



# Data Science: Datenflüsterer entlocken Werbung ihre Geheimnisse

Success Story XING Marketing Solutions GmbH



## Ausgangssituation:

- ⊕ Expertenunterstützung bei der Optimierung eines Empfehlungssystems für das plattformweite Online Advertising

## Probleme:

- ⊕ Ausbau und Optimierung der Ausgangslösung zur Steigerung von Umsatzzielen
- ⊕ Schaffung eines funktionsübergreifenden Teams zur Umsetzung einer datengetriebenen Strategie mit dazugehörigen Methoden und Prozessen
- ⊕ Schaffung von Steuerungsmöglichkeiten, die Geschäftsmetriken an die Optimierungslogik des Empfehlungssystems koppeln

## Lösung:

- ⊕ Wissenstransfer der Spezialkenntnisse der codecentric AG in den Bereichen Data Science und Machine Learning sowie Apache Spark
- ⊕ Beratung und Unterstützung bei der Entwicklung und Umsetzung einer datengetriebenen Strategie
- ⊕ Konzeption und Implementierung eines neuen Empfehlungssystems gemeinsam mit dem Kunden

## Ergebnis:

- ⊕ Befähigung und verbesserte Ownership in den Bereichen Data Science und Data Engineering für das bereichseigene Team
- ⊕ Steigerung des Umsatzes pro Werbe-Einblendung (Revenue per Impression, RPI) im zweistelligen Prozentbereich



Die codecentric AG unterstützte 2016/17 das Unternehmen XING SE am Standort Hamburg in den Bereichen Data Science und Data Engineering. XING ist mit mehr als 13 Millionen Nutzern das führende soziale Netzwerk für berufliche Kontakte im DACH-Bereich. Das 2003 gegründete, börsennotierte Unternehmen gibt auf seiner Plattform Mitgliedern die Möglichkeit, sich in beruflichen Belangen zu informieren, auszutauschen und vor allem zu vernetzen. Gelegenheit hierzu bieten zum Beispiel News-Angebote, Themenportale, Diskussionsgruppen und ein plattformeigener Stellenmarkt.

## Ausgangssituation

Der Geschäftsbereich XING Marketing Solutions verantwortet innerhalb der XING SE die Werbevermarktung. Eine der Kernaufgaben stellt hier das Native Advertising dar: das Einbinden von relevanten Werbeinhalten, die sich dem Erscheinungsbild der Plattform anpassen und redaktionelle Beiträge sowie Beiträge anderer Nutzer ergänzen.

Das bereichseigene Software-Entwicklungsteam hatte zur Umsetzung dieser Aufgabe bereits über längere Zeit eine Lösung entwickelt, die neben einer Architektur am Ende auch aus einem Algorithmus für die Auswahl relevanter Werbemittel bestand. Dieser Algorithmus beruhte auf anonymen Nutzermerkmalen und Statistiken über die Klicks auf Werbemittel. Zum damaligen Zeitpunkt rückte für die XING Marketing Solutions als Hauptverantwortliche für das Online Advertisement der Ausbau des Teams um die Kernkompetenzen Data Science und Data Engineering verstärkt in den Fokus, um innerhalb des Gesamtunternehmens die gesetzten Ziele anhand einer datengetriebenen Strategie und Optimierung eigenständig umsetzen zu können. Neben Neueinstellungen und einer teilweisen Umstrukturierung wurde das bereichseigene Team zusätzlich durch die codecentric AG in diesen Themengebieten unterstützt, um die bestehenden Herausforderungen anzugehen.

Eines der Hauptziele war eine weitere Steigerung des Umsatzes. Hierfür wünschten sich Product Owner und das Management bessere Kontroll- und Steuerungsmöglichkeiten, als sie die bisherige Lösung bot, um Geschäftsmetriken besser mit einer systeminternen Optimierung durch den Algorithmus koppeln zu können. Zudem zeigte sich die Notwendigkeit, im Ausbau der Tätigkeiten eine klar definierte datengetriebene Methodologie einzuführen, um Wirkungen und Änderungen im datengetriebenen Prozess nachweislich messen und von anderen Ursachen unterscheiden zu können. Neben diesen Anforderungen im Bereich Methodologie und Algorithmik



bestanden auch Erfordernisse bei der Anpassung der System-Architektur der Ad Delivery Pipeline. Neben Latenz-Zielen galt es auch in einem funktionsübergreifenden Team aus Software-Entwicklern und Datenwissenschaftlern neben Backend-Komponenten auch dedizierte Systeme für maschinelles Lernen entwickeln und betreiben zu können, die mit den anfallenden großen Datenmengen umgehen konnten.

