

# Zwei-Stufen-XGBoost – Experiment-Report

Experiment-ID: hp\_long\_h4\_thr0p4pct\_hit\_1\_3

Dieses Dokument fasst die wichtigsten Parameter, Datenquellen und Metriken eines Zwei-Stufen-XGBoost-Experiments zusammen.

Stufe 1 (Signal): neutral vs. Bewegung ('move'). Stufe 2 (Richtung): down vs. up – nur an Bewegungstagen.

## Label-Parameter:

- horizon\_days: 4
- up\_threshold: 0.004
- down\_threshold: -0.004
- strict\_monotonic: False
- max\_adverse\_move\_pct: 0.01
- hit\_within\_horizon: True (True = Schwelle reicht, wenn sie irgendwo im Horizont erreicht wird)

## Datensatz & Splits:

- dataset\_path:  
  
data/processed/datasets/eurusd\_news\_training\_\_hp\_long\_h4\_thr0p4pct\_hit\_1\_3.csv
- test\_start: 2025-01-01
- train\_frac\_within\_pretest: 0.8
- feature\_mode: price\_only

## Entscheidungsgrenzen (Modelle):

- SIGNAL\_THRESHOLD (Stufe 1 – move vs. neutral): 0.5 (höher → höhere Precision, niedrigerer Recall).
- SIGNAL\_THRESHOLD\_TRADE (Stufe 1 – Trading): 0.5 (höher → weniger Trades, tendenziell höhere Qualität).
- DIRECTION\_THRESHOLD (Stufe 2 – down vs. up, für Metriken): 0.46 (niedriger → mehr up, höher → weniger up).
- DIRECTION\_THRESHOLDS (Stufe 2 – Trading-Entscheidungen): down, wenn  $P(\text{up}) \leq 0.3$ , up, wenn  $P(\text{up}) \geq 0.7$ .

# Legende & Begriffe (Kurzüberblick)

## Zielvariablen:

- label: 3-Klassen-Ziel auf Basis des 4-Tage-Lookaheads (neutral / up / down).
- signal: 0 = neutral, 1 = Bewegung (up oder down).
- direction: 0 = down, 1 = up; nur definiert, wenn signal == 1.

## Wichtige Metriken:

- precision: Anteil der vorhergesagten positiven Fälle, die wirklich positiv sind.
- recall: Anteil der tatsächlichen positiven Fälle, die erkannt wurden.
- f1: harmonischer Mittelwert aus precision und recall (Balance beider Größen).
- support: Anzahl der Beobachtungen in der jeweiligen Klasse.

## Feature-Abkürzungen (Auswahl, nicht vollständig – vollständige Liste siehe Seite 'Verwendete

### Features'):

- article\_count: Anzahl News-Artikel pro Tag.
- avg\_polarity / avg\_neg / avg\_neu / avg\_pos: durchschnittliche Sentiment-Werte.
- pos\_share / neg\_share: Anteil positiver bzw. negativer Sentiment-Komponente.
- intraday\_range\_pct: (High - Low) / Close – relative Tages-Spanne (Volatilität).
- upper\_shadow / lower\_shadow: obere/untere Dochte der Kerzen (High/Low vs. Körper).
- month / quarter: Kalendermonat und Quartal.

## Modell-Parameter (XGBoost)

### Signal-Modell (Stufe 1):

- objective: binary:logistic
- max\_depth: 3
- learning\_rate: 0.05
- n\_estimators: None
- subsample: 0.9
- colsample\_bytree: 0.9
- scale\_pos\_weight: 0.15201768933112217

### Richtungs-Modell (Stufe 2):

- objective: binary:logistic
- max\_depth: 3
- learning\_rate: 0.05
- n\_estimators: None
- subsample: 0.9
- colsample\_bytree: 0.9
- scale\_pos\_weight: 1.0

## Verwendete Features (FEATURE\_COLS)

#	feature_name	description
0	intraday_range_pct	(High - Low) / Close – relative Tagesvolatilität.
1	upper_shadow	Oberer Kerzendocht: High - max(Open, Close).
2	lower_shadow	Unterer Kerzendocht: min(Open, Close) - Low.
3	price_close_ret_1d	Relativer Schlusskurs-Return gegenüber Vortag: $\text{Close}_t / \text{Close}_{\{t-1\}} - 1$ .
4	price_close_ret_5d	Schlusskurs-Return über 5 Tage: $\text{Close}_t / \text{Close}_{\{t-5\}} - 1$ .
5	price_range_pct_5d_std	Standardabweichung der intraday_range_pct über 5 Tage (Volatilität).
6	price_body_pct_5d_mean	Durchschnittlicher Kerzenkörper-Prozentsatz über 5 Tage.
7	price_close_ret_30d	Schlusskurs-Return über 30 Tage: $\text{Close}_t / \text{Close}_{\{t-30\}} - 1$ .
8	price_range_pct_30d_std	Standardabweichung der intraday_range_pct über 30 Tage.
9	price_body_pct_30d_mean	Durchschnittlicher Kerzenkörper-Prozentsatz über 30 Tage.
10	month	Kalendermonat (1–12).
11	quarter	Kalenderquartal (1–4).
12	cal_dow	Wochentag (0 = Montag, 6 = Sonntag).
13	cal_day_of_month	Kalendertag im Monat.
14	cal_is_monday	Flag: 1 wenn Montag, sonst 0.
15	cal_is_friday	Flag: 1 wenn Freitag, sonst 0.

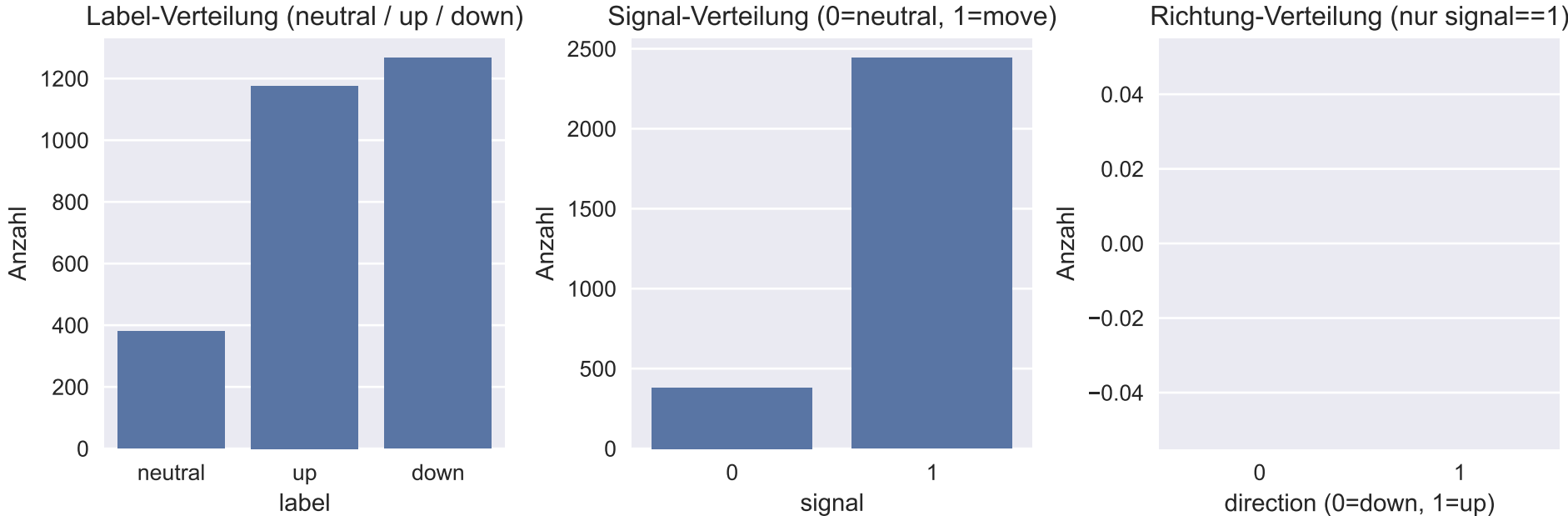
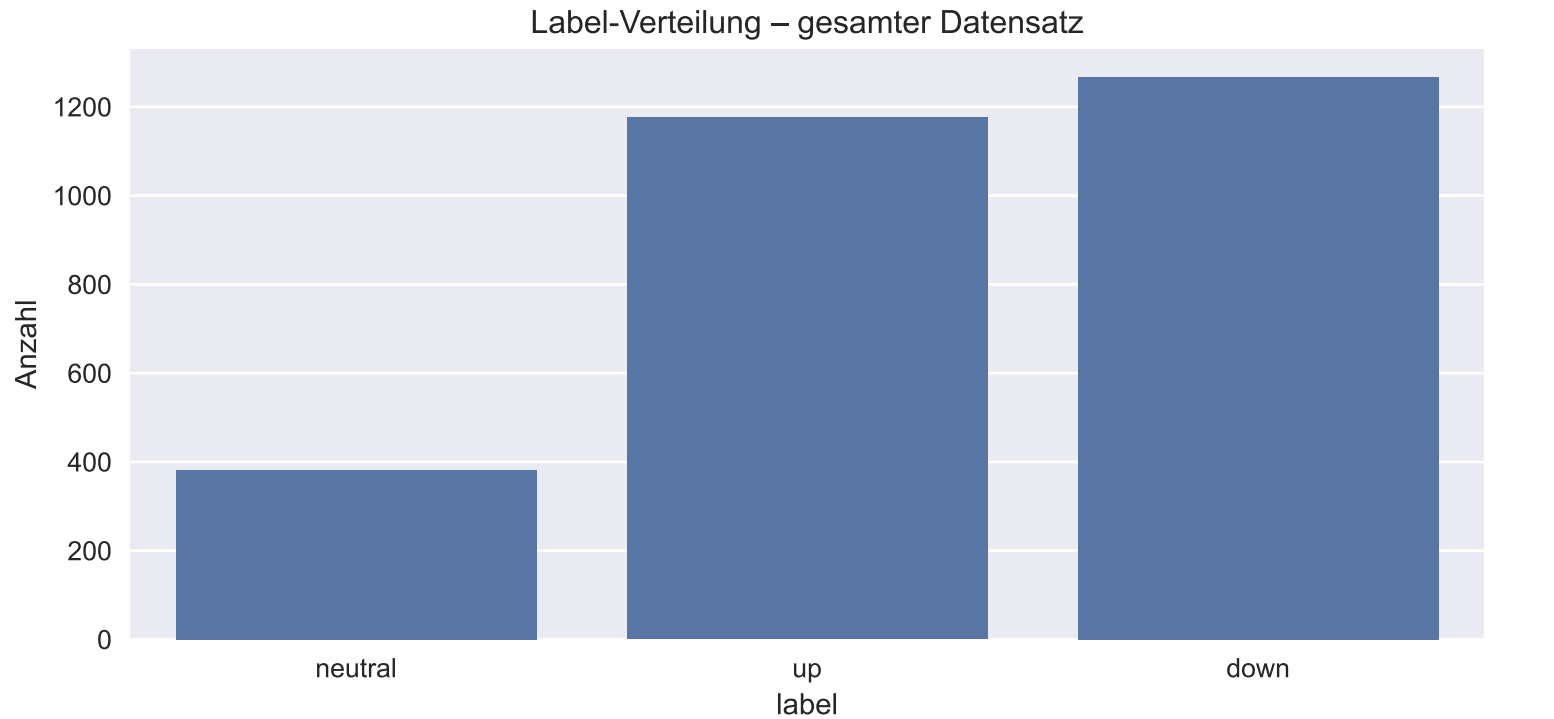


