

Zwei-Stufen-XGBoost – Experiment-Report

Experiment-ID: hp_long_eod_h4_thr0p6pct_hit_6

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: 0.006
- down_threshold: -0.006
- 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_eod_h4_thr0p6pct_hit_6.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): 0.3 (höher → weniger Trades, tendenziell höhere Qualität).
- DIRECTION_THRESHOLD (Stufe 2 – down vs. up, für Metriken): 0.4 (niedriger → mehr up, höher → weniger up).
- DIRECTION_THRESHOLDS (Stufe 2 – Trading-Entscheidungen): down, wenn $P(\text{up}) \leq 0.475$, up, wenn $P(\text{up}) \geq 0.5249999999999999$.

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: 0.5413165266106442

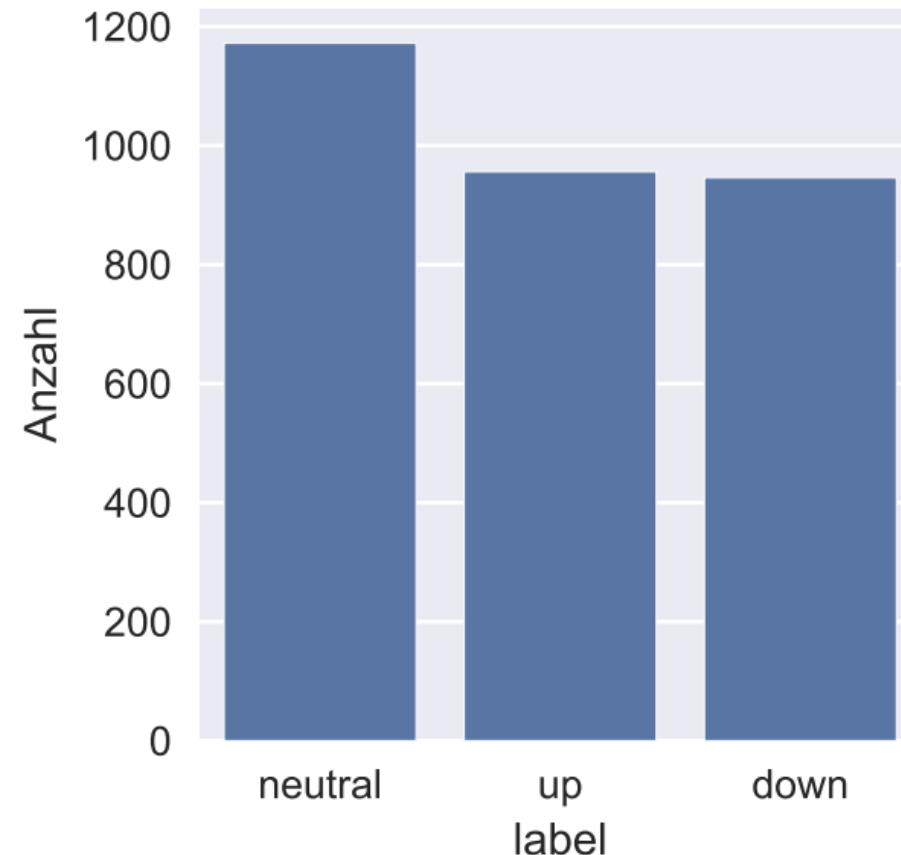
Richtungs-Modell (Stufe 2):

- 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

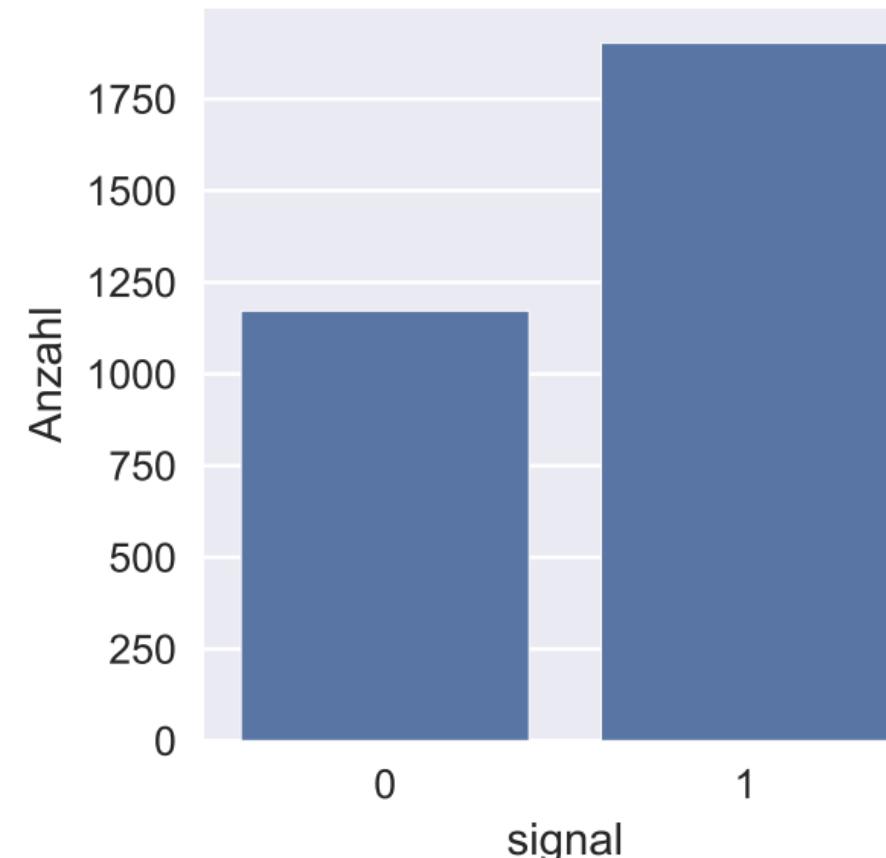
Verwendete Features (FEATURE_COLS)

#	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: Close_t / Close_{t-1} - 1.
4	price_close_ret_5d	Schlusskurs-Return über 5 Tage: Close_t / 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: Close_t / 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.
16	cal_is_month_start	Flag: 1 wenn Monatsanfang, sonst 0.
17	cal_is_month_end	Flag: 1 wenn Monatsende, sonst 0.
18	hol_is_us_federal_holiday	Flag: 1 wenn US-Feiertag, sonst 0.
19	hol_is_day_before_us_federal_holiday	Flag: 1 wenn Tag vor US-Feiertag.
20	hol_is_day_after_us_federal_holiday	Flag: 1 wenn Tag nach US-Feiertag.

Label-Verteilung (neutral / up / down)



Signal-Verteilung (0=neutral, 1=move)



Richtung-Verteilung (nur signal==1)

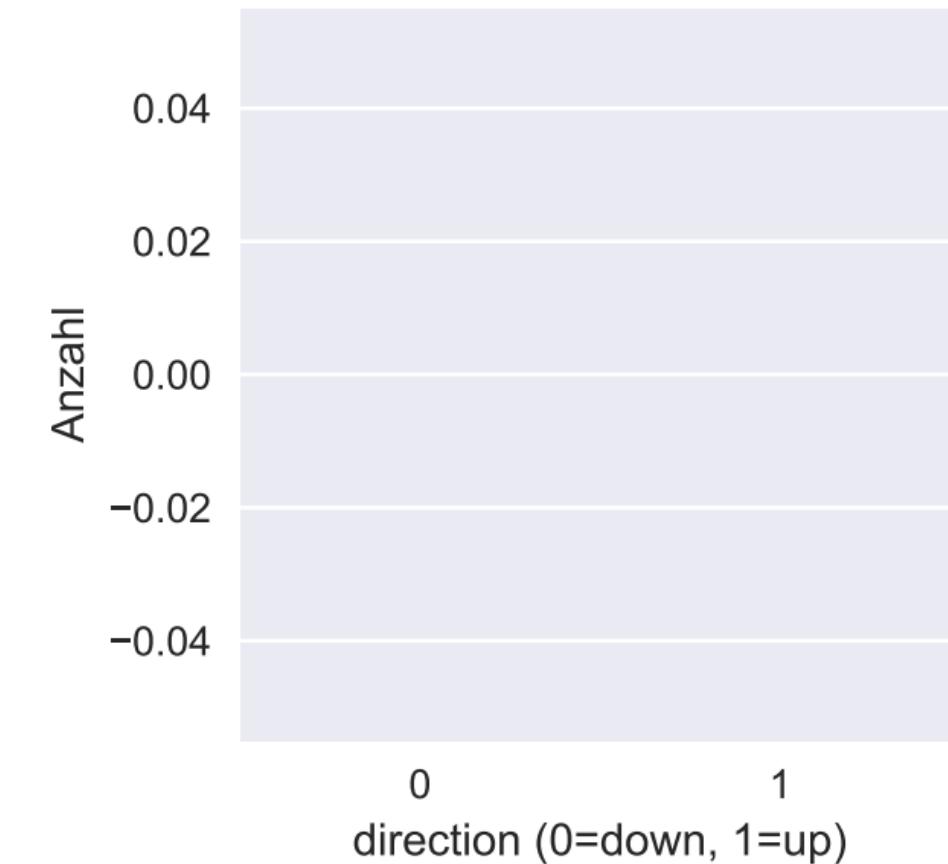
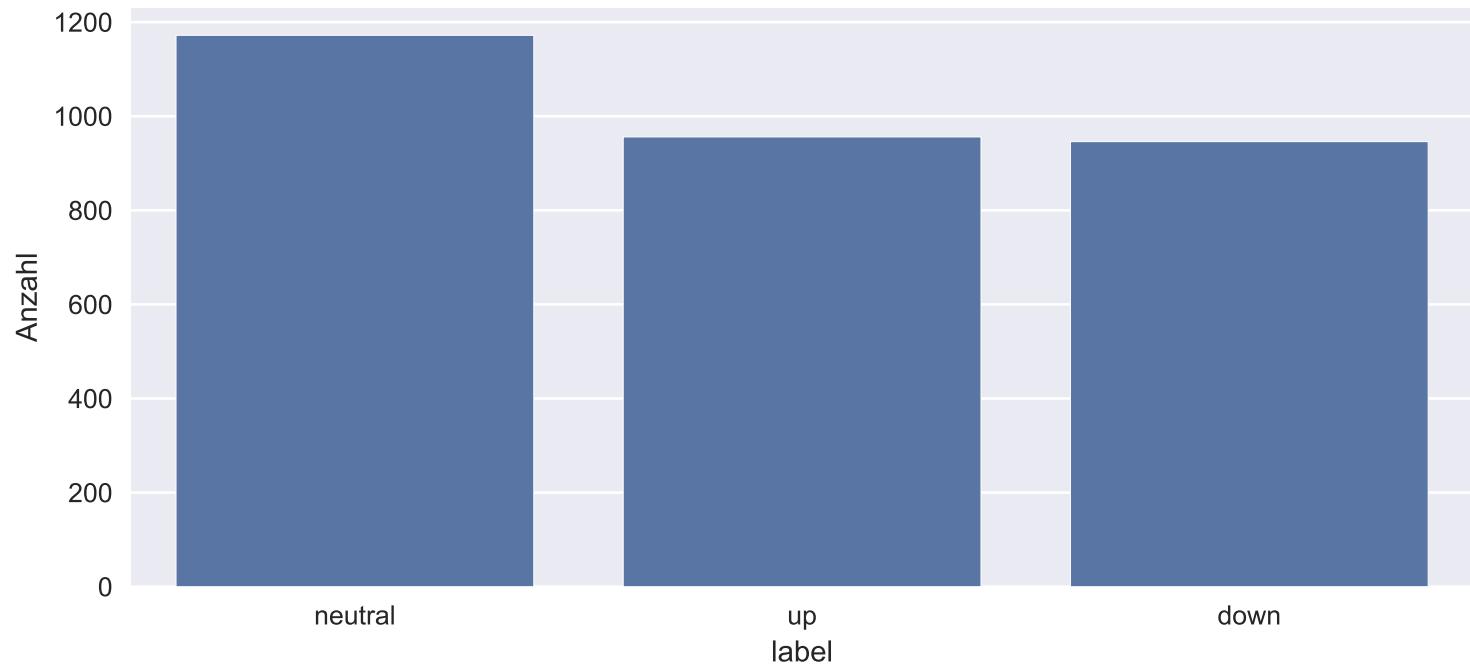


