Carla Krauss, Markus Hausmann, Mathias Steinmann

Data Analytics datenschutzkonform nutzen

Den Datenschatz ohne juristische Fallstricke heben

Im Artikel „Data Analytics und künstliche Intelligenz – Was ist Data Science?“ wurden Einsatzmöglichkeiten und Nutzen von Data Analytics bis hin zur künstlichen Intelligenz vorgestellt (vgl. NEWS 1/2019). In diesem Beitrag steht die datenschutzkonforme Nutzung dieser Daten im Mittelpunkt. Denn der Schutz sensibler Kundendaten muss für Banken – auch um das Vertrauen der Kunden nicht zu verspielen – oberste Priorität haben.

Gerade Banken verfügen in der Regel über weit umfassendere Informationen über ihre Kunden als realwirtschaftliche Unternehmen. Daher bieten ihnen die Methoden und Werkzeuge der Datenanalyse – neben Anwendungsfällen beispielsweise in Risikosteuerung und Betrugserkennung – die Möglichkeit, vertriebliche Potenziale zu heben. Bei einem Einsatz im Vertrieb ist das Ziel häufig, Kunden zu identifizieren, die einen konkreten Produktbedarf beziehungsweise eine Produktaffinität haben und somit ein vertriebliches Potenzial darstellen. Darüber hinaus bietet eine Datenanalyse aber auch neue Möglichkeiten zur Ermittlung der Preissensibilität oder Vertriebswegepräferenzen von Kunden.

DIE RELEVANZ (DATENSCHUTZ-)RECHTLICHER FRAGEN

In diesem Zusammenhang ist wichtig, dass das Verwenden von Kundendaten nicht allein dazu dient, Ertrag für die Bank zu generieren. Auch die Kunden profitieren davon. Denn an die Stelle einer unspezifischen Ansprache „mit der Gießkanne“ rückt eine individuelle Ansprache, die dem konkreten Bedarf des Kunden in seiner aktuellen Lebenssituation entspricht. Der Kunde erhält somit ein seinen Bedürfnissen entsprechendes Angebot. Mithilfe von Data Analytics kann also die Bindung zwischen Bank und Kunde erhöht werden. »

Verfahren zur Dimensionsreduktion – Datenminimierung und qualitativ hochwertige Ergebnisse

Grundsätzlich gibt es Möglichkeiten, um die Daten für maschinelle Lernverfahren und Datenanalysen zu komprimieren. Dabei selektieren Algorithmen aus allen vorhandenen Merkmalen (zum Beispiel Alter, akademischer Grad, Einkommen etc.) unter einem gewissen Informationsverlust Teilmengen. Damit können die für Datenanalysen notwendigen Merkmale bereits vorab reduziert beziehungsweise eingegrenzt werden. Die gewünschte Anzahl der Merkmale kann bei einigen Verfahren direkt vom Anwender vorgegeben werden.

Eine Möglichkeit zur Realisierung einer vorgelagerten Datenkomprimierung sind beispielsweise die sequenzielle Rückwärtsauswahl (Sequentiell Backwards Selection) oder eine Merkmalsauswahl mit Random Forests.

Bei der Merkmalsauswahl mit Random Forests werden die relativen Einflüsse der einzelnen Merkmale ermittelt. Somit kann bereits vorab eingeschätzt werden, welche Merkmale einen relativ hohen Einfluss haben und auf welche Merkmale (ohne allzu großen Informationsverlust) verzichtet werden kann (siehe Abbildung 1).

