

Vorhersagen im Fußball

Deskriptive und normative Aspekte von Vorhersagemodellen im Sport

Christian Gröschner^{1,2} und Markus Raab^{2,3}

¹Freie Universität Berlin, Erziehungswissenschaften und Psychologie,

²Max-Planck-Institut für Bildungsforschung, Zentrum für Adaptives Verhalten und Kognition

³Universität Flensburg, Institut für Bewegungswissenschaften und Sport

Zusammenfassung. Bisherige Studien zum Vergleich von modellbasierten und realen Vorhersagen im Bereich des Sports zeigten, dass menschliches Vorhersageverhalten durch heuristische Modelle sehr gut abgebildet wird. In Studie I wurden erstmals drei heuristische und ein normatives Modell hinsichtlich ihrer Informationsnutzung bei Vorhersagen von Spielergebnissen zweier Vorhersagejahre der 1. Fußball-Bundesliga miteinander verglichen. Die Ergebnisse der Untersuchung zeigen, dass die heuristischen Modelle dem Bayes-Modell bezüglich der Vorhersagegenauigkeit ebenbürtig bzw. überlegen waren, wobei sie bis zu 50 % weniger Informationen nutzten. Als wichtige Informationen im Fußball wurden das Halbzeitergebnis und der Heimvorteil identifiziert. In einer Folgestudie II zeigten Vorhersagen einer Stichprobe ($N = 208$) zur Fußball-Weltmeisterschaft 2002, dass Laien Experten hinsichtlich ihrer Vorhersagegenauigkeit überlegen waren.

Schlüsselwörter: Fußball, Vorhersagen, heuristische Modelle, normative Modelle

Predicting soccer outcomes: Descriptive and normative aspects of prediction models in sport

Abstract. Previous studies comparing real with model-based predictions in the field of sport have shown that human prediction can be modeled quite accurately by heuristic models. Study I compared three heuristic and one normative model in terms of how well they predicted game outcomes over two seasons in the premier German soccer league. Results showed that the prediction accuracies of the heuristic models were just as good if not better than those based on Bayes' model while using up to 50 % less information. The half time result and the home advantage were identified as the most important items of information in soccer. Study 2 compared the predictions of 208 laypersons on the outcome of the soccer world championship in 2002 with those of experts. Results showed that the laypersons produced more accurate predictions than the experts.

Key words: soccer, prediction, heuristic models, normative models

Wer wird wohl gewinnen? – Dieses ist die alles entscheidende Frage bei den meisten Sportwettkämpfen. Infolgedessen sind private und öffentliche Wetten eine tägliche Erscheinung im Leistungssport. Der Internet-Provider Novomatic erwirtschaftete bspw. mit seiner Subcompany „Admiral“, dem österreichischen Marktführer im Filial-Wettgeschäft, im Jahr 2001 einen Zehn-Millionen-Euro-Gewinn. Zum Jahresabschluss 2003 konnte dieser – bei einem Wetterlösumsatz von 93,70 Millionen € – einen Rohertrag von 27,57 Millionen € aufweisen.

Das Faszinierende an Vorhersagen im Sport ist, dass es bislang kein optimales Vorhersagemodell gibt und wahrscheinlich auch nie geben wird. Im Laufe der Jahre sind viele Versuche unternommen worden Systeme und Algorithmen zur erfolgreichen Vorhersage von Sportergebnissen zu entwickeln. Dabei versuchen sowohl Wettanbieter als auch ihre Kunden in einem Wettstreit verschiedene Vorhersagemodelle immer wieder zu verbessern. Überraschend ist hier, dass oftmals Personen, die behaupten, sich gar nicht in einem Sport auszukennen, mit ihren Vorhersagen am besten abschneiden.

Auf diesem Umstand basierend ergeben sich zwei Forschungsfragen bezüglich modellbasierter und realer Vorhersagen im Sport, denen in dieser Untersuchung anhand der Vorhersage von Fußballergebnissen nachgegangen wird:

Wir danken Uwe Czienskowski für die Simulationsprogrammierungen, R. T. Stefani und Gerd Gigerenzer für Diskussionen und die Unterstützung des Max-Planck-Institutes für Bildungsforschung, Zentrum für Adaptives Verhalten und Kognition, zur Finanzierung dieses Projektes.

Frage 1 (Studie 1): Treffen Vorhersagemodelle unter Verwendung von wenigen Informationen bessere Vorhersagen als Modelle, die viele Informationen einbeziehen?

Frage 2 (Studie 2): Treffen Menschen mit weniger Informationen bessere Vorhersagen als Menschen mit vielen Informationen?

Im Folgenden werden wir die Fragen aus kognitionspsychologischer Perspektive hinsichtlich des Entscheidungsverhaltens von Menschen beantworten. Die den Forschungsfragen zugrunde liegende Frage ist, wie viel Wissen für die richtige Vorhersage eines Sportergebnisses notwendig ist, wie viele Informationen der Zuschauer (oder ein Modell) benötigt, um Vorhersagen in einem sportlichen Wettkampf zu erstellen.

Zur Beantwortung der beiden Forschungsfragen werden zwei zentrale Modellvorstellungen zur Beschreibung menschlichen Entscheidungsverhaltens gegenübergestellt: Der Ansatz der bounded rationality („begrenzte Rationalität“) und der Ansatz der unbounded rationality („unbegrenzte Rationalität“) (Gigerenzer, Todd & the ABC Group, 1999). Ein zentraler Ansatz begrenzter Rationalität ist die Adaptive Toolbox, die sich theoretisch – wie wir zeigen werden – besonders gut für die Vorhersage von Sportergebnissen eignen kann. Der aktuelle Stand der Forschung zu Vorhersagen im Sport bzw. im Fußball präzisiert die Verwendung dieser Modellvorstellungen. Die Modell- (Studie I) und Personenvorhersagen (Studie II) werden im Hinblick auf das resultierende Entscheidungsverhalten, die verwendete Informationsmenge, die Art ihrer Verwendung sowie deren Auswahl dargestellt.

In Studie I gehen wir der ersten Frage nach, ob kontraintuitiv Modelle unter Verwendung weniger Informationen bessere Vorhersagen treffen können, als Modelle mit vielen Informationen. Dazu werden in Studie I Modelle der begrenzten und unbegrenzten Rationalität zur Vorhersage zweier Wettspiellunden (Vorhersagejahre 1998/1999 und 1999/2000) der 1. Fußball-Bundesliga eingesetzt. Diese Modelle unterschieden sich hinsichtlich der Anzahl und Art der verwendeten Informationen.

In einer kleinen Folgestudie II wird untersucht, ob auch Menschen mit weniger Informationen bessere Vorhersagen treffen können. Zur Beantwortung der Frage wurde in Zusammenarbeit mit der Zeitschrift „Spektrum der Wissenschaft“ ein Internetspiel zur Vorhersage des Fußballweltmeisters 2002 durchgeführt und analysiert.

Unbegrenzte versus begrenzte Rationalität

Bezüglich der Beschreibung menschlichen Entscheidungsverhaltens existieren zwei grundsätzlich entgegengesetzte Ansätze. Die Annahmen des Ansatzes der „unbegrenzten Rationalität“ setzen voraus, dass der Mensch über hohe mentale Kapazitäten verfügt, die es ihm ermöglichen, in jeder Entscheidungs- oder Beurteilungssituation über sämtliche notwendigen Informationen zu verfügen oder sie zu sammeln. Diese werden anschließend in geeigneter und dem Problem angepasster Weise gewichtet und bewertet. Erst seit einiger Zeit wird dieser Denkweise eine andere Sicht des menschlichen Umgangs mit Unsicherheit in Entscheidungssituationen gegenübergestellt (vgl. Gigerenzer et al., 1999 für einen Überblick verschiedener Ansätze).

Der Ansatz der „begrenzten Rationalität“ betrachtet eingeschränktes Wissen der Menschen als einen ökologisch sinnvollen adaptiven Vorteil. Dabei geht der Ansatz von begrenzten Kapazitäten und Ressourcen des Menschen aus, aufgrund derer er in vielen Situationen versuchen muss, die bestmögliche Lösung durch geeignete Entscheidungshilfen näherungsweise zu erreichen (Simon, 1990). Diese Entscheidungshilfen, so genannte Heuristiken, beinhalten Wiedererkennungsprozesse zur Einleitung einer gerichteten Informationssuche, einfache Regeln zur Steuerung und Beendigung der Informationssuche sowie einfache Entscheidungsalgorithmen, die mit den gefundenen Informationen arbeiten (Gigerenzer & Todd, 1999). Möchte man bspw. einen Roboter entwickeln, der im Fußball einen Pass durch einen Kopfball weiterleiten kann (direkte Ballannahme), so würde man nach dem klassischen Ansatz der „unbegrenzten Rationalität“ versuchen, diesen Roboter mit allen zur Verfügung stehenden Informationen wie den möglichen Flugbahnen des Balles und weiteren Einflussgrößen, wie Wind, Drall, Entfernung etc., zu versorgen, sodass er unter Verwendung eines Hochleistungscomputers den richtigen Ort und Zeitpunkt errechnen kann, um den Ball zu erreichen. Der Ansatz der „unbegrenzten Rationalität“ geht hingegen davon aus, dass ein Fußballspieler für diese Aufgabe bspw. auf die so genannte Blick-Heuristik (gaze-heuristic) zurückgreift. Aufgrund dieser beginnt er zu laufen, während er den Ball visuell fixiert. Die Heuristik besteht darin, die Laufgeschwindigkeit so anzupassen, dass der Blickwinkel immer konstant bleibt (McLeod & Dienes, 1996). Bei Verwendung dieser Heuristik müsste der Roboter also nur den Blickwinkel konstant beibehalten und seine Laufgeschwindigkeit dementsprechend anpassen, alle weiteren Variablen wie Wind, Drall, Entfernung usw. müssten nicht weiter beachtet und verrechnet werden. Die Verwen-

