwickelt, um Inferencing in großem Umfang und mit geringer Latenz für Echtzeit-Workloads durchzuführen. Der Chip unterstützt die Betrugserkennung in Echtzeit selbst in Verarbeitungsumgebungen mit sehr hohen Datenvolumen im Banken-, Karten- und Zahlungssektor.

Um Banken und Zahlungsverarbeitern zu helfen, den potenziellen Mehrwert dieser Innovation für ihre Betrugsbekämpfung zu verstehen, hat Celent Schätzungen über die potenzielle Reduzierung von Betrugsverlusten entwickelt, wenn diese Unternehmen KI-Inferencing auf 100 % ihrer Transaktionen anwenden.

Quantifizierbare Leistung KI-basierter Betrugserkennung auf IBM z16 Mainframes:

Verringerung der Betrugsverluste in der Branche um		Verlustreduzierung pro Bank um		Reduzierung abgelehnter Kartentransaktionen um
USA	Weltweit	Tier-1-Bank, USA	Tier-2-Bank, USA	
5,6 Cent je 100 USD	2,0 Cent je 100 USD	105 Mio. USD	18 Mio. USD	46 %

Nach Schätzungen von Celent könnte die Anwendung fortschrittlicher Inferencing-Modelle auf theoretisch sämtliche Bank-, Karten- und Zahlungstransaktionen auf IBM zSystems Mainframes die weltweiten Betrugsverluste potenziell um 161 Mrd. USD senken. In einem solchen Fall könnten Banken potenziell Verluste in Höhe von 140 Mrd. USD vermeiden, die Karten- und Zahlungsanbieter Verluste in Höhe von 21 Mrd. USD. Allein in den USA ließen sich die Verluste durch Bankenbetrug potenziell um 44 Mrd. USD und bei Karten- und Zahlungsbetrug um 6 Mrd. USD reduzieren.

Natürlich gibt es Hürden bei der Einführung von KI-Inferencing zur Betrugsbekämpfung auf dem Mainframe: Dazu zählen Governance-Fragen im Zusammenhang mit dem Modell, die Kosten eines vollständigen Austauschs der Betrugserkennung, die Verfügbarkeit interner Data-Science-Ressourcen sowie eine überzeugende Kosten-Nutzen-Analyse.

Dennoch bedeutet die Ausführung fortschrittlicher KI-Modelle direkt in der Mainframe-Umgebung eine leistungsfähige Innovation in einer Branche, in der geschätzte 70 % der gesamten weltweiten Transaktionswerte auf IBM Mainframes verarbeitet werden. Die Betrugserkennung ist ein wichtiger Anwendungsfall dieser neuen Funktionalität von IBM, die nachweislich Vorteile für das Geschäftsergebnis und das Kundenerlebnis mit sich bringt.

DIE HOHEN KOSTEN VON BETRUG FÜR DEN BANKEN-, KARTEN- UND ZAHLUNGSSEKTOR

Im Jahr 2021 verursachte Betrug für den Banken-, Karten- und Zahlungssektor weltweit einen geschätzten Schaden in Höhe von 385 Mrd. USD.

Banken- und Zahlungsbetrug nimmt im Privat- und Unternehmenssektor viele Formen an. Bankenbetrug umfasst Kontoübernahmen, APP-Betrug (Authorized Push Payments), Rechnungsbetrug und eine Vielzahl von Phishing- und Social-Engineering-Methoden für illegale Überweisungen oder das Ausspähen von Kontodaten. Auch Karten und der Zahlungsverkehr sind anfällig für Kontoübernahmen und Phishing sowie spezielle Betrugsversuche wie synthetische ID, Bust-out und Man-in-the-Middle.

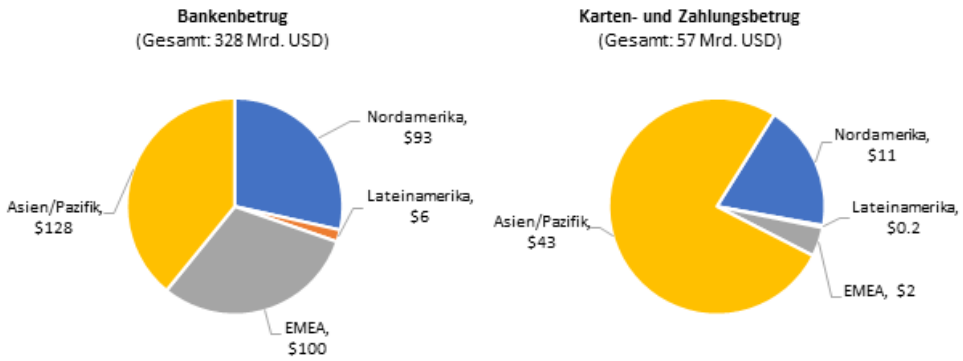
Abbildung 1: Häufige Banken- und Kartenbetrugsmaschen

