

8 Zusammenfassung und Diskussion

Ziel dieser Arbeit war es, verschiedene Ansätze zu betrachten, um die Spielstärke der Fußballbundesligisten in der Saison 2014/2015 zu modellieren. Zunächst wurde hierfür das statische Bradley-Terry Modell behandelt. Nach einer kurzen Vorstellung dieses Modells wurde es zu einem ordinalen Modell erweitert. Zusätzlich wurde noch ein globaler und ein teamspezifischer Heimeffekt eingeführt, welche getrennt voneinander zu betrachten sind. Durch seinen Zusammenhang mit den generalisierten linearen Modellen konnte auf die vorgestellten Modelle die bekannte goodness-of-fit Statistik sowie die Likelihood-ratio-Statistik angewandt werden. Die Anwendung dieser Modelle zeigte, dass eindeutig ein Heimeffekt für die verwendeten Daten existiert. Ein teamspezifischer Heimeffekt brachte im Vergleich zu einem globalen Heimeffekt jedoch keine Verbesserung. Die geschätzten Stärken der Mannschaften spiegelten die tatsächlichen Platzierungen der Mannschaften zwar relativ gut wider, jedoch zeigten die quasi-Standardabweichungen auf, dass sich die meisten Spielstärken nicht signifikant voneinander unterscheiden lassen. Dies warf die Vermutung auf, dass die Spielstärken der einzelnen Mannschaften geclustert sein könnten. Kapitel 4 beschäftigte sich daher mit dem Clustering der Spielstärken durch Regularisierungsmethoden. Zunächst wurde hier das Ranking LASSO vorgestellt. Diese Methodik belegt die Bradley-Terry Likelihood mit einem Strafterm für alle paarweisen Spielstärkekoeffizienten. Dieser Strafterm ist eine Verallgemeinerung des fusionierten LASSO-Strafterms von Tibshirani et al. (2005). Um keine inkonsistenten Schätzer zu erhalten, werden die Inhalte des Strafterms stärker gewichtet als die Abnahme der Effekte. Ein Nachteil dieses Vorgehens besteht darin, dass die Methodik erst angewandt werden kann, sobald jedes Team nicht mehr jedes Spiel gewonnen oder verloren hat, da sonst die Spielstärken dieser Teams als $+\infty$ oder $-\infty$ geschätzt werden. Bei einer Unterscheidung der Ergebnisse in fünf Kategorien schwächt sich diese Bedingung etwas ab. Hier ist eine Modellierung möglich, sobald jedes Team nicht mehr durchgehend dasselbe Ergebnis erzielt hat. Als Auswahlkriterien für das Clustering werden das Akaike und das Bayessche Informationskriterium verwendet. Bei Betrachtung des BIC als Auswahlkriterium und globalem Heimeffekt lassen sich die Mannschaften in sechs verschiedene Cluster einteilen, wobei diese teilweise sehr ähnlich sind, so dass eine Einteilung in vier Cluster vertretbar scheint. Die ersten drei Cluster sind dabei von den vier stärksten Teams der Liga besetzt. Diese Teams hatten sich in der abgelaufenen Saison deutlich von den restlichen Teams abgesetzt, so dass diese Einteilung sinnvoll erscheint.

8 Zusammenfassung und Diskussion

In Kapitel 5 wurde nun versucht herauszufinden, ob sich die Spielstärke der Mannschaften durch einzelne Kovariablen erklären lässt. Dazu wurde das Modell so erweitert, dass sich Kovariablen, welche eindeutig einem Verein zugeschrieben werden können ins Modell mit aufnehmen lassen. Zunächst beschränkte sich diese Erweiterung auf Kovariablen, welche sich für jede Mannschaft über die Saison nicht verändern. Dies waren der Lizenzspielertat und der Marktwert der einzelnen Mannschaften. Später wurde das Modell noch so angepasst, dass auch spieltagsabhängige Kovariablen modelliert werden können. Diese Erweiterung fand Anwendung bei der Betrachtung der durchschnittlichen Marktwerte der Startaufstellungen der Mannschaften. Bei der Analyse der Ergebnisse für die erstgenannte Erweiterung fiel auf, dass trotz Berücksichtigung der Marktwerte oder auch der Vereinsbudgets vier der ersten fünf Mannschaften der Abschlusstabelle als stärkste Mannschaften klassifiziert wurden. Nur der FC Bayern München fällt hier ins Mittelfeld ab. Werden die Marktwerte als unabhängige Variablen und die ermittelten Spielstärken als abhängige Variablen betrachtet und führt mit diesen eine lineare Regression durch, so ist festzustellen, dass sich über 55% der Streuung durch ein lineares Modell und fast 63% der Streuung durch ein additives Modell erklären lassen. Daraus lässt sich schließen, dass die gewählte Kovariable einen erheblichen Teil der Spielstärken der Mannschaften erklären kann. Bei Verwendung der durchschnittlichen Marktwerte der Startaufstellungen als Kovariablen gehört der FC Bayern München überraschenderweise trotzdem zu den stärksten Teams bei den ermittelten Spielstärken. Dies lässt vermuten, dass diese Wahl der Kovariablen eventuell doch nicht so geeignet ist. Bei Betrachtung eines Regressionsmodelles, wie gerade beschrieben, lassen sich jedoch fast 57% bei einem linearen Modell bzw. über 69% bei einem additiven Modell der Streuung des Modells so erklären.