derung der Heuristik verringert sowohl die notwendige Informationsmenge als auch die benötigten Verarbeitungskapazitäten und führt dennoch zum selben Ergebnis unter erheblich geringerem Aufwand. Mit seinem Konzept der „Adaptiven Toolbox“ geht der Ansatz der „unbegrenzten Vernunft“ davon aus, dass der Mensch über verschiedene Heuristiken verfügt, die entsprechend den Umweltafordernungen zum Einsatz kommen. Diese sind schnell (fast), sparsam (frugal) und robust. Robustheit bezieht sich auf die Fähigkeit einer Heuristik nur die wichtigen und zuverlässigsten Informationen innerhalb einer Umwelt zu berücksichtigen. Aufgrund der Einfachheit der Informationssuch-, Stopp- und Entscheidungsregeln ermöglichen es die Heuristiken dem Menschen sich auch in veränderlichen Umwelten schnell zu orientieren. Alle Heuristiken sind aufgrund ihrer individuellen Struktur jeweils für den Einsatz in einer ganz bestimmten Art von Umwelt eines Lebewesens geeignet, die sich durch eine bestimmte Struktur der in ihr enthaltenen Informationen charakterisiert. Somit entscheidet die in der Umwelt zugrunde liegende Informationsstruktur darüber, welche Heuristik unter den gegebenen Umständen angemessen zur Lösung eines Problems beiträgt (ökologische Rationalität, Gigerenzer, 1997). Die Heuristiken der „Adaptiven Toolbox“ unterscheiden sich in der Komplexität ihrer informationsverarbeitenden Struktur und der Art der Informationen, auf die sie zugreifen und im Umfang der von ihnen benötigten Ressourcen. Gemeinsam ist allen Heuristiken der „Adaptiven Toolbox“, dass sie Schritt-für-Schritt-Prozeduren, eine eingeschränkte Informationssuche, beschränktes Wissen über die Umwelt sowie die Struktur der Umwelt nutzen (Gigerenzer, 1997).

Recognition-Heuristik

Die Recognition-Heuristik ist die einfachste und schnellste Heuristik der „Adaptiven Toolbox“. Ihre Entscheidung zwischen zwei Objekten basiert auf dem Prinzip der Wiedererkennung. Notwendige Bedingung dabei ist, dass dem Entscheider das eine Objekt bekannt ist, das andere jedoch nicht. Hierbei entscheidet die Heuristik zugunsten des wieder erkannten Objekts nach der Regel: Das Bekannte (und Bewährte) ist besser als das Unbekannte. Möchte man bspw. auf den Ausgang eines Tennisturniers wetten, bei dem man nur einen der beiden Spieler kennt, so wird man sehr wahrscheinlich auf den bekannten der beiden Spieler tippen. Das Entscheidungskriterium beschränkt sich auf den Hinweisreiz (Cue) Spieler mit den Ausprägungen bekannt-unbekannt. Dieses einfache Vorgehen macht die Heuristik überaus schnell und effektiv, da sie keine weiteren Informationen über die Objekte oder Situationen sucht, beachtet und verarbeitet. Andererseits ist der Einsatz dieser Heuristik auf Entscheidungen be-

grenzt, bei denen eine der Alternativen bekannt, die andere jedoch unbekannt ist. Muss eine Entscheidung zwischen zwei bekannten oder zwei unbekanntem Objekten getroffen werden, kann die Recognition-Heuristik nicht angewendet werden.

Take The Best-Heuristik

Sind beide Objekte bekannt, muss auf andere Informationen bezüglich der Objekte zugegriffen werden. In einigen Umwelten ist dem Entscheider auch die Nützlichkeit oder Gültigkeit (Validität) von Hinweisreizen für eine Entscheidung bekannt. Wurde ein bestimmter Cue oft erfolgreich bei der Lösung eines bestimmten Problems verwendet, liegt eine hohe Validität dieses Cues nahe. Die Take The Best-Heuristik sucht unter den bestehenden (z. B. bekannten oder vorliegenden) Informationen den Cue mit der höchsten Validität und verwendet diese für die Entscheidung. Ist diese Information für jedes der Objekte bekannt, wird die Informationssuche beendet und die Entscheidung zugunsten des Objekts mit dem höheren Wert getroffen, ohne weitere Information zu beachten. Kann diese Information nicht zwischen den Objekten diskriminieren, wird die Information mit der nächst höheren Validität gesucht und verwendet. Sind bspw. Bayern München und Schalke 04 als Mannschaften bekannt und weitere Informationen wie bspw. ihr Rangplatz, ihre Angriffsstärke (Anzahl erzielter Tore), ihre Verteidigungsstärke (Anzahl der Gegentore), ihre Heimspielstärke etc. verfügbar, dann werden diese Cues ihrer Validität nach solange verglichen bis ein Unterschied (oder ein bestimmter Grenzwert eines Unterschiedes) zwischen diesen Mannschaften diskriminiert. Die Ordnung der Cues nach ihrer Validität berücksichtigt dabei keine eventuell bestehenden Zusammenhänge zwischen den einzelnen Cues.

Entscheidend für die Schnelligkeit aller Heuristiken ist also der Umgang mit den benötigten und zur Verfügung stehenden Informationen. Ist eine Information gefunden, die den Ansprüchen des Modells entspricht, wird die Informationssuche sofort beendet und alle weiteren Informationen werden ignoriert. Ebenso wird der Verarbeitungsaufwand minimiert, da die zur Verfügung stehenden Informationen vor der Entscheidung nicht miteinander in Beziehung gesetzt oder kombiniert werden.

Die Anwendbarkeit und Zuverlässigkeit der beschriebenen Heuristiken der „Adaptiven Toolbox“, auch im Vergleich zu klassischen Modellen wie der linearen Regression, wurde von Gigerenzer und der ABC-Gruppe für verschiedene Umweltkontexte gezeigt, wie etwa am Aktienmarkt, der Wohnungssuche oder der Schätzung der Stadtgröße (Czerlinski, Gold-

stein & Gigerenzer, 1997; Goldstein & Gigerenzer, 2002).

Stand der Forschung für Vorhersagen im Sport

Mit der Zeit wurde eine große Anzahl an statistischen Modellen und Ansätzen zur Vorhersage der populärsten Sportarten entwickelt, wie für American Football (Boulier & Stekler, 2003; Harville, 1980; Stefanie, 1987; Stern, 1991) oder Basketball (Boulier & Stekler, 1999; Schwertman, Schenk & Holbrook, 1996). Einige Modelle können sogar zur Vorhersage in mehreren Sportarten verwendet werden, wie etwa für American Football, Basketball und Fußball (Knorr-Held, 1999; Stefani, 1980).

Für Vorhersagen im Fußball werden bei den verschiedenen Turnierformen, wie beispielsweise den World Cup-Ergebnissen (Kuonen & Röhrli, 2000), vorwiegend statistische Vorhersagemodelle eingesetzt. Besonders der Wettkampfmärkte in dieser Sportart ist ein begehrtes Einsatzziel für Vorhersagemodelle (Dixon & Coles, 1997; Forrest & Simmons, 2000; Maher, 1982; Stefanie, 2003; Stern, 1997). Dabei bemüht man sich, unter all den zur Verfügung stehenden Informationen die validesten zu selektieren, um diese in meist hoch komplexen Algorithmen mit dem Ziel einer möglichst exakten Vorhersage zu verarbeiten. Einige Autoren versuchen einfachere Vorhersagemodelle zu entwickeln, die mit Hilfe möglichst weniger Informationen ein möglichst optimales Vorhersageergebnis erzielen (Kuonen & Röhrli, 2000). Ein Vorteil einfacher Modelle, die eine Vorhersage anhand weniger Informationen ermöglichen, liegt in der Robustheit dieser Vorhersagen, da nur aussagekräftige Informationen beachtet werden und so ein Overfitting durch die verzerrende Wirkung irrelevanter Informationen auf das Vorhersageergebnis vermieden wird.

Ein häufig verwendeter Ansatz im Sport ist die Bestimmung von Teamstärkefaktoren für das Abwehr- und Angriffsspiel einer Mannschaft. Diese werden aus Daten über die Anzahl erzielter Tore und erhaltener Gegentore erstellt, mit deren Hilfe die Verteilungen der Sieg- bzw. Niederlagewahrscheinlichkeiten einer Mannschaft bestimmt werden (Dixon & Coles, 1997; Maher, 1982). Dieses Vorgehen wird durch die Einbeziehung des Faktors Zeit und einer zeitabhängigen Gewichtung der Ergebnisse aus den letzten Spielen mit dem Ziel erweitert, eine Verbesserung der Vorhersagegenauigkeit zu erreichen (Knorr-Held, 1999; Stefani, 2003; Stern, 1997). Als weiterer Faktor zur Bestimmung der Mannschaftsleistung wurde vor allem die Rolle des Heimvorteils untersucht (Clarke & Norman, 1995).

Deskriptive Ansätze zu Spielvorhersagen

Für die Beantwortung der beiden Forschungsfragen ist es wichtig zu wissen, welche und wie viele Informationen bei Vorhersagen im Sport benutzt werden. Durch den Vergleich von Experten- und Novizenvorhersagen ist es möglich, Unterschiede in der Art der Verwendung und Menge der zur Vorhersage verwendeten Informationen zu erfassen. Daraus können Rückschlüsse auf die verwendeten Entscheidungsalgorithmen gezogen und diese mit bestehenden Modellen zur Erklärung menschlichen Verhaltens verglichen werden.