Abbildung: Klassenverteilungen für label, signal und direction im vollständigen Trainingsdatensatz.

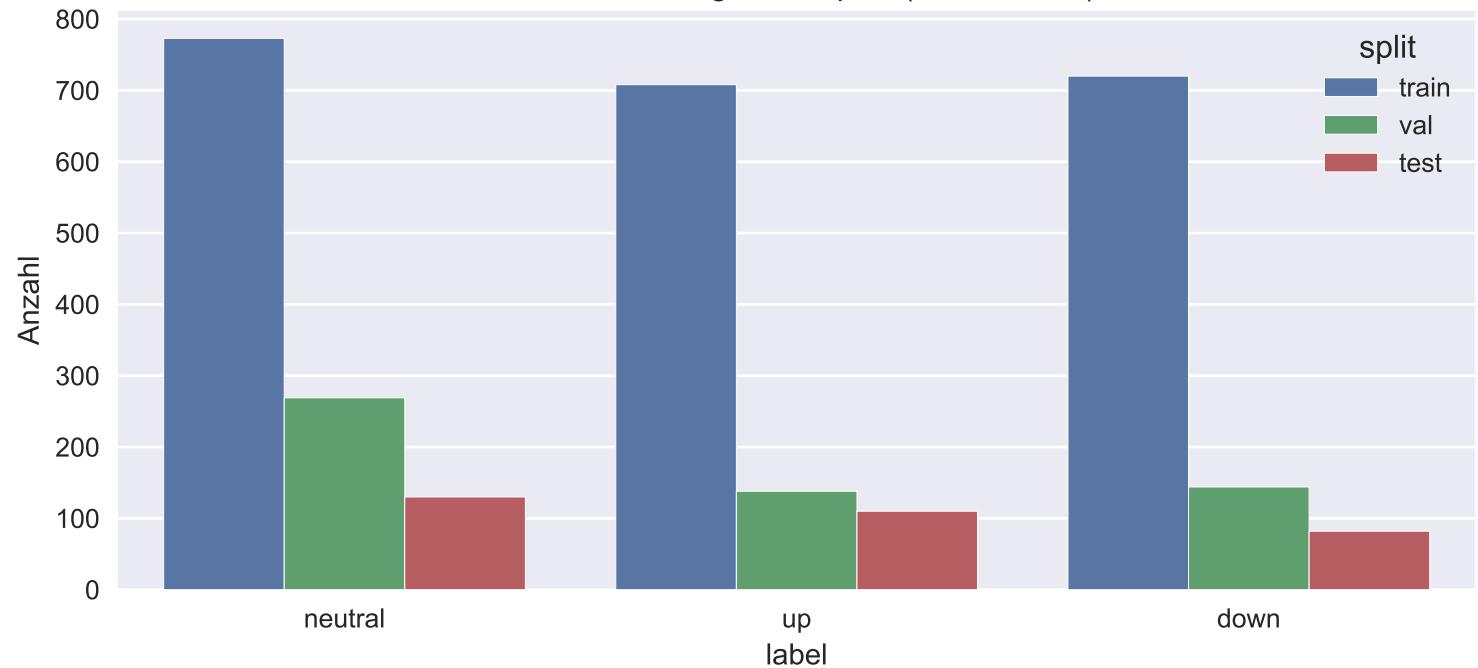
Label-Verteilung – gesamter Datensatz



label	count
neutral	1172
up	956
down	946

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

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



split	neutral	up	down
train	773	708	720
val	269	138	144
test	130	110	82

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

EURUSD-Zeitreihe mit Train/Val/Test-Bereichen

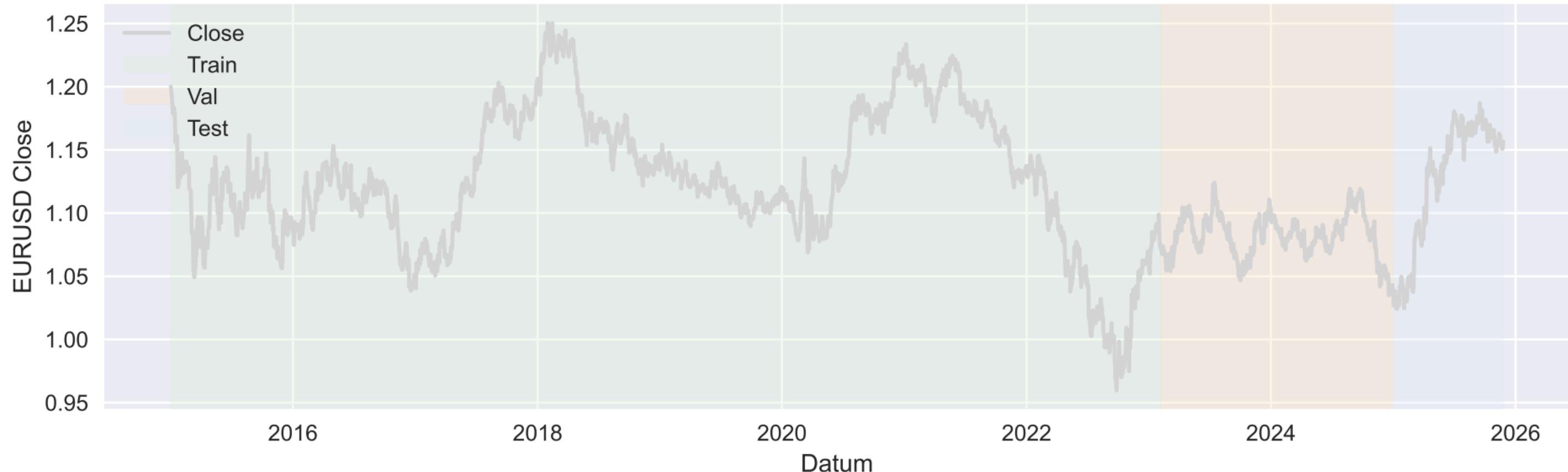


Abbildung: EURUSD-Schlusskurs über den gesamten Zeitraum mit farblich markierten Trainings-, Validierungs- und Testphasen.

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

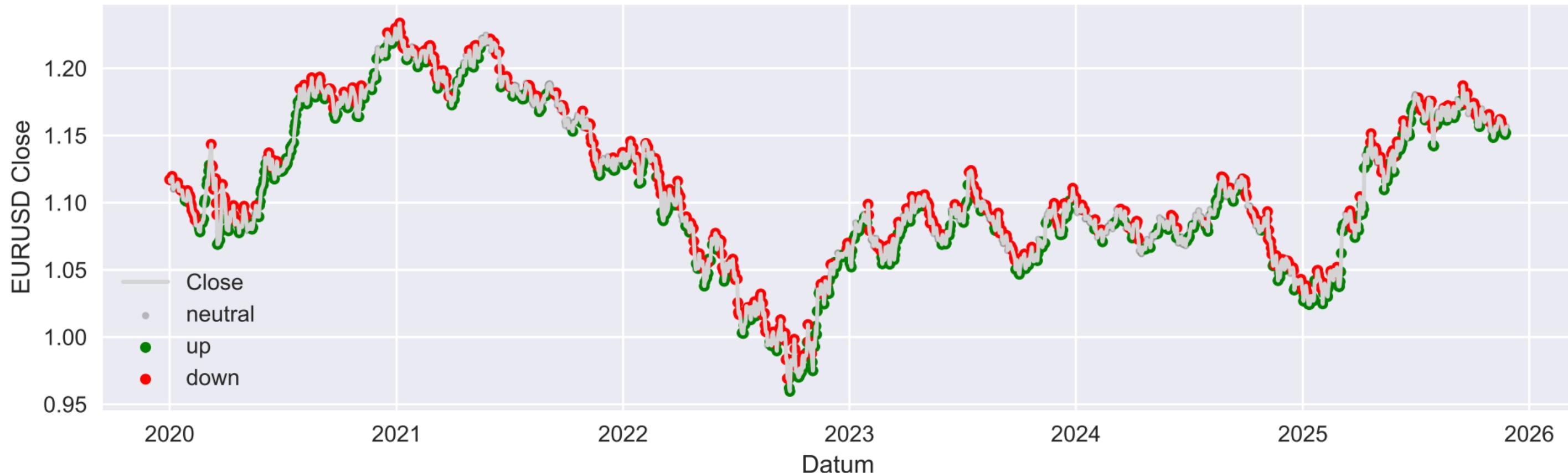


Abbildung: EURUSD-Schlusskurs mit markierten up/down-Tagen im betrachteten Zeitraum.

