

# Vorhersage der Spielersperre beim virtuellen Automatenspiel und bei Sportwetten

Vadim Kufenko



UNIVERSITÄT  
HOHENHEIM



*In Bearbeitung*



- 1 Einleitung
- 2 Daten
- 3 Methoden
- 4 Ergebnisse: virtuelle Automaten
- 5 Ergebnisse: Sportwetten
- 6 Diskussion und Zusammenfassung

## Hintergrund

- Identifizierung des **problematischen und pathologischen Spielverhaltens** ist ein essentieller Bestandteil der **Warn- und Sperrsysteme**.
- Nach dem GlüStV 2021 müssen die Anbieter ein auf **“wissenschaftlichen Erkenntnissen beruhendes, auf Algorithmen basierendes automatisiertes System zur Früherkennung von glücksspielsuchtgefährdeten Spielern und von Glücksspielsucht einsetzen”**.

## Motivation

- Wie können solche Systeme **funktionieren** und wie lassen sich diese **evaluieren**?
- Wie kann ein **anbieterübergreifendes System** funktionieren?

## Unsere Studie

- Wir analysieren die **Fremd-** und **Selbstsperrn** von Spielern anhand eines einzigartigen deutschen anbieterübergreifenden Datensatzes aus dem Safe Server in Schleswig-Holstein.
- Unsere Arbeit wurde von einigen Studien zur Analyse der Sperrn inspiriert: Haefeli et al. (2011), Brosowski et al. (2012), Philander (2014), Percy et al. (2016), Mijić and Varga (2017), Ukhov et al. (2020), Auer and Griffiths (2022).
- Unsere Ziele sind:
  - eine anbieterübergreifende Analyse durchzuführen ✓
  - diverse Modelle und Balancierungs Methoden zu vergleichen ✓
  - zusätzliche Metriken zu testen ✓
  - die **Sperrnlogik** der **SpielerInnen** und der **Anbieter** zu vergleichen ✓

- 1 Einleitung
- 2 **Daten**
- 3 Methoden
- 4 Ergebnisse: virtuelle Automaten
- 5 Ergebnisse: Sportwetten
- 6 Diskussion und Zusammenfassung

## Features

- Aus den Rohdaten wurden Daten für die vier größten Anbieter zusammengestellt.
- **Konto Information** (Kontoerschöpfung, Ein- und Auszahlungen, Bilanz, Limit);
- **Spielbezogene Information** (Einsätze, Gewinne, Verluste, Boni, Spielzyklen und Runden);
- **Timing** (Dauer, Sessions, Pausen und Unterbrechungen);
- **Spielverhalten** (Spielwechsel, Präferenzen, Aktivitäten nachts oder am Wochenende);
- **Anbieterbezogene Information** (Anbieter Indikatoren, Multikonto).

## Abgeleitete Indikatoren

- **Niveaus** (Anzahl, Summe, Mittelwerte);
- **Dynamik** (Wachstumsraten und Veränderungen);
- **Dispersion** (Variationskoeffizient, Schiefe, Kurtosis, MAD, Interquartilbereich).

## Datensatz

- **Aggregierte Querschnittsdaten** für vier Anbieter zwischen November 2019 und Juli 2022.
- Insgesamt 9363 Beobachtungen (Konten) und  $> 100$  Variablen.
- Fremd- und Selbstsperrern als Zielvariable (siehe weiter).

- 1 Einleitung
- 2 Daten
- 3 Methoden**
- 4 Ergebnisse: virtuelle Automaten
- 5 Ergebnisse: Sportwetten
- 6 Diskussion und Zusammenfassung

## Potentielle Herausforderungen und Lösungen

- **Kollinearität** → Dimensionsreduktion und Variablenselektion basierend auf Redundanz (Dormann et al., 2013, Malato, 2021).
- **Unausgewogene Klassenverteilung (Selbstsperrung  $\approx 12,5\%$  und Fremdsperrung  $\approx 4,5\%$ )**  
→ Balancierung der Training Daten beim Training:
  - Zufälliges Over- und Undersampling (**RusRose**, Batista et al., 2004, Menardi and Torelli, 2012, Lemaître et al., 2017),
  - Zufälliges Undersampling und mehrheitsgewichtetes Minoritäts-Oversampling (**MWMote**, Barua et al., 2014, Kovács, 2019)
- **Keine soziodemographischen Daten.**



