



Projektbeispiel

10 Tipps für Erfolg mit Data Science Teams

Fraud Erkennung im Finanz- und Versicherungswesen

Mit künstlicher Intelligenz geht es jetzt Kriminellen im großen Stil an den Kragen: Versicherungsbetrug, Geldwäsche, Kreditkartenbetrug, die Manipulation von Geldautomaten und andere Betrügereien können aufgedeckt und vereitelt werden. Daten und Technologie geben das her, so dass Data Scientist Teams als Datendetektive hier hervorragende Arbeit leisten können.

Wie kann man ein Data Science Team darin unterstützen, effizient und in meist eng gefassten Zeitrahmen innovativ und gleichzeitig erfolgreich zu sein?

BUSINESS.
DATA.
TECHNOLOGY.

CINTELLIC
CONSULTING GROUP

Wie kann man ein Data Science Team darin unterstützen, effizient und in meist eng gefassten Zeitrahmen innovativ und gleichzeitig erfolgreich zu sein? CINETELLIC stellt für Kunden im Finanz- und Versicherungswesen sowie in anderen Branchen Data Science Teams zusammen und begleitet sie von der Aufgabenstellung über die Entwicklung bis zur Erstellung und Anpassung von Datenprodukten mit echtem Mehrwert. CINETELLIC hilft die Potenziale zu heben, die in Datenschätzen schlummern, welche getreu dem Motto „Kenne deinen Kunden“ vor allem in CRM-Bereichen angehäuft wurden.

KI-Anwendungen mögen in aller Munde sein, deren Entwicklung zählt jedoch noch lange nicht zu den typischen Aufgaben der in den CRM-Bereichen von Banken oder Versicherungen verorteten Data Science Teams. Ein Beispiel: Über den Zeitraum von drei Monaten soll eine prototypische KI-Anwendung zur Fraud-Erkennung entwickelt werden. Diese beinhaltet das Finden, Analysieren und Aufbereiten entsprechender Daten, das Ausprobieren neuester Modellierungsansätze und letztlich die Überführung der Erkenntnisse in ein automatisiertes Reporting. Die dabei gesammelten Erkenntnisse sollen möglichst auch auf andere Anwendungsgebiete übertragbar sein, beispielsweise auf eine Customer Journey Analyse.

Eine komplexe Aufgabenstellung, für die es viele unterschiedliche Fähigkeiten und kreativen Raum braucht. Schließlich sind hier Pioniere am Werk, die Methoden ausprobieren und entwickeln, die so im Unternehmen noch nie zuvor angewendet wurden und die zudem schwer kommunizierbar sind.



Weitere Herausforderungen liegen darin, dass kein Data Scientist dem anderen gleicht. Hier treffen verschiedene Fähigkeiten und Persönlichkeiten aufeinander, was einerseits wichtig für die Bewältigung der Aufgabe ist, andererseits Kommunikation und Orchestrierung erfordert. In der Praxis braucht man für das Management eines solchen Teams neben dem fachlichen und technischen Know-how auch viel Kommunikationsgeschick, Disziplin und eine gute Organisation. Eine Kombination an Fähigkeiten, die so nur wenige Manager oder Data Scientists mitbringen. Damit's trotzdem klappt, im Folgenden zehn Praxistipps.

10 Praxistipps

Praxistipp Nr. 1: Die Aufgabenstellung klar definieren

Bevor das Team loslegt, sollten die Fragestellung und das Ziel allen klar sein. Data Scientists haben viele Ideen und sind kreativ. Wenn die Fragestellung nicht eindeutig definiert ist, besteht insbesondere zu Projektbeginn die Gefahr, auf die falschen Methoden und Techniken zu fokussieren oder sich zu verzetteln und damit wertvolle Zeit zu verlieren.

Empfehlung: Zwischen „Ich hab da mal ne Idee“ und „Ich hab da mal eine Woche lang was programmiert, das ich euch unbedingt zeigen möchte“ – Enthusiasmus ist gut, noch besser ist es, man lenkt ihn in Bahnen!