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

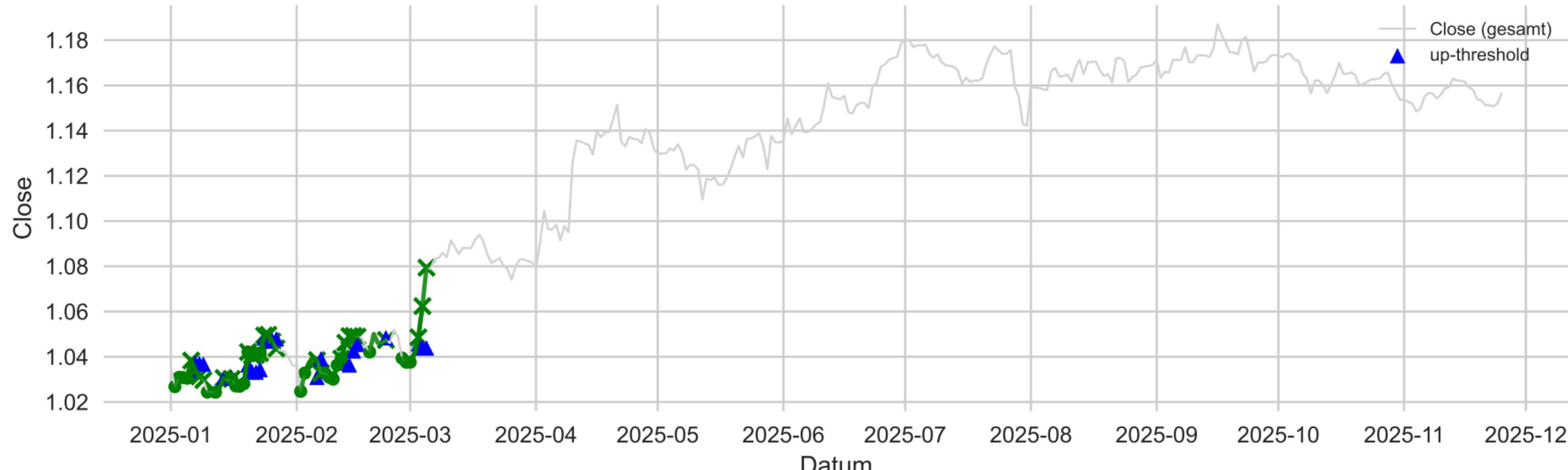


Abbildung: Preis-Segmente t..t+horizon für alle Testtage mit true label 'up'.

