

Zwei-Stufen-XGBoost – Experiment-Report

Experiment-ID: hv1_h4_thr0p3pct_hit

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: 0.003
- down_threshold: -0.003
- 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__hv1_h4_thr0p3pct_hit.csv
- test_start: 2025-01-01
- train_frac_within_pretest: 0.8

Entscheidungsgrenzen (Modelle):

- SIGNAL_THRESHOLD (Stufe 1 – move vs. neutral): 0.55 (höher → höhere Precision, niedrigerer Recall).

Features (FEATURE_COLS): vollständige Liste auf der Feature-Seite weiter unten.

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: 0.07539118065433854

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	article_count	Anzahl News-Artikel an Tag t.
1	avg_polarity	Durchschnittliche Sentiment-Polarity der Artikel an Tag t (VADER).
2	avg_neg	Durchschnittlicher negativer Sentiment-Anteil an Tag t.
3	avg_neu	Durchschnittlicher neutraler Sentiment-Anteil an Tag t.
4	avg_pos	Durchschnittlicher positiver Sentiment-Anteil an Tag t.
5	pos_share	Anteil positiver Sentiment-Komponente: avg_pos / (avg_pos + avg_neg).
6	neg_share	Anteil negativer Sentiment-Komponente: avg_neg / (avg_pos + avg_neg).
7	intraday_range_pct	(High - Low) / Close – relative Tagesvolatilität.
8	upper_shadow	Oberer Kerzendocht: High - max(Open, Close).
9	lower_shadow	Unterer Kerzendocht: min(Open, Close) - Low.
10	price_close_ret_1d	Relativer Schlusskurs-Return gegenüber Vortag: Close_t / Close_{t-1} - 1.
11	price_close_ret_5d	Schlusskurs-Return über 5 Tage: Close_t / Close_{t-5} - 1.
12	price_range_pct_5d_std	Standardabweichung der intraday_range_pct über 5 Tage (Volatilität).
13	price_body_pct_5d_mean	Durchschnittlicher Kerzenkörper-Prozentsatz über 5 Tage.
14	price_close_ret_30d	Schlusskurs-Return über 30 Tage: Close_t / Close_{t-30} - 1.
15	price_range_pct_30d_std	Standardabweichung der intraday_range_pct über 30 Tage.
16	price_body_pct_30d_mean	Durchschnittlicher Kerzenkörper-Prozentsatz über 30 Tage.
17	news_article_count_3d_sum	Summe article_count über die letzten 3 Tage.
18	news_article_count_7d_sum	Summe article_count über die letzten 7 Tage.
19	news_pos_share_5d_mean	Durchschnittlicher pos_share über die letzten 5 Tage.
20	news_neg_share_5d_mean	Durchschnittlicher neg_share über die letzten 5 Tage.
21	news_article_count_lag1	article_count am Vortag.
22	news_pos_share_lag1	pos_share am Vortag.
23	news_neg_share_lag1	neg_share am Vortag.
24	month	Kalendermonat (1–12).
25	quarter	Kalenderquartal (1–4).
26	cal_dow	Wochentag (0 = Montag, 6 = Sonntag).
27	cal_day_of_month	Kalendertag im Monat.
28	cal_is_monday	Flag: 1 wenn Montag, sonst 0.
29	cal_is_friday	Flag: 1 wenn Freitag, sonst 0.
30	cal_is_month_start	Flag: 1 wenn Monatsanfang, sonst 0.
31	cal_is_month_end	Flag: 1 wenn Monatsende, sonst 0.
32	hol_is_us_federal_holiday	Flag: 1 wenn US-Feiertag, sonst 0.
33	hol_is_day_before_us_federal_holiday	Flag: 1 wenn Tag vor US-Feiertag.
34	hol_is_day_after_us_federal_holiday	Flag: 1 wenn Tag nach US-Feiertag.