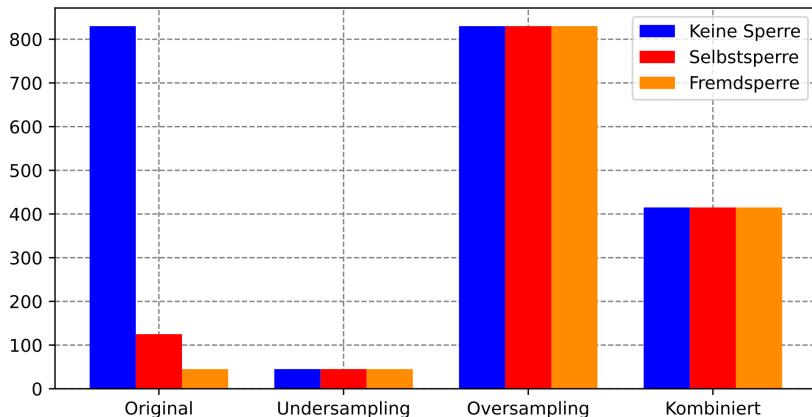
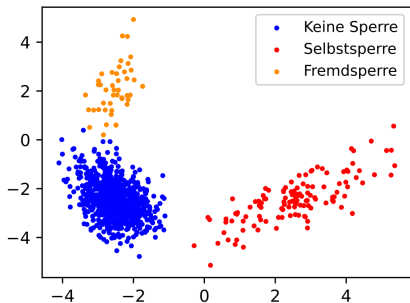
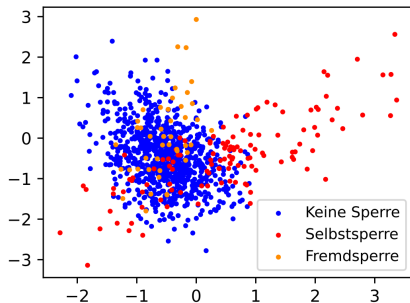


Abbildung (1) Klassenverteilung und Balancierung: Beispiel mit 1000 Beobachtungen



(a) Leichte Klassifikationsaufgabe



(b) Schwierige Klassifikationsaufgabe

Abbildung (2) Verteilung der Klassen

## Modelle und Metriken

- **Logistische Regression** als Benchmark.
- **Random Forest** (RF, Breiman, 2001).
- **Extreme Gradient Boosting** (XGB, Chen and Guestrin, 2016).
- **Categorical ordered Gradient Boosting** (CatGB, Ostroumova et al., 2017).
- Zusätzlich für Sportwetten: **Light Gradient Boosting** (LightGB, Ke et al., 2017).
- Stratifiziertes Train-Test-Split-Verfahren (80/20 für Automaten und zusätzl. 50/50 für Sportwetten), wiederholte stratifizierte Kreuzvalidierung und randomisierte Parameteroptimierung.
- Erweiterte Anzahl der Metriken für die **Test Daten**: AUROC, AUPRC, F1, Precision und Recall. Metriken

- 1 Einleitung
- 2 Daten
- 3 Methoden
- 4 Ergebnisse: virtuelle Automaten**
- 5 Ergebnisse: Sportwetten
- 6 Diskussion und Zusammenfassung

Tabelle (1) Unsere Studie im Kontext der relevanten Literatur

Studie	Methoden	Zeitraum	Größe	Features (Anzahl)	Balancing
Haefeli et al. (2011)	logit	1 M	300	16	zufälliges Sampling
Philander (2014)	logit, GLM (el. net), ANN, SVM, RF	24 M	530	11	–
Percy et al. (2016)	logit, GLM (el. net), ANN, SVM, RF	23 M	1475	38	SMOTE
Haeusler (2016)	logit, ANN	13 M	2696	18	Rus
Mijić and Varga (2017)	RF, GB	24 M	530	–	–
Ukhov et al. (2020)	GB	>36 M	10000	40	–
Finkenwirth et al. (2020)	logit, RF	12 M	19683	20	Rus
Hopfgartner et al. (2022)	AB, DT, ExT, GB, RF	14 M	25720	34	–
Unsere Studie	logit, RF, XGB, CatGB	30 M	9363	74	MWMote Rus, Rose

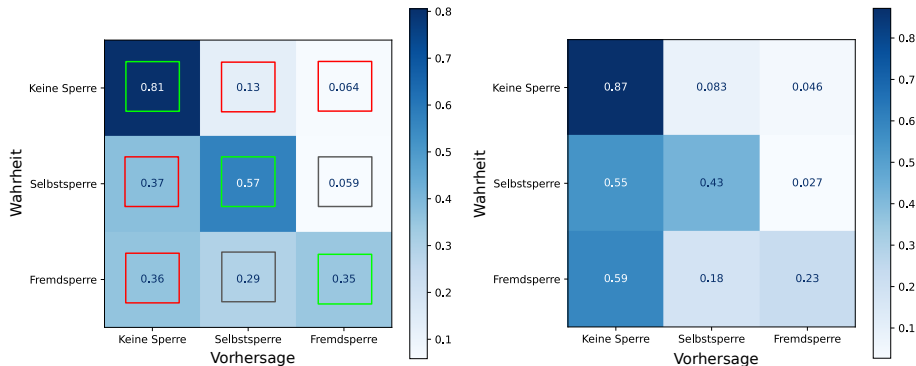
Hinweis: '–' bedeutet nicht explizit erwähnt