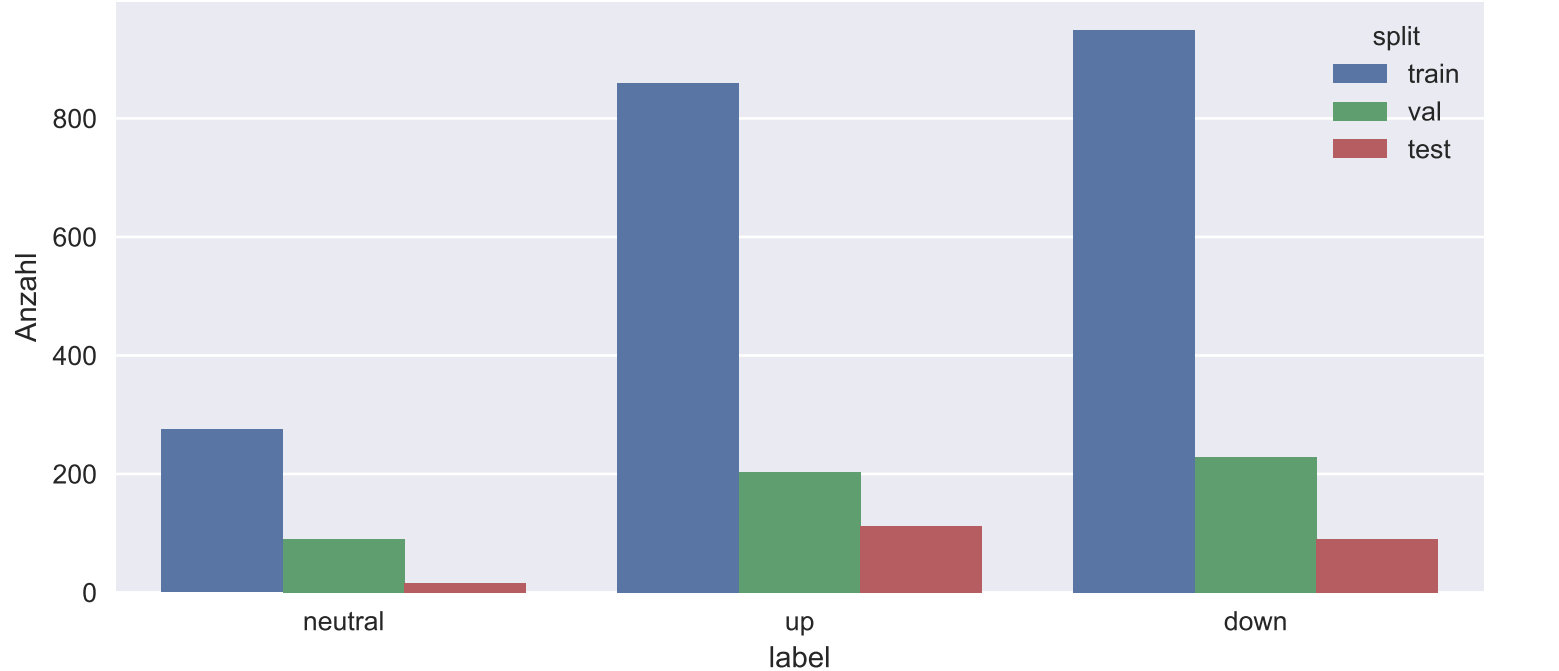
Abbildung: Klassenverteilungen für label, signal und direction im vollständigen Trainingsdatensatz.



label	count
neutral	381
up	1176
down	1267

Abbildung/Tabelle: Verteilung der Zielvariable 'label' (neutral/up/down) im gesamten Datensatz.

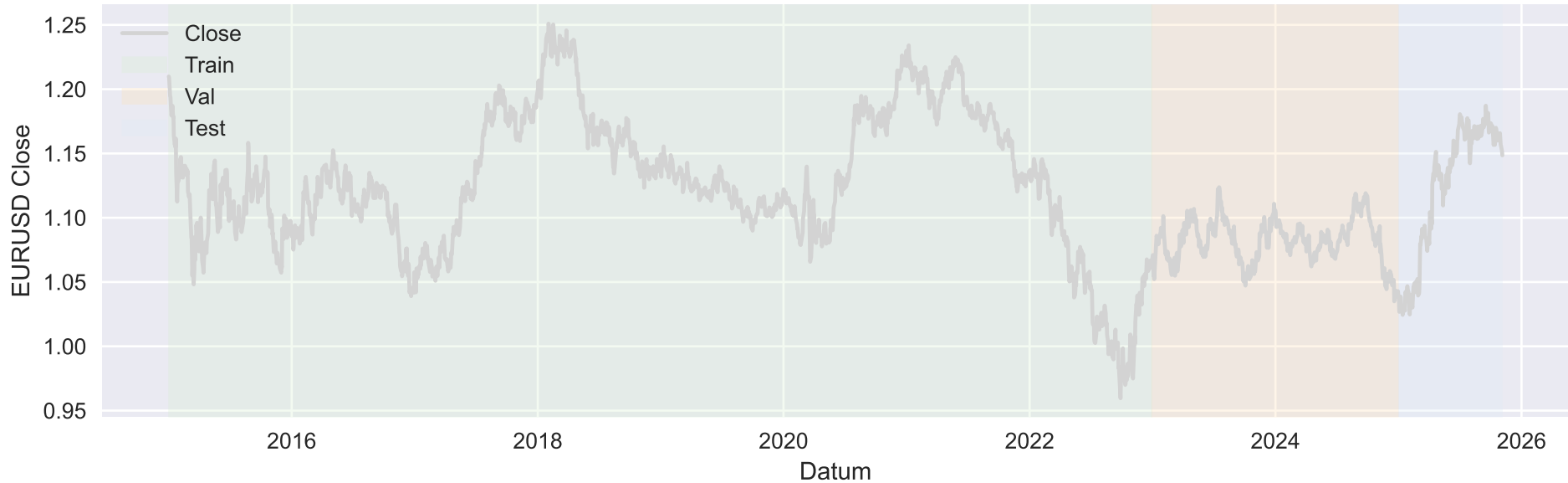
Label-Verteilung nach Splits (train/val/test)



split	neutral	up	down
train	275	860	949
val	90	204	228
test	16	112	90

Abbildung/Tabelle: Label-Verteilung getrennt nach Trainings-, Validierungs- und Test-Split.

EURUSD-Zeitreihe mit Train/Val/Test-Bereichen





EURUSD-Zeitreihe mit hervorgehobenen up/down-Tagen (ab 2020)



Abbildung: EURUSD-Schlusskurs mit markierten up-/down-Tagen im betrachteten Zeitraum.

# EURUSD-Segmente mit label='up' (Test-Split)

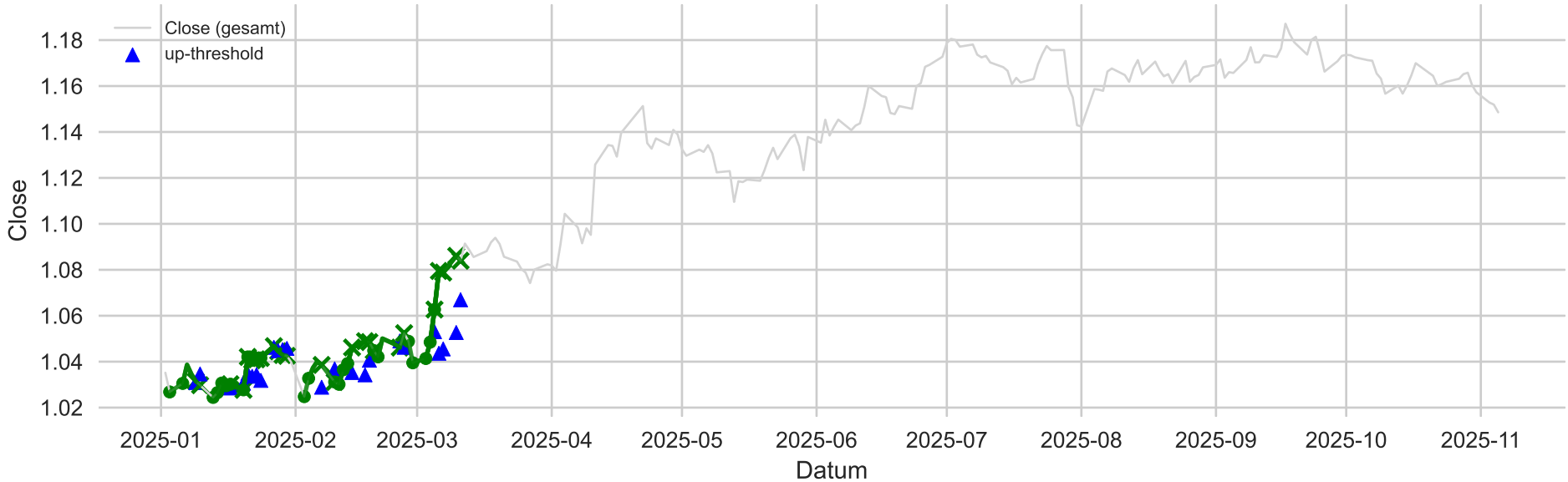


Abbildung: Preis-Segmente  $t..t+\text{horizon}$  für alle Testtage mit true label 'up'.