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

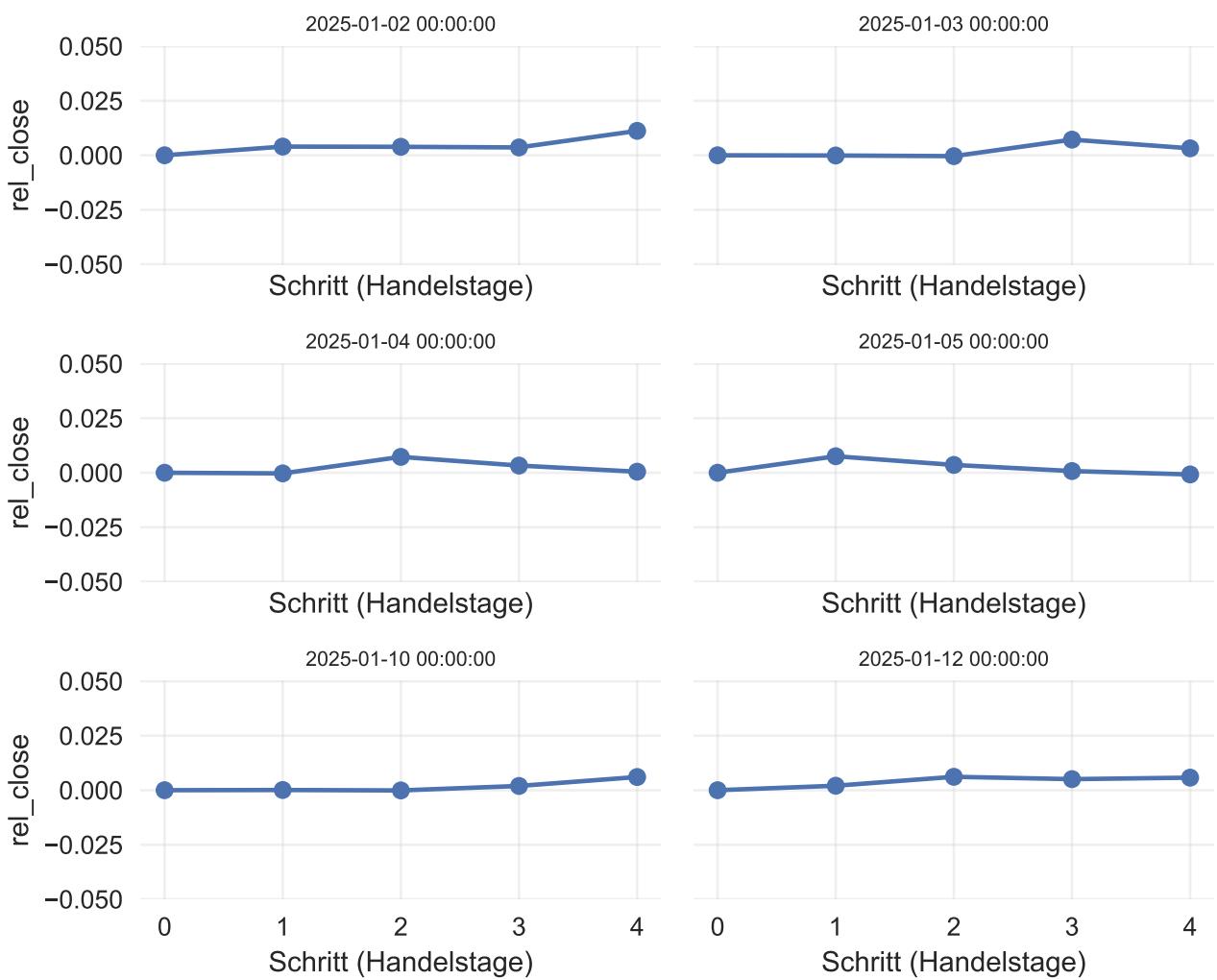


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

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

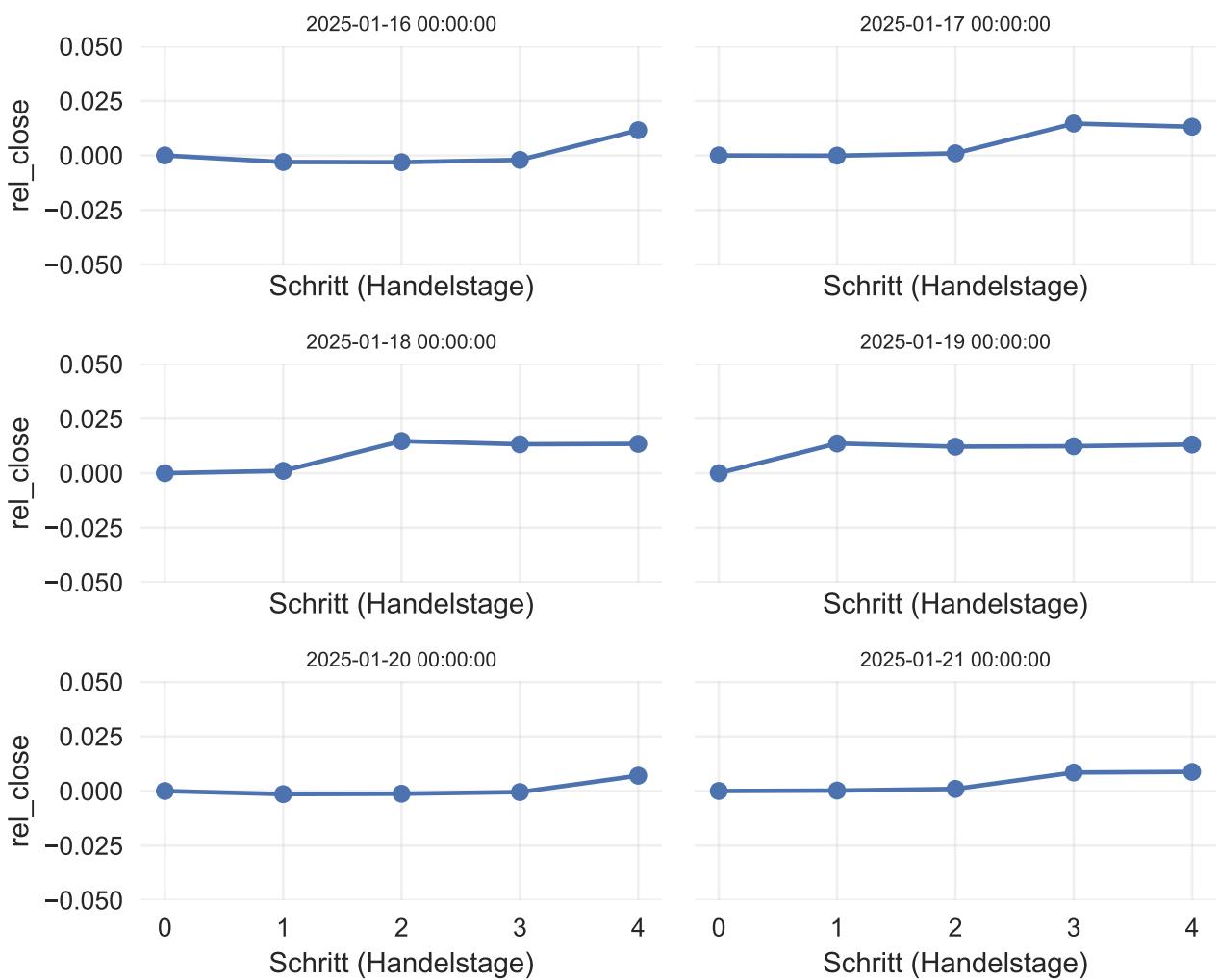


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

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

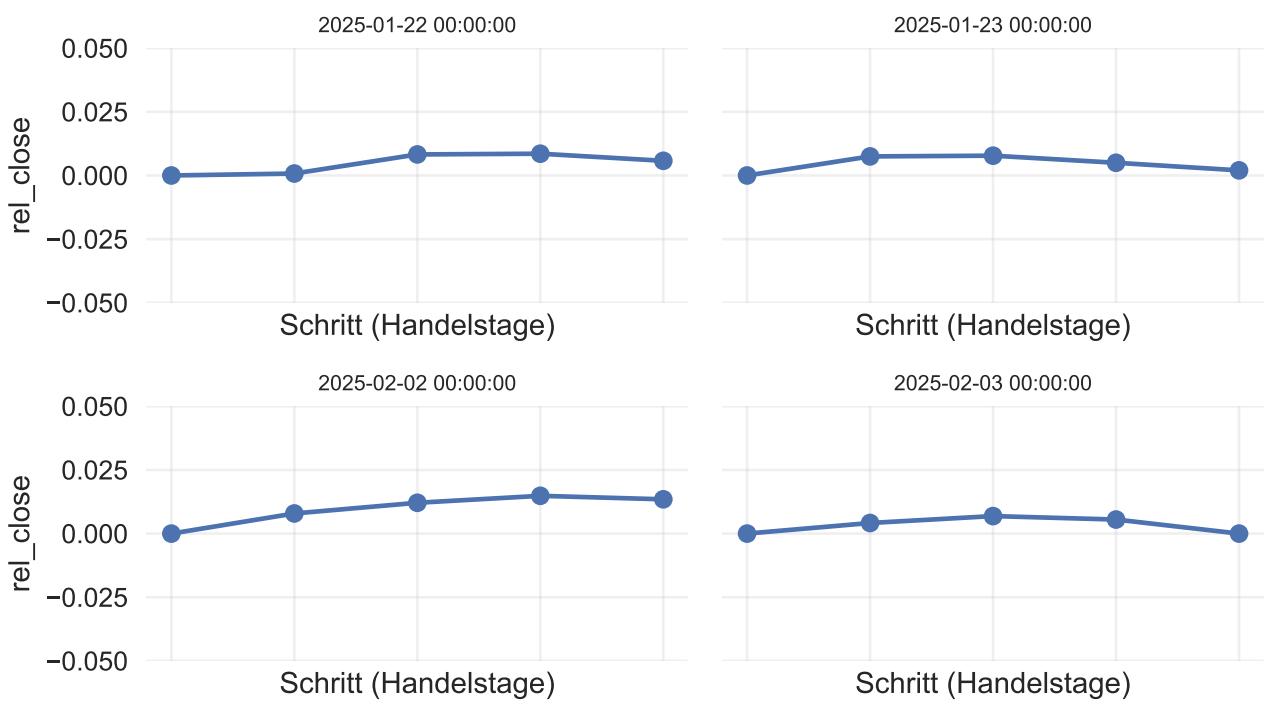


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

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

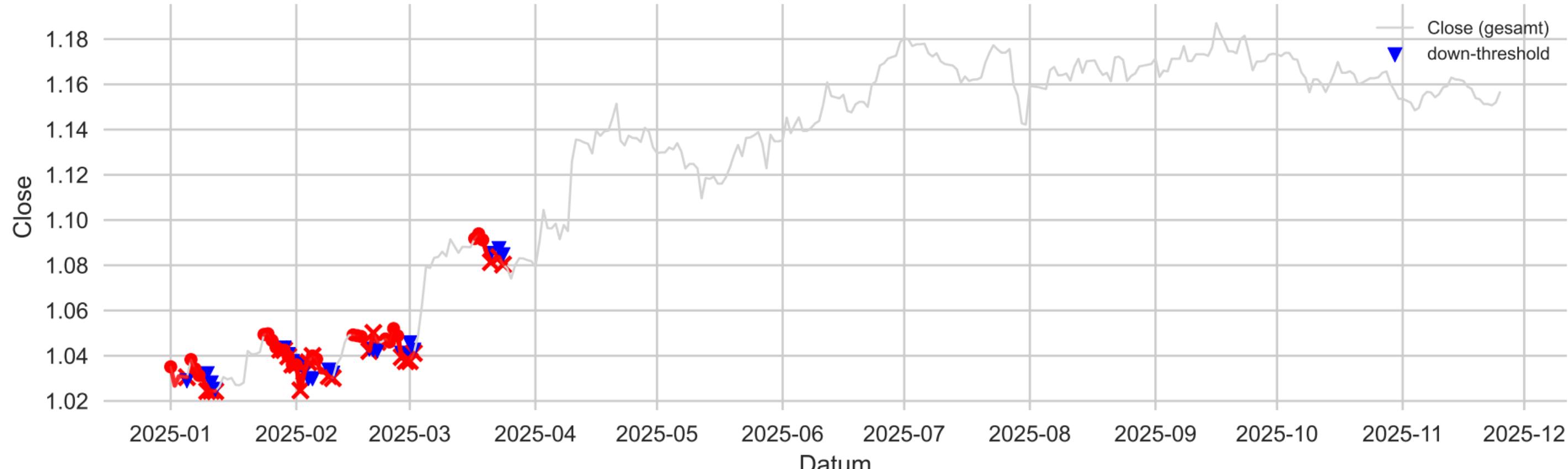


Abbildung: Preis-Segmente t..t+horizon für alle Testtage mit true label 'down'.