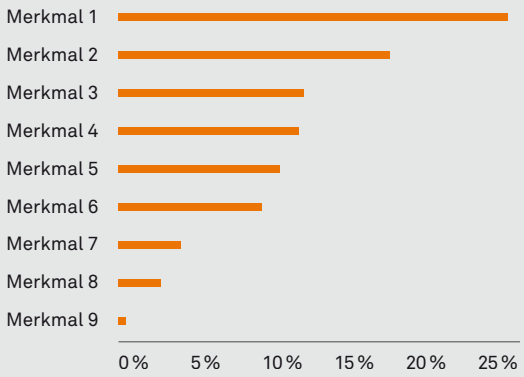

Abbildung 1: Merkmalsauswahl

In der Abbildung ist zu erkennen, dass beispielsweise das Merkmal „Berufsstand" einen sehr geringen relativen Einfluss hat. Möchte man nun den Merkmalsraum, also die Anzahl der Merkmale, verringern, könnte man bei Datenanalysen und maschinellen Lernverfahren auf das Merkmal „Berufsstand" verzichten. Des Weiteren ist erkennbar, dass allein die ersten vier Merkmale einen relativen Einfluss von ca. 60 Prozent aufweisen.

Somit können Verfahren im Bereich der Dimensionsreduktion (der Daten) ideal zur Minimierung (personenbezogener) Daten herangezogen werden.

Gerade das Mengengeschäft entwickelt sich dadurch von einer weitgehend anonymen Beziehung wieder hin zu einer persönlichen Kundenbeziehung.

Aber wer Wissen über den Kunden zum beiderseitigen Nutzen einsetzen kann, kann dies natürlich auch zum Nachteil des Kunden tun. Diese Bedenken müssen ernst genommen werden. Daher hat der Gesetzgeber mit der Datenschutzgrundverordnung (DSGVO) vollkommen zu Recht der Verwertung von Kundendaten Grenzen gesetzt. Der Grund, warum Banken Data Analytics nicht nutzen, liegt daher vielfach nicht in Zweifeln oder Unsicherheiten hinsichtlich der Möglichkeiten von Data Analytics, sondern eher in der Frage nach einem rechtssicheren Einsatz.

Aus den grundlegenden Regelungen der DSGVO ergeben sich eine Reihe von praktischen Fragen. Zu den wichtigsten Fragen gehören insbesondere:

- Welche Daten werden verarbeitet?
- In wessen Verantwortung liegen die Daten?
- Sollten Daten anonymisiert oder pseudonymisiert verarbeitet werden?
- Wie wird die Einwilligung zur Datenauswertung eingeholt?
- Wie muss mit Kunden beziehungsweise dem Geschäft in anderen Rechtsräumen umgegangen werden, das heißt insbesondere in Nicht-EU-Ländern?

Die Gültigkeit der neuen europäischen DSGVO seit Mai 2018 war in den Medien ein sehr prominentes Thema. Im Zusammenhang mit dem Einsatz von Data Analytics hat sie die Frage aufgeworfen, wie Datenschutz und Datenanalyse mit den neuen Anforderungen der DSGVO vereinbar sind. Banken befinden sich in einem Spannungsfeld aus verschärften Datenschutzbestimmungen durch die DSGVO und der Erwartung, mittels neuer Technologien innovative digitale Strategien zum Kundenverhalten entwickeln zu können. Doch die neuen Anforderungen der DSGVO bieten Banken auch Chancen, beispielsweise beim Einsatz von Data Analytics Datenmissbrauch zu vermeiden und so das Vertrauen von Kunden zu stärken.

Ziel der neuen DSGVO ist die Verbesserung des Datenschutzniveaus. Gerade beim Austausch und der Analyse von Kundendaten steht der Schutz der Grundrechte und Grundfreiheiten natürlicher Personen und insbesondere deren Recht auf Schutz personenbezogener Daten im Vordergrund. Dieser Zweck des Datenschutzrechts

