

Topics: natural sciences

Keywords: Datenanalyse, Modellierung

## Vorhersage von Treffern in Biathlonwettkämpfen mit maschinellem Lernen

Thomas Maier<sup>1</sup>, Daniel Meister<sup>2</sup>, Severin Trösch<sup>1</sup>, Jon Peter Wehrli<sup>1</sup>

<sup>1</sup>Eidgenössische Hochschule für Sport Magglingen EHSM, Schweiz; <sup>2</sup>Datahouse AG, Schweiz; [thomas.maier@baspo.admin.ch](mailto:thomas.maier@baspo.admin.ch)

### Theoretischer Hintergrund

Schiessfehler in Biathlonwettkämpfen wirken sich bedeutend auf das Schlussklassement aus. Der erschwerende Effekt der vorangehenden körperlichen Belastung sowie biomechanische Einflussfaktoren auf die Trefferquote wurden bereits untersucht (Hoffman, Gilson, Westenburg, & Spencer, 1992; Sattler, Buchecker, Gressenbauer, Müller, & Lindinger, 2017). Allerdings ist bisher unbekannt, ob und wie genau sich einzelne Treffer und Fehlschüsse vorhersagen lassen. Für automatische Vorhersagen eignen sich Algorithmen aus dem maschinellen Lernen, welche komplexe Zusammenhänge aus gegebenen Daten generalisieren, um Vorhersagen für neue Fälle zu berechnen (Bauer & Kohavi, 1999).

### Fragestellungen

- Welche Faktoren beeinflussen die Trefferquoten in Biathlonwettkämpfen?
- Wie gut lassen sich zukünftige Treffer und Fehlschüsse mit maschinellem Lernen vorhersagen?

### Methode

Wir aggregierten und untersuchten Daten von 118'300 Schüssen aus vier Wettkampfsaisons (2012/13 bis 2015/16) mittels explorativer Datenanalyse und trainierten verschiedene Vorhersagemodelle aus dem maschinellen Lernen mittels zeitsegmentierter Kreuzvalidierung. Als Basismodell diente eine logistische Regression mit einer Eingangsvariablen. Zusätzlich wurden zwei komplexe Modelle verwendet: ein Boosting-Modell basierend auf Entscheidungsbäumen und ein künstliches neuronales Netzwerk. Mit den finalen Modellen berechneten wir Vorhersagen für 34'340 zukünftige Schüsse der nächsten Wettkampfsaison (2016/17).

### Ergebnisse

- Tiefere Trefferquoten zeigten sich in den Disziplinen Sprint und Verfolgung verglichen mit den Disziplinen Einzel und Massenstart ( $p < 0.01$ , Cohen's  $h = 0.14$ ). Im Stehendschiessen war die Trefferquote tiefer als im Liegendschüssen ( $p < 0.01$ ,  $h = 0.15$ ). Im Vergleich zu den anderen Schüssen in der gleichen Position zeigte der erste Liegendschuss und der fünfte Stehendschuss eine tiefere Trefferquote ( $p < 0.01$ ,  $h = 0.08$  und  $p < 0.05$ ,  $h = 0.05$ ).
- Das Boosting-Modell prognostizierte zukünftige Treffer und Fehlschüsse mit einer Fläche unter der ROC-Kurve von 0.62, 95% VI [0.60, 0.63], leicht besser als das künstliche neuronale Netzwerk (0.61, [0.60, 0.62],  $p < 0.01$ ) und die logistische Regression (0.60, [0.59, 0.61],  $p < 0.01$ , Eingangsvariable: vorgängige positionsspezifische Trefferquote).

### Diskussion

Die Trefferquote im Biathlon variiert zwischen den Disziplinen, den Schiesspositionen und den Schussnummern einer Serie. Die vorgängige positionsspezifische Trefferquote eines Athleten oder einer Athletin ist aber der wichtigste Faktor, um einzelne Treffer vorherzusagen. Ein grosser Zufallseffekt bleibt bestehen, den auch komplexe Modelle nicht relevant verringern können: Die Anzahl der Fehlschüsse eines Athleten oder einer Athletin pro Wettkampf variiert zufällig um ca. 1-2 Schüsse.

Im Training sollte somit primär die allgemeine Trefferquote in beiden Schiesspositionen unter wettkampfnahen Bedingungen optimiert werden, evtl. mit einem besonderen Fokus auf dem ersten Liegendschuss und dem letzten Stehendschuss. Echtzeitvorhersagen der Trefferwahrscheinlichkeiten einzelner Schüsse könnten zudem Wettkampfübertragungen bereichern.

#### **Referenzen**

Bauer, E., & Kohavi, R. (1999). An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Machine Learning*, 36(1), 105-139.

Hoffman, M. D., Gilson, P. M., Westenburg, T. M., & Spencer, W. A. (1992). Biathlon shooting performance after exercise of different intensities. *International Journal of Sports Medicine*, 13(3), 270-273.

Sattlecker, G., Buchecker, M., Gressenbauer, C., Müller, E., & Lindinger, S. J. (2017). Factors discriminating high from low score performance in biathlon shooting. *International Journal of Sports Physiology and Performance*, 12(3), 377-384.