Tabelle (2) Unsere Ergebnisse (Test Daten) im Kontext der relevanten Literatur

Studie	Accuracy	AUROC	Precision	Recall/Sensitivity	Specificity
Haefeli et al. (2011)	0.766	–	–	0.782	0.746
Philander (2014)	0.621–0.665	0.491–0.551	0.333–0.571	0.018–0.291	0.811–1.000
Percy et al. (2016)	0.650–0.700	0.680–0.760	–	0.620–0.710	0.660–0.690
Hausler (2016)	0.721	–	–	0.800	0.662
Mijić and Varga (2017)	0.630–0.654	–	–	–	–
Ukhov et al. (2020)	–	0.870–0.920	0.270–0.450	0.420–0.600	–
Finkenwirth et al. (2020)	–	0.390–0.760	–	0.030–0.740	0.560–0.760
Hopfgartner et al. (2022)	–	0.579–0.787	–	–	–
Unsere Studie					
Binär	0.764-0.831	0.791-0.798	0.360-0.405	0.573-0.663	0.793-0.835
Multiclass (Selbstsperre)	0.74-0.828	0.775-0.794	0.347-0.436	0.428-0.595	0.779-0.872

Weitere Metriken

# Ergebnisse: virtuelle Automaten

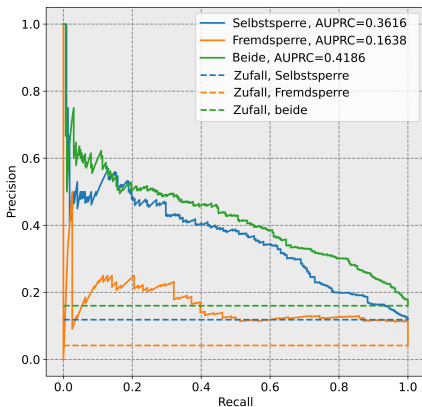


(a) RusRose RF

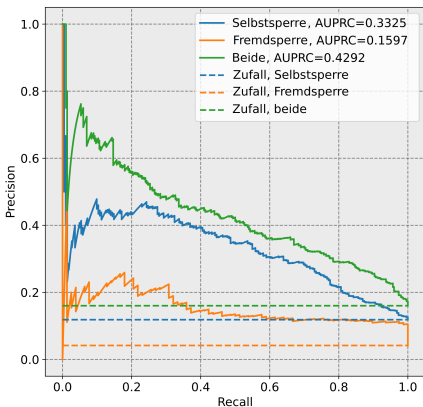
(b) MWMote RF

Abbildung (3) Konfusionsmatrix für RF

# Ergebnisse: virtuelle Automaten



(a) RusRose RF

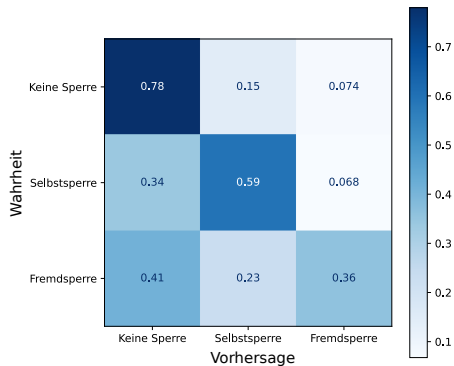


(b) MWMote RF

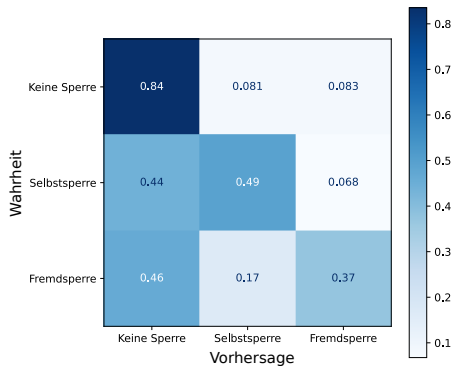
Abbildung (4) Precision-Recall Kurve für RF