Relativer Verlauf der Segmente (label='up') – Seite 1

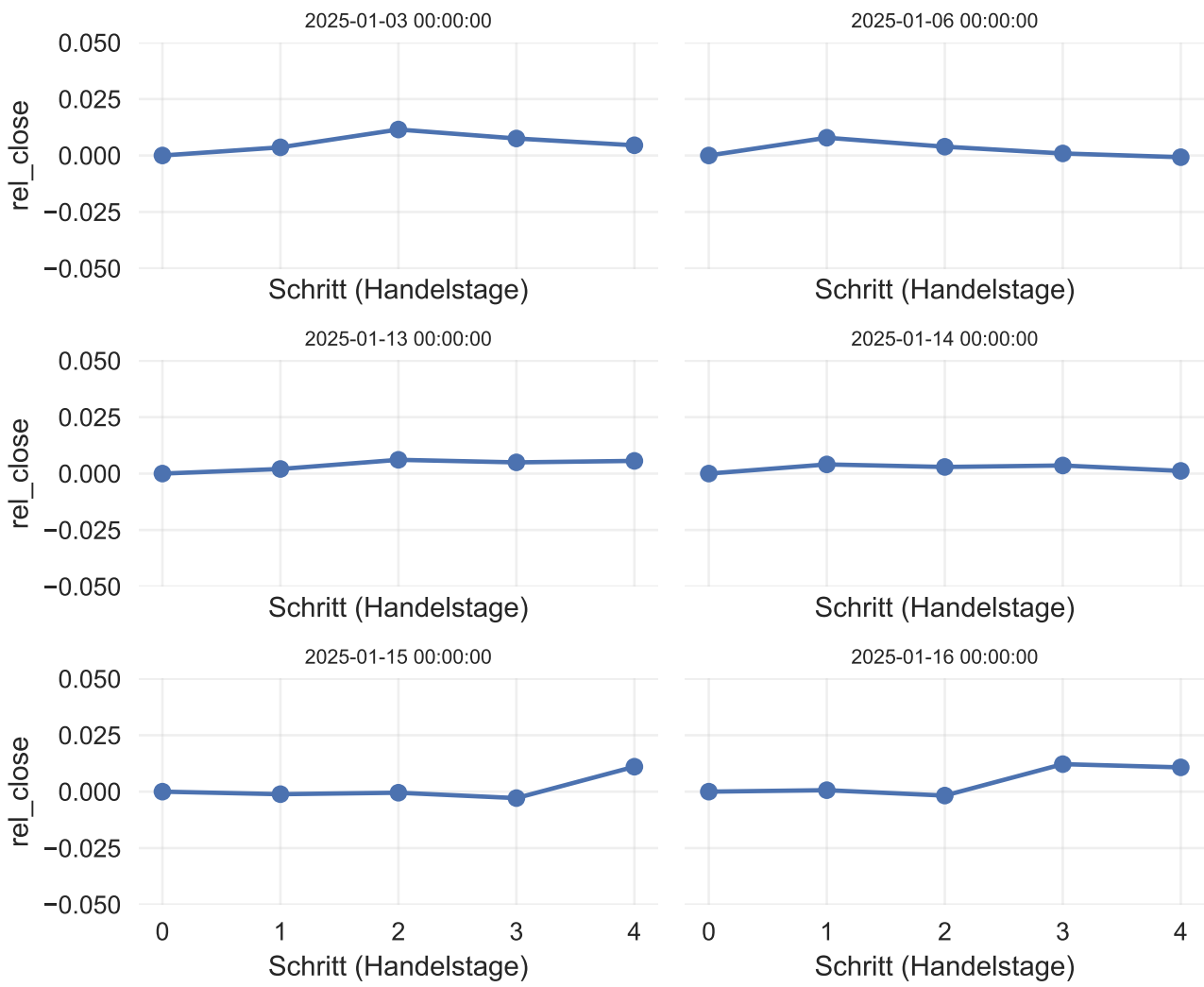


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

Relativer Verlauf der Segmente (label='up') – Seite 2

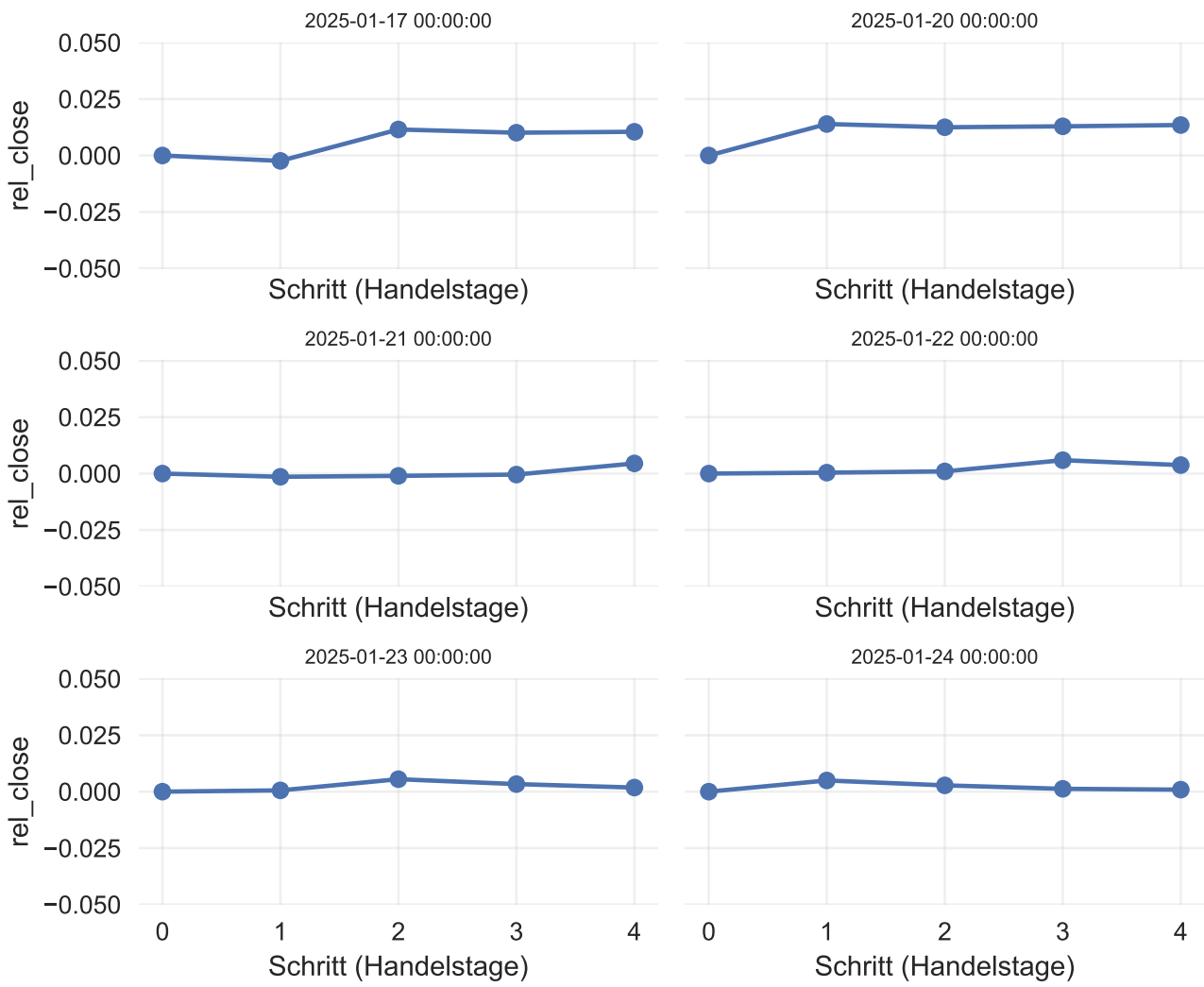


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

Relativer Verlauf der Segmente (label='up') – Seite 3

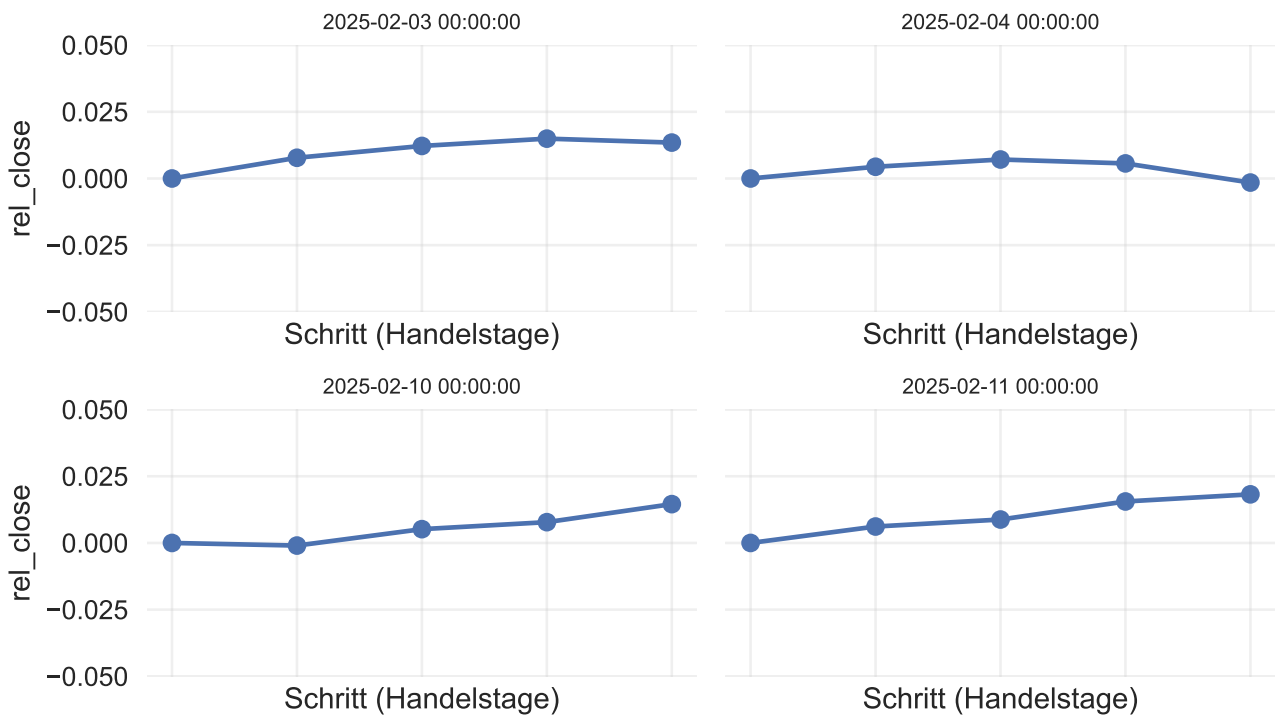
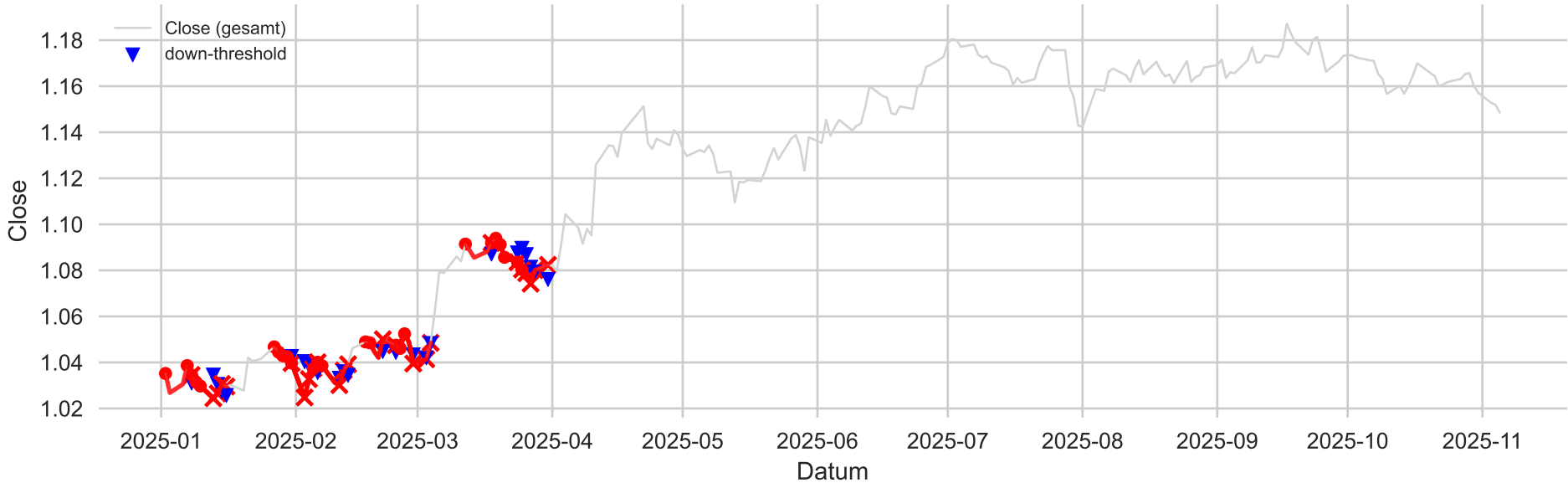