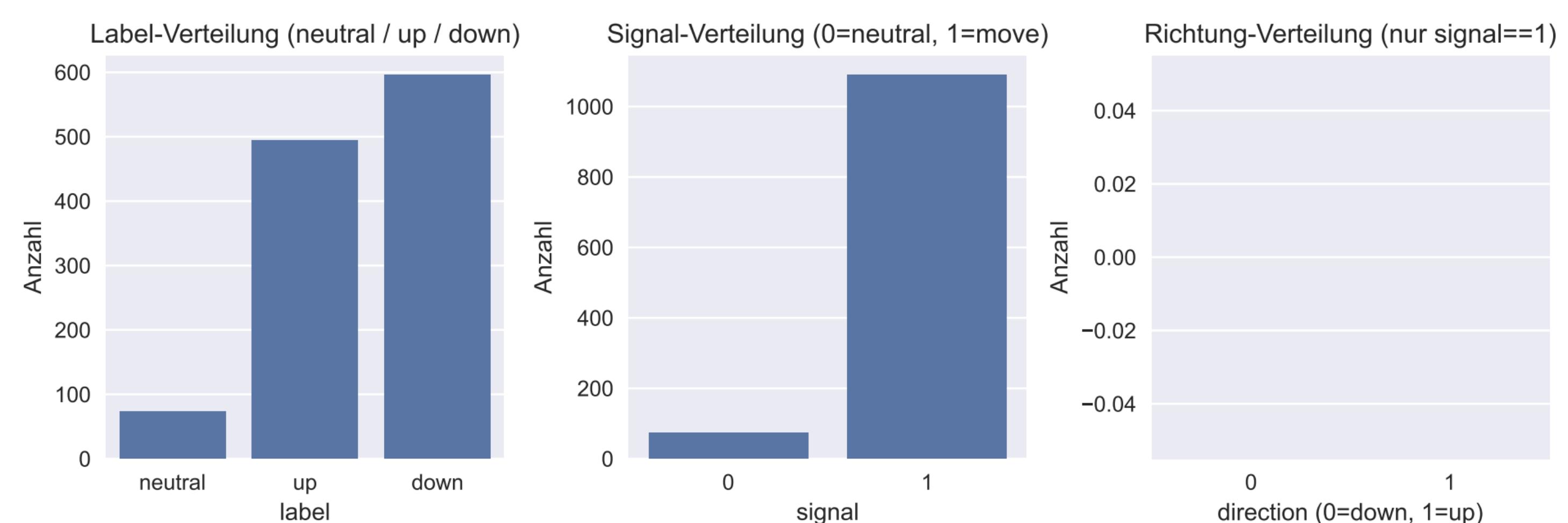
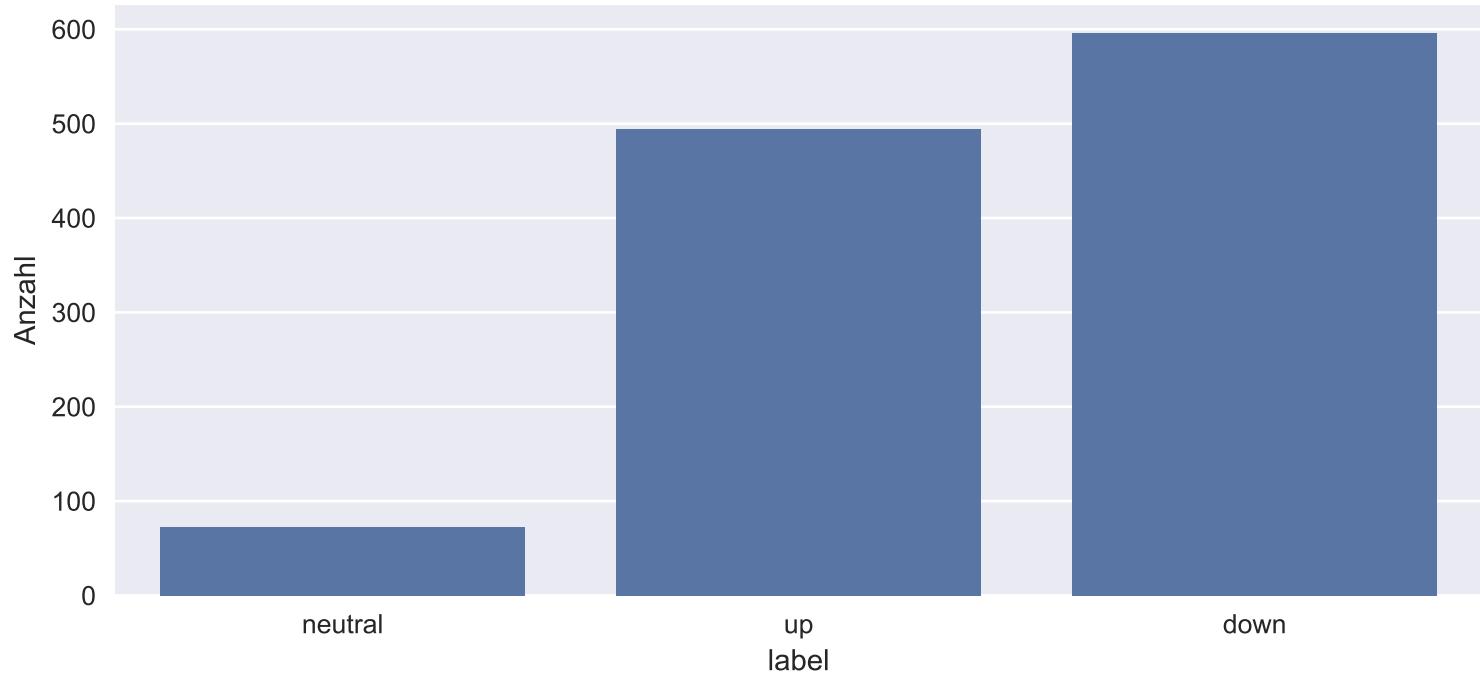