Bankenbetrug		Kartenbetrug	
	Kontenübernahme		Anwendungsbetrug
	APP-Betrug		Bust-out-Betrug
	Scheckbetrug		Man-in-the-Middle
	Rechnungsbetrug		Phishing
	Social Engineering		Synthetische ID

Quelle: Celent

Diese und andere auf Bankkonten, Karten und Zahlungsverkehr abzielenden Betrugsmethoden stellen Finanzinstitute vor gravierende Probleme. Nach Schätzungen von Celent belaufen sich die jährlichen Verluste durch Betrug für eine Tier-1-Bank (Gesamtvermögen über 100 Mrd. USD) in den USA durchschnittlich auf 209 Mio. USD und für eine Tier-2-Bank (Gesamtvermögen zwischen 50 und 100 Mrd. USD) auf 35 Mio. USD. Im Jahr 2021 erlitten Banken weltweit Betrugsverluste in Höhe von 328 Mrd. USD. Der Karten- und Zahlungssektor häufte Verluste in Höhe von weiteren 57 Mrd. USD an. Im Jahr 2021 verursachte Betrug im Banken-, Karten- und Zahlungssektor weltweit einen geschätzten Schaden in Höhe von 385 Mrd. USD.

Abbildung 2: Betrugsverluste im Banken-, Karten- und Zahlungssektor im Jahr 2021



Quelle: Schätzungen von Celent auf Basis von BIZ-Transaktionsdaten und betrugsbezogenen Daten der Zentralbanken.
Hinweis: Bankenbetrug umfasst Überweisungen, Lastschriften und Schecks. Karten- und Zahlungsbetrug umfasst Kredit- und Debitkarten, elektronische und andere Zahlungen.

Obwohl Banken und Zahlungsabwickler seit Jahrzehnten versuchen, den Betrug mittels Erkennungssystemen und chipbasierter Kartensicherheit einzudämmen, steigen die Verluste weiter an, da die Betrüger durch die Entwicklung neuer technologischer und Social-Engineering-Tricks immer einen Schritt voraus sind.

Die Corona-Pandemie hat die Betrugszahlen in die Höhe getrieben. Eine erhebliche Quelle für Banken sind Phishing- und Social-Engineering-Methoden, die Ängste und medizinische Bedürfnisse im Zusammenhang mit der Pandemie ausnutzen. Im Hinblick auf Kartentransaktionen hat die Pandemie zu einer Zunahme von digitalem Banking und E-Commerce geführt, da die Verbraucher die Filialen und Geschäfte mieden. Da der Löwenanteil der Kartenbetrügereien – etwa 65 % – auf Card Not Present-Transaktionen (CNP) entfällt, sind die Verluste durch Kartenbetrug gestiegen.

Hilfe naht: Betrugsmodelle auf Basis von Deep Learning

Durch Fortschritte im Bereich der künstlichen Intelligenz, wie Deep Learning, können Banken jetzt Betrug viel effektiver bekämpfen: Sie analysieren Daten in großem Umfang, um auf Betrug hindeutende Muster zu finden, einschließlich neuer, bisher noch nicht beobachteter Typologien.

Deep Learning ist eine Art Modell des maschinellen Lernens und basiert auf einem tiefen neuronalen Netz (DNN). Ein DNN besteht aus Knoten, den Neuronen, die mithilfe von fortlaufenden Gewichtungen bestimmte Verknüpfungen zwischen diesen Knoten verstärken. Die Knoten sind dabei auf verschiedenen Schichten angeordnet – so entsteht ein „tiefes“ Netz, was die Leistungsfähigkeit und Lernrate des Modells erhöht. Deep-Learning-Modelle werden anhand von bestehenden Daten trainiert, im Falle von Betrugsmodellen beispielsweise mit Daten älterer Transaktionen. Das so trainierte Modell wird dann auf aktive Daten wie beispielsweise Echtzeit-Transaktionen angewendet, um ein Ergebnis oder eine Inferenz zu erzeugen. Bei Betrugsmodellen ist diese Inferenz typischerweise ein numerischer Wert, der ausdrückt, mit welcher Wahrscheinlichkeit es sich um eine betrügerische Transaktion handelt.

Basierend auf Gesprächen mit der Branche und Studienergebnissen schätzt Celent, dass KI-Inferencing auf Grundlage von Deep-Learning-Modellen die Genauigkeit der Betrugserkennung gegenüber bisherigen Modellen um 60 % erhöhen kann.