Die Auswahl der relevanten Information hängt u.a. von der Art der betrachteten Sportart und der Form des Turniers ab. Gigerenzer, Hell und Blank (1988) zeigten, dass bei Vorhersagen von Fußballspielen die Basisrate, ein Quotient aus der Anzahl der Siege und der Anzahl insgesamt gespielter Spiele innerhalb des betrachteten Zeitraums (z.B. eine Saison), der wichtigste Hinweisreiz ist. Dieses gilt nach Todorov (2001) auch im Basketball, wo die Basisrate und das Halbzeitergebnis von Basketballfans als die wichtigsten Informationen für eine Vorhersage beschrieben werden. Im Tennis wurde von Richardson, Adler und Hanks (1988) und Silva, Hardy und Grace (1988) gezeigt, dass zwischen dem Gewinn des ersten Satzes und dem Gewinn des Matches ein hoher Zusammenhang besteht.

Eine Reihe der folgenden Untersuchungen aus dem Bereich der Experten-Novizen-Forschung sind nicht im Fußball durchgeführt worden und unterstützen damit nur indirekt die Hypothese, dass Laien im Fußball bessere Vorhersagen machen. Beispielsweise zeigten Heit, Price und Bower (1994) im Basketball, dass Novizen, die nur über sehr wenig oder gar kein Hintergrundwissen in dieser Sportart verfügen, genauso gute Vorhersagen über den Ausgang und den Punkteabstand der Teams machen konnten, wie Experten im Basketball, die über die in amerikanischen Zeitungen existierenden Statistiken der Mannschaften verfügen. Serwe und Frings (2003) zeigten im Tennis, dass auf Rekognitionsdaten von Nicht-Tennispielern und Amateur-Tennispielern basierende Vorhersagen bei Turnieren, wie etwa Wimbledon, ebenso erfolgreich waren, wie Vorhersagen, die aufgrund von Expertenratings (ATP und SEED) erstellt wurden. Dabei folgte das Vorhersageverhalten in fast 90% der Fälle dem Vorgehen der Recognition-Heuristik. Weiterhin sagte Todorov (2001) die Ergebnisse von insgesamt 1.187 Spielen der NBA einer Saison mithilfe von zwei Versionen eines auf der Take The Best-Heuristik basierenden Modells und einem auf dem Bayes-Theorem (vgl. Anhang) basierenden Modell vorher. Der Vergleich der Modellvorhersagen zeigte, dass die

Take The Best-Heuristik ebenso genaue Vorhersagen trifft wie das Bayes-Modell. In einer zweiten Studie wurden Studenten um die Vorhersage von zufällig ausgewählten Spielen aus dem in der ersten Studie verwendeten Datenpool gebeten. Die Teilnehmer der zweiten Studie von Todorov (2001) stuften die von allen Modellen verwendeten Informationen Basisrate und Halbzeitergebnis ebenfalls als die wichtigsten Informationen im Basketball ein. Dabei zeigte sich, dass die verwendeten heuristischen Modelle das Vorhersageverhalten der Teilnehmer sehr gut approximieren.

Andersson, Ekman und Edman (2003) verglichen im Zusammenhang mit der Fußball-Weltmeisterschaft 2002 in Südkorea und Japan die Vorhersagen von Experten und Laien hinsichtlich der Art und Anzahl der verwendeten Informationen. Die Vorhersagen der Laien waren, wenn auch nicht signifikant, besser als die der Experten. Es zeigte sich, dass die Laien ihre Vorhersagen meist aufgrund sehr weniger Informationen erstellten. Dabei verwendeten sie auch eine der Recognition-Heuristik entsprechende Strategie, indem sie sich eher für die Mannschaft entschieden, die ihnen im Zusammenhang mit Fußball bekannter war.

Vorhersagen mit heuristischen Modellen im Fußball sind bislang kaum in Fachzeitschriften veröffentlicht. Ayton und Önkal (1997) untersuchten bspw. in einer unveröffentlichten Studie das Vorhersageverhalten von Laien und Experten bei der Vorhersage der Spiele eines englischen Fußball-Cups. Dabei sagte die naive Gruppe in 95% der Fälle entsprechend der Recognition-Heuristik voraus, d. h. immer wenn eine Versuchsperson nur eine Mannschaft kannte, dass die ihr vom Namen bekannte Mannschaft gewinnen würde. Die Experten, die oft beide Mannschaften kannten, benutzten weitere Informationen für ihre Vorhersagen, konnten aber keine besseren Vorhersagen erreichen als die naive Gruppe.

Studie I: Erstellen Vorhersagemodelle mit weniger Informationen bessere Vorhersagen als Modelle mit vielen Informationen?

Insgesamt werden vier Vorhersagemodelle, von denen drei auf der Take The Best-Heuristik und eins auf dem Bayes-Theorem basieren, verwendet. Für die Vorhersagen werden zwölf Informationen als Prädiktoren (Cues) benutzt, wie sie üblicherweise in den Fachzeitschriften des Fußballs vorliegen.

Daten

Für die Erstellung der Vorhersagen von zwei Wettkampfrunden wurden die Daten aus drei Saisons der

1. Fußball-Bundesliga von 1997/98 bis 1999/2000 verwendet. Damit kann aus den Daten der Saison 1997/98 die Saison 1998/99 und aus der Saison 1998/99 die Saison 1999/2000 vorhergesagt werden. Die Tabellen und Spielergebnisse wurden über die Webseite www.kicker.de bezogen.

Die Simulationen beschränken sich auf die Vorhersage von je 210 Spielen der ersten 15 Mannschaften der Saisons 1998/99 und 1999/2000. Da die letzten drei Mannschaften der 1. Fußball-Bundesliga nach Ende der Saison in die 2. Bundesliga wechseln und wiederum die ersten drei Mannschaften der 2. Bundesliga in die 1. Bundesliga aufsteigen, war es nicht möglich, die Spiele dieser Mannschaften zur Vorhersage in der nächsten Saison mit einzubeziehen.

Die Spiele der Saison 1997/98 wurden zur Bestimmung der notwendigen Cue-Validitäten und der vom Bayes-Modell benötigten Kennwerte, wie der Basisrate und die Likelihood ratio, verwendet. Hierzu wurden die 210 Spiele der Mannschaften, die in der folgenden Saison mitspielten, in die Kalkulationen einbezogen. Auf der Basis dieser Validitäten und Kennwerte wurden die Vorhersagen der Spielergebnisse der Saison 1998/99 erzeugt. Für die Vorhersagen der Saison 1999/2000 wurden die hierzu notwendigen Kennwerte und Validitäten aus den Daten der zuvor vorhergesagten 210 Spiele der Saison 1998/99 gewonnen.

Prädiktoren (Cues)

Die Daten der Abschlusstabellen des letzten Spieltes einer Saison dienten jeweils als Prädiktoren für die Vorhersagen zu den Spielen der nächsten Saison. Die Auswahl der Prädiktoren basiert auf den vor einer Saison in Fußballfachzeitschriften (z. B. Kicker) in Tabellen zusammengefassten Informationen über das Abschneiden der Mannschaften in der letzten Saison. Drei weitere Prädiktoren (Basisrate, Halbzeitergebnis und Stadtgröße) wurden ergänzt, da sie bereits in vorherigen Untersuchungen verwendet wurden (z. B. Gigerenzer, Hell & Blank, 1988; Todorov, 2001). Folgende zwölf Prädiktoren standen für die Vorhersagen zur Verfügung: Siege, Unentschieden, Niederlagen, Punkte, Rang in der Tabelle, erzielte Tore, Gegentore, Tordifferenz, Heimvorteil, Halbzeitergebnis, Basisrate, Einwohnerzahl der Stadt.

Bei der Bestimmung der Güte einzelner Hinweisreize in den verschiedenen Vorhersagejahren (vgl. Tabelle 1) zeigte sich, dass die einzelnen Cues über die Saisons hinweg in ihrer Güte variieren. Die Validität eines Cues wird aus dem Quotienten „Anzahl richtiger Vorhersagen aufgrund dieses Cues und der Ge-

samtzahl an Vorhersagen“ bestimmt. Je zuverlässiger ein Cue vorhersagt, desto valider ist er. Das Halbzeitergebnis hatte in beiden Saisons die höchste Validität. Dabei war der Abstand in Saison 1998/99 mit etwa 10% und in Saison 1999/2000 mit 14% zum jeweils zweitbesten Cue erheblich.

Tabelle 1. Cue-Validitäten

Cue	Saison 1998/1999		Saison 1999/2000	
	Validität	Cue	Validität	Cue
Halbzeitergebnis	0.55	Halbzeitergebnis	0.57	
Erzielte Tore	0.45	Erzielte Tore	0.43	
Rang	0.45	Punkte	0.40	
Gegentore	0.44	Rang	0.40	
Tordifferenz	0.44	Niederlagen	0.39	
Niederlagen	0.44	Gegentore	0.37	
Punkte	0.44	Heimvorteil	0.37	
Siege	0.43	Tordifferenz	0.36	
Basisrate	0.43	Siege	0.36	
Heimvorteil	0.41	Basisrate	0.36	
Stadtgröße	0.40	Stadtgröße	0.35	
Unentschieden	0.34	Unentschieden	0.30	

Die nachfolgenden Cues variierten über die Saisons in ihrer Reihenfolge, wobei sich die Varianz der Validitäten der fünf nachfolgenden Cues nur in einem Bereich von durchschnittlich 1% bzw. 6% bewegte. Diese geringe Varianz basiert dabei auf der hohen Abhängigkeit der Cues untereinander.