## Lösung

Zu Projektbeginn existierte bereits eine Tracking-Lösung, die wesentliche Eckdaten und Performance-Kennzahlen des die Werbung ausspielenden Algorithmus erfasste und in einem Dashboard visualisierte. Ein erstes Ergebnis war eine Verbesserung des Trackings und die Identifikation geeigneter Metriken. Hierdurch wurde es möglich, besser zwischen Ursache und Wirkung unterscheiden zu können, ohne dass diese in den Schwankungen der Key-Performance-Indikatoren (KPIs) untergingen.

Um Symptom und Ursache besser voneinander abgrenzen zu können und die Effekte von System-Änderungen messbar nachweisen zu können, wurden zudem die technischen Möglichkeiten von klar definierten A/B-Tests geschaffen. Hierdurch bestand nun die Möglichkeit, die Performance des Werbung ausspielenden Algorithmus objektiv zu messen und mit anderen Lösungsansätzen zu vergleichen.

Im Rahmen dieser Schritte erfolgte zeitgleich mit dem kundeneigenen Team ein Review der historisch gewachsenen Architektur und Implementierung samt des Datenerhebungsprozesses in Hinblick auf Data Quality Assurance und Data Readiness. Hierdurch gelang es auch, ein besseres gemeinsames Verständnis im gesamten Team über die bestehende Lösung sowie die Ursachen von Problemen, die mit ihr zusammenhängen, zu schaffen.

Durch intensive Zusammenarbeit und kontinuierlichen Austausch mit dem Team wurde eine Verbesserung für das Verständnis der Anforderungen und Abläufe eines datenwissenschaftlichen Prozesses sowie bessere Priorisierung von Data Science gegenüber anderen Zielen im Entwicklungsprozess erreicht. Zudem wuchs hierdurch die Ownership und Accountability für das Thema im Team.

Zu diesem Zeitpunkt wurde die Entscheidung getroffen, das bestehende System schrittweise durch ein neues abzulösen. Dadurch wurde es auch möglich, einen Toolstack für Data-Science- und Machine-Learning-Aufgaben zu integrieren, mit dem es möglich war, neue Algorithmen zu erproben, um die Verbesserungen der KPIs und damit Umsatzziele zu erreichen. Durch Schulung der Mitarbeiter und Team-Programming wurden Wissensgrundlagen zur Arbeit mit dem Toolstack vermittelt.



## Das Ergebnis

Es wurden erfolgreich Lösungen umgesetzt und Grundlagen geschaffen, um für die Produktweiterentwicklung anhand intelligenter, selbstlernender Algorithmen neue datengetriebene Produktfeatures zu entwickeln bzw. die bestehenden weiter zu verbessern, in Produktion zu bringen und quantitativ messbar zu evaluieren. Hierauf aufbauend konnte XING Marketing Solutions für den pro Einblendung gemachten Umsatz (RPI) eine Steigerung im zweistelligen Prozentbereich und eine Reduktion der Latenzen erreichen.

Das Beispiel der plattformweiten Personalisierung von Inhalten steht stellvertretend für einen breiten Kreis von Anwendungsfällen, in denen automatisierte Entscheidungen aufgrund einer dynamischen Datenlage anhand von Eigenschaften verschiedener, ein selbstlernendes und adaptives System durchströmender, Datenpunkte getroffen werden müssen.

Die codecentric AG unterstützt mit ihren Experten in den Bereichen Data Science and Data Engineering bei der Auswahl geeigneter Techniken des maschinellen Lernens und der technischen Umsetzung von Datenverarbeitungsstrecken, die selbstlernende Komponenten enthalten. Wir begegnen mit Ihnen den Herausforderungen bei der Entwicklung datengetriebener Lösungen. Dabei ist es egal, ob es sich um das Kerngeschäft, unterstützende Services oder ein neuartiges digitales Geschäftsmodell handelt, dessen Herzstück Daten sind.



**Dr. Michael Plümacher**, *Data Scientist*  
Hamburg  
michael.pluemacher@codecentric.de

”

*„Die XING Marketing Solutions GmbH konnte dank der professionellen Unterstützung der codecentric AG im Bereich Data Science und Big-Data-Technologien binnen kurzer Zeit ein modernes Recommender-System zur Optimierung der Werbeanzeigen in Betrieb nehmen. Unsere Datenwissenschaftler können nun effizienter Hypothesen testen und neue Algorithmen validieren.“*

*Mark Weber, Director Engineering, XING Marketing Solutions GmbH*

”