Allerdings wird das Potenzial des Inferencing zur Verbesserung der Betrugserkennungsquoten drastisch eingeschränkt, weil solche Modelle in volumenstarken Mainframe-Umgebungen häufig nur auf einen Bruchteil aller Transaktionen – weniger als 10 % – angewendet werden. Gründe sind Probleme durch Latenz, Kosten und Kundenunzufriedenheit. Das bedeutet, dass rund 90 % aller potenziell vermeidbaren Betrugsangriffe nach wie vor unentdeckt bleiben. So können Banken die Fortschritte in der KI bei der Eindämmung von Betrugsverlusten nur sehr begrenzt nutzen.

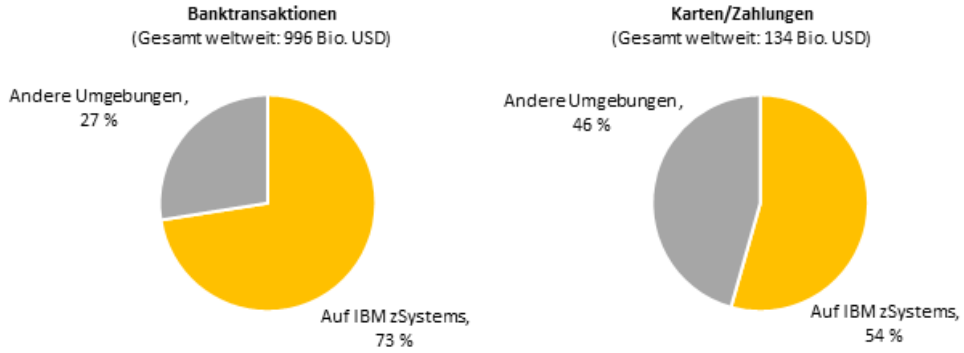
Doch die Hürden Latenz und Kosten, deretwegen nicht alle Banken- und Kartentransaktionen fortschrittliche Erkennungsmodelle durchlaufen, gehören nun möglicherweise der Vergangenheit an. Der neue IBM z16 Telum-Prozessor beinhaltet einen KI-Beschleuniger, der – erstmals bei IBM zSystems – KI-Modelle in Echtzeit direkt auf dem Chip ausführen kann. Datendurchsatz und Reaktionszeiten verbessern sich exponentiell und erstmals können Deep-Learning-basierte Betrugserkennungsmodelle praktisch auf sämtliche Transaktionen angewendet werden.

DIE GRENZEN DER DERZEITIGEN BETRUGSERKENNUNG

Technologisch und operativ erfolgt die Betrugserkennung in Mainframe-Umgebungen in der Regel in Systemen außerhalb der Plattform für ausgewählte Transaktionen bzw. im Anschluss an eine Transaktion. Dies schränkt die Möglichkeiten von Banken und Zahlungsverarbeitern erheblich ein, fortschrittliche KI-Modelle auf alle Transaktionen anzuwenden.

Viele große Banken und Zahlungsabwickler betreiben ihre Kernsysteme auf Großrechnern. Nach Einschätzung von IBM arbeiten 45 der 50 größten Banken der Welt mit IBM zSystems Mainframes. Auch die meisten großen Karten- und Zahlungsabwickler nutzen diese Plattform. Weltweit arbeiten nach Schätzungen von Celent 70 % der Banken-, Karten- und Zahlungstransaktionen auf IBM zSystems-Umgebungen.

Abbildung 3: Banken-, Karten- und Zahlungstransaktionen auf IBM zSystems



Quelle: Celent

Die Latenzzeit zwischen Kernsystemen und Erkennungssystemen außerhalb der Plattformen kann für einige Transaktionen toleriert werden. Bei datenintensiven KI-Inferenzroutinen für Echtzeit-Transaktionen – wie Echtzeit-Zahlungen, Kartentransaktionen und digitale Banktransaktionen – ist es wegen der Latenz jedoch unpraktisch, in volumenstarken Umgebungen alle Transaktionen durch eine KI-Erkennungsplattform zu leiten. Werden Kernsystemtransaktionen vom Mainframe zur Echtzeitanalyse an ein Erkennungssystem außerhalb der Plattform gesendet, betragen die Antwortzeiten zwischen 50 und 80 Millisekunden – während die Transaktionen warten. Dies verlangsamt die Genehmigungszeiten für Transaktionen, was insbesondere bei Kartentransaktionen zu Unzufriedenheit bei den Kunden führen kann.
Ganz grundsätzlich kann es wegen hoher Latenz unmöglich sein, alle Transaktionen eine Betrugserkennung außerhalb der Plattform durchlaufen zu lassen. Die Latenz zwischen Kernsystem und Erkennungssoftware kann den Erhalt der Ergebnisse so weit verzögern, dass das Zeitlimit von Echtzeit-Transaktionen überschritten wird. Deswegen führen manche Banken ihre Deep-Learning-Modelle zur Betrugserkennung erst nach der Transaktion aus.