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

# EURUSD-Segmente mit label='down' (Test-Split)



## Relativer Verlauf der Segmente (label='down') – Seite 1

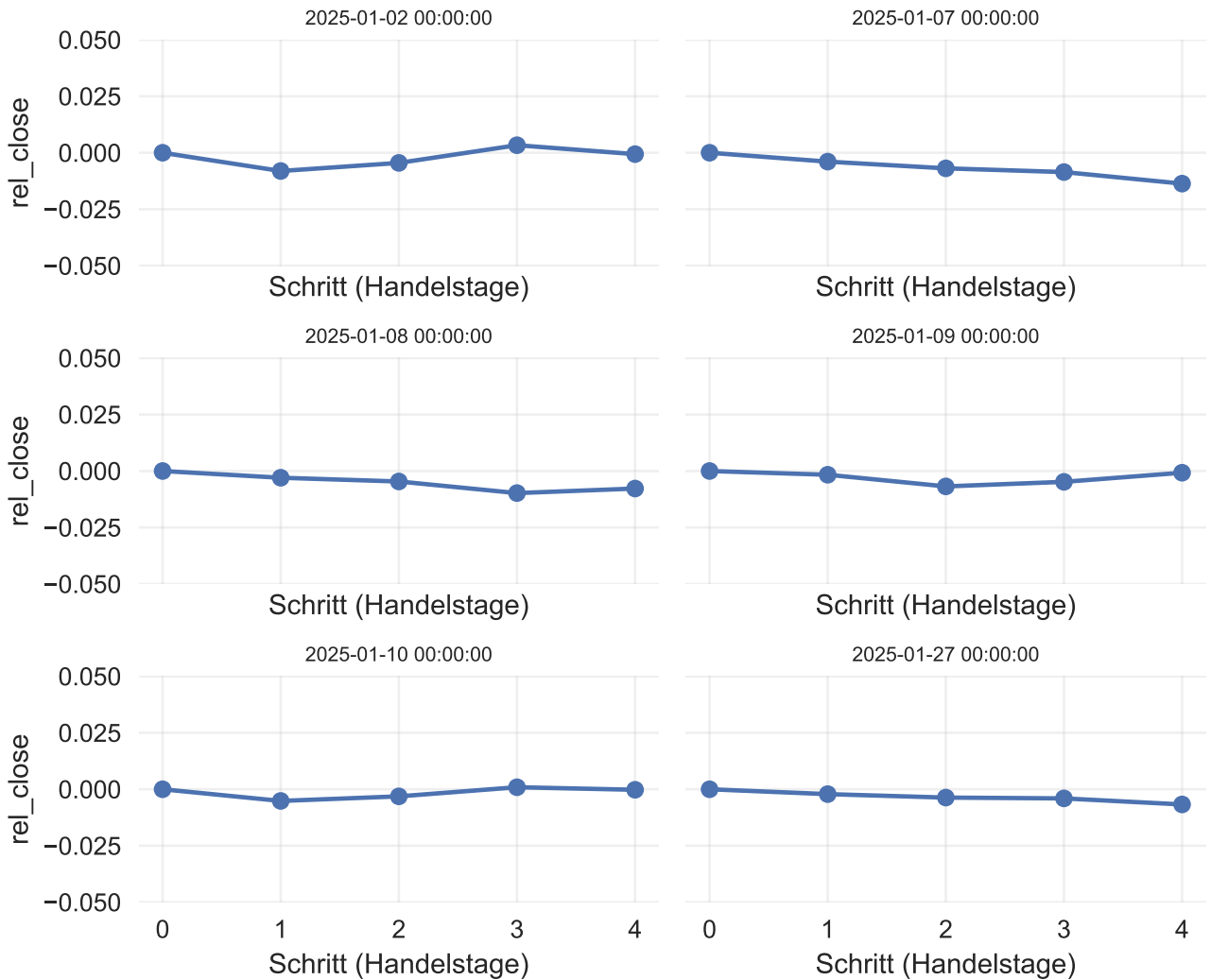


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

## Relativer Verlauf der Segmente (label='down') – Seite 2

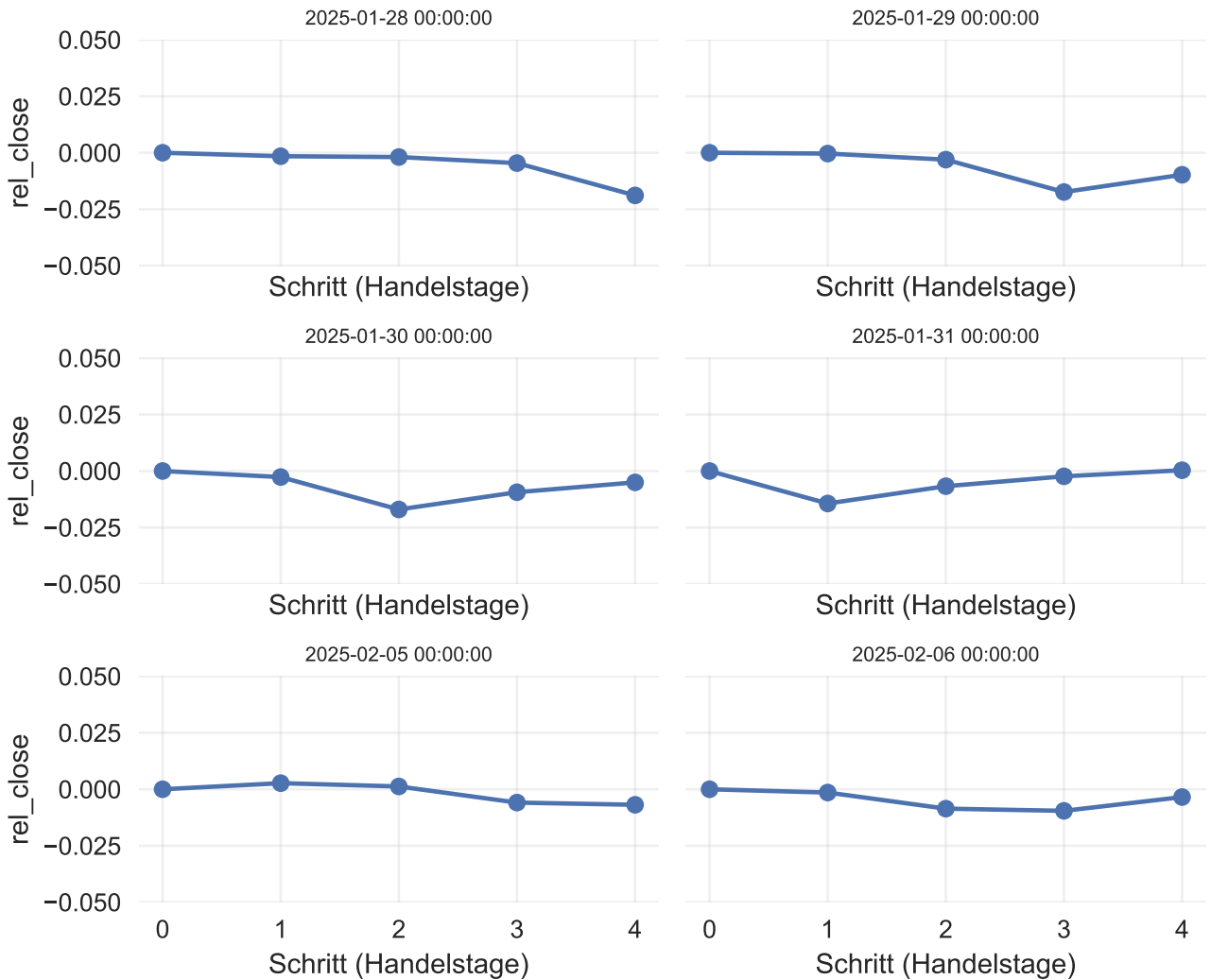


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.



## Relativer Verlauf der Segmente (label='down') – Seite 3

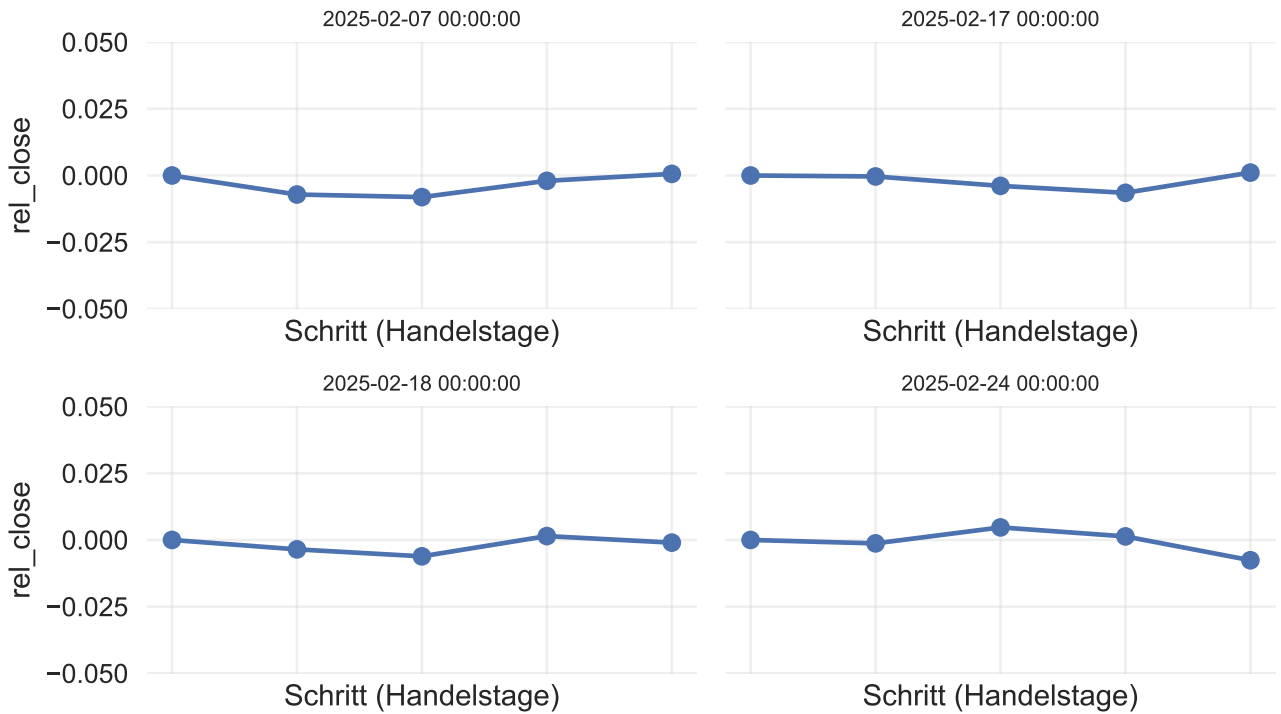