Vorhersagemodelle

420 Spiele aus zwei Jahren der 1. Fußball-Bundesliga wurden mithilfe von vier Modellen vorhergesagt; drei Varianten der Take The Best-Heuristik (TTB) aus der Adaptiven Toolbox (TTB-Classic, TTB-Soccer, TTB-Grenzwert) und ein Bayes-Modell. Die drei Varianten der Take The Best-Heuristik werden benutzt, um unterschiedlich viele Informationen für Vorhersagen zu nutzen. Zudem ist die klassische TTB-Heuristik nicht in der Lage Unentschieden vorherzusagen und muss so durch die weiteren Varianten ergänzt werden. Mit der zusätzlichen Vorhersage durch das Bayes-Modell ist außerdem gewährleistet, dass Ergebnisse früherer Forschung integriert werden können (Todorov, 2001). Neu ist damit, dass mit Vorhersagen im Fußball direkte Vergleiche zwischen normativen und verschiedenen Heuristik-Modellen realisiert werden können, während bislang die Forschung nur *ein* heuristisches Modell gegen *ein* normatives Modell verglichen hat.

Heuristik: TTB-Classic

Die klassische Take The Best-Heuristik kann auf Fußballvorhersagen übertragen werden, jedoch können keine Unentschieden vorhergesagt werden. Das Modell kann lediglich eine Entscheidung zwischen Sieg und Niederlage der Teams treffen. Zur Erstellung der Vorhersagen durch das TTB-Modell wurden die Daten der Tabelle der vorhergehenden Saison nach ihrer Validität sortiert. Zusätzlich wurde das Halbzeitergebnis in die Vorhersage einbezogen. In die Berechnung ging das Merkmal mit der höchsten Validität ein. Unterschieden sich die Werte der beiden gegeneinander antretenden Mannschaften bei dieser Information, entschied der Algorithmus für Sieg der Mannschaft mit dem besseren Wert.

Diskriminierte der Wert bei einem Cue nicht zwischen den beiden Mannschaften, da der Wert des Cues für beide Mannschaften gleich groß war, so wurde die Information mit der nächst besten Validität gesucht und anhand dieser entschieden.

Heuristik: TTB-Soccer

Da mit TTB-Classic keine Unentschieden vorhergesagt werden können, wurde die Modell-Variante „TTB-Soccer“ hinzugezogen. Diese verwendete für ihre Vorhersagen nur eine einzige Information, nämlich das Halbzeitergebnis eines Spiels. Steht ein Spiel zur Halbzeit unentschieden, so wird nicht wie bei TTB-Classic eine weitere Information gesucht, sondern ein Unentschieden vorhergesagt. Somit war es möglich, alle drei Ergebniszustände vorherzusagen. Für die Entscheidungen wurde von diesem Modell sonst keine weitere Information verwendet. Dabei ist anzumerken, dass in den betrachteten Vorhersagejahren zur Halbzeit etwa 44% der Spiele unentschieden standen, wobei von diesen nur 40% auch unentschieden ausgingen. Dennoch war das Halbzeitergebnis den Validitätsanalysen nach über die betrachteten Saisons hinweg die valideste Information.

Heuristik: TTB-Grenzwert

Das TTB-Grenzwert-Modell kann im Gegensatz zum TTB-Classic das Halbzeitergebnis dann einsetzen, wenn der Unterschied in der Leistung zwischen zwei Mannschaften einen bestimmten Grenzwert unterschritten hat. Das Modell bildet damit menschliche Entscheidungen ab, die bspw. erst bei leistungsähnlichen Mannschaften weitere Informationen hinzuziehen. Als Prädiktoren für das TTB-Grenzwert-Modell stehen zwei Informationen zur Verfügung: Die Basisraten der Mannschaften und das Halbzeitergebnis des

vorherzusagenden Spiels. Die Basisraten der Mannschaften errechnen sich aus dem Quotienten von Anzahl der Siege und Anzahl aller Spiele einer Mannschaft innerhalb der Saison. Hat also eine Mannschaft innerhalb einer Saison 34 Spiele absolviert und von diesen zehn gewonnen, so errechnet sich die Basisrate wie folgt: $10/34 = .29$. Für den Vergleich der Mannschaften werden Baserate-Odds (Basisraten-Quoten) gebildet, für deren Dichotomisierung ein Grenzwert bestimmt wurde. Dieser Grenzwert gibt den Unterschied zwischen den Basisraten an, ab dem das Modell für Sieg einer der Mannschaften entscheidet. Wird dieser Grenzwert nicht überschritten, so wird eine weitere Information, in diesem Fall das Halbzeitergebnis, betrachtet.

Im ersten Schritt der Datenverarbeitung wird ein Quotient aus den Basisraten der gegen einander spielenden Teams gebildet, die Baserate Odds (Ω_{BR} = Basisraten-Quoten). Im zweiten Schritt kommt der zur Dichotomisierung verwendete Grenzwert zur Anwendung. Der Grenzwert beschreibt einen Wert der Baserate Odds, ab dem diese Information nicht mehr ausreichend ist, um zwischen den beiden Teams zu diskriminieren. Überschreitet also der Wert der Baserate Odds einen vorher festgelegten Grenzwert, stoppt TTB die Suche nach Informationen und entscheidet zugunsten des Teams mit der höheren Baserate. Wird dieser Grenzwert unterschritten, wird das Halbzeitergebnis für die Entscheidung herangezogen. Zur Bestimmung der endgültigen Grenzwerte für die beste Performanz des Modells wurden die Vorhersagen der TTB-Modelle mit verschiedenen Grenzwerten generiert. Diese variierten zwischen Baserate Odds von 0.4 bis 2.0 in Schritten von 0.1.

Das normative Bayes-Modell

Das Bayes-Modell ist ein normatives Modell der Wahrscheinlichkeitstheorie. Dieses ermöglicht die Bestimmung der Auftretenswahrscheinlichkeit eines Ereignisses (Posterior-Wahrscheinlichkeit) aufgrund allgemein bestehender Informationen zu diesem Ereignis unter Berücksichtigung einer aktuellen Stichprobeninformation. Das Modell bedient sich zum einen Informationen aus der Vergangenheit (Prior-Wahrscheinlichkeit: unbedingte Wahrscheinlichkeit des Auftretens dieses Ereignisses aufgrund bisher bestehender Informationen) und aus Informationen einer aktuell bezogenen Stichprobe (Likelihood ratio: bedingte Wahrscheinlichkeit für das Auftreten des aktuellen Ereignisses bezüglich der untersuchten Stichprobe).

Zur Bestimmung der Posterior-Wahrscheinlichkeit für einen Sieg der Mannschaft A [Wahrscheinlichkeit

P (Team A gewinnt)] werden in diesem Fall zunächst die Baserate Odds (Ω_{BR}) der Mannschaften zur Bestimmung der Prior-Wahrscheinlichkeit (unbedingte Wahrscheinlichkeit) berechnet. Die Likelihood ratio (Wahrscheinlichkeitsrate) des aktuellen Halbzeitergebnisses liefert die bedingte Wahrscheinlichkeit für das Auftreten dieses Halbzeitergebnisses, basierend auf den Daten der aktuellen Erhebung, also des vorherzusagenden Spiels. Die Likelihood ratio eines Halbzeitergebnisses gibt somit an, wie wahrscheinlich das Auftreten dieses Halbzeitergebnisses ist, unter der Bedingung das Mannschaft A gewinnt. Aus dem Produkt der Prior-Wahrscheinlichkeit und der Likelihood-Wahrscheinlichkeit ergeben sich die Posterior Odds (Ω_{post}). Diese beschreiben die Wahrscheinlichkeit für das Auftreten des betrachtenden Ereignisses, nachdem die vergangenen und aktuellen Daten betrachtet wurden. Den Posterior Odds für das betrachtete Ereignis (Sieg Mannschaft A) wird im Bayes-Theorem die Gesamtwahrscheinlichkeit der vorliegenden Daten gegenübergestellt, die angibt, wie wahrscheinlich die vorliegenden Daten überhaupt sind. Diese berechnet sich aus der Summe der Posterior Odds des Auftretens des Ereignisses (Sieg Mannschaft A) und den Posterior Odds des Gegenereignisses (Nicht-Sieg Mannschaft A), also der Wahrscheinlichkeit für das Nicht-Auftreten des Ereignisses unter Berücksichtigung der vergangenen und aktuellen Daten (vgl. Anhang für eine formale Beschreibung).

Das Ergebnis dieser Gegenüberstellung ist die Posterior-Wahrscheinlichkeit, also die Wahrscheinlichkeit des Auftretens des Ereignisses „Sieg für Mannschaft A, gegeben der Daten“. Die Siegwahrscheinlichkeit der Mannschaft B lässt sich unmittelbar aus der Siegwahrscheinlichkeit der Mannschaft A ableiten durch: P (Team B gewinnt) = $1 - P$ (Team A gewinnt). Sollten die beiden Ereignisse (Sieg der Mannschaft A und Sieg der Mannschaft B, gegeben der Daten) eine gleich große Wahrscheinlichkeit aufweisen, wird dieses als Unentschieden gewertet.

Bereits nach der allgemeinen Beschreibung des Bayes-Modells ist erkennbar, dass dessen Datenverarbeitung komplexer gestaltet ist als bei den TTB-Modellen. Zudem verwendet das Bayes-Modell für jede seiner Vorhersagen alle zur Verfügung stehenden Informationen, die der Basisraten und die des Halbzeitergebnisses. Um die Halbzeitergebnisse für das Modell nutzbar zu machen, müssen für alle in der Saison auftretenden Halbzeitergebnisse die Likelihood ratios berechnet werden. Die Likelihood ratio ergibt sich aus dem Quotienten der Anzahl der Spiele, die das zur Halbzeit führende Team gewonnen hat, und der Anzahl der Spiele, die das zur Halbzeit zurückliegende Team gewonnen hat. Somit stellt die Likelihood ratio eines Halbzeitergebnisses eine Angabe über die Wahrscheinlichkeit des Sieges der jeweils be-

trachteten Mannschaft dar, wenn dieses Halbzeitergebnis vorliegt.