Banken senden also lediglich einen Bruchteil ihrer Transaktionen – weniger als 10 % – in Echtzeit durch die Betrugserkennung. Diese Herangehensweise hat schwerwiegende Folgen. Deep-Learning-Modelle ermöglichen inzwischen erheblich bessere Erkennungsraten von 60 %. Banken profitieren jedoch nicht in vollem Umfang davon, da nur ein Teil der Transaktionen diese Modelle durchlaufen. Das bedeutet, dass ein größerer Teil der Betrugsfälle nicht entdeckt wird und betrugsbedingte Verluste steigen. Da Betrug immer mehr in den Fokus der Finanzkriminalität rückt, sind Banken möglicherweise auch einem regulatorischen Risiko ausgesetzt, wenn sie nicht alle Transaktionen durch die Betrugserkennung leiten können.

Altlasten in einer Tier-1-Bank in den USA

Eine Tier-1-Bank in den USA, die ihr Kernsystem auf einer IBM zSystems-Plattform betreibt, hat außerhalb der Plattform eine KI-basierte Betrugserkennung implementiert. Aufgrund von Kosten- und Latenzproblemen durchlaufen nur sehr risikoreiche Transaktionen das KI-System. Ein Großteil der Transaktionen durchläuft der Praktikalität halber ein regelbasiertes Scoring und wird erst nachträglich einer Analyse unterzogen. Die Vorteile der KI kommen dadurch nicht voll zum Tragen, weil die Modelle nicht auf alle Transaktionen angewendet werden können.

REDUZIERUNG VON BETRUGSVERLUSTEN DURCH KI-INFERENCING AUF DEM MAINFRAME

IBM hat für seinen IBM z16 Mainframe-Computer einen mit einem KI-Beschleuniger ausgestatteten Prozessor entwickelt, der direkt auf dem Chip erweiterte Inferenzverfahren auch im großen Maßstab ausführen kann. Nach Einschätzung von Celent unterstützt der neue IBM z16 Prozessor eine Deep-Learning-basierte Betrugserkennung für praktisch alle Transaktionen und könnte so die weltweiten Verluste durch Betrug im Banken-, Karten- und Zahlungssektor um 161 Mrd. USD senken.

Deep-Learning-Algorithmen sind tendenziell rechenintensiver als herkömmliche Betrugserkennungsmodelle. Wenn Banken Deep-Learning-basiertes KI-Inferencing zur Betrugsbekämpfung implementieren, müssen sie diese geschäftskritischen Workloads auch verwalten. Erfolgt die Erkennung auf Systemen außerhalb der Plattformen, kann es zu Reaktionszeiten von mehr als 80 Millisekunden und Durchsatzraten im Bereich von 1.000–1.500 Transaktionen pro Sekunde (TPS) kommen.

Aufgrund dieser Latenz- und Durchsatzbeschränkungen kommt es bei Banken zu Überschreitungen des Zeitlimits von Transaktionen, während auf die Ergebnisse der Erkennung gewartet wird. Dies und andere Probleme führen dazu, dass Banken nur einen Bruchteil ihrer Transaktionen – weniger als 10 % – durch ihre Erkennungssysteme schicken.

Deep Learning auf dem Mainframe

Auf Basis eines Deep-Learning-Modells zum Kreditkartenbetrug können 32 IBM Telum-Chips, die auf einem einzigen Server laufen, bis zu 3,5 Mio. Inferenzen pro Sekunde mit einer durchschnittlichen Reaktionszeit von 1,2 Millisekunden bearbeiten.

Quelle: IBM Microbenchmark, August 2021

HAFTUNGSAUSSCHLUSS: Das Leistungsergebnis wurde aus internen Tests von IBM extrapoliert.

IBM hat einen Beschleuniger für den IBM z16 Mainframe-Computer entwickelt, der KI-Inferenzmodelle direkt auf dem Chip ausführen kann. Laut IBM reichen der Durchsatz und die Verbesserungen bei der Ausführung von KI-Modellen auf dem Mainframe aus, um Betrugsanalysen in Echtzeit für praktisch alle Transaktionen zu unterstützen – selbst in volumenstarken Umgebungen von Banken, Karten- und Zahlungsabwicklern.