und der Einsatz von Datenanalysen stehen sich dabei natürlich kontrovers gegenüber. Das Spannungsverhältnis zeigt sich insbesondere mit Blick auf die datenschutzrechtlichen Grundsätze der Datenminimierung und Transparenz nach Art. 5 DSGVO. Diese Grundsätze besagen zum einen, dass bei der Datenverarbeitung nur die personenbezogenen Daten gesammelt werden sollen, die für die jeweilige Anwendung unbedingt notwendig sind und zum anderen, dass Kunden über die Verarbeitung ihrer personenbezogenen Daten und über die Daten verarbeitenden Stellen informiert werden müssen. Das erscheint besonders bei der Nutzung von Data Analytics kontrovers, da gerade hier große Datenmengen benötigt werden, um entsprechende Analysen durchführen zu können. Wie Verfahren zur Dimensionsreduktion dabei helfen können, den Grundsatz der Datenminimierung zu erfüllen, veranschaulicht die Infobox.

Darüber hinaus hat der Gesetzgeber in der DSGVO die Verarbeitung personenbezogener Daten beim Einsatz von Data Analytics nicht explizit geregelt, sondern stellt auf die allgemeinen und sehr abstrakten Regelungen der Datenschutzgrundverordnung ab.

» Die Verarbeitung personenbezogener Daten zu Analysezwecken muss nach den datenschutzrechtlichen Regelungen rechtmäßig erfolgen.

Der Gesetzgeber ordnet dem Einsatz von Data Analytics den Begriff Profiling zu. Artikel 22 DSGVO regelt die Anforderungen an die Verwendung von automatisierten Verarbeitungsvorgängen, die eine spezielle Form der Verarbeitung personenbezogener Daten darstellt, mit dem Ziel, bestimmte persönliche Aspekte der Person zu bewerten und zu analysieren. Dabei handelt es sich im Rahmen von Datenanalyse um Aspekte wie Arbeitsleistung, wirtschaftliche Lage, Gesundheit, persönliche Vorlieben oder Interessen, Zuverlässigkeit oder Verhalten und Aufenthaltsort sowie die Vorhersage des Kaufverhaltens oder die Beurteilung der Kreditwürdigkeit einer Person (vgl. Art. 4 Nr. 4 DSGVO). Profiling schließt die Verwendung von Techniken zur Verarbeitung personenbezogener Daten ein, durch die von einer natürlichen Person ein Profil erstellt wird, das insbesondere die Grundlage für sie betreffende Entscheidungen bildet oder anhand dessen ihre persönlichen Vorlieben, Verhaltensweisen oder Gepflogenheiten analysiert oder vorausgesagt werden sollen. Welche Möglichkeiten das datenschutzrechtliche Rahmenwerk der DSGVO trotzdem für einen rechtssicheren Einsatz von Data Analytics bietet, wird im Folgenden näher beleuchtet.

Zunächst muss die Frage geklärt werden, was nach der DSGVO die Verarbeitung personenbezogener Daten genau ist. Denn zum einen sind im Rahmen des Einsatzes von Data Analytics bei der Auswertung großer Datenmengen auch regelmäßig personenbezogene Daten vorhanden. Zum anderen ist dies wichtig, da, wie bereits oben gesagt, die Regelungen der DSGVO nur für die Verarbeitung personenbezogener Daten gelten. Sofern beim Einsatz von Datenanalysen und bei der Auswertung großer Datenmengen keine personenbezogenen Daten verarbeitet werden, sind die Regelungen der DSGVO nicht einschlägig. Eine weitere Ausnahme von der Anwendung der Vorschriften der DSGVO bei der Datenanalyse besteht, wenn die personenbezogenen Daten in einer anonymisierten Weise verarbeitet werden. Nachfolgend werden bestimmte Techniken beziehungsweise technisch-organisatorische Maßnahmen untersucht, die im Kontext von Data Analytics eingesetzt werden können, um die Datensicherheit im Rahmen von Analyse-Bewertungsverfahren zu gewährleisten. Weiterhin wird kurz dargestellt, wer für die personenbezogenen Daten verantwortlich ist, wenn große Datenmengen anhand aktueller und historischer Fakten analysiert werden, um Vorhersagen über ein mögliches Kundenverhalten zu treffen.