# Ergebnisse: virtuelle Automaten



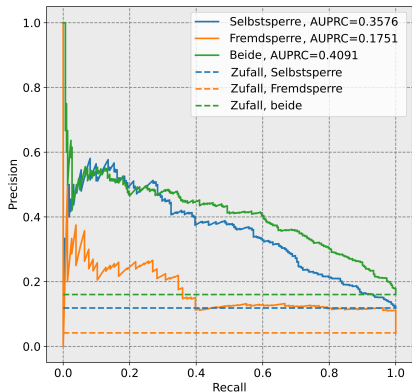
(a) RusRose XGB



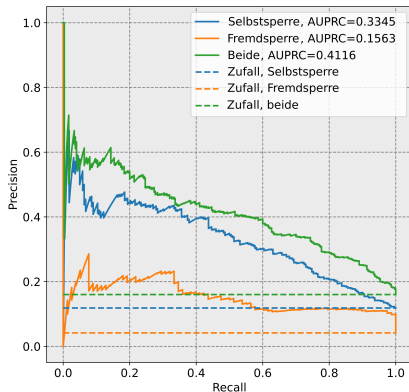
(b) MWMote XGB

Abbildung (5) Konfusionsmatrix für XGB

# Ergebnisse: virtuelle Automaten



(a) RusRose XGB



(b) MWMote XGB

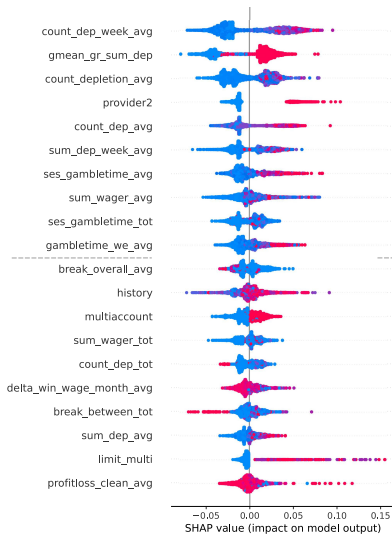
Abbildung (6) Precision-Recall Kurve für XGB

## Warum gibt es Unterschiede in der Vorhersagekraft?

- Mit den gleichen Features und gleichen Modellen sind die **Fremdsperren** viel schwieriger zu erklären.
- Potentielle Erklärungen wären:
  - Weniger Fremdsperren;
  - Besondere Verteilung der Fremdsperren (Borderline, Kluster, Ausreißer Brzezinski et al., 2021);
  - ...
  - Unterschiede in der Logik? → Interpretation der Modelle.

# Ergebnisse: virtuelle Automaten

## Shapley values for self-exclusion



## Shapley values for provider-exclusion

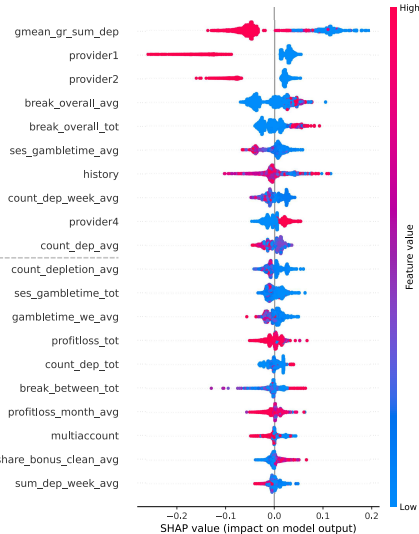
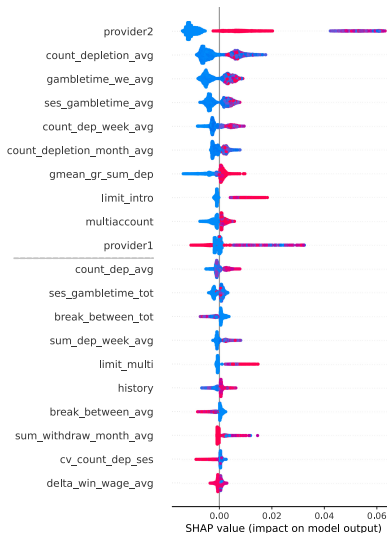