Beispiel: Ein Halbzeitergebnis mit einer Tordifferenz von einem Tor kam während einer Saison 50-mal vor. Dabei siegte in 27 Fällen das führende Team und in 23 Fällen das zurückliegende Team. Daraus ergibt sich für ein Halbzeitergebnis mit einer Tordifferenz von eins eine Likelihood ratio von $27/23 = 1.17$. Dies bedeutet, dass die Siegchance für das mit einem Tor führende Team bei 1.17 und für das zurückliegende Team bei 0.85 liegt. Ein Wert von 1 bedeutet, dass die Siegwahrscheinlichkeit für beide Teams gleich groß ist.

Die Berechnung der Vorhersagen durch das Bayes-Modell erfolgt in vier Schritten:

Im ersten Schritt werden aus den Basisraten die Baserate Odds ermittelt, indem ein Quotient aus den Basisraten der beiden Teams gebildet wird.

Im zweiten Schritt wird dieser Quotient mit der Likelihood ratio des in diesem Spiel vorkommenden Halbzeitergebnisses multipliziert und so die Posterior Odds gebildet.

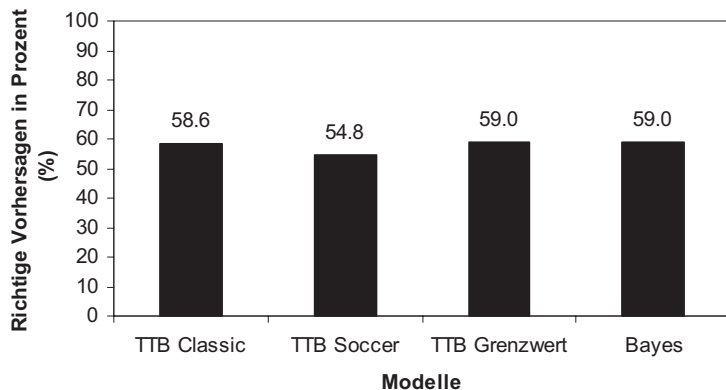


Abbildung 1a. Anzahl richtiger Vorhersagen der Modelle in Saison 1998/99.

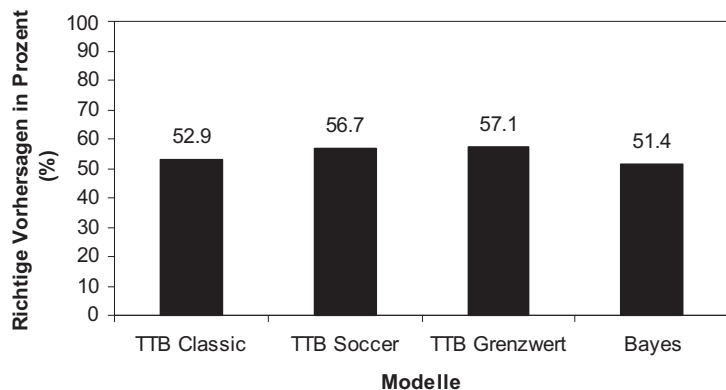


Abbildung 1b. Anzahl richtiger Vorhersagen der Modelle in Saison 1999/2000.

Im dritten Schritt wird die Posterior-Wahrscheinlichkeit des Sieges von Mannschaft A ermittelt, aus der sich gleichzeitig auch die Gewinnwahrscheinlichkeit von Team B ableiten lässt.

Die Posterior-Wahrscheinlichkeit ermittelt sich aus dem Quotienten: $P(\text{Team A gewonnen}) = \text{Posterior Odds} / (1 + \text{Posterior Odds})$.

Im vierten Schritt wird die Entscheidung getroffen. Liegt der Wert der Posterior-Wahrscheinlichkeit über 0.5, so entscheidet das Modell zugunsten von Team A. Liegt der Wert unter 0.5, so entscheidet es für einen Sieg von Team B.

Ergebnisse der Vorhersagen

Modellqualität

Wir fragten, ob Vorhersagemodelle unter Nutzung weniger Informationen ebenso gute Vorhersagen machen, wie Modelle unter Verwendung von vielen Informationen. Wie aus den Abbildungen 1a und 1b ersichtlich ist, sind die Heuristiken dem Bayes-Modell in beiden Saisons ebenbürtig (Saison 1998/99: Bayes-Modell 59%; TTB-Grenzwert 59%, TTB Classic 58.6%, TTB-Soccer 54.8%) bzw. überlegen (Saison 1999/2000: Bayes 51.4%, TTB-Grenzwert 57.1%, TTB Classic 52.9%, TTB-Soccer 56.7%). Dabei ist in beiden Fällen der TTB-Grenzwert unter den Heuristiken die überlegene Modellvariante.

Bei der Vorhersage der Spiele der Saison 1999/2000 aufgrund der Daten der Tabelle des letzten Spieltages der Saison 1997/98 zeigt sich eine deutliche Überlegenheit der Heuristiken TTB-Soccer (57.1%) und TTB Grenzwert (56.4%) gegenüber dem Bayes-Modell (49.4%). Diese Vorhersagen sollten zeigen, inwieweit die im Fußball relevanten Daten und Prädiktoren über die Zeit konsistent sind und wie gut die Modelle über verschiedene Saisons hinweg vorhersagen. Dabei konnten 54 Spiele nicht einbezogen werden, da die betreffenden Mannschaften in der Saison 1997/98 nicht in der ersten Bundesliga spielten und somit für sie keine Ergebnistabelle vorhanden waren. Die guten Vorhersagen von TTB-Soccer sind dadurch zu erklären, dass TTB-Soccer für seine Vorhersagen ausschließlich auf das Halbzeitergebnis zurückgreift, welches von den Prädiktoren der Tabelle der Saison 1997/98

unabhängig ist. TTB-Grenzwert verwendet als erste Information die Basisrate der Saison 1997/98, greift jedoch gegebenenfalls auf das Halbzeitergebnis zurück. Im vorliegenden Fall entschied das Modell nur bei 11.5 % der Spiele im Sinne der Basisrate. Bei allen weiteren Vorhersagen nutzte es das Halbzeitergebnis und war dadurch ebenfalls unabhängig von den Prädiktoren der Tabelle. Das Bayes-Modell verwendete für alle Vorhersagen die Basisraten aus der Saison 1997/98 und die Likelihood ratio der aktuellen Halbzeitergebnisse. TTB-Classic nutzte als ersten Cue das Halbzeitergebnis für seine Vorhersagen. Falls dieses nicht diskriminierte, griff es auf die Daten der Saison 1997/98 zurück. In 49.4 % der Fälle verwendete das Modell die Prädiktoren, die sich aus den Daten der Abschlusstabelle des 34. Spieltags der Saison 1997/98 zusammensetzten. Insgesamt verwendete TTB-Classic durchschnittlich 1.5 Informationen für die Vorhersagen. Damit war es nur 2.8 % (vier Vorhersagen) schlechter als das erheblich komplexere Bayes-Modell.

Informationsnutzung

Bezüglich der Frage, wie viele und welche Informationen bei der Vorhersage von Fußballergebnissen wichtig und notwendig sind um eine gute Vorhersage zu machen, ist ein Vergleich der von den Modellen genutzten Informationen notwendig. TTB Soccer verwendete, wie aus der Modellbeschreibung ersichtlich, ausschließlich den Cue Halbzeitergebnis. TTB-Grenzwert nutzte in 70.5 % (Saison 1998/99) bzw. 79 % (Saison 1999/2000) der Fälle das Halbzeitergebnis und bei 29.5 % (Saison 1998/99) bzw. 21 % (Saison 1999/2000) die Basisrate zur Erstellung der Vorhersagen. TTB-Classic verwendete in 56.6 % (Saison 1998/99) bzw. 52.4 % (Saison 1999/2000) der Fälle das Halbzeitergebnis und entsprechend in 43.3 % (Saison 1998/99) bzw. in 47.6 % (Saison 1999/2000) der Fälle einen Cue der Tabelle. Das Bayes-Modell verwendete beide Informationen bei 100 % der Vorhersagen.

Vergleicht man TTB-Soccer und TTB-Classic miteinander, zeigt sich, dass die zusätzliche Nutzung einer weiteren Information neben dem Halbzeitergebnis im Fall der Saison 1998/99 eine Steigerung der Vorhersagegenauigkeit von 3.8%, und für die Saison 1999/2000 eine Verschlechterung von 3.8% brachte. Die Vorhersagegenauigkeit von Entscheidungsstrategien variiert und kann – wie in dem vorherigen Beispiel demonstriert – sich sogar umkehren.

Ergebnisse zur Validität des „Faktors Heimvorteil“

Hinsichtlich der Anforderung Modellvorhersagen zu generieren, die zum einen menschliches Entscheiden

abbilden und zum anderen eine für reale Vorhersagen zur Verfügung stehende Information nutzen, wurden mit TTB-Classic Vorhersagen unter Einbeziehung des Heimvorteils generiert. Die Simulation berechnet die Vorhersage mit dem Faktor Heimvorteil und weicht damit von der klassischen TTB-Heuristik ab. Dieses Vorgehen ist durch die Annahme motiviert, dass Personen, die im Sport auf der Grundlage eines heuristischen Modells Vorhersagen erstellen, sich des Heimvorteils bewusst sind und diesen berücksichtigen.