Die Verarbeitung personenbezogener Daten zu Analysezwecken muss nach den datenschutzrechtlichen Regelungen rechtmäßig erfolgen. Daher ist es unerlässlich, auf die Rechtsgrundlagen bei der Verarbeitung personenbezogener Daten bei der Nutzung von Data Analytics einzugehen. Abschließend folgt vor dem Hintergrund eines internationalen Datenaustausches eine Erläuterung, was bei der Übermittlung von personenbezogenen Daten in Staaten außerhalb der EU zu beachten ist.

WELCHE DATEN WERDEN VERARBEITET UND IN WESSEN VERANTWORTUNG LIEGEN DIE DATEN?

Ziel der neuen Regelungen der DSGVO ist, den Schutz natürlicher Personen bei der Verarbeitung ihrer personenbezogenen Daten zu stärken. Vor diesem Hintergrund ist es also wichtig, bei der Umsetzung und Implementierung angemessener Datenschutzmaßnahmen immer zu differenzieren, ob die Daten personenbezogen sind oder nicht. »

Liegt kein Personenbezug vor, greifen die Regelungen der DSGVO nicht. Dabei sind personenbezogene Daten nach Art. 4 Nr. 1 DSGVO alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen. Eine Person kann zum Beispiel anhand ihres Namens, ihrer Anschrift, ihres Geburtsdatums etc. identifiziert werden. Darüber hinaus werden Personen als identifizierbar angesehen, die zum Beispiel anhand der Zuordnung einer Kennnummer, Online-Kennung, Kfz-Zeichens etc. identifiziert werden können. Insofern wird bei der Nutzung von Data Analytics regelmäßig eine Vielzahl von Kundendaten auch einen Personen-

» Data Analytics bietet Banken große Chancen zur Steigerung von Kundenorientierung und Erträgen.

bezug aufweisen. Daneben gibt es auch noch besonders sensible personenbezogene Daten. Das sind solche personenbezogene Daten, wie ethnische Herkunft, politische Meinungen etc. Aufgrund der besonderen Sensibilität dürfen diese personenbezogenen Daten im Rahmen des Profilings nicht verwendet werden. Soweit ein Personenbezug bei Daten hergestellt werden kann, müssen diese personenbezogenen Daten verarbeitet werden. Der Begriff der Verarbeitung wird nach der DSGVO weit gefasst. Zur Verarbeitung gehören nach Art. 4 Nr. 2 DSGVO das Erheben, Erfassen, Organisieren, Speichern, Anpassen und Verändern, Auslesen, Abfragen und Verwenden, Übermitteln, Einschränken, Löschen und Vernichten von personenbezogenen Daten. Verantwortlich bei der Verarbeitung dieser Daten ist nach Art. 4 Nr. 7 DSGVO derjenige, der über die Zwecke und Mittel der Verarbeitung von personenbezogenen Daten entscheidet. Er muss im Rahmen eines risikoorientierten Datenschutzes auch sicherstellen, dass die Rechte betroffener Personen in allen Verfahren und Prozessen berücksichtigt werden. Das bedeutet, dass auch im Rahmen des Einsatzes von Data Analytics die Rechte der Betroffenen nach Art. 12 ff. DSGVO berücksichtigt werden. Insbesondere betrifft dies das Recht auf Information, Auskunft, Berichtigung, Datenportabilität, Einschränkung der Verarbeitung, Löschung und Widerspruch gegen die Verarbeitung personenbezogener Daten. Alle Informationen über die Rechte müssen den betroffenen Personen dabei in transparenter Form und in einer leicht verständlichen Sprache übermittelt werden.

SOLLTEN DATEN ANONYMISIERT ODER PSEUDONYMISIERT VERARBEITET WERDEN?