Ein anderer Ansatz legte einen exponential-gewichteten moving-average Prozess zur Modellierung der Spielstärke der einzelnen Mannschaften zugrunde und wurde in Kapitel 6 beschrieben. Bei einem solchen Ansatz werden wesentlich weniger Parameter geschätzt als bei der zuvor vorgestellten Methodik. Daher ist das Ziel eines solchen Modells, Ergebnisse zu erhalten, welche möglichst nahe an denen des klassischen Bradley-Terry Modells liegen. Der Vorteil dieses Ansatzes liegt darin, dass dynamische Spielstärken für jede Mannschaft erhalten werden. Für jede Mannschaft werden getrennte Spielstärken für Heim- und Auswärtsspiele modelliert. Die Heimspielstärke lässt sich mit Hilfe eines Glättungsparameters sowie der vergangenen Ergebnisse der Heimspiele einer Mannschaft iterativ ermitteln. Die Auswärtsspielstärke ergibt sich äquivalent. Bei der Wahl von gemeinsamen Startwerten für jede Mannschaft lassen sich bei den ermittelten Stärkekurven die Veränderungen während der Saison gut erkennen. Bei individuellen Startwerten für jede Mannschaft sind

8 Zusammenfassung und Diskussion

diese Verläufe nicht zu erkennen. Die Spielstärke ergibt sich vor allem aus den Endplatzierungen der Vorsaison. Zum Vergleich der Modelle mit dem klassischen Bradley-Terry Ansatz wurde der rank probability score eingeführt. Dieser gibt an, wie gut die geschätzten Ergebnisse auf die tatsächlichen Daten passen. Für beide EWMA-Modelle finden sich sowohl Argumente für als auch Argumente gegen ihre Anwendung. Ein Nachteil dieser Betrachtungsweise ist die strikte Trennung zwischen Heim- und Auswärtsspielen. Es wäre wünschenswert, wenn sich die getrennten Betrachtungen zu einem gemeinsamen Modell zusammenfügen ließen. Ein weiterer Ansatz, um ein dynamisches Modell zu erhalten, ist dem vorgestellten klassischen Ansatz aus Kapitel 3 sehr ähnlich. Durch Modellierung der Spielstärken für jeden Spieltag einzeln ergibt sich bereits ein dynamisches Modell. Zusätzlich wurden aber noch die Likelihoods mit Hilfe einer Kernfunktion gewichtet. Somit sollen weiter zurückliegende Spiele weniger Einfluss auf die Spielstärke erhalten. Um zu entscheiden, welche Bandbreite für die Kernfunktion gewählt werden soll, wurde ein prädiktives Modell auf die Daten der Saison 2013/2014 angewandt. Durch den nun gewählten Ansatz lässt sich der rank probability score deutlich verbessern. Auch die Spielstärkenentwicklung der einzelnen Vereine über die Saison ist bei diesem Modell gut erkennbar. Hier ist jedoch wieder zu beachten, dass sich vor allem zu Beginn der Saison aufgrund der geringen Anzahl an Daten sehr hohe quasi-Standardabweichungen ergeben. Eine signifikante Unterscheidung der Spielstärke der Mannschaften ist daher meist nicht möglich.