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

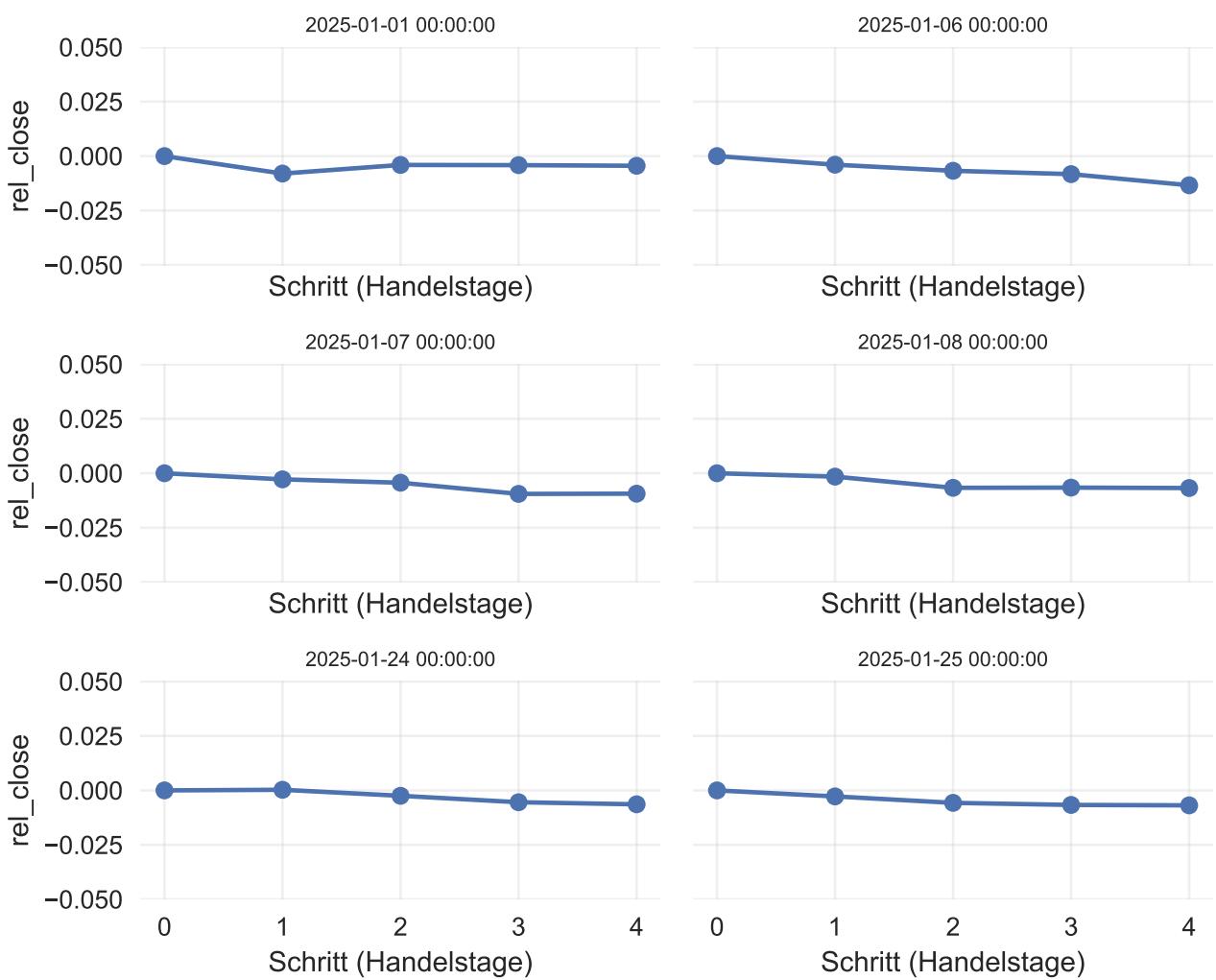


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

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

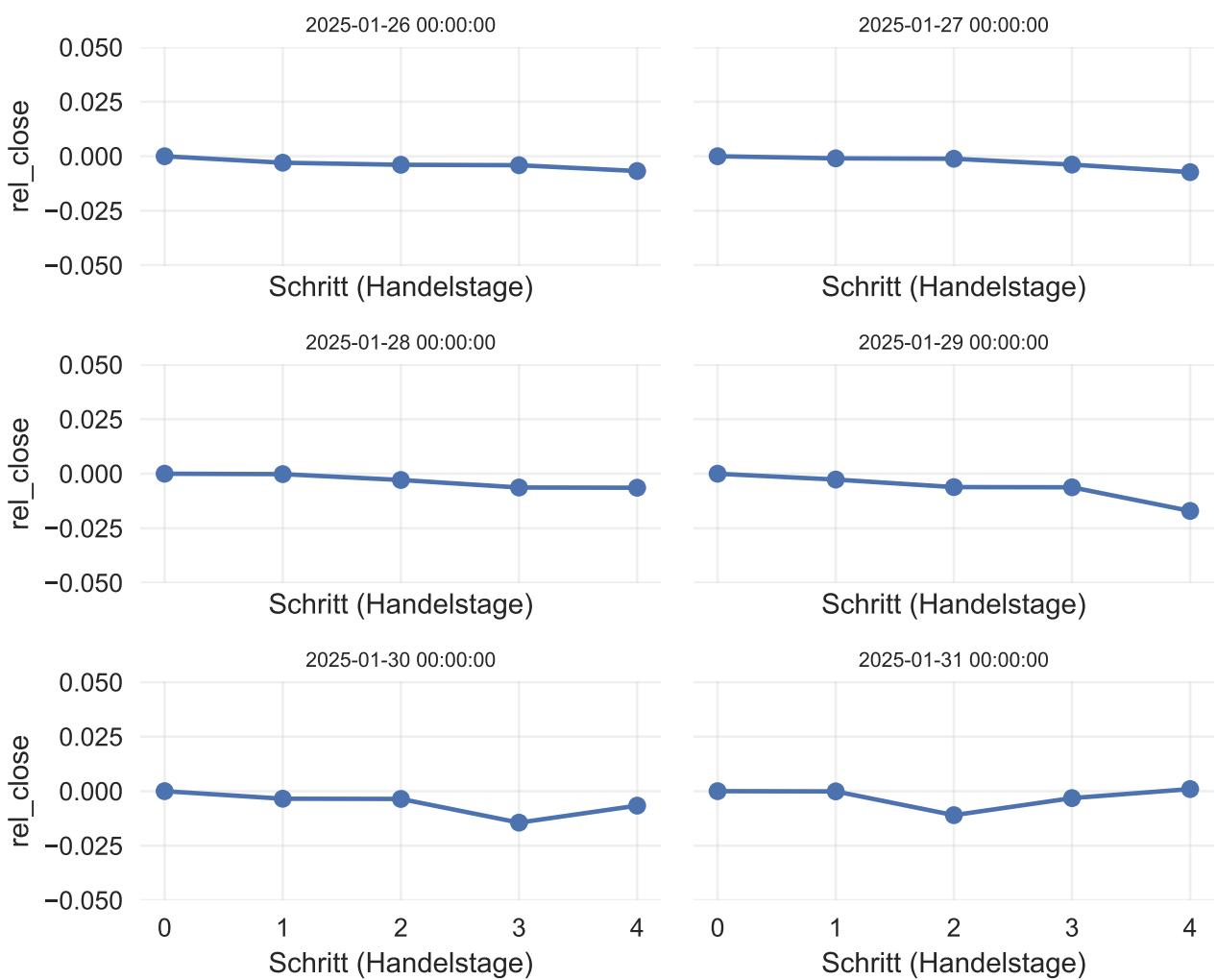


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

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

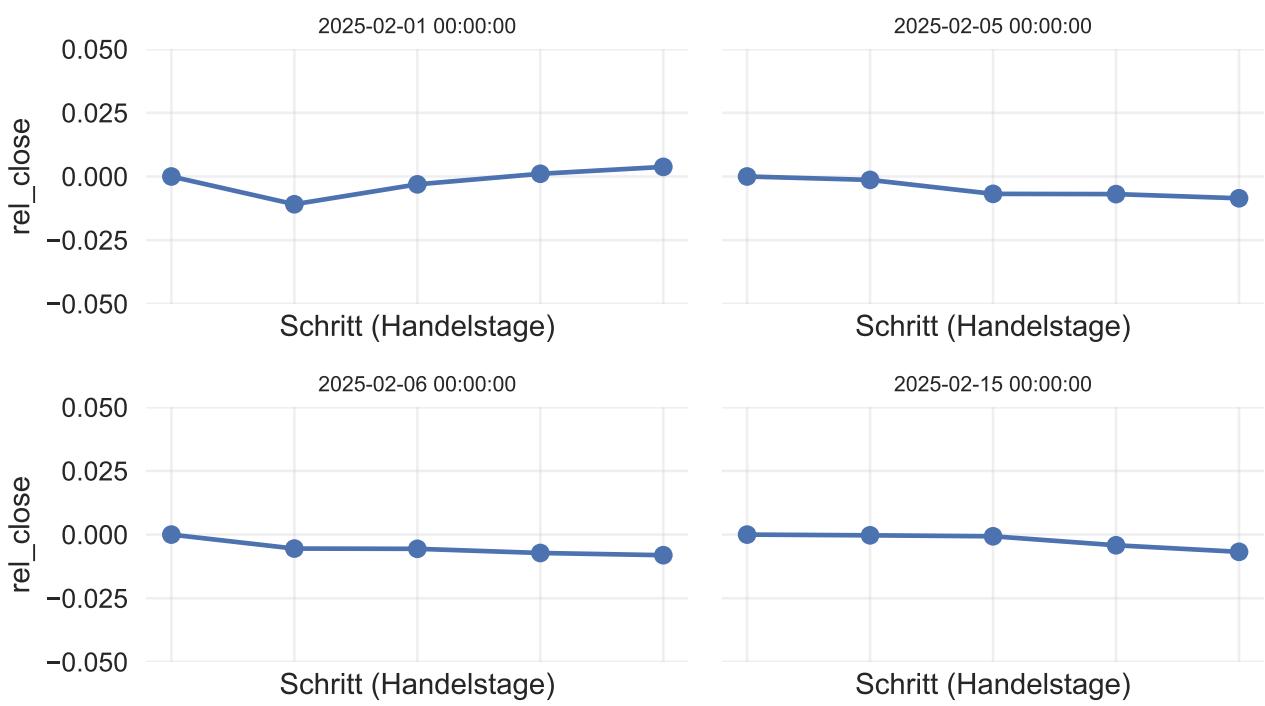


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

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

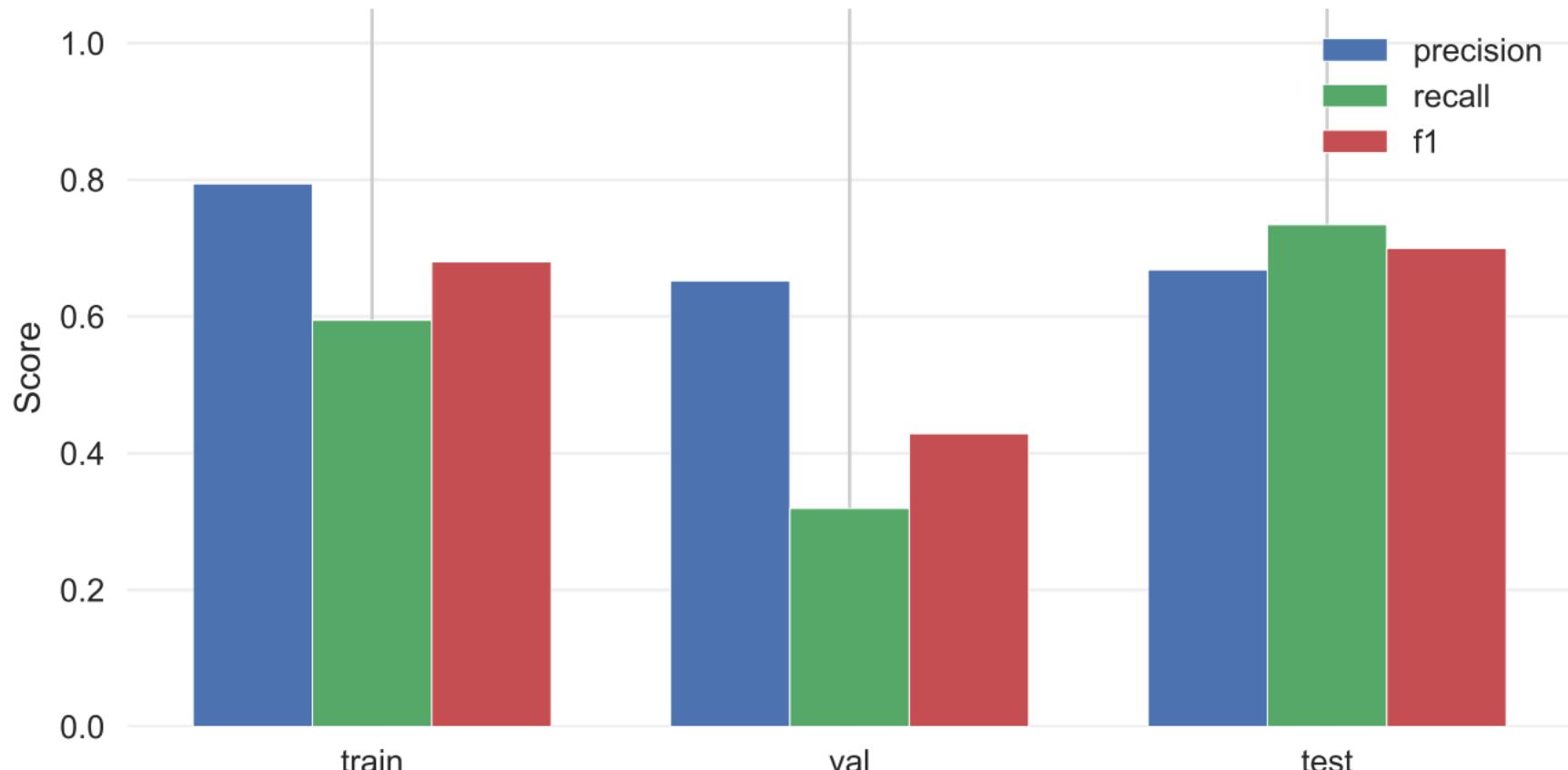


Abbildung: Precision, Recall und F1 der positiven Klasse je Split (train/val/test).