Da die Anforderungen der DSGVO nicht für anonyme Informationen gelten, ist der Einsatz von Anonymisierungstechniken eine vielfach genutzte Möglichkeit beim Umgang mit personenbezogenen Daten im Rahmen von Datenanalysen. Anonyme Informationen sind dabei solche Informationen, die sich nicht auf eine identifizierte oder identifizierbare natürliche Person beziehen, oder personenbezogene Daten, die in einer Weise anonymisiert worden sind, dass die betroffene Person nicht oder nicht mehr identifiziert werden kann. Der Vorgang der Anonymisierung soll dabei nicht mehr oder nur mit unverhältnismäßigem Aufwand umkehrbar sein. Im Sinne der Dokumentationspflichten des Datenschutzrechtes müssen die getroffenen Maßnahmen hierzu hinreichend dokumentiert werden. Das heißt, die getroffenen Maßnahmen müssen genau beschrieben werden, um gegebenenfalls nachweisen zu können, welche Datenschutzmaßnahmen implementiert wurden, wie sie die Datensicherheit gewährleisten und wer für die regelmäßige Kontrolle verantwortlich ist.

Beim Ansatz einer pseudonymisierten Verarbeitung müssen – im Unterschied zur anonymisierten Verarbeitung – die Anforderungen der DSGVO erfüllt werden. Dieser Ansatz der Handhabe personenbezogener Daten ermöglicht es, personenbezogene Daten durch ein Pseudonym zu ersetzen. Pseudonymisierung ist nach Art. 4 Nr. 5 DSGVO die Verarbeitung personenbezogener Daten in der Art, dass die personenbezogenen Daten ohne Hinzuziehen zusätzlicher Informationen nicht mehr einer spezifischen betroffenen Person zugeordnet werden können. Dadurch können Risiken bei der Verarbeitung personenbezogener Daten für die betroffene Person reduziert werden.

WIE WIRD DIE EINWILLIGUNG ZUR DATENAUSWERTUNG EINGEHOLT?

Der Umgang mit personenbezogenen Daten muss nach Art. 6 DSGVO rechtmäßig sein. Grundsätzlich gilt dabei das „Verbot mit Erlaubnisvorbehalt". Das bedeutet, dass der Umgang mit personenbezogenen Daten grundsätzlich verboten und nur unter bestimmten Voraussetzung erlaubt ist. Voraussetzung für die Rechtmäßigkeit der Datenverarbeitung ist daher das Vorliegen einer Einwilligung. Eine Einwilligung ist jede für einen bestimmten Fall abgegebene Willensbekundung der betroffenen Person. Diese muss in informierter Weise und unmissverständlich abgegeben werden und eindeutig belegen, dass die betroffene Person mit der Verarbeitung der sie betreffenden personenbezogenen Daten einverstanden ist. Die Einwilligung in die Verarbeitung personenbezogener Daten im Rahmen von Data Analytics kann etwa in Form einer schriftlichen Erklärung erfolgen. Dabei ist zu beachten, dass im Fall einer online abgegebenen Willenserklärung umstritten ist, wann diese wirksam beziehungsweise nachweisbar von Kunden abgegeben worden ist. Die Einwilligung des Kunden muss sich auf einen konkreten Zweck beziehen und auf die Verarbeitung personenbezogener Daten im Rahmen von Data Analytics eingehen. Dient die Einwilligung mehreren Zwecken, muss sie für jeden Zweck eingeholt werden. Dem Kunden muss durch die Ausgestaltung der Einwilligung klar sein, dass seine personenbezogenen Daten für die Zwecke von Datenanalysen verwendet werden. Die Zweckbindung gilt auch, wenn personenbezogene Daten zur Durchführung einer vertraglichen Beziehung zwischen der Bank und dem Kunden verarbeitet werden. Besteht der Zwecke der Vertragserfüllung nicht mehr und werden personenbezogene Daten zu Marketing- und Werbezwecken erhoben und verarbeitet, ist dies jedoch nicht mehr von der Erlaubnisvorschrift umfasst.