Der für die Simulationen verwendete Faktor bildet den durchschnittlichen Heimvorteil über alle Mannschaften innerhalb einer Saison ab. Er wurde anschließend mit den spezifischen Cue-Werten der Heimmannschaften verrechnet. Dazu wurde der Wert des allgemeinen Heimvorteils mit dem jeweiligen Cue-Wert der Heimmannschaft multipliziert und der Wert des Produktes zum ursprünglichen Cue-Wert hinzuaddiert (formale Darstellung des Berechnungsverfahrens siehe Anhang). Dieses Verfahren wurde in Anlehnung an das Vorgehen von Stefani (Stefani, 2003; Stefani & Clarke, 1992) verwendet. Unter Benutzung dieses Faktors und der beschriebenen Verrechnungsformel wurden die Validitäten der einzelnen Cues neu ermittelt und die Vorhersagen unter Verwendung des Modells TTB-Classic erstellt. Bei den zuvor beschriebenen Vorhersagesimulationen verwendete TTB-Classic in über 50 % der Fälle den Cue „Halbzeitergebnis“. Angesichts der Tatsache, dass die meisten „realen Vorhersagen“ aber im Vorfeld eines Spiels gemacht werden und zu diesem Zeitpunkt das Halbzeitergebnis noch gar nicht bekannt ist, wurden bei den hier beschriebenen Vorhersagen von TTB-Classic der Cue „Halbzeitergebnis“ nicht mehr zugelassen. Durch den Ausschluss des Halbzeitergebnisses als Cue und die gleichzeitige Einbeziehung eines Faktors für den Heimvorteil soll ein Vorhersageverhalten simuliert werden, welches dem Verhalten einer Person in einer realen Wettsituation am nächsten kommt. Das Vorgehen basiert auf der Annahme, dass eine vorhersagende Person eine starke Gewichtung des Informationswertes des Cues „Heimmannschaft“ gegenüber dem der Gastmannschaft aufgrund des Heimvorteils vornimmt. Bei der Bestimmung der Validitäten unter Verwendung des Heimvorteils hatte der Cue „Punkte“ in beiden Vorhersagejahren die höchste Validität. Es wurde daher in die Simulation aufgenommen.

Nach Hinzunahme eines Faktors für den Heimvorteil zeigt sich, dass die Validität des Cues „Punkte“ in Saison 1998/99 um bis zu 7.1% und in Saison 1999/2000 um bis zu 6.4% ansteigt (Saison 1998/99: 47.1%; Saison 1999/2000: 51.4%). Damit erreicht diese Information noch nicht die Vorhersagegenauigkeit des Halbzeitergebnisses. Jedoch bedeuten diese Werte, dass sich die Vorhersagen aufgrund nur dieser

einen Information und unter Hinzunahme eines Faktors für den Heimvorteil bei 306 Spielen um bis zu 22 Spiele gegenüber der alleinigen Verwendung des Prädiktors „Punkte“ verbessern.

In dieser Simulation verwendete das TTB-Modell ausschließlich eine Information für die Vorhersagen, da die Werte, die durch die Verrechnung mit dem Faktor entstanden, immer diskriminierten. Es ist davon auszugehen, dass das Modell bei einer Verwendung von gerundeten Werten oder unter Einbeziehung eines Grenzwertes auch weitere Informationen hinzuziehen würde. Dieses könnte durchaus zu einer weiteren Verbesserung der Vorhersagen führen. In diesem Fall stand jedoch der Versuch im Mittelpunkt, die generellen Auswirkungen der Einbeziehung eines Faktors „Heimvorteil“ zu demonstrieren.

Diskussion der Studie I

Der Vergleich der Vorhersagen von drei heuristischen Modellen und einem auf dem Bayes-Theorem basierenden Modell über zwei Vorhersagejahre der 1. Fußball-Bundesliga zeigt, dass die heuristischen Modelle unter Verwendung von weniger Informationen dem statistischen Modell ebenbürtig bzw. ihm sogar in den meisten Fällen überlegen waren. Dabei verwendeten die heuristischen Modelle nicht nur weniger Informationen, sondern verarbeiteten diese unter Aufwendung viel geringerer zeitlicher und rechnerischer Ressourcen. Diese Ergebnisse bestätigen die Ergebnisse von Todorov (2001), der den Modellvergleich zwischen der Take The Best-Heuristik und dem Bayes-Modell im Basketball durchführte. Ein zusätzlicher Vergleich der drei Heuristiken untereinander zeigt, dass TTB-Grenzwert mit seiner Informationsauswahl am besten auf die Umwelt des Fußballs ausgelegt ist. Diese Variante ist speziell auf die Umwelt Sport erstellt worden und verwendet dabei die zur Verfügung stehenden Informationen in einem sehr guten Verhältnis, obwohl die von ihr verwendeten Informationen der Basisrate und des Halbzeitergebnis nicht in der Reihenfolge ihrer Validität genutzt werden. Diese „spezielle“ Heuristik ist dem weitaus komplexeren statistischen Modell, das die gleichen Informationen verwendet, in allen Bereichen überlegen bzw. zumindest ebenbürtig.

Hinsichtlich der Frage nach den bedeutsamsten Informationen für Vorhersagen im Sport zeigen die Ergebnisse, dass es im Fußball wichtig ist, statt der Basisrate das Halbzeitergebnis eines Spieles zu berücksichtigen. Denn im Fußball ist das Halbzeitergebnis zeitlich stabiler als bspw. im Basketball (Todorov, 2001) und stellt somit eine validere Information da. Besonders in den Fällen, in denen zur Halbzeit ein von Unentschieden abweichendes Ergebnis besteht,

ist eine Vorhersage auf Grundlage des Halbzeitstandes angezeigt. Da aber zur Halbzeit noch 44% der Spiele unentschieden stehen, ist die Wahl des Halbzeitergebnis als einzige Information in diesen Fällen nicht sehr hilfreich, denn von diesen Spielen gehen wiederum mehr als 60% mit einem Sieg einer der Mannschaften aus. Bei einer eingehenden Prüfung der in die Untersuchung einbezogenen Daten ist eine Tendenz zu erkennen, dass über beide Vorhersagejahre im Mittel 65.4% der Siege durch die Heimmannschaft erzielt werden. Der Anteil an Heimsiegen der einzelnen Mannschaften variiert dabei zwischen 40% und 88%. Betrachtet man die Spiele, die zur Halbzeit noch unentschieden stehen, ist zu erkennen, dass im Mittel 62.1% dieser Spiele durch die Heimmannschaft gewonnen werden. Die erhöhte Heimsiege-rate, die in den betrachteten Vorhersagejahren gefunden wurde, unterstützt die Ergebnisse zahlreicher Studien zum Heimvorteil (Clarke & Norman, 1995). Die Verwendung eines Faktors für den Heimvorteil bei der Vorhersage der Spielergebnisse der zwei Vorhersagejahre zeigt, dass dieser die Vorhersagegenauigkeit des Prädiktors „Punkte“ um bis zu 7.1% verbesserte. Dieser Wert kann durch die Verwendung von gerundeten Werten für den Faktor „Heimvorteil“ und unter Verwendung weiterer Heimvorteils-Varianten noch verbessert werden.

Studie II: Treffen Menschen mit weniger Informationen bessere Vorhersagen als Menschen mit vielen Informationen?

Ziel der Folgestudie II war die Untersuchung der Frage, ob auch im Fußball Laien unter Verwendung von wenigen unspezifischen Informationen ebenso gute Vorhersagen machen können wie Experten. Dazu wurde im Zuge der Fußball-Weltmeisterschaft 2002 eine Erhebung von Vorhersagen im Internet durchgeführt.

Versuchsteilnehmer

An dieser Erhebung nahmen 208 Personen teil, von denen sich 62 (29.8%) als Experten und 146 (70.2%) als Laien einstufen.

Aufgabe und Versuchsdurchführung

In Zusammenarbeit mit „Spektrum der Wissenschaft“ wurde im Internet eine freie Befragung zum Ausgang der Fußball-Weltmeisterschaft 2002 durchgeführt.

Die Teilnehmer wurden auf der Homepage von „Spektrum der Wissenschaft“ gebeten, aus einer Aufzählung aller Mannschaften die Teilnehmer der zwei Halbfinalspiele auszuwählen und den Weltmeister zu bestimmen. Dabei konnten sie aus einer Scroll-up-Liste aller beteiligten Mannschaften die Plätze 1 bis 4 auswählen.

Im Anschluss wurden sie gebeten, eine Selbsteinschätzung ihres Wissens im Bereich Fußball abzugeben. Dazu standen zwei Optionen, Experte oder Laie, zur Verfügung. Weitere Angaben über verwendete Informationen oder Strategien wurden aufgrund der Konzeption der Befragung nicht erfasst.

Auswertungsprozeduren

Vor der Besprechung der Ergebnisse ist eine nähere Erläuterung zur Bewertung der Selbsteinschätzungen der Teilnehmer notwendig. Im Rahmen dieser Erhebung wird davon ausgegangen, dass selbstbezeichnete Experten sich für Fußball interessieren und den Verlauf der Entwicklungen in der Bundesliga, der Qualifikationsrunden vor großen Turnieren, wie EM oder WM und weiterer Turniere der FIFA, regelmäßig beobachten. Ebenso nehmen sie an dem Geschehen im Umfeld der Mannschaften, Spieler und Trainer teil. Sie besitzen also durchaus einen Überblick innerhalb dieser Sportart und verfügen über Hintergrundinformationen. Hinsichtlich der Selbsteinschätzung der Teilnehmer als Experten ist zu vermuten, dass sie aufgrund von Erfahrungen bei vorangegangenen Vorhersagen über ein Wissen auf metakognitiver Ebene bezüglich ihrer Vorhersageleistung verfügen, dass diese Selbsteinschätzung rechtfertigt.