Abbildung (7) RusRose RF, Dotplot

# Ergebnisse: virtuelle Automaten

## Shapley values for self-exclusion



## Shapley values for provider-exclusion

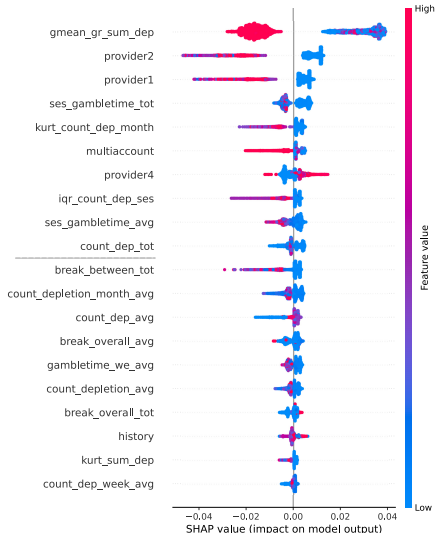


Abbildung (8) MWMote XGB, Dotplot

- 1 Einleitung
- 2 Daten
- 3 Methoden
- 4 Ergebnisse: virtuelle Automaten
- 5 Ergebnisse: Sportwetten**
- 6 Diskussion und Zusammenfassung

## Features und Abgeleitete Indikatoren

- Aus den Rohdaten wurden Daten für zwei Anbieter zusammengestellt.
- **Features aus virt. Automaten:** Konto Information und Timing.
- **Spezifische Features:** Wettverhalten und wettbezogene Information.
- Die abgeleitete Indikatoren sind ähnlich (Niveaus, Dynamik und Dispersion).

## Methoden

- Aus den Rohdaten wurden Daten für zwei Anbieter zusammengestellt.
- **Methoden aus virt. Automaten:** Balancierung; Modelle (RF, XGB, CatGB); Metriken.
- **Unterschiede:** LightGB, 80/20 und 50/50 Split (Joseph, 2022).

## Datensatz

- **Aggregierte Querschnittsdaten** für zwei Anbieter zwischen Januar 2020 und Juli 2021.
- Insgesamt 16089 Beobachtungen (Konten) und  $> 100$  Variablen.
- Fremd- und Selbstsperrern als Zielvariable  
(Selbstsperrere  $\approx 2,29\%$  und Fremdsperrere  $\approx 2,26\%$ )