Praxistipp Nr. 2: Im Kick-off auf eine gemeinsame Arbeitsweise einigen

Das gemeinsame Arbeiten an einem Datenprodukt ist für viele Data Scientists ungewohnt oder gar fremd. Gleich am Anfang sollten daher Standards festgelegt werden, um die Zusammenarbeit zu erleichtern. Dazu gehören beispielsweise die einheitliche Benennung von Variablen und Dateien sowie für alle verbindliche Basisstrukturen und Mindestanforderungen an eine Dokumentation der Codes. Auch sollte geklärt sein, wie sich das Team organisieren möchte. Soll Code versioniert werden? Wenn ja, wie? Wie sollen Aufgaben festgehalten werden? Nach der Kanban-Methode, mit Jira oder in einer Excelliste? Wie diese Fragen beantwortet werden, hängt vom Projektumfang und von den Fähigkeiten im Team ab.

Empfehlung: Egal für welches Hilfsmittel zur Nachverfolgung von Aufgaben man sich entscheidet, es sollte praktikabel sein, dem Projektumfang angemessen und verbindlich als zentrales Dokument oder Board (bei Kanban) von allen im Team genutzt werden.

Praxistipp Nr. 3: Den Überblick behalten mit Hilfe des Backlogs

Gleich zu Beginn des Projektes wird mit der Arbeitsumgebung auch ein Backlog eingerichtet. Dieses Backlog ist der gemeinsame Informationspool, in dem alle Informationsfäden zusammenlaufen und verknüpft werden: Hier werden alle Aufgaben, Ideen und Konzepte hinterlegt und deren Entwicklung kontinuierlich eingepflegt. Im Backlog werden Themen priorisiert, den Teammitgliedern Aufgaben zugewiesen und der Fortschritt jederzeit nachvollziehbar dokumentiert. Alles in einem einzigen, für alle verbindlichen Pool.

 **Empfehlung:** Aufgaben von anfänglichen Konzepten und Ideen zu trennen ist wichtig für die Übersicht im Arbeitsalltag. Dennoch sollte die Nachverfolgung im selben Tool erfolgen, egal ob man sich auf Kanban, Jira, Excel oder etwas anderes geeinigt hat. Abhängig vom gewählten Tool bieten sich zum Beispiel unterschiedliche Farben von Stiften oder Haftnotizen an einer Kanban-Tafel an, oder entsprechende Filter in einer Excel-Liste. Hauptsache übersichtlich, nachvollziehbar und einfach zu verwalten!

Praxistipp Nr. 4: Für den kontinuierlichen Austausch im Team sorgen


Meetings geben für den Projektfortschritt den Takt vor. Die Zielsetzung und der Rahmen verschiedener Meetings werden festgelegt. So werden Meetings nicht als Zeitfresser erlebt, sondern als wirkungsvolle Grundlage für den nächsten Projektschritt.

1. Stand-up: Täglich 15 Minuten für den Überblick, Arbeits-Zwischenstände und das Ansprechen auftretender Probleme. Das Meeting bleibt im zeitlichen Rahmen, bietet also nicht den Raum, Probleme zu lösen, wohl aber dafür, dass sich Teammitglieder für die Lösungsfindung verabreden. Die Aufgabenliste (siehe Praxistipp Nr. 2) dient für diese Meetings als roter Faden. Umgekehrt zeigt sich beim Stand-up ganz schnell, wenn diese Aufgabenliste nicht sauber erstellt ist, das heißt, wenn noch etwas unklar geblieben ist und damit die Effektivität des Stand-up Meetings leidet.

2. Regelmäßig Details besprechen: Wie oft sich das Team trifft, hängt von der Projektphase ab (s. Praxistipp Nr. 6) und von der Komplexität der Aufgabenstellung. Zwei bis drei einstündige Meetings pro Woche reichen in der Regel, damit sich das Team über Ergebnisse, Konzepte und Ideen im Detail austauschen kann. Dieser Austausch ist wichtig um Themen zu besprechen, die den Rahmen des täglichen Stand-up Meetings sprengen würden.


3. Boxenstopp: In größeren Abständen, beispielsweise alle zwei Wochen, nehmen wir uns Zeit für die Retrospektive und den Soll-Ist-Vergleich.

Sind wir auf dem richtigen Weg (s. Praxistipp Nr. 1)? Läuft alles wie vereinbart, werden Code-Standards eingehalten und ist alles ordnungsgemäß dokumentiert (s. Praxistipp Nr. 2)? Der Boxenstopp ist vor allem in der kreativen Phase zu Anfang eines Projekts wichtig, in der noch für die richtige Herangehensweise experimentiert wird.