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

split	precision	recall	f1	support
train	0.794	0.595	0.680	1428.000
val	0.652	0.319	0.429	282.000
test	0.668	0.734	0.700	192.000

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

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

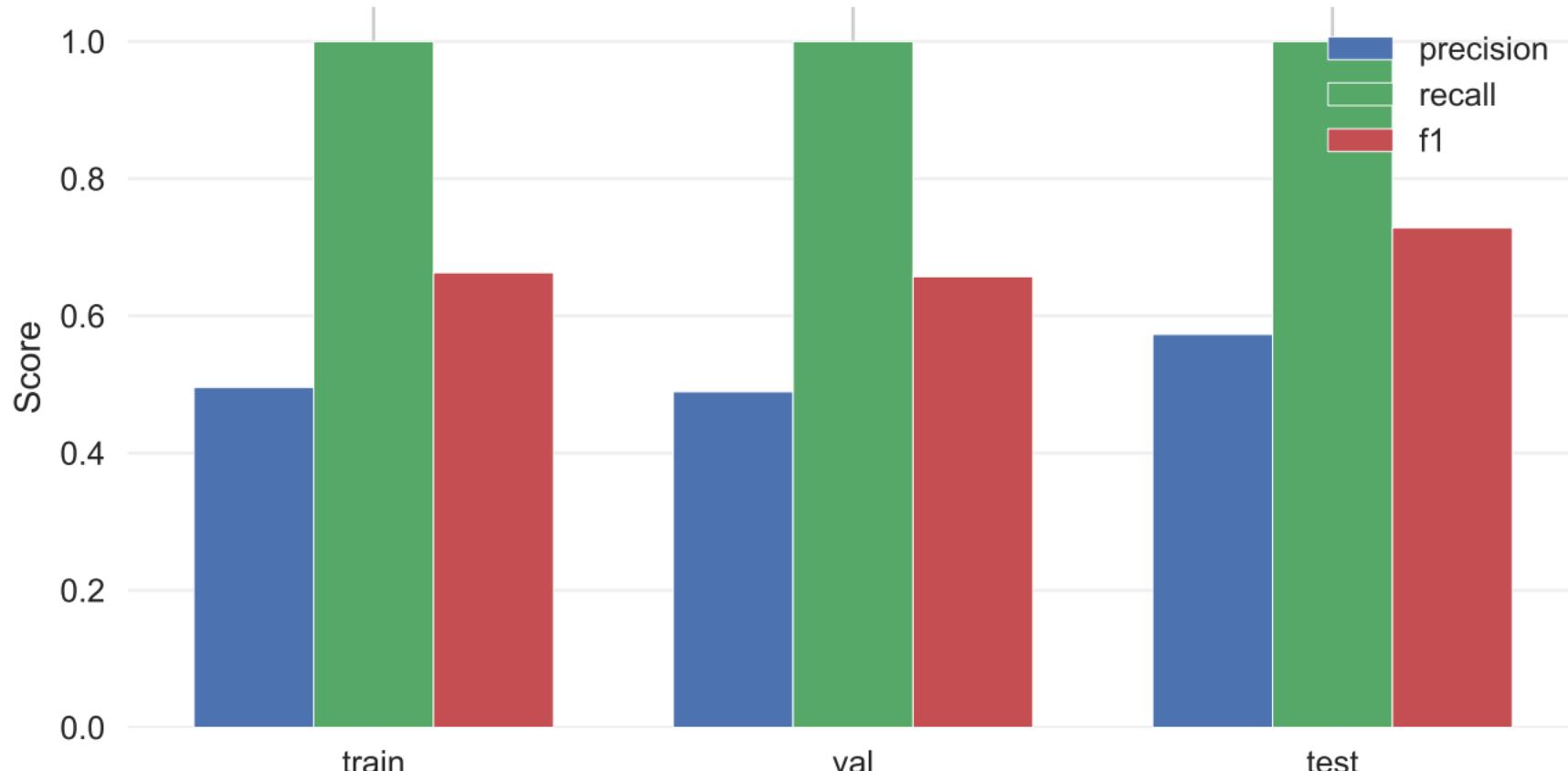


Abbildung: Precision, Recall und F1 der positiven Klasse je Split (train/val/test).

Richtungs-Modell – Tabelle (Klasse 'up')

split	precision	recall	f1	support
train	0.496	1.000	0.663	708.000
val	0.489	1.000	0.657	138.000
test	0.573	1.000	0.728	110.000

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

Kombinierte Test-Auswertung – neutral / up / down

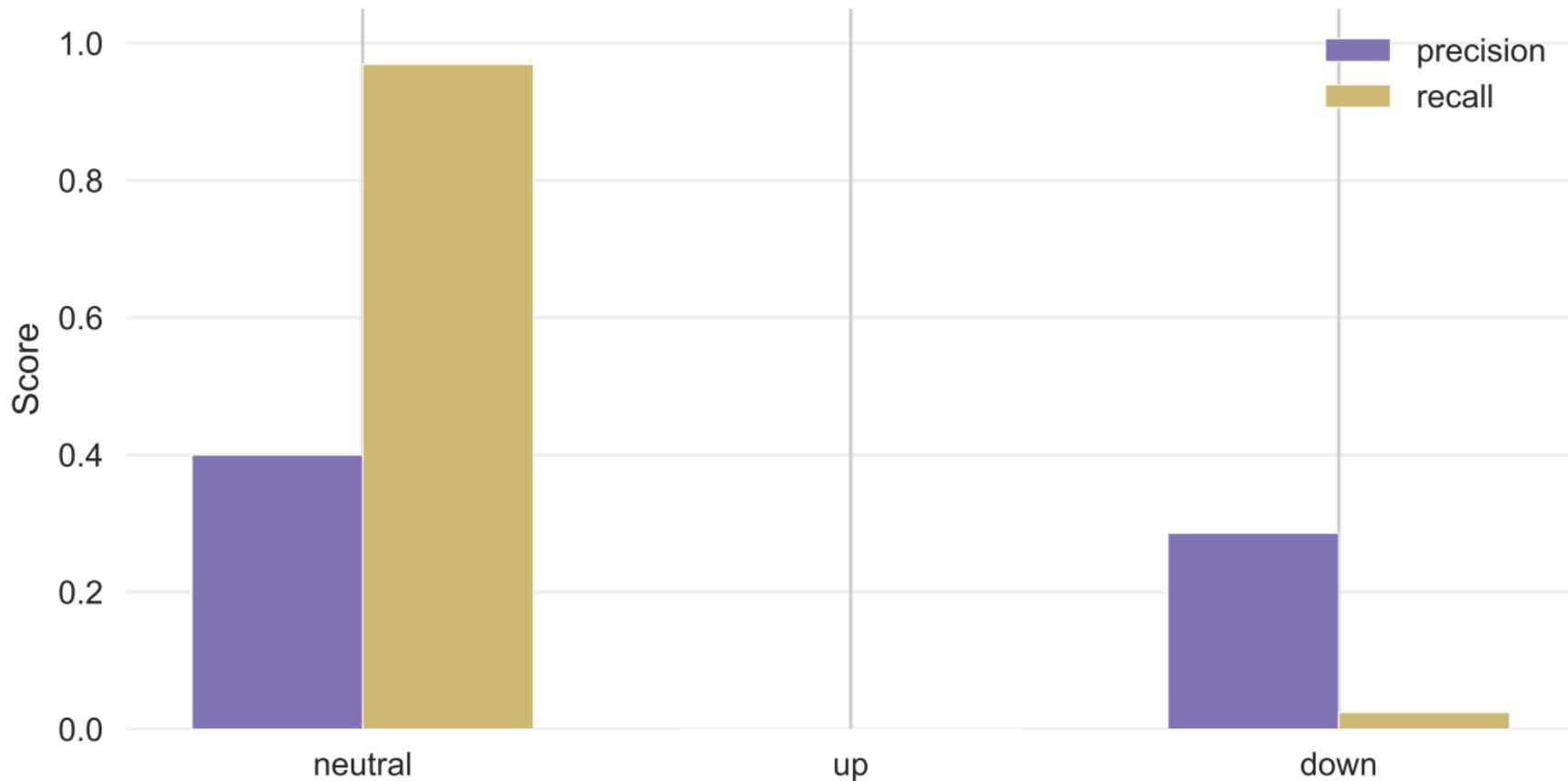


Abbildung: Precision und Recall der kombinierten 3-Klassen-Vorhersage (neutral/up/down) auf dem Test-Split.

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

klasse	precision	recall	f1	support
neutral	0.4	0.969	0.566	130
up	0.0	0.0	0.0	110
down	0.286	0.024	0.045	82

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

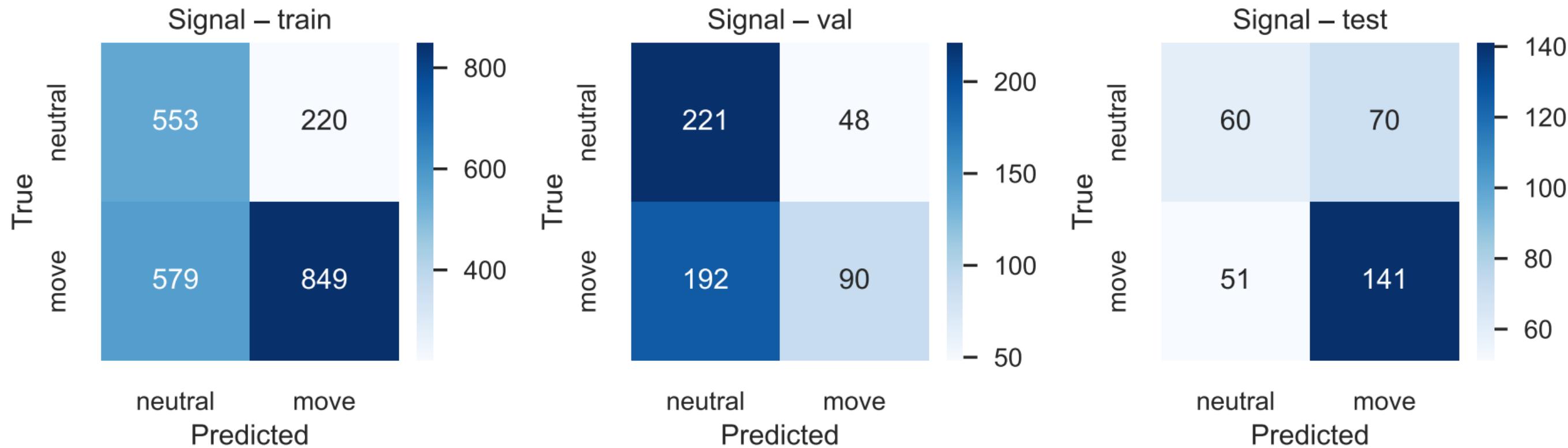


Abbildung: Confusion-Matrizen des Signal-Modells (neutral vs move) für Train-, Validierungs- und Test-Split.