Bei den selbsternannten „Laien“ wird davon ausgegangen, dass sie sich in ihrer Freizeit nur in geringem Maß oder überhaupt nicht mit den Geschehnissen im Bereich des Fußballs auseinandersetzen. Somit verfügen sie über keine besonderen Hintergrundinformationen, wie etwa die Leistung einer Mannschaft in den letzten Spielen, Trainerwechsel, Fähigkeiten der einzelnen Spieler u. a. Dennoch ist zu vermuten, dass diese Personen durch die Medien in der Zeit vor der Weltmeisterschaft diverse Informationen über den Rangplatz in der Weltrangliste, Mannschaften, Spieler oder Spielverläufe in den Qualifikationsspielen erhalten haben. Diese ermöglichen zwar eine Wiedererkennung der Ländernamen, aber weitere Informationen zur Leistung der jeweiligen Nationalmannschaft wurden nicht so differenziert aufgenommen, wie dieses bei den Experten zu erwarten ist.

Nach diesen Annahmen basieren die von den Laien gemachten Vorhersagen auf wenigen und unspezifischen Informationen. Die Experten konnten

auf ein umfassenderes und differenzierteres Grundwissen zurückgreifen, das im Vorfeld der WM durch aktuelle Informationen angereichert wurde.

Ergebnisse

Von den 208 Teilnehmern sagten insgesamt 37 (17.8%) den Weltmeister richtig vorher. Von den 37 Personen mit einer korrekten Weltmeistervorhersage hatten sich sechs (16.2%) als Experten und 31 (83.8%) als Laien im Bereich des Fußballs eingestuft. Der zweite Platz wurde von insgesamt zwölf Teilnehmern richtig vorhergesagt. Von diesen hatten sich zwei als Experten eingeschätzt. Platz drei und vier wurden nur von je einem Teilnehmer richtig vorhergesagt, die sich beide als Laien eingestuft hatten. Zwei Teilnehmer sagten beide Teilnehmer des Finales und ihre spätere Platzierung richtig vorher. Auch diese stuften sich als Laien ein. Somit sagten 9.7% der 62 an der Befragung teilnehmenden Experten den Weltmeister richtig vorher. Von den 146 Laien tippten 21.2% auf den richtigen Weltmeister. Bei den Ergebnissen dieser Umfrage ist zu beachten, dass die Verteilung von Experten vs. Laien sehr unausgeglichen war und daher die absoluten Zahlen der Ergebnisse täuschen können. Die signifikanten Ergebnisse bestätigen die Vermutung, dass die Laien zu ebenso guten bzw. sogar besseren Vorhersagen im Sport fähig sind wie Experten ($\chi^2(1, N = 208) = 3.97, p < .03$).

Diskussion der Studie II

Anhand der Daten ist gut zu erkennen, wie überraschend die Ergebnisse dieser Weltmeisterschaft für alle waren. Nur 18% der Teilnehmer sagten den Weltmeister richtig vorher. Dabei sagten weniger als 10% der selbstbezeichneten Experten den Weltmeister richtig vorher. Von den Laien lagen über 20% mit ihrem Tipp richtig.

Zusammenfassend unterstützen die Ergebnisse die Hypothese, dass bei der Vorhersage von Fußballergebnissen Laien dazu ebenso gut in der Lage sind, wenn nicht sogar besser, die Spielergebnisse vorherzusagen, wie Experten. Wir vermuten, dass die Art und Anzahl der verwendeten Informationen als auch kognitive Illusionen der Experten diesen Befund erklären können. Diese Ergebnisse bestätigen und erweitern die bereits in anderen Sportarten gemachten Beobachtungen (Heit et al., 1994; Todorov, 2001; Andersson et al., 2003; Serwe & Frings, 2003). Weitere Experimente im Fußball sollten mögliche Ursachen der Expertise auf die Vorhersageleistung systematisch analysieren (vgl. Plessner, Trefzer, Hess & Fischbach, 2005).

Generelle Diskussion

Die hier vorgestellte Hauptstudie I und deren Folgestudie II wurden durch die Frage motiviert, wie viele Informationen für die Vorhersage von Spielergebnissen im Fußball notwendig sind. In diesem Zusammenhang wurden die Spielergebnisse zweier Vorhersagejahre der ersten Fußballbundesliga unter Verwendung von drei heuristischen Modellen, basierend auf der TTB-Heuristik, und einem auf dem Bayes-Theorem basierenden Modell vorhergesagt. Diese wurden in Bezug auf ihre Vorhersagegenauigkeit, die Zahl der verwendeten Informationen sowie deren Verarbeitung miteinander verglichen. Mit dieser Untersuchung wurde der Vergleich von heuristischen und normativen Vorhersagemodellen auf den Fußball übertragen. Zudem wurden im Gegensatz zu vorherigen Studien (Todorov, 2001) mehrere heuristische Modelle mit einem normativen Modell verglichen. Des Weiteren wurden die Simulationen anhand mehrerer Vorhersagejahre durchgeführt, wobei die Spiele einer Saison unter Verwendung der Prädiktoren aus der vorherigen Saison vorhergesagt wurden. Es zeigte sich, dass die Modell-Varianten der TTB-Heuristik dem Bayes-Modell hinsichtlich der Vorhersagegenauigkeit ebenbürtig und in einer Saison sogar überlegen waren. Diese Leistung erreichen diese Modelle dabei unter Verwendung von 25% bis 50% weniger Informationen und einem deutlich geringeren Rechenaufwand. Die Ergebnisse der Studie I erweitern die Aussagen anderer Studien im Bereich des Sports (Todorov, 2001), dass auch im Fußball einfache heuristische Modelle der Adaptiven Toolbox komplexen normativen Modellen, wie dem Bayes-Modell, in ihrer Vorhersagegenauigkeit gewachsen und sogar überlegen sind. Die Untersuchung der Ergebnisstruktur der für diese Simulationen betrachteten Vorhersagejahre zeigte, dass 44% aller Spiele zur Halbzeit noch unentschieden stehen, jedoch nur etwa ein Viertel aller Spiele unentschieden endeten. Im Fall eines Unentschieden zur Halbzeit ist die Einbeziehung des Heimvorteils durchaus angeraten, da 62% der Spiele durch die Heimmannschaft gewonnen wurden. Zudem ist die Einbeziehung des Halbzeitergebnisses bei den meisten Fußballwettsystemen, wie etwa Oddset, nicht zugelassen. Um sich der Simulation von realen Vorhersagesituationen im Fußball anzunähern, wurden die 420 Spiele der Saisons 1998/99 und 1999/2000 unter Ausschluss des Halbzeitergebnisses, jedoch unter Berücksichtigung des Heimvorteils erneut vorhergesagt. Die Verwendung eines Faktors für den Heimvorteil in Modellvorhersagen wird durch die Tatsache gestützt, dass diesem Phänomen in zahlreichen Untersuchungen im Sport ein bedeutender Stellenwert zugesprochen (Clarke & Norman, 1995) und auch in vielen statistischen Modellen zur Ergebnisvorhersage hinzugezogen wird (Dixon & Coles, 1997; Stefani, 2003;

Stefani & Clarke, 1992). Die Ergebnisse der Simulationen zeigten, dass die Vorhersagegenauigkeit des Cues „Punkte“ durch die Einbeziehung des Heimvorteils um 7.1% bzw. um 6.4% verbessert wurde. Dieses entspricht einer Verbesserung von bis zu 22 Vorhersagen bei 306 Spielen. Dabei konnte das Modell der TTB-Heuristik aus simulationstechnischen Gründen nur eine Information für die Vorhersagen verwenden. Es ist zu vermuten, dass die Vorhersagen weiter verbessert werden, wenn durch eine Anpassung des Simulationsmodells und der Datenverrechnung ein Zugriff auf wenige aber relevante Informationen möglich wird.

In der Folgestudie II wurde der Frage nachgegangen, ob Laien im Fußball ebenso gute Vorhersagen machen können wie Experten in dieser Sportart. Die in der zweiten Studie gezeigten Ergebnisse bestätigen die von Andersson et al. (2003) im Fußball, von Heit et al. (1994) im Basketball sowie von Serwe und Frings (2003) im Tennis gemachten Beobachtungen, dass Laien aufgrund weniger unspezifischer Informationen durchaus bessere Vorhersagen erzeugen können als Experten. Im Verlauf der Studie wurde mehrfach versucht, die Vorhersagen von Personen zu Fußballspielen über die staatliche Lotto-Toto-GmbH und das Oddset-Team der staatlichen Lotteriegesellschaft in München zu beziehen, um diese in einen direkten Vergleich mit modellbasierten Vorhersagen stellen zu können. Jedoch konnten diese Daten, auch nach mehrfacher Anfrage auf verschiedenen Ebenen der Geschäftsleitungen aus Datenschutzgründen nicht zur Verfügung gestellt werden. Nach Angaben der verantwortlichen Stellen hätte dieses zudem einen zu hohen Verarbeitungsaufwand für die bestehenden Rohdaten von Seiten der Gesellschaften mit sich gebracht. Dabei ist allerdings die Frage, ob in diesem Fall nicht auch eventuell die Befürchtung dahinter stand, dass mit einer Freigabe der Daten womöglich das Auffinden eines einfachen Vorhersagemodells unterstützt werden würde, das jedermann zugänglich ist und somit die Gewinnspannen solcher Unternehmen schmälern könnte. Immerhin suggerieren die Betreiber solcher Spielsysteme, dass sie ihre Quoten aufgrund unglaublicher Datenmengen präzise errechnen und somit nur ein Wettender, der ebenfalls auf einen möglichst großen Datenpool zurückgreift, eine Chance auf einen guten Tipp haben kann.

Wer wird nächster Deutscher Meister im Fußball? Wir wissen es nicht; es ist aber auf der Grundlage dieser und vergleichbarer Ergebnisse anzuraten, den Vorhersagen von Laien mehr zu vertrauen als bisher angenommen.