Außerdem kann dies nahezu ohne Beeinträchtigung der Transaktionsverarbeitungszeiten erfolgen. IBM gibt an, dass IBM Integrated Accelerator for AI im neuen IBM Telum-Prozessor KI-Modelle auf dem Mainframe mit sehr schnellen Reaktionszeiten von nur 1,2 Millisekunden pro Inferenzanfrage ausführen kann.
Im konkreten Fall der Erkennung von Kartenbetrug deuten erste Benchmarks darauf hin, dass eine Konfiguration aus 32 Telum-Chips bis zu 3,5 Mio. Inferenzen pro Sekunde unterstützt.

Das reicht aus, um auch Höchstwerte bei den Transaktionsströmen zu bewältigen, so dass Banken und Zahlungsverarbeiter praktisch alle Transaktionen über Deep-Learning-Modelle abwickeln können.

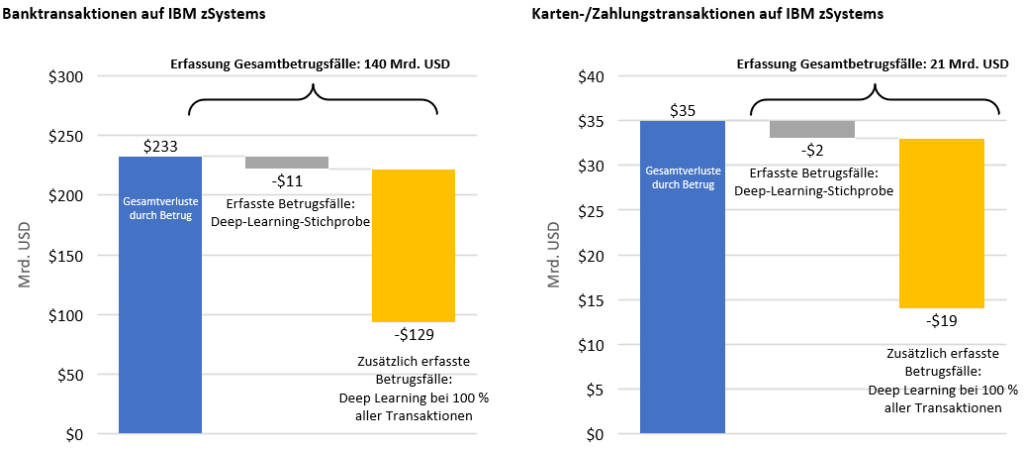
Banken sowie Karten- und Zahlungsverarbeiter können das volle Potenzial der modernen Inferenztechnologie ausschöpfen, indem sie fortschrittliche Modelle auf alle Transaktionen anwenden. Nach Schätzungen von Celent würde die Anwendung fortschrittlicher

Inferenzmodelle auf alle Transaktion die Betrugsverluste weltweit potenziell um 2,0 Cent pro 100 USD (2,0 Basispunkte) senken.

In den USA, wo die Betrugsraten höher sind als im weltweiten Durchschnitt – 9,3 Cent pro 100 USD gegenüber weltweit 3,7 Cent –, könnten die Betrugsverluste um 5,6 Cent je 100 USD reduziert werden. Dies entspricht bei einer durchschnittlichen Transaktion im Wert von 2.375 USD einer Einsparung für die Bank von 1,33 USD.

Nach Schätzungen von Celent könnte die Anwendung von Deep-Learning-Modellen auf sämtliche gegenwärtig auf IBM zSystems abgewickelte Transaktionen die Verluste durch Betrug weltweit um theoretisch 161 Mrd. USD senken. Banken könnten Betrugsverluste in Höhe von 140 Mrd. USD vermeiden. Bei Karten- und Zahlungsabwicklern werden 21 Mrd. USD Ersparnisse durch Betrugsvermeidung geschätzt. Allein in den USA beträgt das Potenzial zur Reduzierung von Betrugsverlusten 44 Mrd. USD bei Banken und 6 Mrd. USD bei Karten- und Zahlungsabwicklern.

Abbildung 4: Potenzielle Verminderung des Betrugsschadens mit Deep-Learning-Modellen