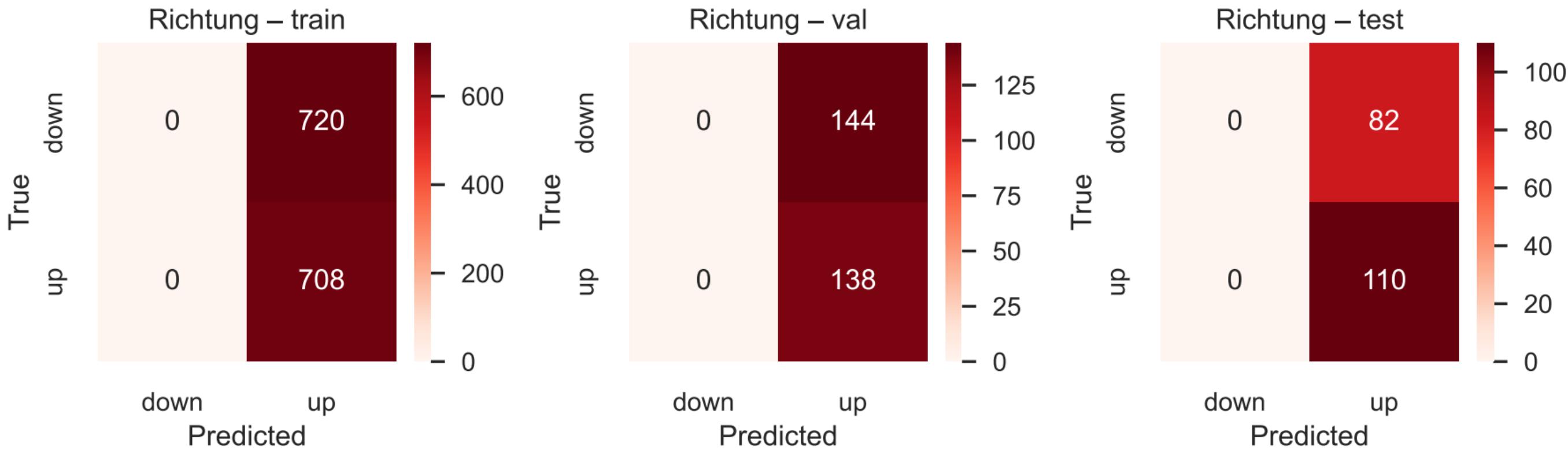


Abbildung: Confusion-Matrizen des Richtungs-Modells (down vs up) für Train-, Validierungs- und Test-Split.

Confusion Matrix – Test (neutral / up / down)

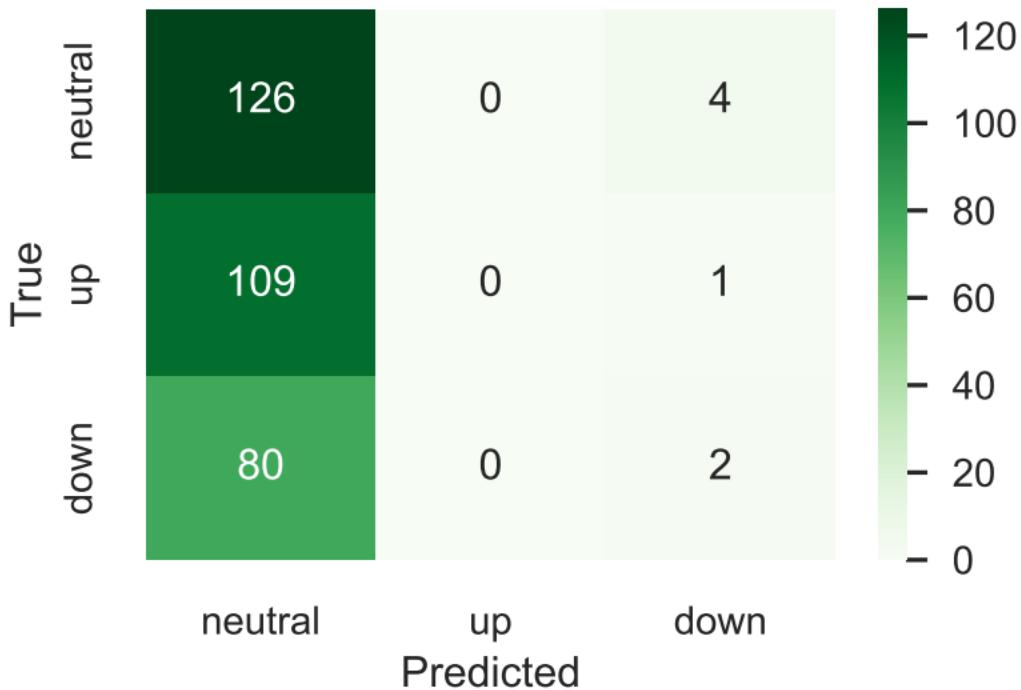


Abbildung: Confusion-Matrix des kombinierten Modells (neutral/up/down) auf dem Test-Split.

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

modell	split	TN	FP	FN	TP
signal	train	553	220	579	849
signal	val	221	48	192	90
signal	test	60	70	51	141
direction	train	0	720	0	708
direction	val	0	144	0	138
direction	test	0	82	0	110

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	0	-
combined	down	5	neutral:4, up:1
signal	move	70	neutral:70

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

Keine Fälle: true=neutral, predicted=up im kombinierten Test.

Abbildung: Es gibt keine Testtage, an denen ein neutraler Tag fälschlich als 'up' klassifiziert wurde.

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

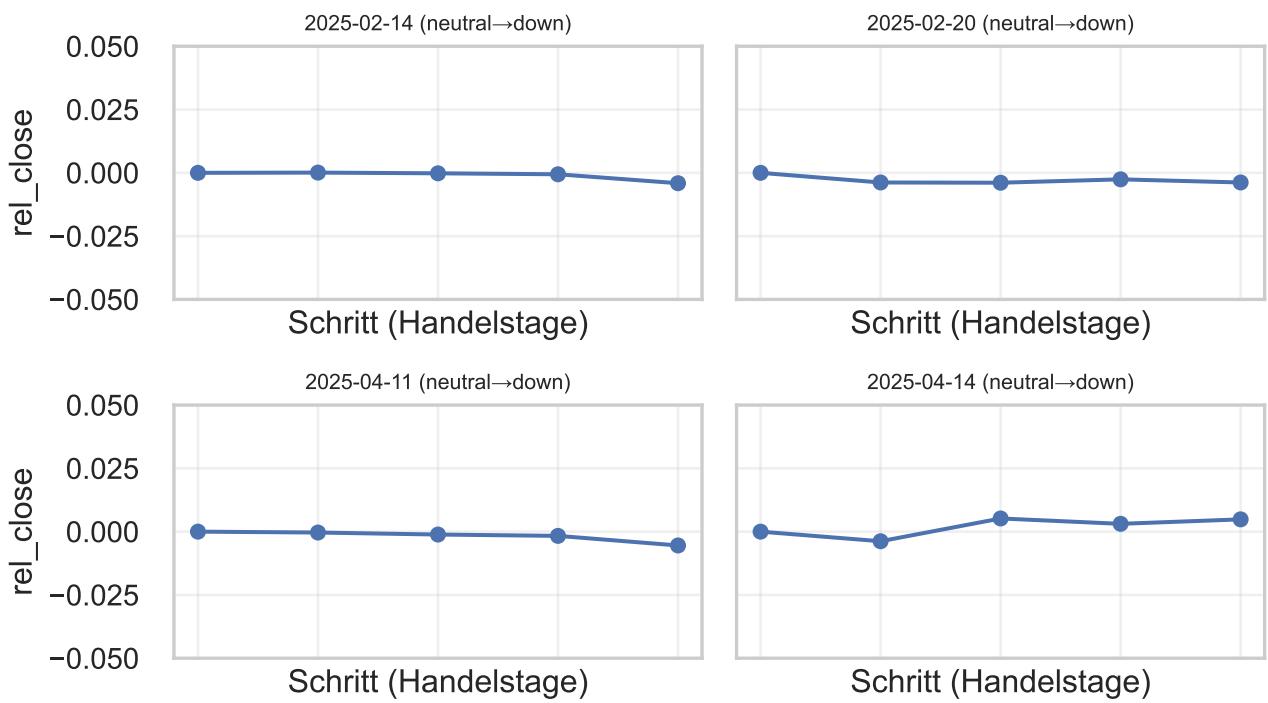


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'neutral', die im kombinierten Test als 'down' klassifiziert wurden.