Literatur

- Admiral Sportwetten GmbH. Presseinformation [03.06.2003]. *Admiral Sportwetten schafft Turnaround im Rekordjahr 2002*. Verfügbar unter: http://www.admiral.ag/040presse/020Presse_Archiv/PA_Admiral_BPK030603_final.pdf. [25.06.2004].
- Andersson, P., Ekman, M. & Edman, J. (2003). Forecasting the fast and frugal way: A study of performance and information-processing strategies of experts and non experts when predicting the World Cup 2002 in soccer. *SSE/EFI Working Paper Series in Business Administration*, No. 2003.
- Ayton, P. & Önkal, D. (1997). *Forecasting football fixtures: Confidence and judged proportion correct*. Unpublished manuscript, Department of Psychology, City University, London.
- Boulier, B. L. & Stekler, H. O. (1999). Are sport seedings good predictors? An evaluation. *International Journal of Forecasting*, 15, 83–91.
- Boulier, B. L. & Stekler, H. O. (2003). Predicting the outcomes of National Football League games. *International Journal of Forecasting*, 19, 257–270.
- Clarke, S. R. & Norman, J. M. (1995). Home ground advantage of individual clubs in English soccer. *The Statistician*, 44, 509–521.
- Czerlinski, J., Goldstein, D. & Gigerenzer, G. (1997). Information environments and algorithms that exploit them. *Swiss Journal of Economics and Statistics*, 133, 201–218.
- Dixon, M. J. & Coles, S. G. (1997). Modelling association football scores and inefficiencies in a football betting market. *Journal of the Royal Statistical Society: Series C Applied Statistics*, 46, 265–280.
- Forrest, D. & Simmons, R. (2000). Forecasting sport: the behaviour and performance of football tipsters. *International Journal of Forecasting*, 16, 317–331.
- Gigerenzer, G. (1997). Bounded Rationality: Models of fast and frugal inference. *Swiss Journal of Economics and Statistics*, 133, 201–218.
- Gigerenzer, G., Hell, W. & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 513–525.
- Gigerenzer, G. & Todd, P. M. (1999). Fast and frugal heuristics: The Adaptive Toolbox. In G. Gigerenzer, P. Todd & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 3–34). Oxford: Oxford University Press.
- Gigerenzer, G., Todd, P. & the ABC Research Group (Eds.) (1999). *Simple heuristics that make us smart*. Oxford: Oxford University Press.
- Goldstein, D. G. & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109, 75–90.
- Harville, D. (1980). Predictions for National Football League games via linear-model methodology. *Journal of the American Statistical Association*, 75, 516–524.
- Heit, E., Price, P. C. & Bower, G. H. (1994). A model for predicting the outcomes of basketball games. *Applied Cognitive Psychology*, 8, 621–639.
- Knorr-Held, L. (1999). Conditional prior proposals in dynamic models. *Scandinavian Journal of Statistics*, 26, 129–144.
- Kuonen, D. & Röhrl, A. S. (2000). Was France's World Cup win pure chance? *Student*, 3, 153–166.
- Maher, M. J. (1982). Modelling association football scores. *Statistica Neerlandica*, 36, 109–118.
- McLeod, P. & Dienes, Z. (1996). Do fielders know where to go to catch the ball or only how to get there? *Journal of Experimental Psychology: Human Perception and Performance*, 22, 531–543.
- Plessner, H., Trefzer, I., Hess, P. & Fischbach, J. (2005). The influence of intuition, reflection, and knowledge on the accuracy of predicting football results. In EAES (Eds.), *Programme book of the European Congress of Social Cognition in Würzburg 2005* (p. 256). Würzburg: Universitätsverlag.
- Richardson, P. A., Adler, W. A. & Hanks, D. (1988). Game, set, match: psychological momentum in tennis. *The Sport Psychologist*, 2, 69–76.
- Schwartzman, N. C., Schenk, L. & Holbrook, B. C. (1996). More probability models for the NCAA Regional Basketball tournaments. *The American Statistician*, 50, 34–38.
- Serwe, S. & Frings, C. (2003). *Who will win Wimbledon 2003? The recognition heuristic in predicting sports event*. Manuscript submitted for publication.
- Silva, J. M., Hardy, C. J. & Grace, R. K. (1988). Analysis of psychological momentum in intercollegiate tennis. *Journal of Sport and Exercise Psychology*, 10, 346–354.
- Simon, H. A. (1990). Invariants of human behaviour. *Annual Review of Psychology*, 49, 1–19.
- Stefani, R. T. (1980). Improved least square football, basketball and soccer predictions. *IEEE Transactions on Systems, Man, and Cybernetics*, 10, 116–123.
- Stefani, R. T. (1987). Applications of statistical methods to American football. *Journal of Applied Statistics*, 14, 61–73.
- Stefani, R. T. (2003, 13–20 August). *Estimating the probability of a home win, draw and away win in soccer*. Paper published in Proceedings of the International Statistical Institute, Berlin: Germany.
- Stefani, R. T. & Clarke, S. (1992). Predictions and home advantage for Australian rules football. *Journal of Applied Statistics*, 19, 251–259.
- Stern, H. (1991). On the probability of winning a football game. *The American Statistician*, 45, 179–183.
- Stern, H. (1997). How accurately can sport outcomes be predicted? *Chance*, 10, 19–23.
- Todorov, A. (2001). *Predicting real outcomes: When heuristics are as smart as statistical models*. Unpublished manuscript, New York University.

Christian Gröschner

University Fribourg
 Department of Psychology
 Rue de Faucigny 2
 1700 Fribourg
 Schweiz
 E-Mail: christian.groeschner@unifr.ch

Anhang

Die Schritte der Datenverarbeitung für das TTB-Grenzwert-Modell und das Bayes-Modell und die Berechnung des Heimvorteils sind für den besseren Textfluss im Anhang für Replikationen formal dargestellt.

Schritte der Datenverarbeitung des TTB-Grenzwert-Modells

- Schritt eins: Berechnung der Baserate Odds (Ω_{BR}): BR_{TeamA} / BR_{TeamB} .
- Schritt zwei: Vergleich der Baserate Odds mit dem Grenzwert.
- Schritt drei: Der Wert der Baserate Odds liegt über dem Grenzwert \rightarrow Entscheidung für den Sieg der Mannschaft mit der höheren Baserate.
- Schritt vier: Der Wert der Baserate Odds liegt unter dem Grenzwert \rightarrow Einbeziehung des Halbzeitergebnis und Entscheidung für Sieg der führenden Mannschaft bei diesem Ergebnis.

| Siege Team A – Siege Team B | \geq Grenzwert \rightarrow Sieg des Teams mit mehr Siegen

| Siege Team A – Siege Team B | $<$ Grenzwert \rightarrow Sieg für Team mit Führung im Halbzeitergebnis

Für jeden der verwendeten Datensätze wurde im Vorfeld ein Threshold (Grenzwert) für die Baserate Odds berechnet. Dieser lag für die verschiedenen Saisons bei unterschiedlichen Basisraten-Differenzen.

Formale Darstellung des Bayes-Theorems

$$P(H1/D) = \frac{P(H1) * P(D/H1)}{P(H1) * P(D/H1) + P(H2) * P(D/H2)}$$

$P(H1/D)$: Posterior-Wahrscheinlichkeit für das Eintreten des Ereignisses, gegeben der Daten

$P(H1)$: Prior-Wahrscheinlichkeit, allgemeine Wahrscheinlichkeit des Auftretens des Ereignisses

$P(D/H1)$: Bedingte Wahrscheinlichkeit (Likelihood ratio), Auftretenswahrscheinlichkeit der aktuellen Daten unter der Bedingung des Auftretens des Ereignisses

$P(H2)$: Prior-Wahrscheinlichkeit des Gegenereignisses (oder des Nicht-Eintreffens des Ereignisses)

$P(D/H2)$: Bedingte Wahrscheinlichkeit (LR), Auftretenswahrscheinlichkeit der aktuellen Daten unter der Bedingung des Auftretens des Gegenereignisses

Schritte der Datenverarbeitung des Bayes-Modells

- Erster Schritt: Berechnen der Baserate Odds (Ω_{BR})
 $\Omega_{BR} = BR_{Team A} / BR_{Team B}$
- Zweiter Schritt: Berechnen der Posterior Odds (Ω_{post})
 $\Omega_{post} = \Omega_{BR} * LR$
- Dritter Schritt: Berechnung der Gewinnwahrscheinlichkeit für Team A (gegeben der Daten)
 $P(A \text{ won}/D) = (\Omega_{post}) / (1 + \Omega_{post})$
- Vierter Schritt: Vorhersage:
Wenn $P(A \text{ won}/D) > 0,5 \rightarrow$ Sieg für Team A;
Wenn $P(A \text{ won}/D) < 0,5 \rightarrow$ Sieg für Team B.
Wenn $P(A \text{ won}/D) = 0,5 \rightarrow$ Unentschieden.

Basisraten (Baserates = BR): Anzahl der Siege/Anzahl der Niederlagen eines Teams

Baserate Odds (Ω_{BR}): $BR \text{ Team A} / BR \text{ Team B}$

Likelihood ratio (Aufretenswahrscheinlichkeit = LR): Anzahl der Siege Team A bei vorliegendem Halbzeitergebnis / Anzahl der Siege Team B bei vorliegendem Halbzeitergebnis

Posterior Odds (Ω_{post}): $\Omega_{BR} * LR$

Berechnungsformel für Cuewert mit Heimvorteil

Neuer Cue-Wert Heimteam = Cue-Wert Heimteam + (Cue-Wert Heimteam * h Heimteam)