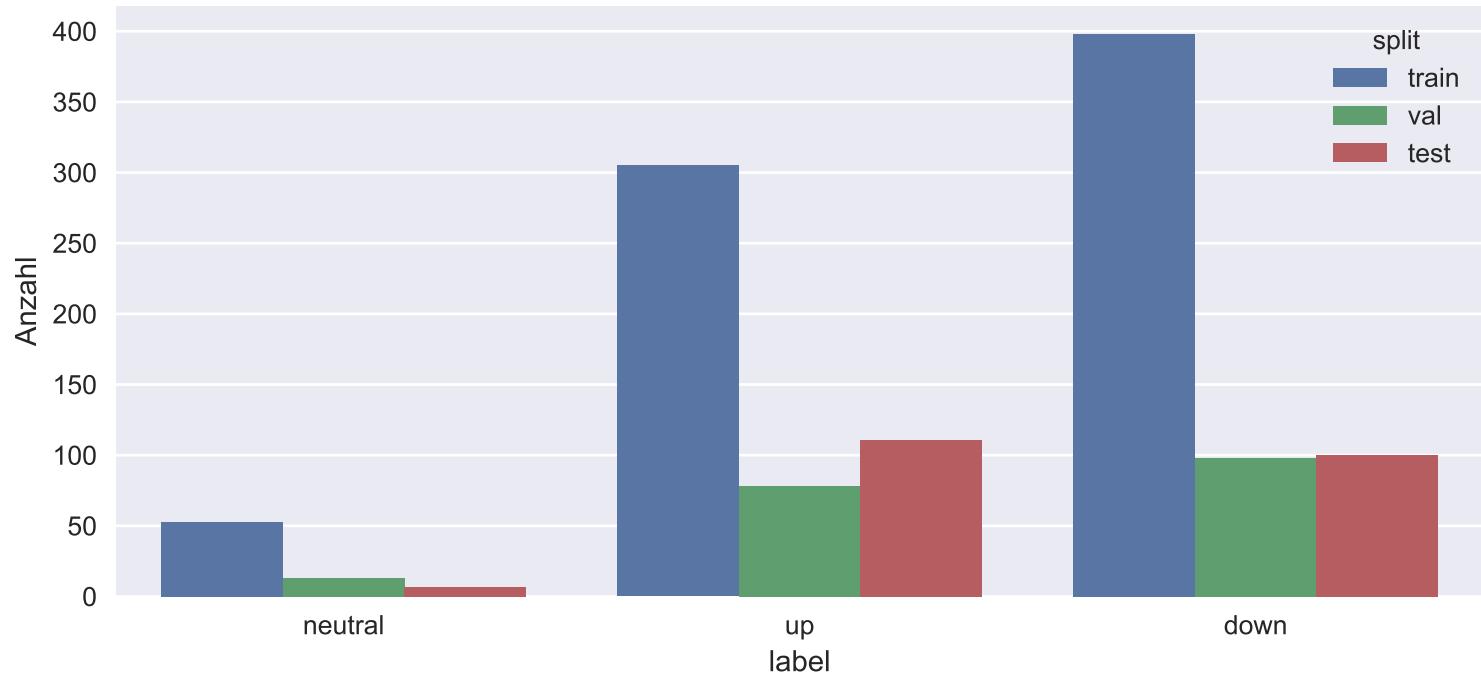
Abbildung: Klassenverteilungen für label, signal und direction im vollständigen Trainingsdatensatz.