Tradesimulation – Strategien A und B (Test-Split)

Strategy	Kennzahl	Wert
A (fixer Einsatz)	Anzahl Trades	7
A (fixer Einsatz)	Einsatz up / down (CHF)	100 / 100
A (fixer Einsatz)	Trades up / down	0 / 7
A (fixer Einsatz)	Gewinner / Verlierer	2 / 5
A (fixer Einsatz)	Gesamt-P&L (CHF)	-3.80
A (fixer Einsatz, Hebel 20)	Gesamt-P&L (CHF)	-76.00
B (10% vom Kapital)	Startkapital (CHF)	1000.00
B (10% vom Kapital)	Endkapital (CHF)	996.20
B (10% vom Kapital)	Minimum Kapital (CHF)	996.20
B (10% vom Kapital, Hebel 20)	Endkapital (CHF)	925.75
B (10% vom Kapital, Hebel 20)	Minimum Kapital (CHF)	925.75

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	0.0
neutral	down	-1.0
up	neutral	0.0
up	up	0.0
up	down	-1.0
down	neutral	0.0
down	up	0.0
down	down	0.6

Tabelle: durchschnittliche Kosten (CHF) pro Fall für jede Kombination aus wahrem 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	126	0.0
neutral	up	0	0.0
neutral	down	4	-4.0
up	neutral	109	0.0
up	up	0	0.0
up	down	1	-1.0
down	neutral	80	0.0
down	up	0	0.0
down	down	2	1.2

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)

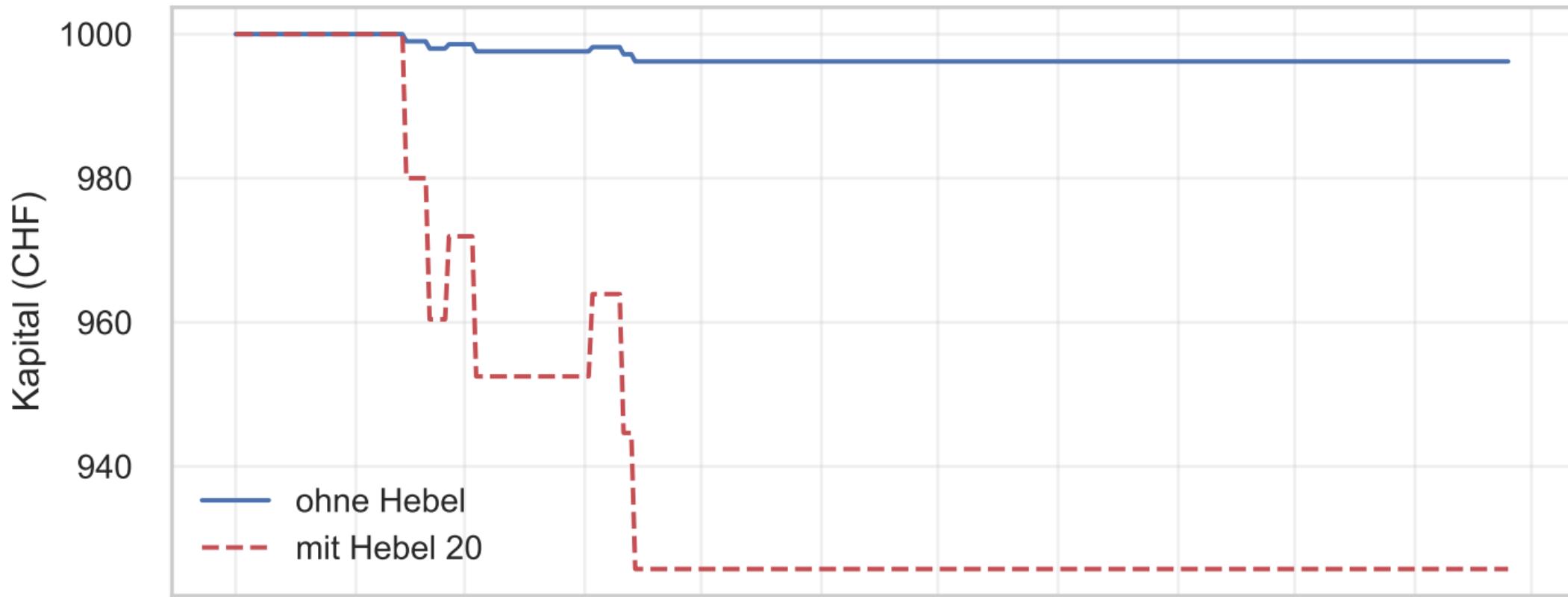


Abbildung: Equity Curve für Strategie B (10 % des aktuellen Vermögens pro Trade)
auf dem Test-Split. 2025-01 2025-02 2025-03 2025-04 2025-05 2025-06 2025-07 2025-08 2025-09 2025-10 2025-11 2025-12

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



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

mit und ohne Hebel 2025-01-01 2025-02-01 2025-03-01 2025-04-01 2025-05-01 2025-06-01 2025-07-01 2025-08-01 2025-09-01 2025-10-01 2025-11-01 2025-12-

Feature Importance – Signal-Modell

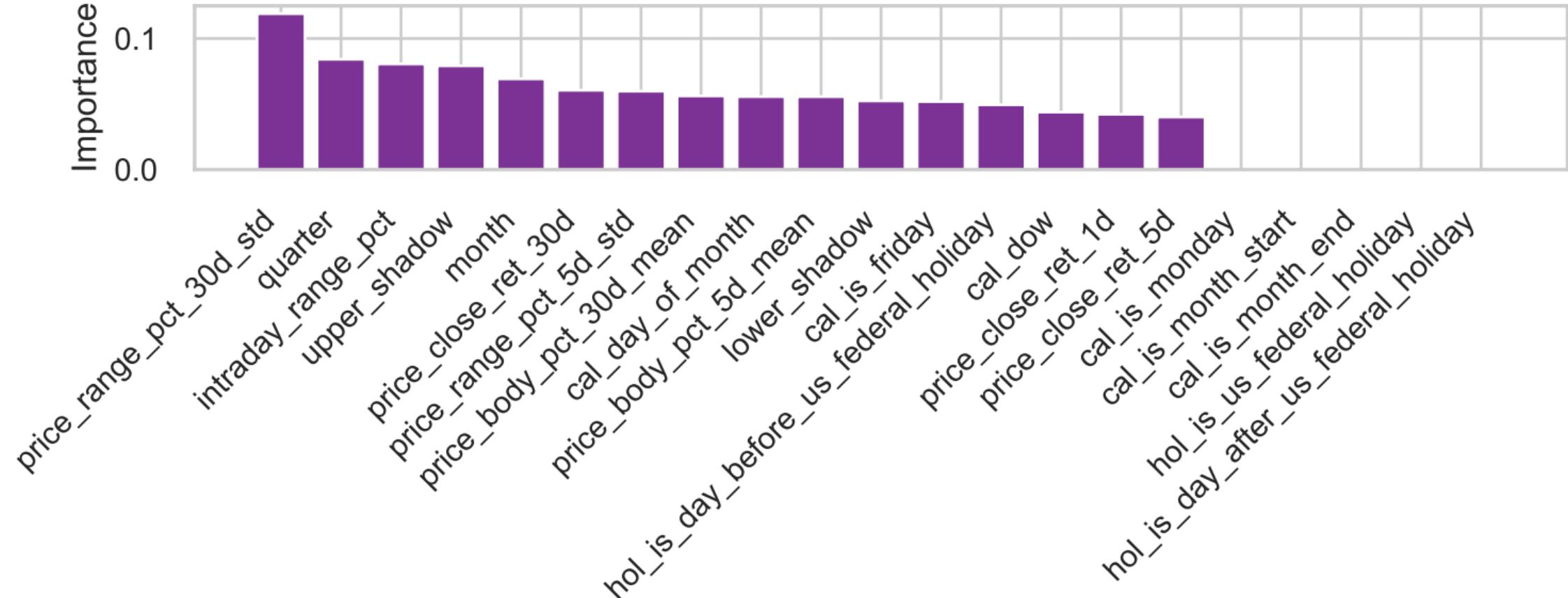


Abbildung: Wichtigkeit der Features für das Signal-Modell (neutral vs move).

Feature Importance – Richtungs-Modell

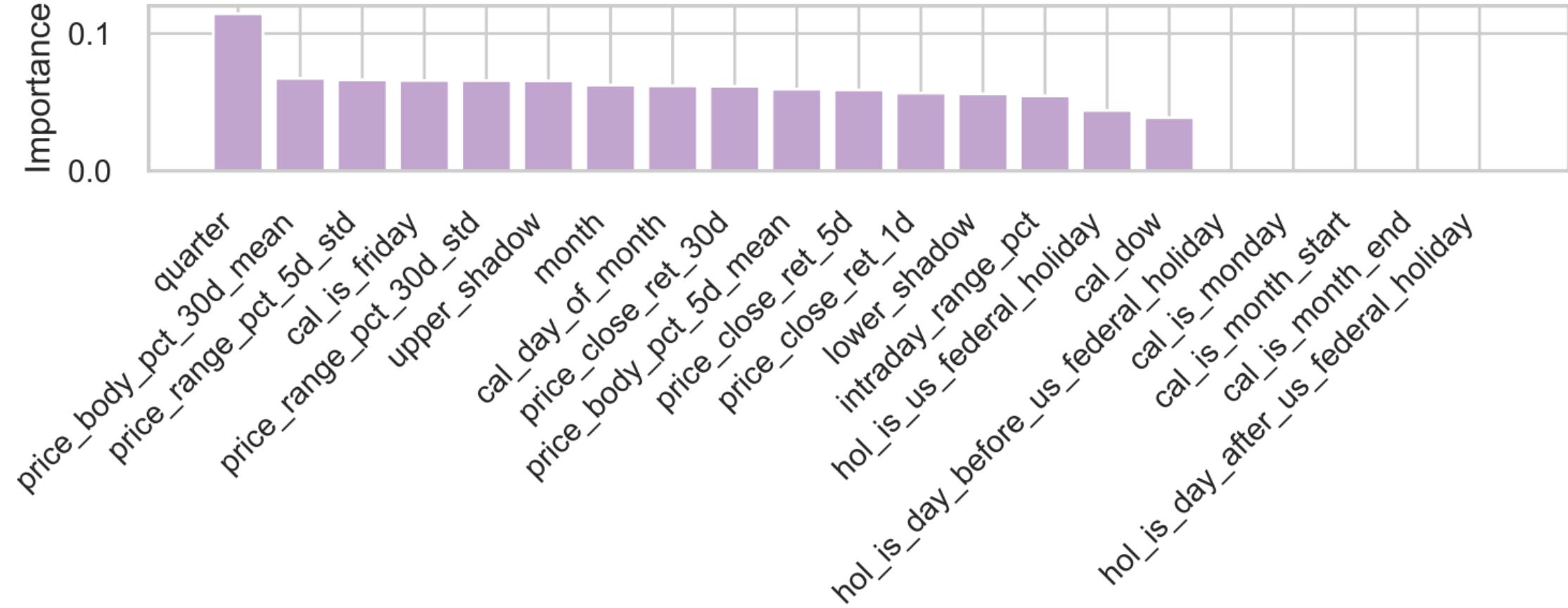


Abbildung: Wichtigkeit der Features für das Richtungs-Modell (down vs up).