Quelle: Celent

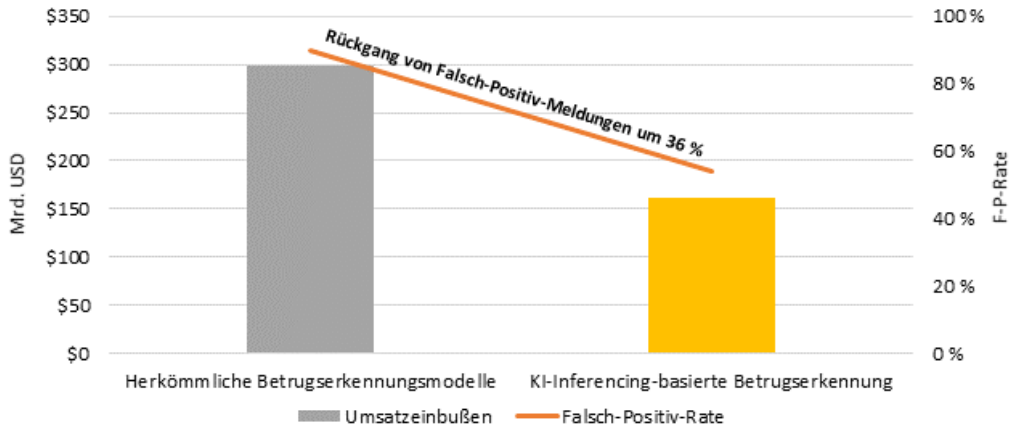
Laut Schätzungen von Celent kann eine Tier-1-Bank mit IBM z16, die alle Transaktionen durch fortschrittliche Inferenzmodelle laufen lässt – im Vergleich zur derzeit praktizierten Methode von KI-Modellen für nur etwa 10 % der Transaktionen – Betrugsverluste um zusätzliche 105 Mio. USD reduzieren könnte. Eine Tier-2-Bank könnte zusätzliche Verluste in Höhe von 18 Mio. USD vermeiden. Gleichzeitig würde der Einsatz fortschrittlicher Modelle bei sämtlichen Transaktionen auch die Modelle selbst weiter verbessern. Mehr Transaktionen bedeuten mehr Daten zum Trainieren der Modelle, was wiederum zu genauerer Betrugserkennung führt.

REDUZIERUNG VON FALSCH-POSITIV-ERGEBNISSEN ZUR VERMINDERUNG VON KUNDENABWANDERUNG

Herkömmliche Betrugsbekämpfungsmodelle weisen sehr hohe Falsch-Positiv-Raten auf – in der Regel 90 % aller gemeldeten Transaktionen oder mehr – so dass Banken eigentlich legitime Transaktionen ablehnen. Die hohe Falsch-Positiv-Rate und verweigerte Transaktionen sorgen nicht nur für Unzufriedenheit bei den Kunden, sondern führen auch zu handfesten finanziellen Einbußen, da Kunden beim nächsten Kauf einfach eine andere Kredit- oder Debitkarte nutzen. Celent geht davon aus, dass der Branche durch abgelehnte Kreditkartentransaktionen weltweit Gebührenumsätze in Höhe von 298 Mrd. USD entgehen.

Banken begrenzen ihre Betrugserkennungsroutinen auch deshalb auf Stichproben, weil sie ein Gleichgewicht zwischen Betrugsbekämpfung und Kundenzufriedenheit finden müssen. Falsch-Positiv heißt, dass die Erkennungssoftware legitime Transaktionen irrtümlich als betrügerisch einstuft. Da Deep-Learning-Modelle genauer sind, können sie die sehr hohen Falsch-Positiv-Raten der Branche signifikant verringern. Dies würde wiederum die Zahl der fälschlicherweise abgelehnten Transaktionen reduzieren. Das verbessert das Kundenerlebnis und verringert Umsatzverluste aufgrund von Kundenabwanderung. Es bedeutet auch, dass Banken alle ihre Transaktionen durch die Betrugserkennung leiten können, ohne dass es zu Kundenunzufriedenheit kommt.

Abbildung 5: Deep-Learning-Modelle verbessern die Falsch-Positiv-Rate



Quelle: Celent

Die Anwendung von Deep-Learning-Modellen auf alle Kartentransaktionen könnte die Falsch-Positiv-Rate auf etwa 55 % verbessern. Das ist zwar immer noch sehr viel, doch könnten die entgangenen Gebühreneinnahmen weltweit um 137 Mrd. USD auf 161 Mrd. USD reduziert werden.

Weniger Falsch-Positiv-Meldungen hätten noch weitere Vorteile: Betrugsanalytiker müssten weniger Warnungen bearbeiten, was die Kosten für nachträgliche Untersuchungen senkt. Weniger Kundenunzufriedenheit und -frustration würde sich positiv auf den guten Ruf der Bank und das Vertrauen der Kunden auswirken.