Label-Verteilung – gesamter Datensatz



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

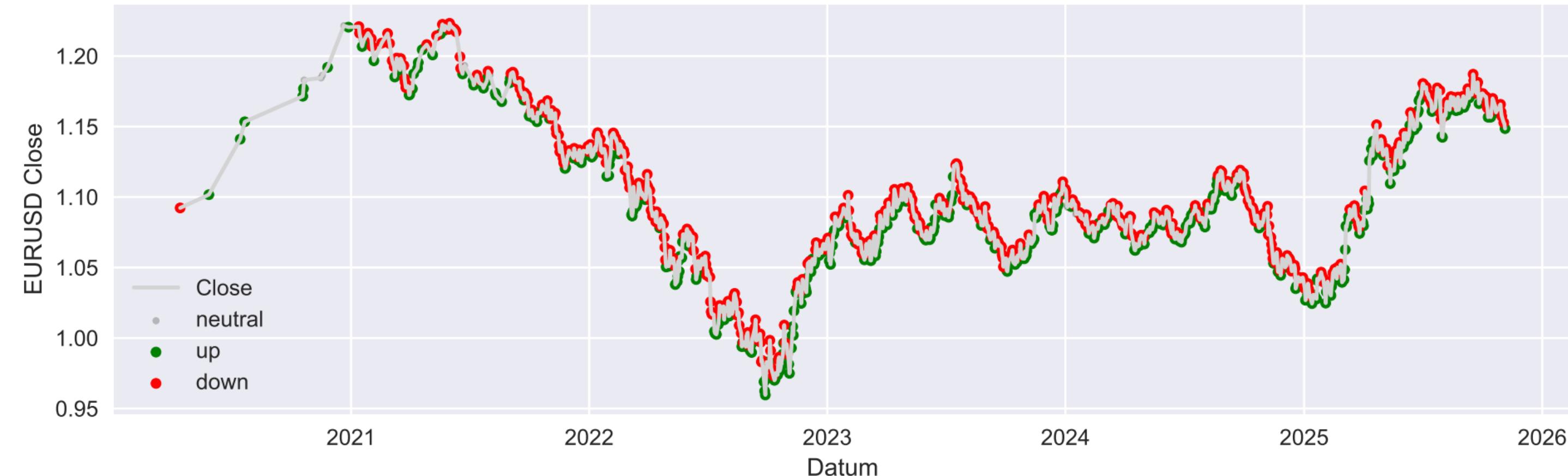
Label-Verteilung nach Splits (train/val/test)



split	neutral	up	down
train	53	305	398
val	13	78	98
test	7	111	100

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

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



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

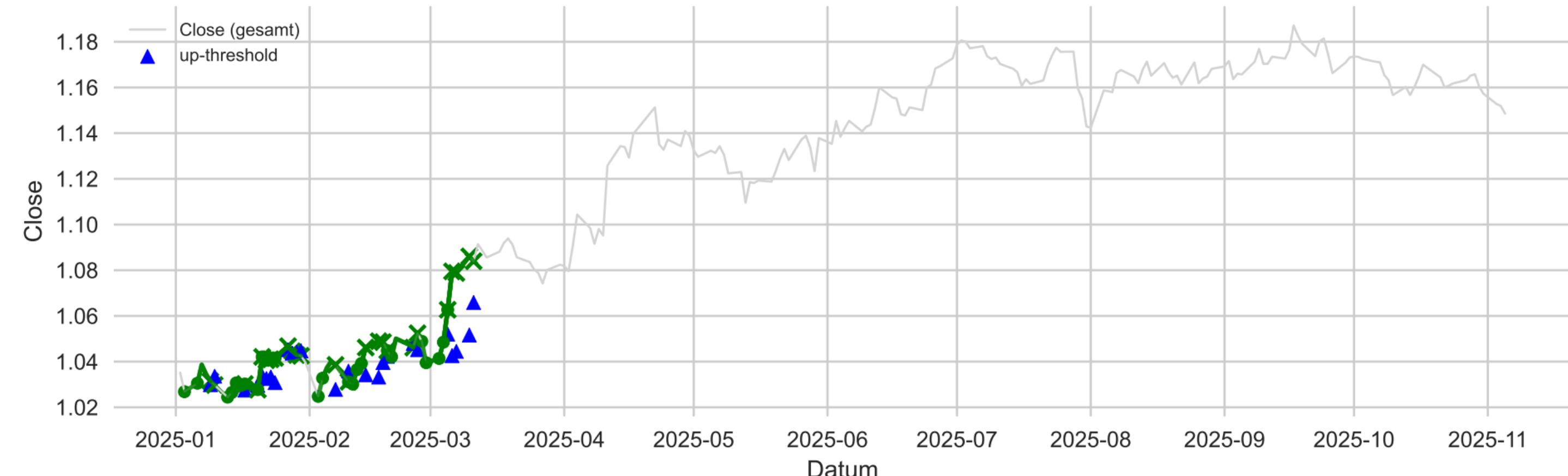


Abbildung: Preis-Segmente t..t+horizon für alle Testtage mit true label 'up'.