# Ergebnisse: Sportwetten

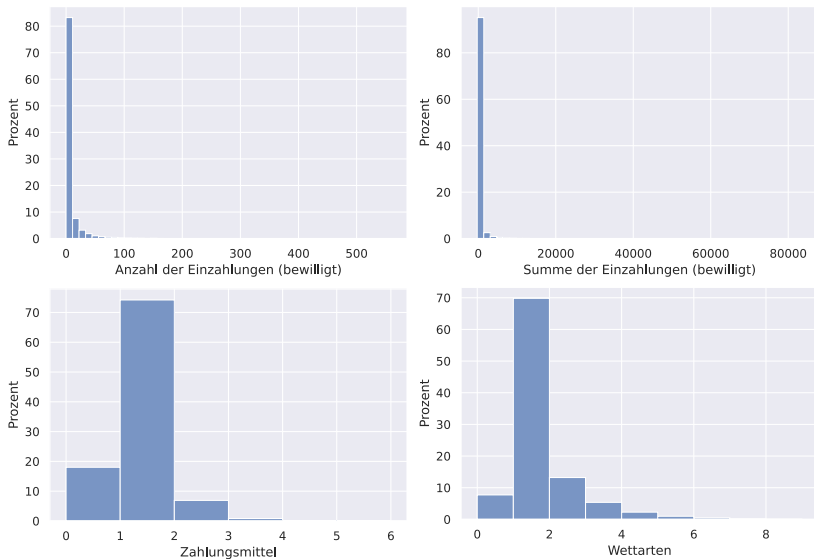


Abbildung (9) Deskriptive Statistik



# Ergebnisse: Sportwetten

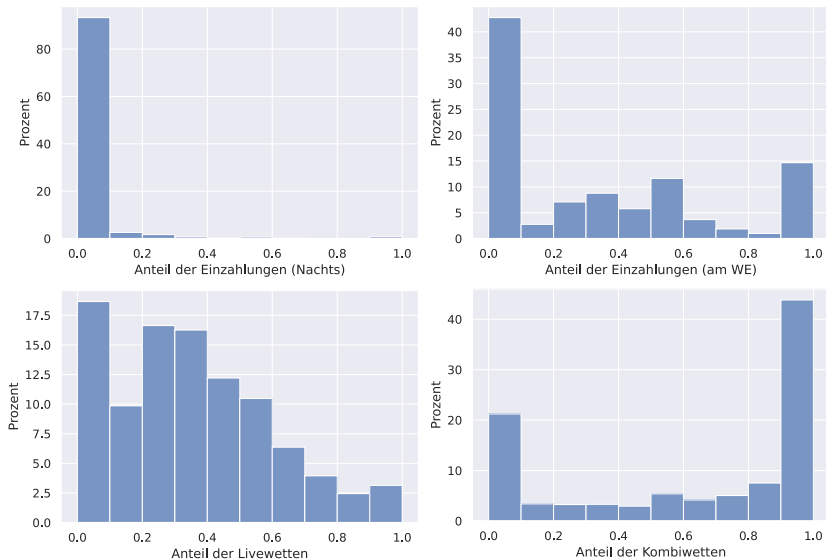


Abbildung (10) Deskriptive Statistik

Tabelle (3) Vorläufige Ergebnisse

Balancierung / Model	f1 Sperre	Precision	Recall/Sensitivity
RusRose RF			
Selbssperre	0.1508	0.1016	0.2923
Fremdsperre	0.0360	0.0526	0.0274
RusRose XGB			
Selbssperre	0.1574	0.1000	0.3692
Fremdsperre	0.0690	0.0694	0.0685
MWMote RF			
Selbssperre	0.1505	0.1157	0.2154
Fremdsperre	0.0438	0.0469	0.0411
MWMote XGB			
Selbssperre	0.1711	0.1494	0.2000
Fremdsperre	0.0364	0.0541	0.0274

## Ähnlichkeiten mit den virtuellen Automaten

- Deutlich weniger Sperren als bei den virtuellen Automaten.
- Die Selbst- und Fremdsperren sind fast im Gleichgewicht: beide  $\approx 2,3\%$ .
- Insgesamt ist die Genauigkeit der Modelle schlechter.
- Genau wie bei den virtuellen Automaten ist die **Fremdsperre schwieriger zu erklären**.

- 1 Einleitung
- 2 Daten
- 3 Methoden
- 4 Ergebnisse: virtuelle Automaten
- 5 Ergebnisse: Sportwetten
- 6 Diskussion und Zusammenfassung**

## Überblick

- Für die anbieterübergreifende Modellierung der Sperrereignisse liefern RF, XGB und CatGB mit kombinierten Balancierungsstrategien relativ gute Ergebnisse.
- RF und XGB liefern eine gute Vorhersagekraft und sind schneller als CatGB.
- Ein breiteres Spektrum der Metriken für unausgewogenen Daten muss verwendet werden.
- Die Selbstsperrung kann relativ gut und intuitiv mit den aus der Literatur bekannten Features (z.B. Verluste, Spielzeit, Limits, Einzahlungen) erklärt werden.
- Die Logik der Fremdsperrungen ist nicht immer interpretierbar und nachvollziehbar.

## Politische Implikationen

- Eine **analytische Evaluierung** der Frühwarn- und Sperrsysteme ist **notwendig**.
- Die **Logik der Fremdsperrungen sollte idealerweise der Logik der Selbstsperrungen entsprechen**.

Danke für Ihre Aufmerksamkeit!  
Für Rückfragen stehe ich gerne zur Verfügung.

[vadim.kufenko@uni-hohenheim.de](mailto:vadim.kufenko@uni-hohenheim.de)

- Auer, M. and Griffiths, M. D. (2022). An empirical attempt to operationalize chasing losses in gambling utilizing account-based player tracking data. *Journal of Gambling Studies*, in print.
- Barua, S., Islam, M. M., Yao, X., and Murase, K. (2014). MWMOTE—majority weighted minority oversampling technique for imbalanced data set learning. *IEEE Transactions on Knowledge and Data Engineering*, 26(2):405–425.
- Batista, G. E. A. P. A., Prati, R. C., and Monard, M. C. (2004). A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explorations Newsletter*, 6(1):20–29.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.
- Brosowski, T., Meyer, G., and Hayer, T. (2012). Analyses of multiple types of online gambling within one provider: an extended evaluation framework of actual online gambling behaviour. *International Gambling Studies*, 12(3):405–419.
- Brzezinski, D. W., Minku, L. L., Pewinski, T., Stefanowski, J., and Szumaczk, A. (2021). The impact of data difficulty factors on classification of imbalanced and concept drifting data streams. *Knowledge and Information Systems*, 63:1429–1469.
- Chen, T. and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 785–794, New York, NY, USA. ACM.
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lefourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., and Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1):27–46.
- Finkenwirth, S., Macdonald, K., Deng, X., Lesch, T., and Clark, L. (2020). Using machine learning to predict self-exclusion status in online gamblers on the playnow.com platform in british columbia. *International Gambling Studies*, 21:220 – 237.
- Haefeli, J., Lischer, S., and Schwarz, J. (2011). Early detection items and responsible gambling features for online gambling. *International Gambling Studies*, 11(3):273–288.
- Hausler, J. (2016). Follow the money: using payment behaviour as predictor for future self-exclusion. *International Gambling Studies*, 16(2):246–262.
- Hopfgartner, N., Auer, M., Griffiths, M. D., and Helic, D. (2022). Predicting self-exclusion among online gamblers: An empirical real-world study. *Journal of Gambling Studies*.
- Joseph, V. R. (2022). Optimal ratio for data splitting. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 15(4):531–538.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 3149–3157, Red Hook, NY, USA. Curran Associates Inc.
- Kovács, G. (2019). smote-variants: a Python implementation of 85 minority oversampling techniques. *Neurocomputing*, 366:352–354.
- Lemaître, G., Nogueira, F., and Aridas, C. K. (2017). Imbalanced-learn: A python toolbox to tackle the curse of imbalanced