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

Signal-Modell – Kennzahlen für Klasse 'move' (train/val/test, thr=0.50)

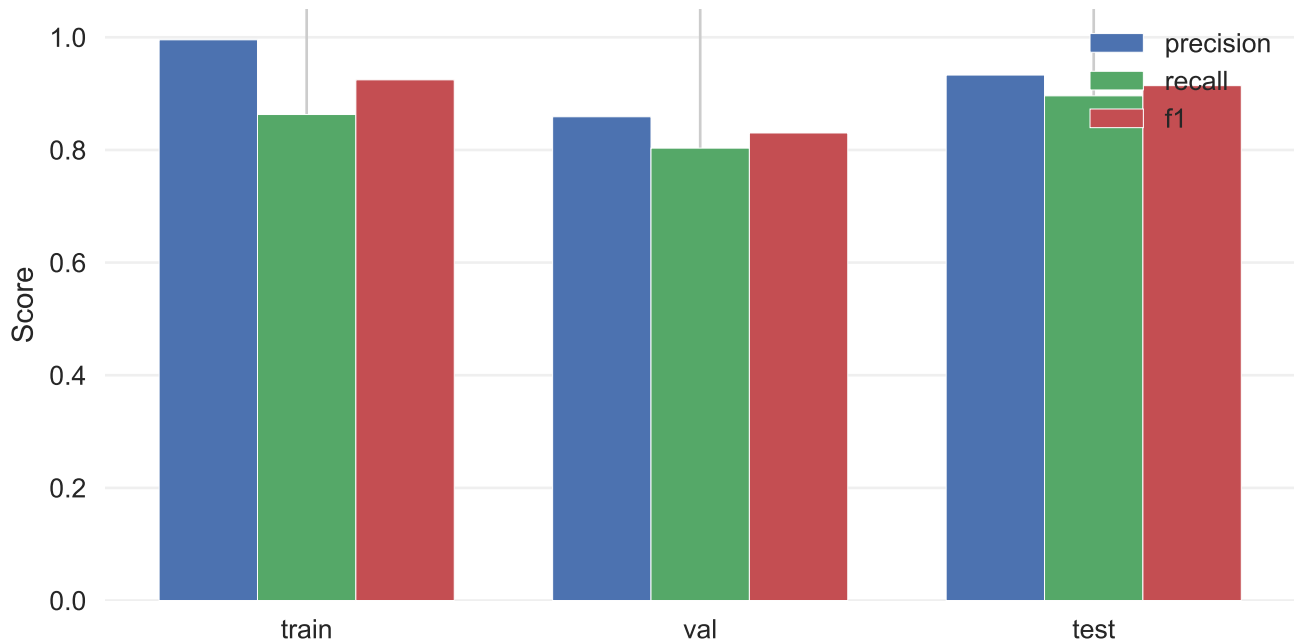


Abbildung: Precision, Recall und F1 der positiven Klasse je Split (train/val/test).

## Signal-Modell – Tabelle (Klasse 'move', thr=0.50)

split	precision	recall	f1	support
train	0.996	0.863	0.924	1809.000
val	0.859	0.803	0.830	432.000
test	0.933	0.896	0.914	202.000

Tabelle: Kennzahlen der positiven Klasse (precision/recall/F1/support) für train/val/test.