Außerdem können fortschrittliche Modelle verdächtiges, auf Geldwäsche hindeutendes Verhalten besser aufdecken. Der Bank Secrecy Act in den USA, die EU-Richtlinien zur Bekämpfung der Geldwäsche und andere Vorschriften führen dazu, dass die Programme von Banken zur Bekämpfung der Geldwäsche (AML) von den Aufsichtsbehörden genau unter die Lupe genommen werden. Die Aufsichtsbehörden in den USA sind besonders aktiv, wenn es darum geht, Banken wegen unzureichender AML-Programme zu belangen, und für einige Banken wurden Geldstrafen von über 1 Mrd. USD verhängt. AML-Operationen leiden außerdem unter sehr hohen Falsch-Positiv-Raten, die in der Regel über 95 % liegen, was für Banken eine hohe operative Belastung darstellt. Darüber hinaus wird die AML-Überwachung in der Regel nach der Transaktion durchgeführt, was Banken einem erhöhten Risiko aussetzt. Die Nutzung von KI-basierten Modellen für AML-Operationen kann bei solchen Problemen helfen, indem sie die Erkennung von AML-Verhalten genauer macht und falsch-positive Ergebnisse reduziert.

WEG NACH VORN

Unsere Analyse deutet auf signifikante, quantifizierbare Vorteile hin, wenn Deep-Learning-Modelle auf bis zu 100 % aller Transaktionen angewendet werden. IBM gibt an, dass sein neuer Beschleuniger diese Anwendung für auf IBM z16 Mainframes laufende Transaktionen unterstützen kann – selbst in Umgebungen mit sehr hohen Datenvolumen. Es bleiben jedoch eine Reihe von Faktoren, die Banken und Zahlungsverarbeiter, die den Sprung wagen wollen, bedenken müssen.

Banken sowie Karten- und Zahlungsabwickler sollten die Vorteile einer Deep-Learning-basierten Betrugserkennung auf dem Mainframe abwägen und dabei nach Empfehlung von Celent u. a. folgende Aspekte berücksichtigen:

- **Modell-Governance.** Aufsichtsbehörden und interne Prüfer fordern eine überzeugende Governance für Betrugsmodelle. Das bedeutet, dass KI-Modelle transparent und erklärbar sein müssen. Während die Anbieter von KI-Plattformen immer weniger „Blackbox“-Ansätze verwenden, bleibt die Governance von KI-Modellen komplex.
- **Regulatorischer Widerstand.** Die Aufsichtsbehörden sind mit der herkömmlichen regelbasierten Erkennung vertraut, aber weniger mit fortgeschrittenen Deep-Learning-Techniken. Banken, Data Scientists und ihre Anbieter müssen in einigen Fällen die Aufsichtsbehörden über die Wirksamkeit und Zuverlässigkeit fortschrittlicher KI aufklären.
- **Kosten für den Systemaustausch.** Viele Einrichtungen haben bereits KI-basierte Betrugserkennung implementiert. Diese Unternehmen müssen eine Kosten-Nutzen-Analyse für die Verlagerung der Erkennung auf den Mainframe entwickeln. Dazu gehört auch die Entscheidung, ob die bestehenden Systeme in irgendeiner Form beibehalten werden sollen – beispielsweise zur Unterstützung nachgelagerter Analysen oder für kleinere Geschäftsbereiche – oder ob sie ganz stillgelegt werden.
- **Data-Science-Ressourcen.** IBM Integrated Accelerator for AI ist zur Ausführung von Modellen optimiert, einschließlich Modellen, die mit Open-Source-Frameworks wie Pytorch oder TensorFlow entwickelt wurden. Bislang wurde jedoch noch nicht gezeigt, ob er Softwarepakete zur Betrugserkennung unterstützt – obwohl wir davon ausgehen, dass einige Anbieter von Betrugserkennungssoftware mit der Zeit Pakete anbieten werden, die auf dem Beschleuniger laufen können. In jedem Fall benötigen Einrichtungen, die KI-basierte Betrugserkennung auf IBM z16 umstellen, Data-Science-Fähigkeiten, um fortschrittliche Deep-Learning-Modelle zur Betrugserkennung zu entwickeln und zu unterstützen, sei es intern oder durch spezialisierte Modellanbieter.

Finanzinstitute sollten diese Faktoren sorgfältig abwägen – und den neuen KI-Beschleuniger von IBM genau unter die Lupe nehmen. Dennoch sind die potenziellen Vorteile in Form von geringeren Verlusten durch Betrug und abgelehnte Transaktionen sowie verbesserte Kundenzufriedenheit und ein verbessertes Kundenerlebnis überzeugend. Unternehmen mit IBM zSystems sollten sich genau ansehen, welche Vorteile die Verlagerung der Betrugserkennung auf den Mainframe mit sich bringt.

NUTZUNG DER EXPERTISE VON CELENT

Wenn dieser Bericht für Sie interessant war, sind für Sie möglicherweise die kundenindividuellen Analysen und Recherchen von Celent von Interesse. Mit unseren gesammelten Erfahrungen und dem Wissen, das wir bei der Arbeit an diesem Bericht gewonnen haben, können Sie die Erstellung, Verfeinerung oder Umsetzung Ihrer Strategien optimieren.