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

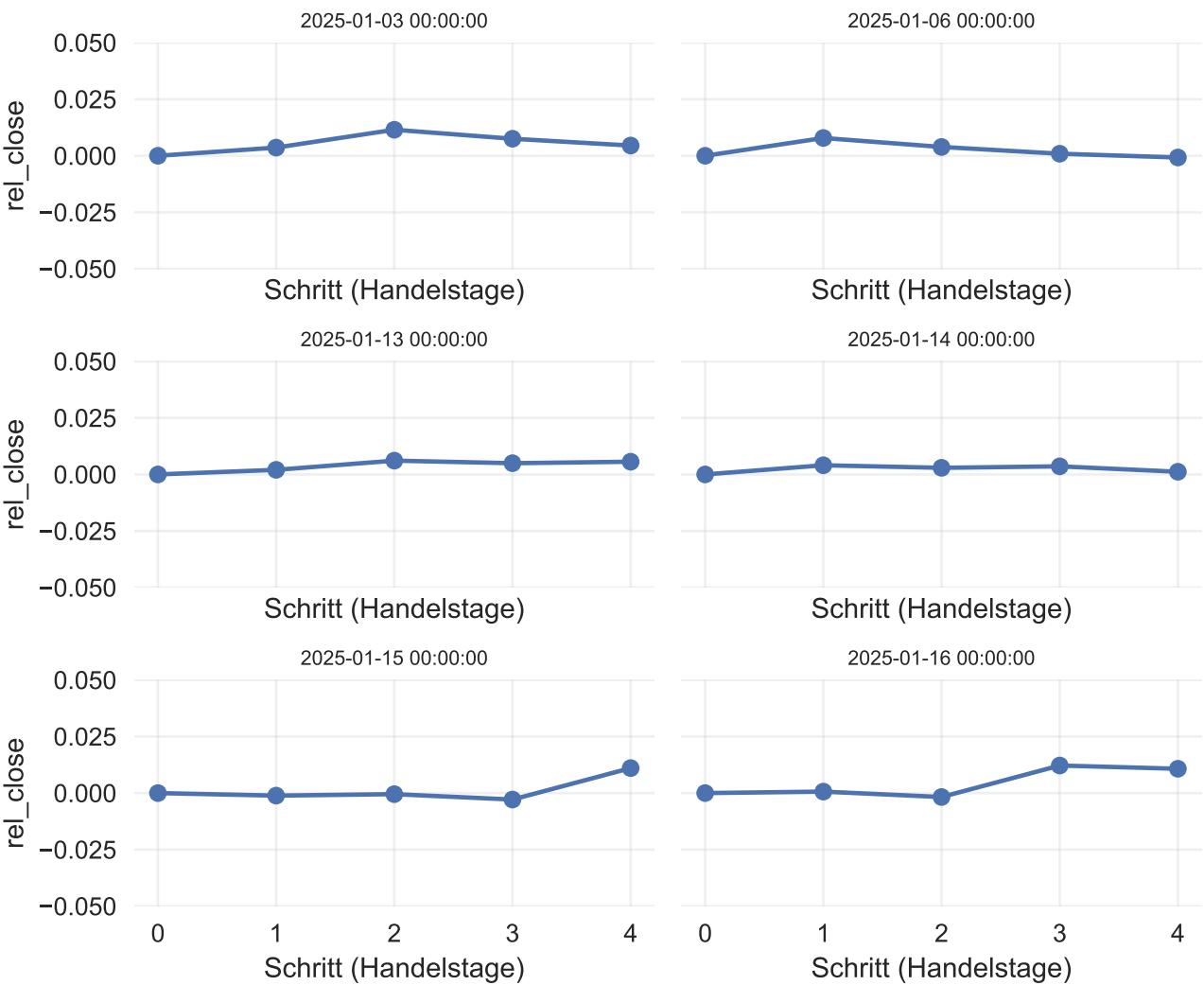


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

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