Richtungs-Modell – Kennzahlen für Klasse 'up' (train/val/test)

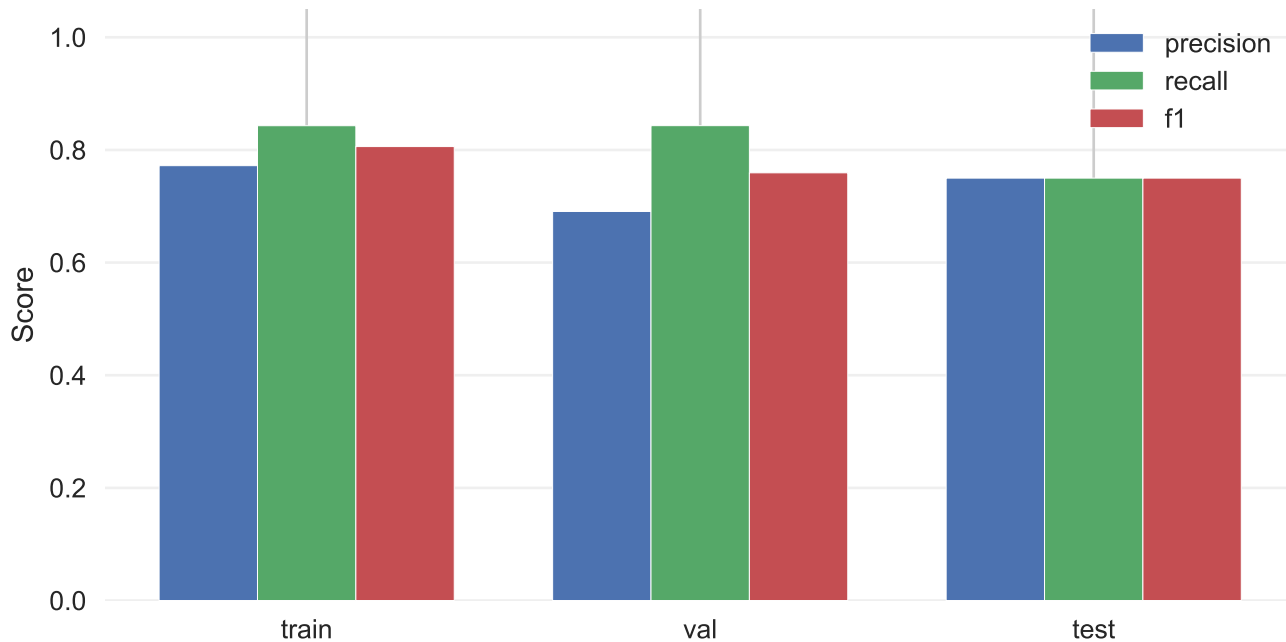


Abbildung: Precision, Recall und F1 der positiven Klasse je Split (train/val/test).

## Richtungs-Modell – Tabelle (Klasse 'up')

split	precision	recall	f1	support
train	0.772	0.843	0.806	860.000
val	0.691	0.843	0.759	204.000
test	0.750	0.750	0.750	112.000

Tabelle: Kennzahlen der positiven Klasse (precision/recall/F1/support) für train/val/test.

Kombinierte Test-Auswertung – neutral / up / down

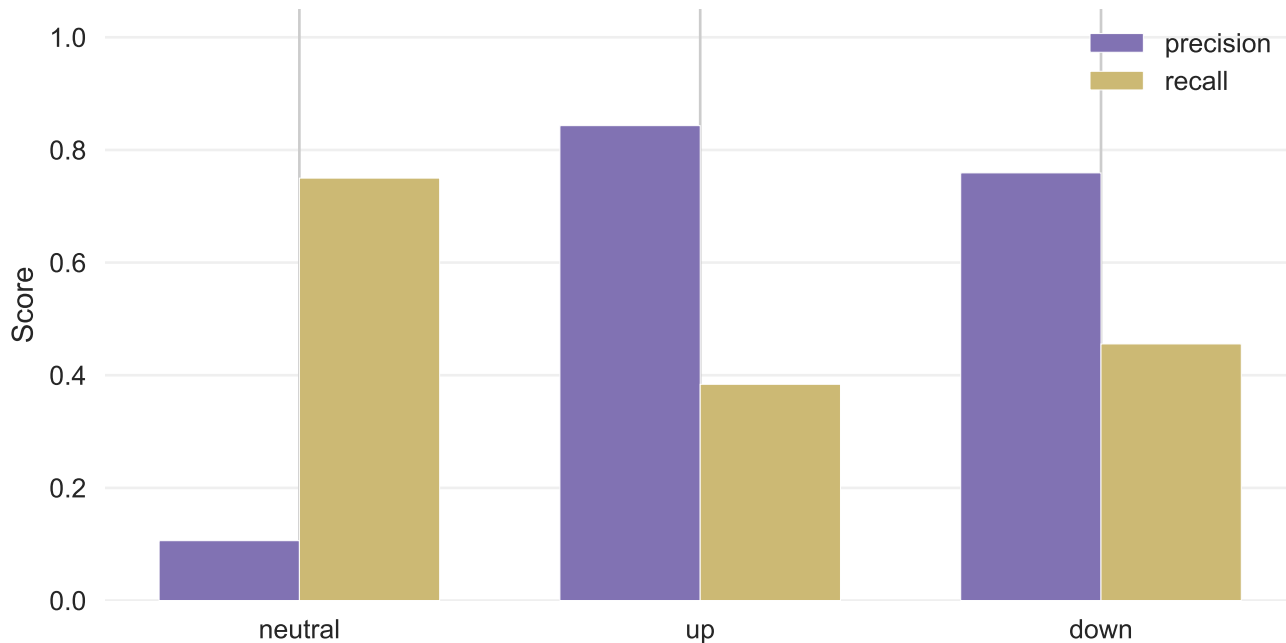


Abbildung: Precision und Recall der kombinierten 3-Klassen-Vorhersage (neutral/up/down) auf dem Test-Split.

## Kombiniertes Modell – Tabelle (Test, neutral/up/down)

klasse	precision	recall	f1	support
neutral	0.106	0.75	0.186	16
up	0.843	0.384	0.528	112
down	0.759	0.456	0.569	90

Tabelle: Kennzahlen der drei Klassen (neutral/up/down) des kombinierten Modells auf dem Test-Split.

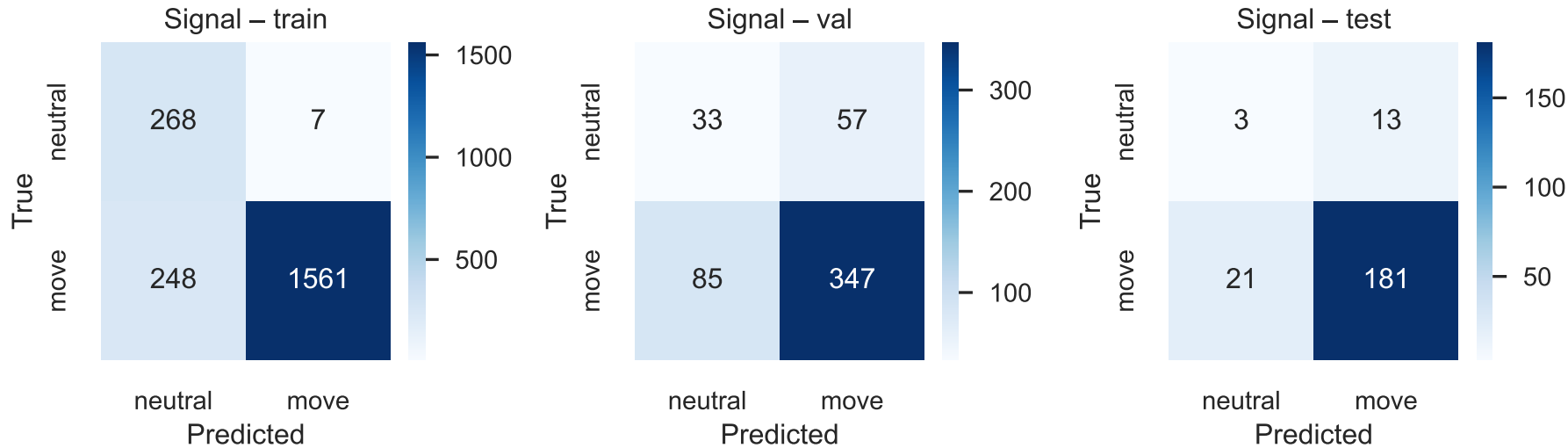


Abbildung: Confusion-Matrizen des Signal-Modells (neutral vs move) für Train-, Validierungs- und Test-Split.



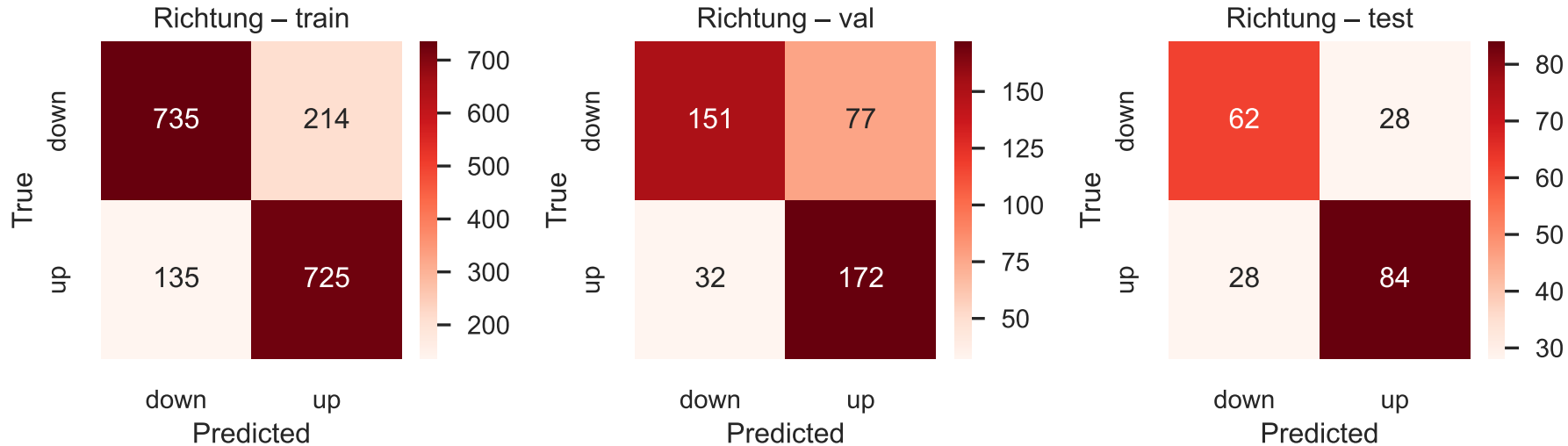


Abbildung: Confusion-Matrizen des Richtungs-Modells (down vs up) für Train-, Validierungs- und Test-Split.