 **Empfehlung:** Scrum bietet einen geeigneten Rahmen für das Projektmanagement über Meetings, auch wenn das Team nicht in dem Sinn agil in Sprints an einer Produktvision arbeitet.


Praxistipp Nr. 5: Die Entwicklung transparent dokumentieren

Es lohnt sich, trotzdem wird es oft vernachlässigt: das Protokollieren von Entscheidungen, Ideen oder Aufgaben. Jedes Meeting sollte dokumentiert werden. Am schnellsten und sichersten ist das Anlegen neuer Einträge oder das Aktualisieren vorhandener Einträge im Backlog.

 **Empfehlung:** Den Hergang eines Projekts zu protokollieren, ist nicht nur wichtig für die Übersicht, Nachvollziehbarkeit und Transparenz, sondern auch für den Ansporn im Team!

Praxistipp Nr. 6: Verantwortlichkeiten und Rollen zuweisen

In einem Data Scientist Team, das eine komplexe Aufgabenstellung zu bewältigen hat, fallen viele unterschiedliche Rollen oder Verantwortlichkeiten an, die gleich zu Beginn der Zusammenarbeit verteilt werden sollten: der Datenanalyst (z.B. für explorative Aufgaben), der Entwickler/ITler (für technische Fragestellungen), der Projektleiter (trifft Entscheidungen, schlichtet Differenzen, für die Stakeholder-Kommunikation – bringt technisches und unternehmerisches Verständnis mit), der Fachexperte (im Fall einer Aufgabenstellung für eine Bank mit entsprechendem Finanz-Hintergrund) und der Protokollant/Schrittführer (auch im Wechsel, jedes Teammitglied ist mal dran). All diese Aufgaben erfordern unterschiedliche Fähigkeiten, und jeder Data Scientist bringt sein eigenes Skill-Set mit. Bei einem Data Science Team handelt sich um eine sehr heterogene Truppe, dessen muss man sich bewusst sein.

 **Empfehlung:** Die Taxonomie des Autors Michael Hochster zur Unterscheidung von Typ A Data Scientists (Schwerpunkt Statistik und Analyse) und Typ B Data Scientists (Schwerpunkt IT und Programmierung) kann bei der Zusammenstellung und der richtigen Rollenverteilung im Data Science Team hilfreich sein.

Praxistipp Nr. 7: Aus der kreativen Phase zur Herausarbeitung einer gemeinsamen Data Science Pipeline gelangen

Ein Datenprodukt, beispielsweise ein Dashboard oder ein Reporting, entwickelt man in mehreren aufeinanderfolgenden Arbeitsschritten. Am Anfang steht die Datenbeschaffung, dann werden die Daten aufbereitet, damit man etwas damit anfangen kann. Hier kommt es darauf an, eine klare Vorstellung davon zu entwickeln, wohin man will.

Darauf folgen die Analyse, die Modellierung und die Ergebnisaufbereitung, wie dies in **Abbildung 1** (Die verschiedenen Arbeitsschritte in einer Data Science Pipeline) dargestellt ist. In dieser Phase werden verschiedene Wege verfolgt, nicht zuletzt abhängig von den Erfahrungen, den Fähigkeiten und der Kreativität der einzelnen Teammitglieder. In dieser Phase ist der regelmäßige Austausch untereinander wichtig (s. **Praxistipp Nr. 4**). Es ist eine Phase, in der man sich schnell verzettelt oder einen Holzweg einschlägt, aber auch jene Phase, in der innovative Ideen und Ansätze ausprobiert werden und enorme Potenziale liegen. In der kreativen Phase arbeiten die Teammitglieder jeder für sich, deshalb besteht die Gefahr, dass jeder seine eigene Pipeline entwickelt. Das ist jedoch nicht sinnvoll. Die Herausforderung für das Team besteht darin, eine gemeinsame Pipeline zu erarbeiten und sich dafür auf verbindliche Basisstrukturen zu verständigen (s. **Praxistipp Nr. 2**).

Die gemeinsame Data Science Pipeline bildet den Rahmen, hier werden die verschiedenen Stränge, an denen das Team gearbeitet hat, zusammengeführt. Sie deckt den Datenfluss von Anfang (Datenbeschaffung) bis Ende (Datenprodukt) vollständig ab. Von nun an kann das Team zielgerichtet an die Ausarbeitung gehen, was die Steuerung für die verbleibende Projektlaufzeit erheblich vereinfacht.