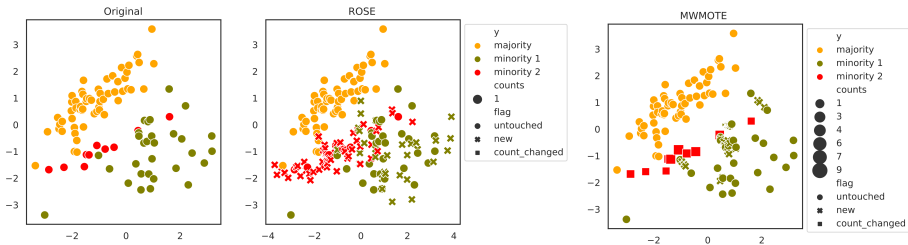


Abbildung (11) Simuliertes Beispiel für Oversampling (see Kovács, 2019).

Zurück

- Specificity =  $WN / (WN + FP)$
- Precision =  $WP / (WP + FP)$
- Recall (or Sensitivity) =  $WP / (WP + FN)$
- f1 Score =  $2 * (Recall * Precision) / (Recall + Precision)$
- Balanced Accuracy (BACC) =  $(Sensitivity + Specificity) / 2$
- Accuracy (ACC) =  $(WP + WN) / (WP + FP + FN + WN)$

Zurück



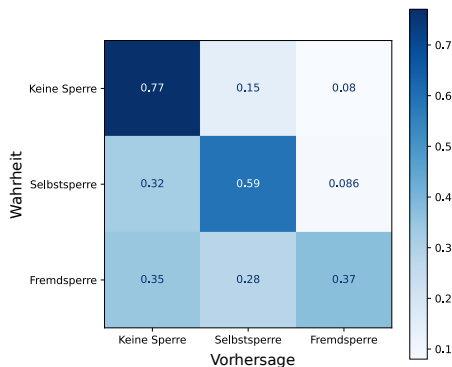
Tabelle (4) Detaillierter Überblick über die Metriken (Test Daten)

Balancierung / Modell	AUROC	AUPRC	f1 Sperre	Precision	Recall/Sensitivity
Logit					
Sperre (binär)	0.4873	0.1513	0.0125	0.0952	0.0067
Selbssperre (multi)	0.5248	0.1349	0.0632	0.2581	0.0360
Fremdsperre (multi)	0.4498	0.0551	0.0225	0.0909	0.0128
RusRose RF					
Sperre (binär)	0.7941	0.4186	0.4587	0.3640	0.6200
Selbssperre (multi)	0.7872	0.3616	0.4375	0.3559	0.5676
Fremdsperre (multi)	0.8546	0.1638	0.2466	0.1915	0.3462
RusRose XGB					
Sperre (binär)	0.7977	0.4091	0.4839	0.4054	0.6000
Selbstsperre (multi)	0.7895	0.3576	0.4378	0.3465	0.5946
Fremdsperre (multi)	0.8531	0.1751	0.2363	0.1761	0.3590
RusRose CatGB					
Sperre (binär)	0.7979	0.4213	0.4693	0.3631	0.6633
Selbstsperre (multi)	0.7872	0.3890	0.4295	0.3376	0.5901
Fremdsperre (multi)	0.8546	0.1637	0.2302	0.1667	0.3718

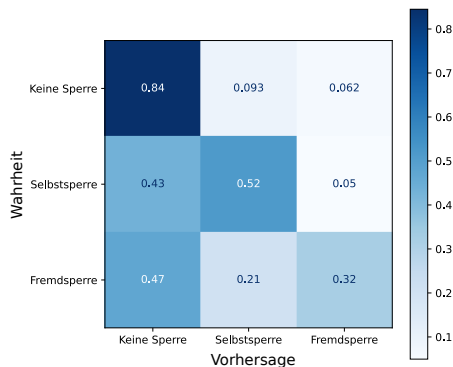
Tabelle (5) Detaillierter Überblick über die Metriken (Test Daten)