Confusion Matrix – Test (neutral / up / down)

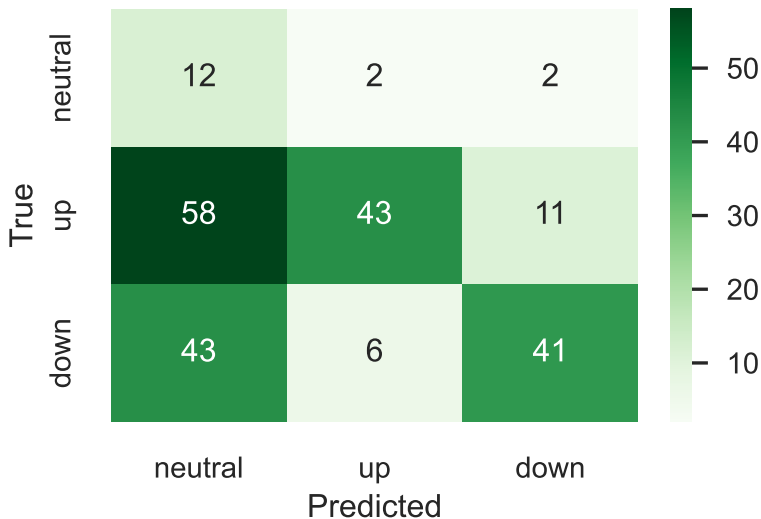


Abbildung: Confusion-Matrix des kombinierten Modells (neutral/up/down) auf dem Test-Split.

## Konfusionsmatrizen – Zählwerte (TN/FP/FN/TP)

modell	split	TN	FP	FN	TP
signal	train	268	7	248	1561
signal	val	33	57	85	347
signal	test	3	13	21	181
direction	train	735	214	135	725
direction	val	151	77	32	172
direction	test	62	28	28	84

Tabelle: Zählwerte der Konfusionsmatrizen (TN/FP/FN/TP)  
für Signal- und Richtungs-Modell je Split.

## Fehlklassifikationen – Übersicht (False Positives)

task	predicted	total_fp	true_label_breakdown
combined	up	8	neutral:2, down:6
combined	down	13	neutral:2, up:11
signal	move	13	neutral:13

Tabelle: Zusammenfassung der wichtigsten False-Positive-Fälle für kombinierten Test (neutral/up/down) und Signal-Test (neutral vs move).

Relativer Verlauf der Segmente (label='neutral→up')

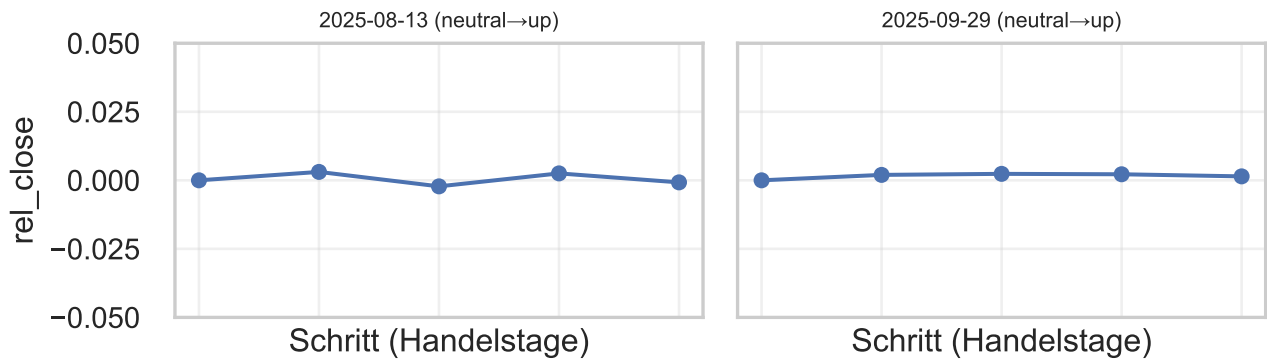


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'neutral', die im kombinierten Test als 'up' klassifiziert wurden.

## Relativer Verlauf der Segmente (label='neutral→down')

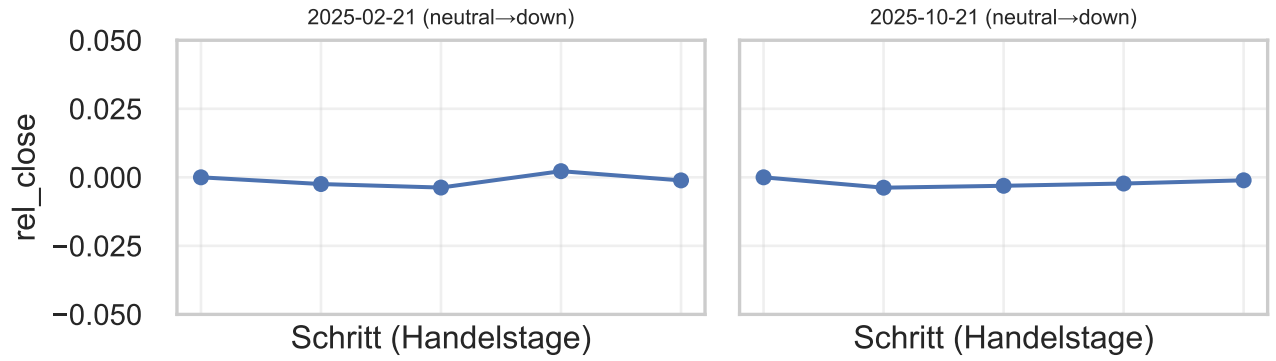


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'neutral', die im kombinierten Test als 'down' klassifiziert wurden.

## Tradesimulation – Strategien A und B (Test-Split)

Strategy	Kennzahl	Wert
A (fixer Einsatz)	Anzahl Trades	105
A (fixer Einsatz)	Einsatz up / down (CHF)	100 / 100
A (fixer Einsatz)	Trades up / down	51 / 54
A (fixer Einsatz)	Gewinner / Verlierer	88 / 17
A (fixer Einsatz)	Gesamt-P&L (CHF)	22.29
A (fixer Einsatz, Hebel 20)	Gesamt-P&L (CHF)	445.73
B (10% vom Kapital)	Startkapital (CHF)	1000.00
B (10% vom Kapital)	Endkapital (CHF)	1022.52
B (10% vom Kapital)	Minimum Kapital (CHF)	998.90
B (10% vom Kapital, Hebel 20)	Endkapital (CHF)	1553.78
B (10% vom Kapital, Hebel 20)	Minimum Kapital (CHF)	977.12

Tabelle: Zusammenfassung der Tradesimulation auf dem Test-Split.  
 Strategie A: fixer Einsatz pro Trade (100 CHF bei up, 100 CHF bei down).  
 Strategie B: 10 % des aktuellen Vermögens pro Trade (optional mit Hebel 20).

## Kostenmatrix – durchschnittliche Kosten pro Fall (Strategie A, Test-Split)

label_true	combined_pred	mean_chf
neutral	neutral	0.0
neutral	up	-1.0
neutral	down	-1.0
up	neutral	0.0
up	up	0.3999999999999999
up	down	-0.43224942384814435
down	neutral	0.0
down	up	-0.42646322818600124
down	down	0.39999999999999997

Tabelle: durchschnittliche Kosten (CHF) pro Fall für jede Kombination aus wahren Label und vorhergesagtem Label (Strategie A, fixer Einsatz).



## Kostenmatrix – Gesamtkosten und Anzahl Trades (Strategie A, Test-Split)

label_true	combined_pred	count	sum_chf
neutral	neutral	12	0.0
neutral	up	2	-2.0
neutral	down	2	-2.0
up	neutral	58	0.0
up	up	43	17.199999999999996
up	down	11	-4.754743662329588
down	neutral	43	0.0
down	up	6	-2.5587793691160075
down	down	41	16.4

Tabelle: Anzahl Fälle und Gesamt-P&L (CHF) auf dem Test-Split  
für jede Kombination aus wahrem Label und vorhergesagtem Label (Strategie A).