Empfehlung: Die Data Science Pipeline kann beispielsweise auch eine Verkettung verschiedener Modelle oder die Zwischenspeicherung berechneter Eingangsvariablen und Modellparameter in Tabellen beinhalten.

Die Arbeitsschritte in einer Data Science Pipeline:

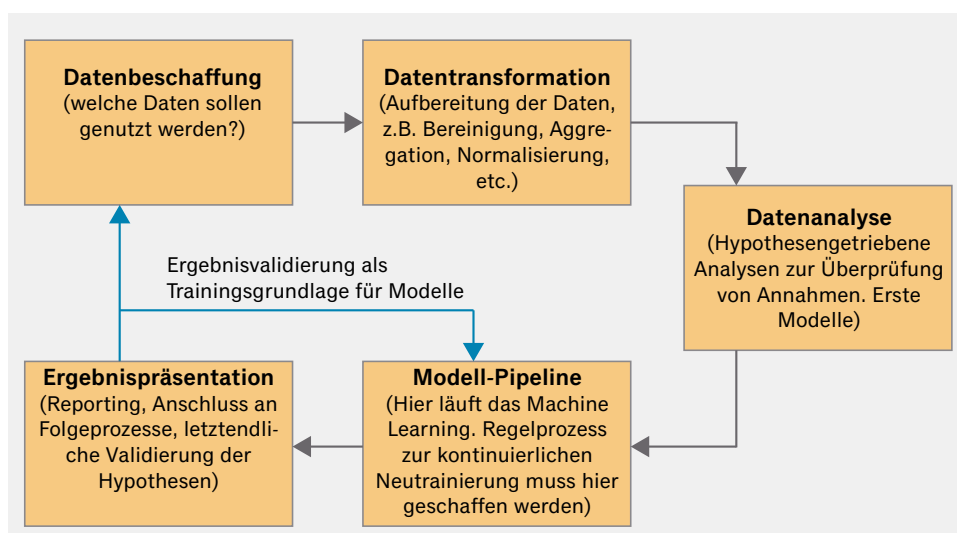


Abbildung 1:
Einfache Data Science Pipeline (in
Anlehnung an: Ojeda et al. (2014) -
Practical Data Science Cookbook).

Praxistipp Nr. 8: Die Ergebnisse in die Praxis überführen

Grau ist alle Theorie: Das Ergebnis der Data Science Pipeline sind Datenprodukte, zum Beispiel ein monatliches Reporting. Ein solches Produkt basiert auf Modellen, die Wahrscheinlichkeiten berechnen und Prognosen abgeben: Wie groß ist die Affinität eines Kunden zu einem bestimmten Produkt? Wie groß ist die Wahrscheinlichkeit, dass ein Kunde dieses oder jenes Produkt kauft? Dies sind Beispiele aus einfachen Modellen. Für die eingangs beschriebene Aufgabenstellung sind die Modelle komplexer und die Ergebnisfindung nicht ganz so leicht. Die Ergebnisse der Data Science Pipeline in die Praxis zu überführen ist daher eine anspruchsvolle Aufgabe für ein Data Science Team. Das Vorstellen, das Bereitstellen auf dem Server, das kontinuierliche Trainieren der Modelle und deren Überarbeitung für neue Umgebungen sind der Praxistest. Teil- oder gar vollautomatisierte Prozesse sind das Ziel.

Empfehlung: Modellergebnisse nützen wenig, wenn sie nicht beispielsweise in Form eines regelmäßigen Reporting weitergenutzt werden. Die kontinuierliche Neuberechnung von Modellvariablen oder das kontinuierliche „Trainieren“ von Modellparametern, das heißt, die Modelle müssen immer wieder aktualisiert werden (z.B. aufgrund veränderten Kundenverhaltens), sollten bedacht werden, damit der Prototyp vom Start weg Produktionsreife besitzt.

Praxistipp Nr. 9: Den Auftraggeber vom Ergebnis begeistern

Was erwartet der Auftraggeber von der Präsentation des Data Science Teams? In der Regel ist der Auftraggeber kein Data Scientist, das heißt, die Ergebnisse müssen erstens so aufbereitet werden, dass sie verständlich sind, und zweitens muss der Wert, der Nutzen der Ergebnisse einleuchtend dargestellt werden. In der Präsentation geht es – sorry, liebe Data Scientists! – weniger um technische Raffinesse, als um unternehmerische Mehrwerte und den Nutzen der Erkenntnisse für weitere Anwendungsfälle.