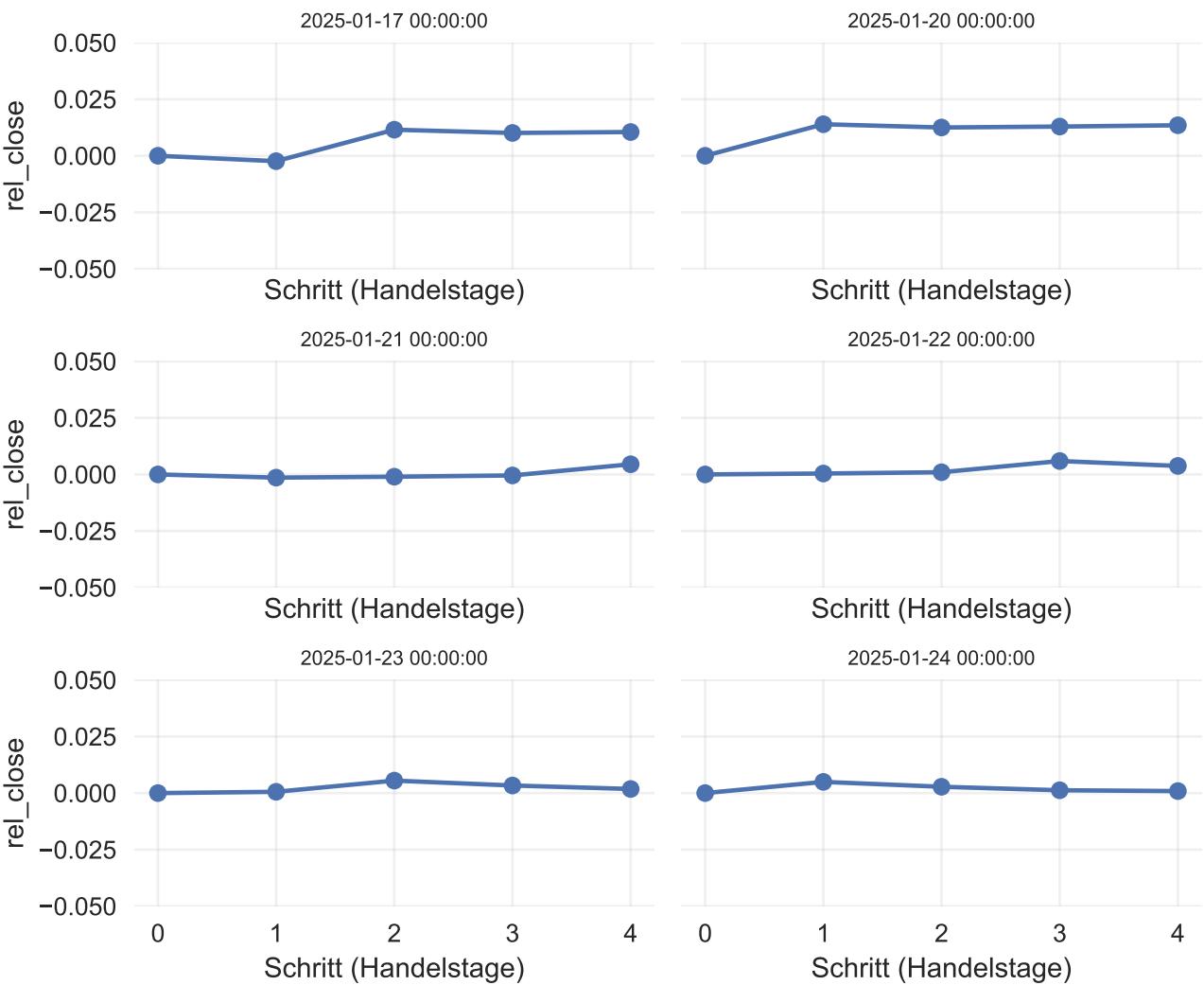


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

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

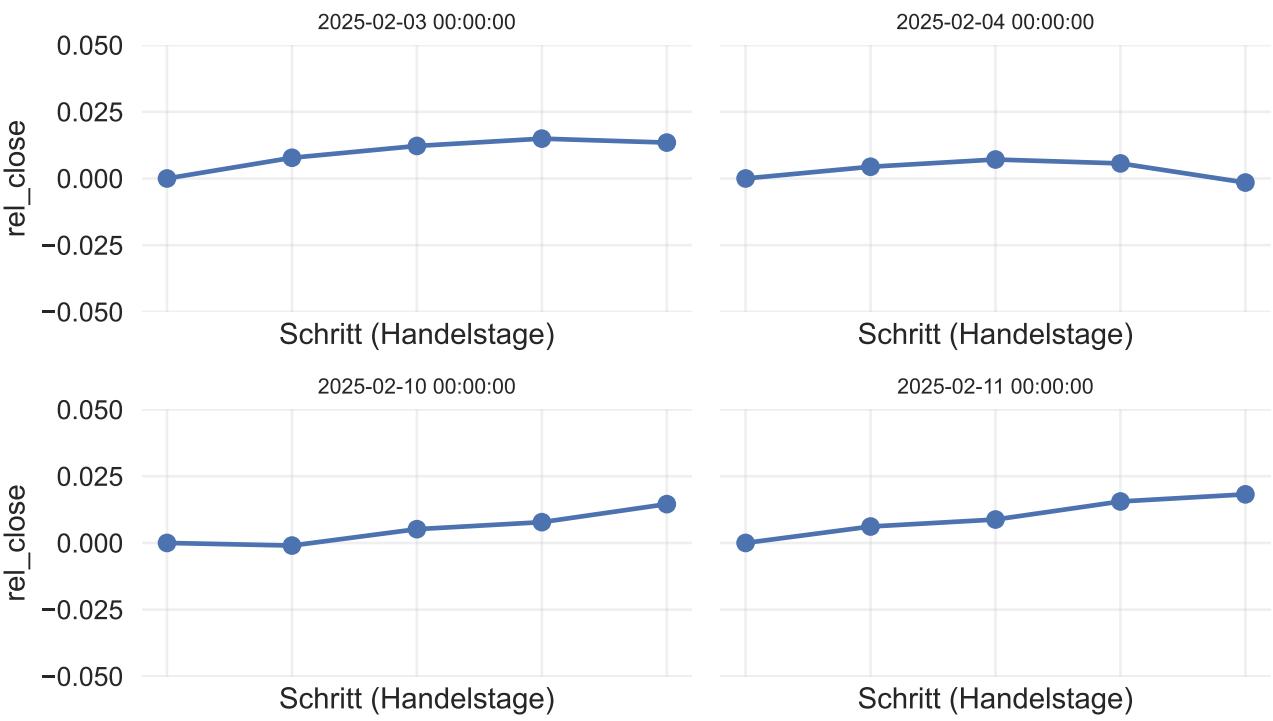