Für einen Anwender ist von Interesse, wie gut die Modelle prognostizieren. Kapitel 7 beschäftigte sich daher mit der Prognose von Ergebnissen. Zunächst wurde der jeweilige rank probability score ausgewählter Modelle miteinander verglichen. Hierbei konnten nur kleine Unterschiede zwischen den Modellen festgestellt werden. Daher wurde anschließend für diese Modelle überprüft, wie gut sie Spielausgänge vorhersagen. Bei einer Simulationsstudie schnitt dabei das Modell mit den gewichteten Likelihoods am besten ab. Bei ausschließlichen Tipps auf die Heimmannschaft liegt die Zahl der durchschnittlich getippten Spiele jedoch deutlich höher als bei diesem Modell. Aufgrund dessen wurde getestet, wie hoch die Erfolgsquote bei Tipps auf das wahrscheinlichste vom Modell prognostizierte Ergebnis liegt. Dadurch wurde erreicht, dass sich die Trefferquote auf über 50% erhöht. Als Letztes wurden zusätzlich noch die Wettquoten für alle Spiele betrachtet. Da die Quoten für Heimsiege geringer sind als für Auswärtsspiele, wird vermutet, dass sich trotz einer geringeren Anzahl richtig getippter Spiele durch gezielt platzierte Wetten höhere Gewinne erzielen lassen. Bei einer Simulation von 10000 Durchgängen konnte man feststellen, dass sich nur bei drei Modellen im Schnitt ein geringer Gewinn

8 Zusammenfassung und Diskussion

erzielen lässt. Die Schwankungen der Ergebnisse sind zusätzlich sehr hoch. Tipps auf den wahrscheinlichsten Spielausgang lieferten bessere Ergebnisse. Bei vier Modellen beträgt die Rendite etwa 7%. Ausschließliches Tippen auf Heimsiege führt zu einer Rendite von 12.5%. Wird zusätzlich die Rentabilität eines Tipps berücksichtigt, lassen sich die Renditen deutlich steigern. Das beste in dieser Arbeit vorgestellte Modell weist eine Rendite von über 16% auf. Bei ausschließlichen Tipps auf einen Heimsieg unter Berücksichtigung der Rentabilität liegt die Rendite bei 28.9%. Eine Wettstrategie, bei der die Wetten nur nach der größten Rentabilität platziert werden, verschlechtert für die meisten Modelle den erwarteten Gewinn. Nur beim gewichteten dynamischen Modell kann so die Rendite auf über 13% gesteigert werden. Abschließend sollen die Ergebnisse nochmals hinterfragt und diskutiert werden.

Für die vorliegenden Daten lassen sich mit Hilfe der vorgestellten Modelle leichte Gewinne durch gezielte Wetten erzielen. Da auch Wettanbieter ihre Quoten über statistische Modelle modellieren, welche wesentlich komplexer sind als die in dieser Arbeit vorgestellten Modelle mit maximal einer Kovariablen, kann angezweifelt werden, dass die Modelle allgemeingültig sind und auch für zukünftige Partien gute Ergebnisse liefern. Es ist jedoch trotzdem überraschend, dass das Modell mit Berücksichtigung der Startaufstellungen keine besseren Ergebnisse liefert als die anderen Modelle, da hier eigentlich für jedes Spiel zusätzliche Information zur Verfügung steht. Dies könnte wohl vor allem an der Wahl des BIC-Kriteriums bei der Clusterbildung liegen. Dieses ist sehr restriktiv und wird daher nur wenige Cluster erzeugen. Durch die Einteilung in Cluster weist das Modell sehr ausgeglichene Ergebniswahrscheinlichkeiten aus. Daher beraubt sich das Modell klarer Aussagen. Ebenso kann die Wahl der Kovariablen hinterfragt werden. Die in dieser Arbeit benutzten Kovariablen beruhten hauptsächlich auf Schätzungen, welche hohe Fehler aufweisen können. Interessant wäre ebenfalls, eine Verbindung zwischen den einzelnen Modellen herzustellen, wie zum Beispiel zwischen den gewichteten Likelihoods und verschiedener zusätzlicher Kovariablen.

Die Modellierung von Sportereignissen ist ein sehr weites Feld, in welchem sich ständig neue Möglichkeiten ergeben und welches sich in den nächsten Jahren wohl stetig weiterentwickeln wird. Besitzt man noch zusätzliches Insiderwissen wie Matthew Benham, so kann mit statistischer Modellierung im Sportbereich auch sehr viel Geld verdient werden.