Im eingangs erwähnten Projektbeispiel, bei der Fraud-Erkennung, ging es darum, Verhaltensmuster und Abweichungen vom normalen Transaktionsverhalten aufzuspüren. Neueste Verfahren, wie zum Beispiel Word-Embeddings, die hier in einem neuen Kontext angewendet wurden und funktionieren, können perspektivisch auch für Customer Journey Analysen eingesetzt werden.



Empfehlung: Bei aller Detailverliebtheit, für die Data Scientists ja auch bezahlt werden: den Auftraggeber niemals mit technischen Details langweilen!

Praxistipp Nr. 10: Zum Ende kommen

Zum Projektende hin fällt es einem Data Science Team oft schwer, einen Schlusstrich zu ziehen. Denn ein Prototyp hat immer den Charakter eines noch nicht zu 100 Prozent fertigen Produktes. Seien es zusätzliche Modellvariablen, weitere Modelle oder das Feintuning von Modellparametern, es gibt endlos Verbesserungsmöglichkeiten. Dann aber bitte nicht in Schönheit sterben! Natürlich können Modelle immer weiter verbessert werden und damit zu immer treffsichereren Prognosen kommen. Der Erfolg an der sechsten Nachkommastelle ist jedoch irgendwann nicht mehr zu rechtfertigen.



Empfehlung: Am Ende strikt den Auftragsumfang für die letzten Meter definieren, Deadlines setzen und diese auch einhalten!

Fazit

Zur Lösung einer komplexen Aufgabenstellung bieten sich Data Science Teams an. Die Teammitglieder, Data Scientists mit unterschiedlichen Fähigkeiten und Arbeits-Schwerpunkten, bringen vielfältige Sichtweisen auf die Aufgabenstellung ein, was ein großer Vorteil gegenüber der Beauftragung einer einzelnen Person ist. Eine gute, heterogene Zusammensetzung und die straffe Organisation des Teams sind nötig, um bei aller Komplexität den Kurs zum Ziel immer im Auge zu behalten.

Was CINTELLIC für Ihr Unternehmen und Ihre Aufgabenstellung tun kann

CINTELLIC stellt für Kunden im Finanz- und Versicherungswesen sowie in anderen Branchen Data Science Teams zusammen und führt sie von der Aufgabenstellung über die Präsentation bis zu Anpassung von Anwendungen in der Praxis. Auch komplexe Aufgabenstellungen wie die eingangs erwähnten lösen unsere Data Science Teams mit viel Erfahrung, Branchen- und Technologieverständnis.

von **Sven Langhoff**,
CINTELLIC Consulting Group

Ansprechpartner



Dr. Jörg Reinnarth
Geschäftsführer
CINTELLIC Consulting Group
joerg.reinnarth@cintelllic.com



Stephan Klöckner
Senior Manager
CINTELLIC Consulting Group
stephan.kloeckner@cintelllic.com

Über CINTELLIC

Die 2010 gegründete CINTELLIC Consulting Group ist eine auf digitales Kundenmanagement spezialisierte Unternehmensberatung, die ihre Klienten vom ersten Konzept bis zur Umsetzung in der Praxis ganzheitlich begleitet. An den Standorten in Bonn, Frankfurt am Main und München arbeiten rund 60 Mitarbeiterinnen und Mitarbeiter.

Zu den Klienten zählen DAX-Konzerne, führende mittelständische Unternehmen und insbesondere zahlreiche sogenannte „Hidden Champions“ mit den Branchenschwerpunkten Banken und Versicherungen, Telekommunikation, IT, Medien, Unterhaltung, Handel, E-Commerce, Versorger und Logistik.

www.cintelllic.com

#jointheteam

CINTELLIC befindet sich auf Wachstumskurs. Vielleicht mit Ihnen? Jetzt Stellenanzeigen entdecken und bewerben!

<https://www.cintelllic.com/stellenangebote/>

Cintelllic im Social Web



Cintelllic GmbH

Remigiusstraße 16
53111 Bonn
t +49 228 92 18 20
info@cintelllic.com
www.cintelllic.com