Balancierung / Modell	AUROC	AUPRC	f1 Sperre	Precision	Recall/Sensitivity
Logit					
Sperre (binär)	0.4873	0.1513	0.0125	0.0952	0.0067
Selbstsperre (multi)	0.5248	0.1349	0.0632	0.2581	0.0360
Fremdsperre (multi)	0.4498	0.0551	0.0225	0.0909	0.0128
MWMote RF					
Sperre (binär)	0.7939	0.4292	0.4522	0.3604	0.6067
Selbstsperre (multi)	0.7832	0.3325	0.4121	0.3975	0.4279
Fremdsperre (multi)	0.8499	0.1597	0.2069	0.1875	0.2308
MWMote XGB					
Sperre (binär)	0.7912	0.4116	0.4699	0.3981	0.5733
Selbstsperre (multi)	0.7745	0.3345	0.4619	0.4360	0.4910
Fremdsperre (multi)	0.8481	0.1563	0.2292	0.1657	0.3718
MWMote CatGB					
Sperre (binär)	0.7942	0.4418	0.4767	0.3969	0.5967
Selbstsperre (multi)	0.7936	0.3724	0.4600	0.4137	0.5180
Fremdsperre (multi)	0.8503	0.1557	0.2370	0.1880	0.3205

# Ergebnisse: virtuelle Automaten



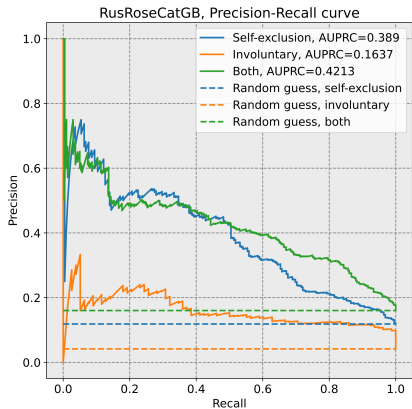
(a) RusRose CatGB



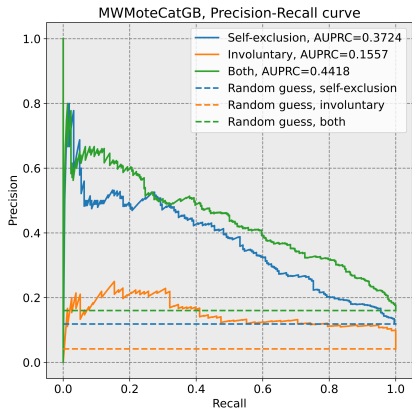
(b) MWMote CatGB

Abbildung (12) Konfusionsmatrix für CatGB

# Ergebnisse: virtuelle Automaten



(a) RusRose CatGB

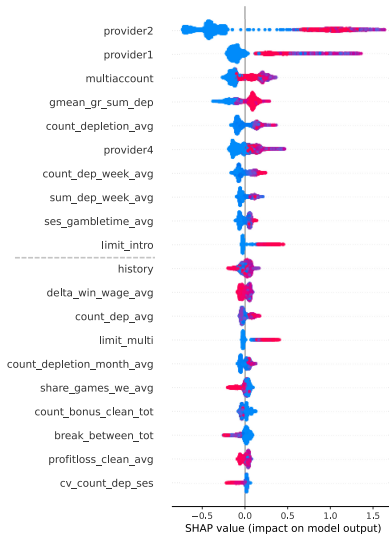


(b) MWMote CatGB

Abbildung (13) Precision-Recall Kurve für CatGB

# Ergebnisse: virtuelle Automaten

## Shapley values for self-exclusion



## Shapley values for provider-exclusion

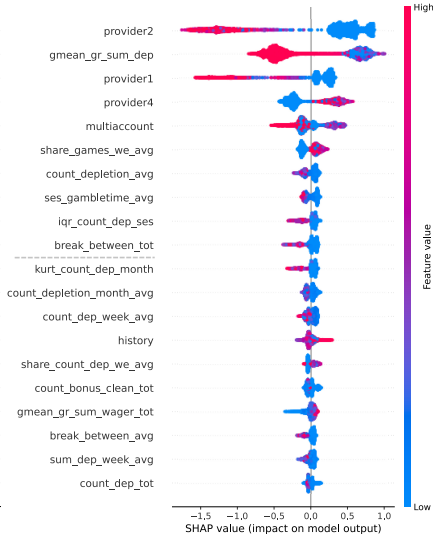


Abbildung (14) MWMote CatGB, Dotplot