

Sportwetten & Data Mining - Eine Erfolgsstory?!

© by KMD's



„Data Mining ist für jeden Sportwetter die reale Chance, regelmäßige Gewinne zu erzielen - für den Buchmacher die Möglichkeit, seine Erträge zu maximieren.“

Einfach erklärt ist es eine Form der Statistik, die einen Datenbestand durch verschiedene Methoden mit dem Ziel der Mustererkennung untersucht, also kurzum eine intelligente Datenanalyse.

Bei Sportwetten liegen Spielergebnisse und Quoten, sowie die statistischen Ausgangslagen (Heim-/Auswärtsbilanzen etc.) der Mannschaften von Spielpaarungen vor. Eine entwickelte, methodisch gut durchdachte Data Mining - Software auf dem Gebiet des Sports, berechnet ganz einfach gesagt, mit welchen Wetten, unter welchen Ausgangsbedingungen der Spiele, in der Vergangenheit maximaler Gewinn erzielt worden wäre, speichert die Aspektkonstellationen als statistische Wettsysteme (Strategien) ab und wendet diese auf zukünftige Spielpaarungen mit gleichen Schemas an.

Es ist der (einzig vertretbare) Ansatzpunkt: Was in den vorangegangenen Spielserien über viele Ligen bestmöglich funktioniert hat, kann in der Zukunft nicht gänzlich falsch sein (Man trifft zum Beispiel auch schon mal das korrekte Endresultat von 2:0 am 28.02.09 zwischen Middlesbrough FC und Liverpool FC mit der Quote 42 !!!, weil dies in der England Premier bereits in ähnlichen Konstellationen des öfteren vorkam).

Für den Kunden sind hier speziell Wettbörsen interessant, da einerseits die Quoten stets nahe 100% (1/x) liegen, sowie es andererseits die Möglichkeit gibt, das Ereignis auch Lay zu spielen.

Für den Buchmacher besteht die Möglichkeit, bei den Quoten die ein oder andere Abstufung vorzunehmen, um seine Gewinne zu maximieren, sowie in bestimmten Fällen attraktivere Odds anzubieten.

Als Datenlieferant der Spiele fungiert die Ligaverwaltung Odds-Liga 7.0 Professional mit über 250000 erfassten Resultaten und ein extra Tool, welches jede Spielpaarung in ca. 700 verschiedene Ausgangsaspekte aufschlüsselt (ein Aspekt ist beispielsweise die Punkte pro Spiel der letzten 10 Heimspiele der Heimmannschaft, ein anderer die Anzahl zurückliegender Spielpaarungen der Auswärtsmannschaft, die diese für 5 Siege benötigte).

Bei den Quoten kommt das Datenarchive der weltgrößten Wettbörse Betfair zum Zuge 2004(05) - 2009(10). Hier wurden alle Quoten ausgelesen, welche die meisten Umsätze nahe der 100% -Marke erzielten. Also die Generierung der zuverlässigsten Back-Quoten im Sportwettenbereich, die es gibt.

Unter diesen Voraussetzungen liefert die Data - Mining Software nach Berechnungen eine Anzahl von Aspektkombinationen, die wenn diese so gespielt worden wären, einen maximal möglichen Gewinn in der Vergangenheit je Wettart / Quotenbereich garantierten.

Ein ganz einfaches Beispiel, ohne zunächst überhaupt Ausgangslagen von Spielpaarungen zu berücksichtigen: Nehmen wir Deutschland die 1. Bundesliga und die Wette auf das korrekte Endresultat 1:0 (Stand 2009). Wenn man alle Spiele der letzten fünf Spielserien komplett Back oder Lay auf das Ergebnis 1:0 gespielt hätte, wo erzielte man den Gewinn - waren die Back oder Lay - Quoten Value? In diesem Fall hatte man einen Yield (durchschnittlicher Gewinn/ pro Wette) von rund 5,0% auf Lay 1:0.

Untersucht man dazu die Quotenbereiche, findet man schnell einen Bereich bei den 1:0 Odds zwischen 10.5-16.0, der den Yield schon auf ca. 15% steigert.

Bringt man einen einzigen Ausgangsaspekt mit ein (zum Beispiel Bereich Prozent Heim-/Auswärts Remis beider Mannschaften im Schnitt zwischen 20% und 45%) steigerte sich der Yield schon auf gut 30% pro Wette. Also hat man statistisch immer nach 4 Wochenenden, Monat für Monat sein Kapital stets mehr als verdoppelt.

Diese Konstellation spielt man eben in der Zukunft bzw. der Buchmacher muss eine Quotenkorrektur vornehmen.

Dies ist wirklich nur ein ganz einfaches Beispiel. Es soll die reichhaltigen Möglichkeiten des Data Minings näher bringen und kann in gleicher Form auf alle erdenklichen Wettarten angewendet werden (auch kann man die Quotenbereiche zum Beispiel weglassen, obwohl es sinnvoll ist, bestehende Wahrscheinlichkeitstheorien des Buchmachergewerbes mit zu berücksichtigen).

Hierzu noch ein paar Anmerkungen:

Anzahl Verletzungen/ Gesperrte Spieler (Ersatzspieler bei guter Motivation gleichwertig?!), Spieleinkäufe, Abgänge vor/in der Saison, Trainerwechsel etc. sind alles subjektive Faktoren - in welchem Maße und in welche Richtung diese einzelne Resultate beeinflussen, ist realistisch auf statistischen Gebiet nicht im entferntesten zuverlässig überprüfbar (Eine B-Mannschaft von Deutschland gewinnt ja auch 2007 in England mit 2:1). Die statistischen Daten, auf deren Grundlage die Berechnungen durchgeführt werden, sind zuverlässig, es sind die, die eine Mannschaft als ganzes Objekt mit ihren ganzen Unwägbarkeiten in der jeweiligen Liga bisher abgeliefert hat. Also die einzig vernünftige Grundlage für eine sinnvolle Datenanalyse und eine daraus folgende konstante Generierung von Tipps bzw. Quotenoptimierungen.

So findet man zum Beispiel auch heraus, unter welchen Ausgangsbedingungen von Spielen durchaus mit einem Auswärtssieg mit einer Quote von über 5.0 zu rechnen ist. So waren am 16.05.09 die Siege von Absteiger Karlsruhe in Bremen, sowie Köln in Hamburg keinesfalls überraschend.

Kann man den gleichen Yield auch für die Zukunft erwarten, weniger oder sogar mehr? Eigentlich alles drei. Welche berechneten Konstellationen bzw. welche Wettarten sich letztendlich als zuverlässig erweisen, kann sicher erst nach einer gewissen Zeit gesagt werden. Auch spielt der Gesichtspunkt eine Rolle, ob über eine Vielzahl von europäischen Ligen gleiche Ergebnisse zu beobachten sind.

Zur Zeit können Sie auf der Webseite www.sportwetten-xxl.net die aktuelle Version der Data Mining Software Odds-Liga L/TE System anfordern.

U. Peter / K. Peter
(Das KMD's - Team)