Außerhalb der Vorschriften der DSGVO wird in § 31 BDSG n. F. als Erlaubnisvorschrift explizit das Scoring geregelt. Ein Scoring liegt vor, wenn der Verantwortliche Wahrscheinlichkeitswerte über ein bestimmtes zukünftiges Verhalten einer natürlichen Person zum Zweck der Entscheidung über die Begründung, Durchführung oder

Beendigung eines Vertragsverhältnisses mit dieser Person verwendet. Werbe-Scoring beziehungsweise Scoring-Verfahren außerhalb der Begründung, Durchführung und Beendigung eines Vertragsverhältnisses sind durch diese Vorschrift jedoch nicht erlaubt. Die Erlaubnisvorschrift nach § 31 BDSG n. F. zu Score-Verfahren ist nur dann gegeben, wenn die Rechtmäßigkeit der Verarbeitung an die Vorbereitung beziehungsweise die Durchführung eines Vertrages geknüpft ist.

WIE IST MIT KUNDEN BEZIEHUNGSWEISE DEM GESCHÄFT IN ANDEREN RECHTSRÄUMEN UMZUGEHEN, INSBESONDERE IN NICHT-EU-LÄNDERN?

Unabhängig von den spezifischen Anforderungen nach Art. 44 ff. DSGVO an die Übermittlung personenbezogener Daten in einen Drittstaat müssen die allgemeinen Bestimmungen an die Datenverarbeitung im Sinne der DSGVO eingehalten werden. Werden im Rahmen von Datenanalysen personenbezogene Daten in ein Drittland (außerhalb der EU) übermittelt, müssen neben den allgemeinen Bestimmungen an die Datenverarbeitung im Sinne der DSGVO die spezifischen Anforderungen an die Übermittlung beachtet werden. Eine rechtmäßige Übermittlung von personenbezogenen Daten in einen Drittstaat kann nur erfolgen, wenn in diesem Drittstaat ein angemessenes Datenschutzniveau gewährleistet ist. Dies ist dann der Fall, wenn die Angemessenheit des Niveaus der Datenschutzgesetzgebung eines Landes, Gebiets oder Sektors von der EU-Kommission anerkannt ist. Die Angemessenheit der Datenschutzgesetzgebung kann auch durch eine Vereinbarung des Landes mit der EU gewährleistet sein (vgl. Art. 44 DSGVO). Beim Datentransfer in ein Drittland kommt eine weitere Schwierigkeit hinzu. Denn in einem solchen Fall muss der Betroffene zusätzlich ausdrücklich und umfassend über die Risiken der Übermittlung seiner Daten in ein Land ohne ausreichendes Datenschutzniveau informiert werden. Erforderlich ist also die Transparenz gegenüber Kunden bezüglich der Schutzmaßnahmen beziehungsweise Datenschutzgarantien bei der empfangenden Stelle oder im Empfängerland.

FAZIT

Data Analytics bietet Banken große Chancen zur Steigerung von Kundenorientierung und Erträgen. Mit der DSGVO sind der Datennutzung jedoch auch klare rechtliche Rahmenbedingungen gesetzt. Es ist daher unabdingbar, sich intensiv mit diesen Vorgaben auseinanderzusetzen. Dann stellt die DSGVO kein Hindernis dar, sondern bietet klare und verlässliche Leitplanken, innerhalb derer die Bank Kundendaten nutzen darf. Zudem stehen unter anderem mit Pseudonymisierung und Anonymisierung sehr gute Verfahren zur Verfügung, die es erlauben, wesentliche Grundsätze der DSGVO einzuhalten. ■

Ansprechpartner:

Center of Competence
Regulatory Reporting
& Compliance

Carla Krauss
Senior Business Consultant
carla.krauss@msg-gillardon.de

Mathias Steinmann
Partner
mathias.steinmann@msg-gillardon.de