## Strategie B – Verlauf des Kapitals (Test-Split)

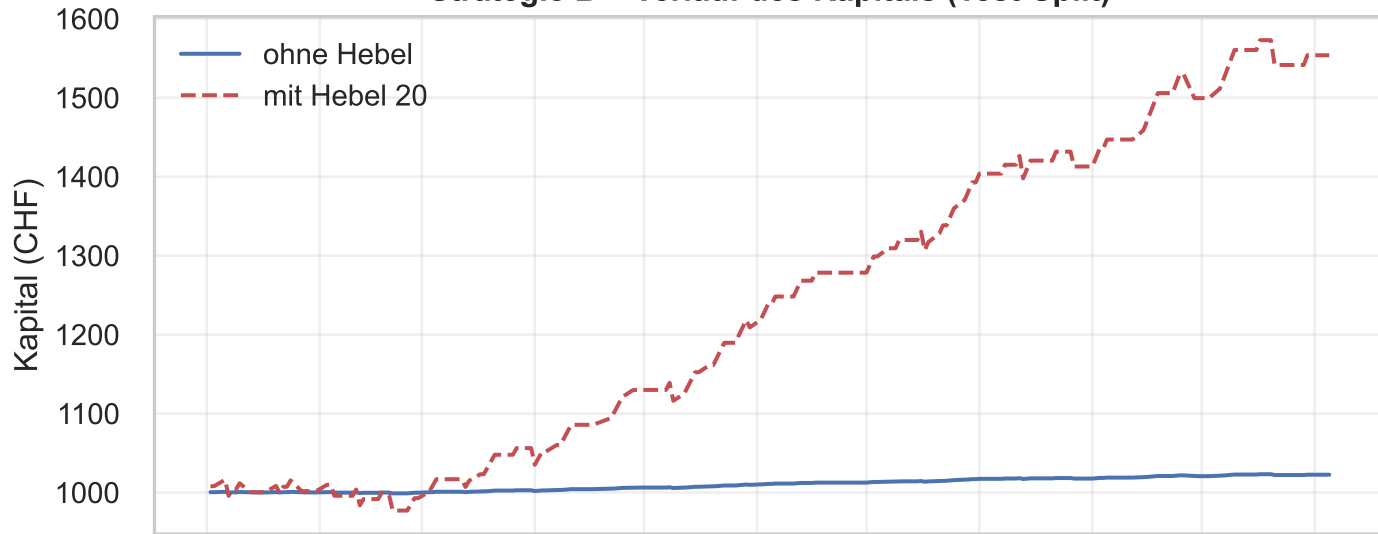


Abbildung: Equity Curve für Strategie B (10 % des aktuellen Vermögens pro Trade)

auf dem Test-Split. 2025-01 2025-02 2025-03 2025-04 2025-05 2025-06 2025-07 2025-08 2025-09 2025-10 2025-11

## Strategie A – kumulierter P&L (Test-Split)

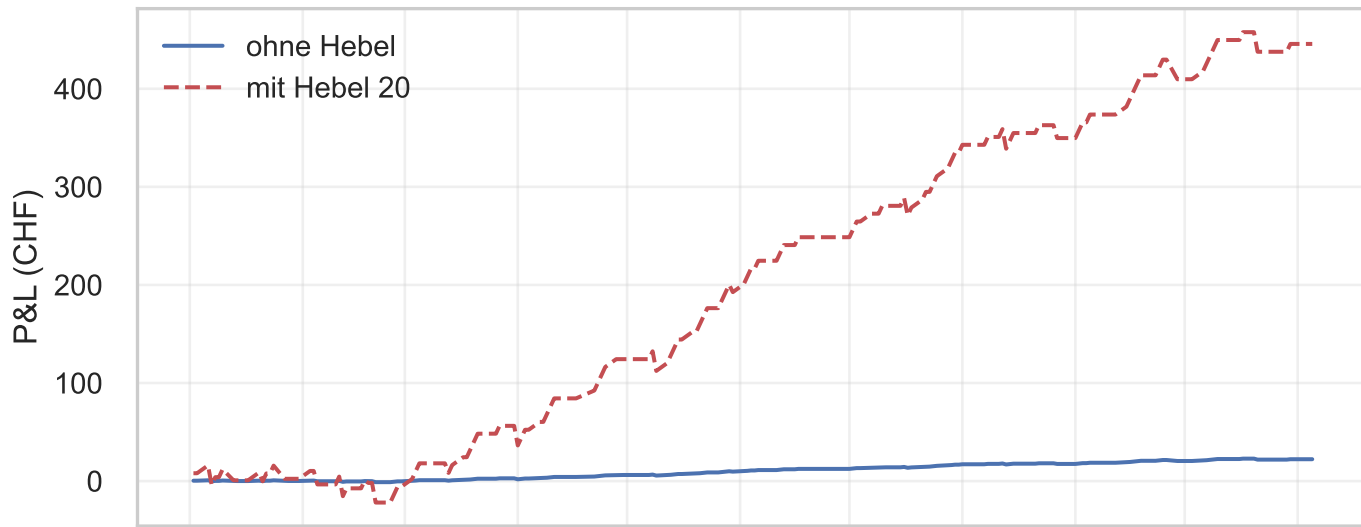


Abbildung: kumulierter Gewinn/Verlust für Strategie A (fixer Einsatz)

mit und ohne Hebel 20 auf dem Test-Split

Feature Importance – Signal-Modell

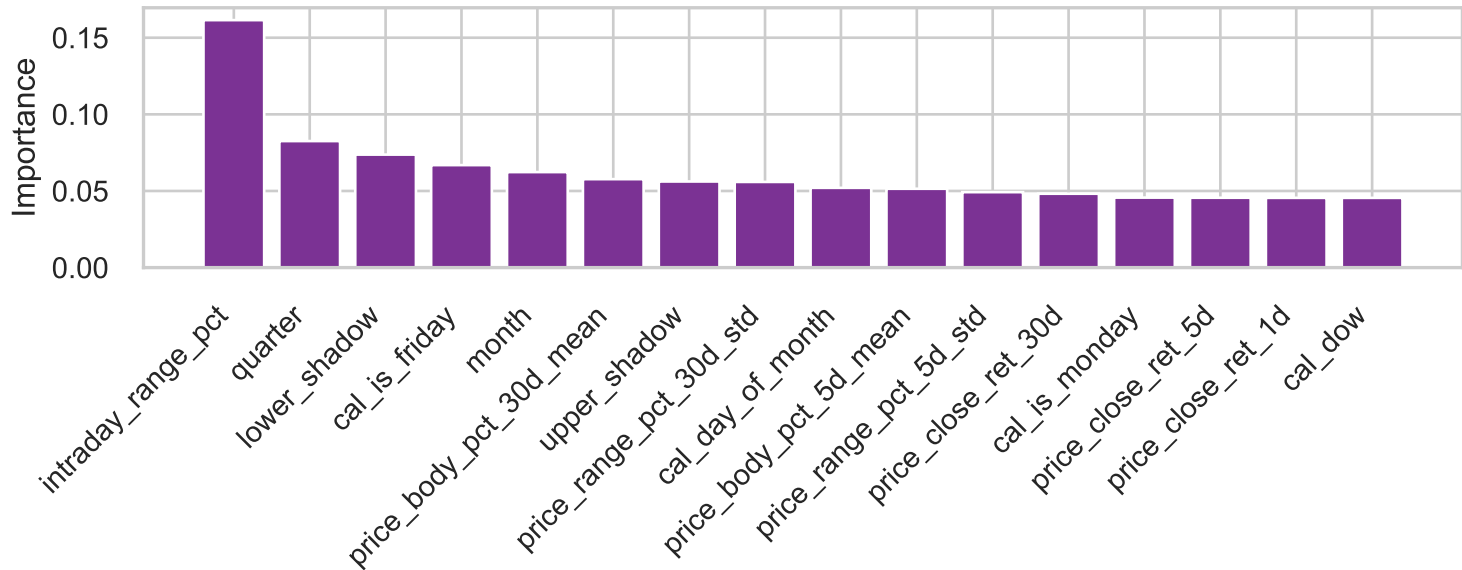


Abbildung: Wichtigkeit der Features für das Signal-Modell (neutral vs move).

Feature Importance – Richtungs-Modell

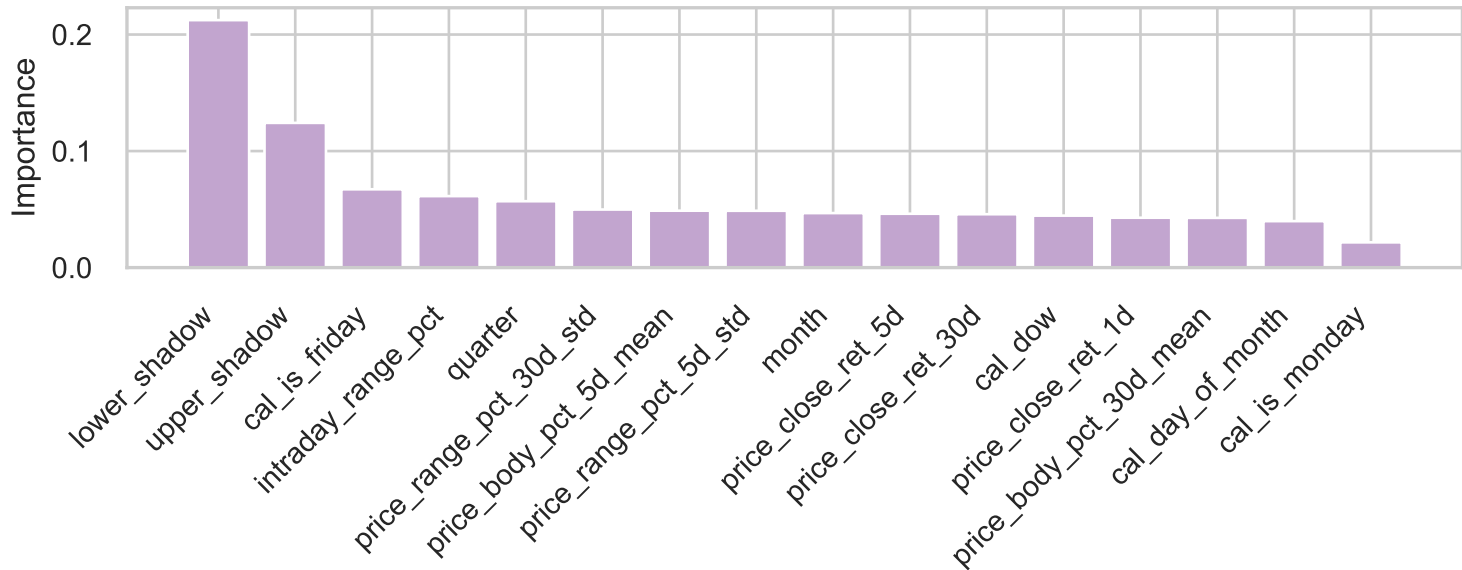


Abbildung: Wichtigkeit der Features für das Richtungs-Modell (down vs up).