Unterstützung für Finanzinstitute

Typische Projekte, die wir unterstützen, sind:

Vorauswahl und Auswahl von Anbietern. Wir führen eine auf Sie und Ihr Unternehmen zugeschnittene Untersuchung durch, um Ihre individuellen Bedürfnisse besser zu verstehen. Anschließend erstellen und organisieren wir eine kundenspezifische Anfrage an die ausgewählten Anbieter, um Sie bei einer schnellen und präzisen Auswahl der Anbieter zu unterstützen.

Bewertung geschäftsrelevanter Prozesse und Verfahren. Wir nehmen uns Zeit zur Bewertung Ihrer Geschäftsprozesse und Anforderungen. Auf Grundlage unserer Marktkenntnisse identifizieren wir mögliche prozess- oder technologiebedingte Einschränkungen und liefern klare Erkenntnisse, mit denen Sie die bewährten Methoden der Branche umsetzen können.

Entwicklung von IT- und Unternehmensstrategien. Wir sammeln die Sichtweisen Ihres Führungsteams, Ihrer Geschäfts- und IT-Mitarbeiter sowie Ihrer Kunden. Anschließend analysieren wir Ihre aktuelle Position, Ihre institutionellen Möglichkeiten und Ihre Technologie im Hinblick auf Ziele. Falls erforderlich, helfen wir Ihnen bei der Neuformulierung Ihrer Technologie- und Geschäftspläne, um kurz- und langfristige Anforderungen zu erfüllen.

Unterstützung für Anbieter

Unsere Leistungen unterstützen Sie bei der Verfeinerung Ihres Produkt- und Leistungsangebots. Einige Beispiele:

Evaluierung Ihrer Produkt- und Servicestrategie. Wir unterstützen Sie bei der Bewertung Ihrer Marktposition im Hinblick auf Funktionalität, Technologie und Services. Unsere Strategie-Workshops helfen Ihnen, gezielt die richtigen Kunden anzusprechen und Ihre Angebote auf deren Bedürfnisse abzustimmen.

Überprüfung von Marktbotschaften und Werbematerial. Auf Grundlage unserer umfassenden Erfahrungen mit Ihren potenziellen Kunden bewerten wir Ihre Marketing- und Verkaufsmaterialien, einschließlich Ihrer Website und aller Begleitmaterialien.

ZUGEHÖRIGE CELENT-STUDIEN

Remaking Risk: A Taxonomy of Regtech
Oktober 2021

Technology Trends Previsory: Risk, 2022 Edition
Oktober 2021

IT and Operational Spending in AML-KYC: 2021 Edition
Dezember 2021

IT and Operational Spending on Fraud: 2021 Edition
Februar 2021

Innovation In Risk: A Snapshot Through the Lens of Model Risk Manager 2021
April 2021

Fino Payments Bank: Remote Implementation of Enterprise-Wide Fraud Management During the Pandemic
März 2021

Swedbank: Modernizing Card Fraud Management and Improving Customer Experience
März 2021

COPYRIGHTVERMERK

Weitere Informationen erhalten Sie von info@celent.com oder:

Neil Katkov nkatkov@celent.com

Nord- und Südamerika

USA
99 High Street, 32nd Floor
Boston, MA 02110-2320, USA

+1.617.424.3200

USA
1166 Avenue of the Americas
New York, NY 10036, USA

+1.212.345.8000

USA
Four Embarcadero Center
Suite 1100
San Francisco, CA 94111, USA

+1.415.743.7800

Brasilien
Rua Arquiteto Olavo Redig
de Campos, 105
Edifício EZ Tower – Torre B – 26º andar
04711-904 – São Paulo, Brasilien

+55 11 3878 2000

EMEA

Schweiz
Tessinerplatz 5
8027 Zürich, Schweiz

+41,44.5533,333

Frankreich
1 Rue Euler
75008 Paris, Frankreich

+33 1 45 02 30 00

Italien
Galleria San Babila 4B
20122 Mailand, Italien

+39,02.305,771

Vereinigtes Königreich
55 Baker Street
London W1U 8EW, UK

+44.20.7333.8333

Asien/Pazifik

Japan
Midtown Tower 16F
9-7-1, Akasaka
Minato-ku, Tokyo 107-6216, Japan

+81,3.6871,7008

Hongkong
Unit 04, 9th Floor
Central Plaza
18 Harbour Road
Wanchai, Hongkong

+852 2301 7500

Singapur
138 Market Street
#07-01 CapitaGreen
Singapore 048946, Singapur

+65 6510 9700