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'up'.

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

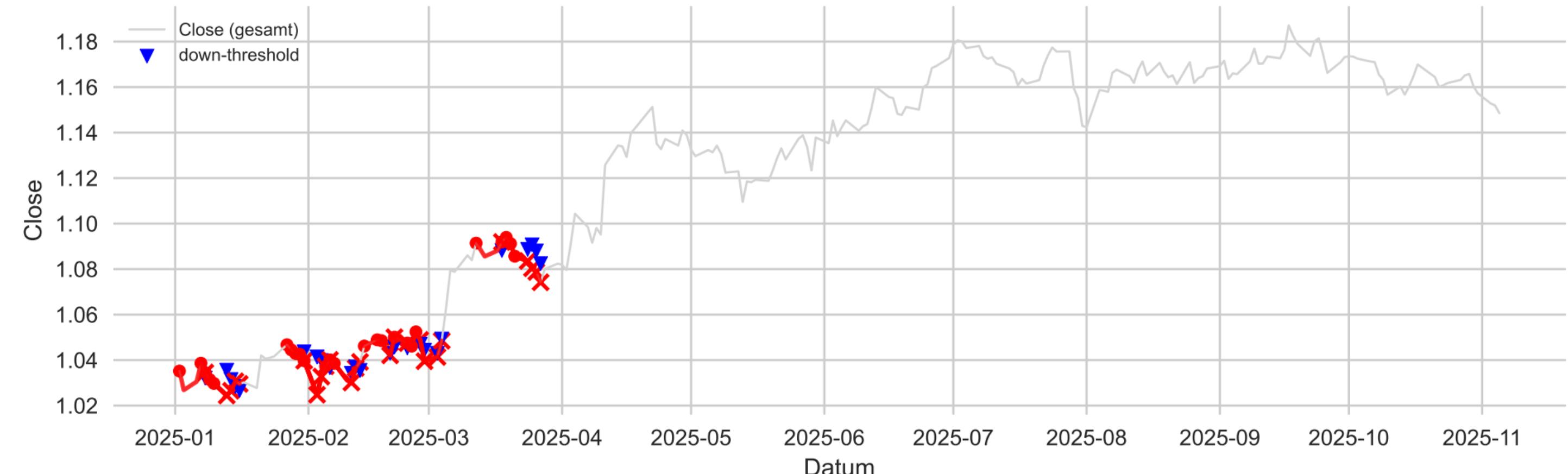


Abbildung: Preis-Segmente t..t+horizon für alle Testtage mit true label 'down'.

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

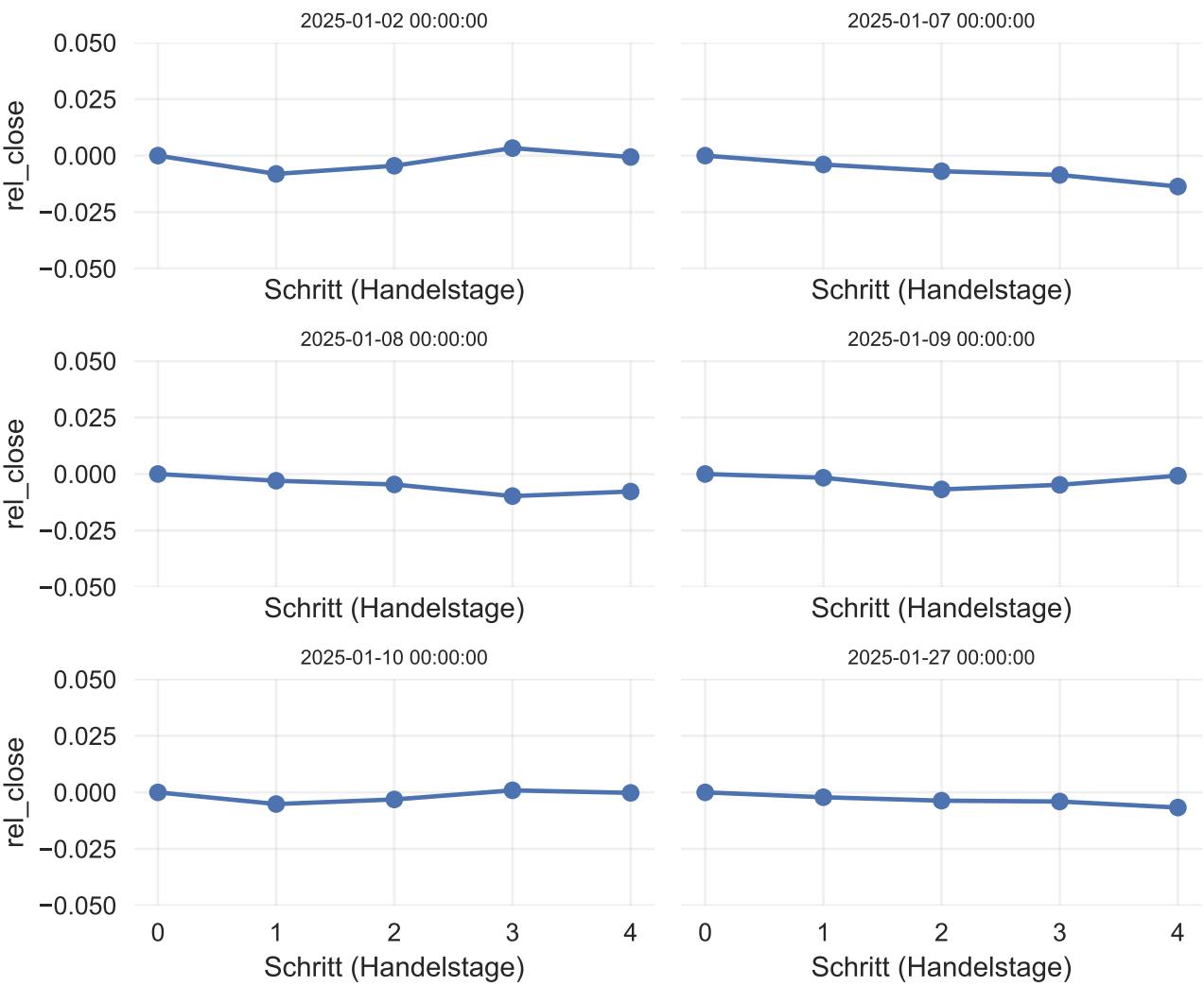


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

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

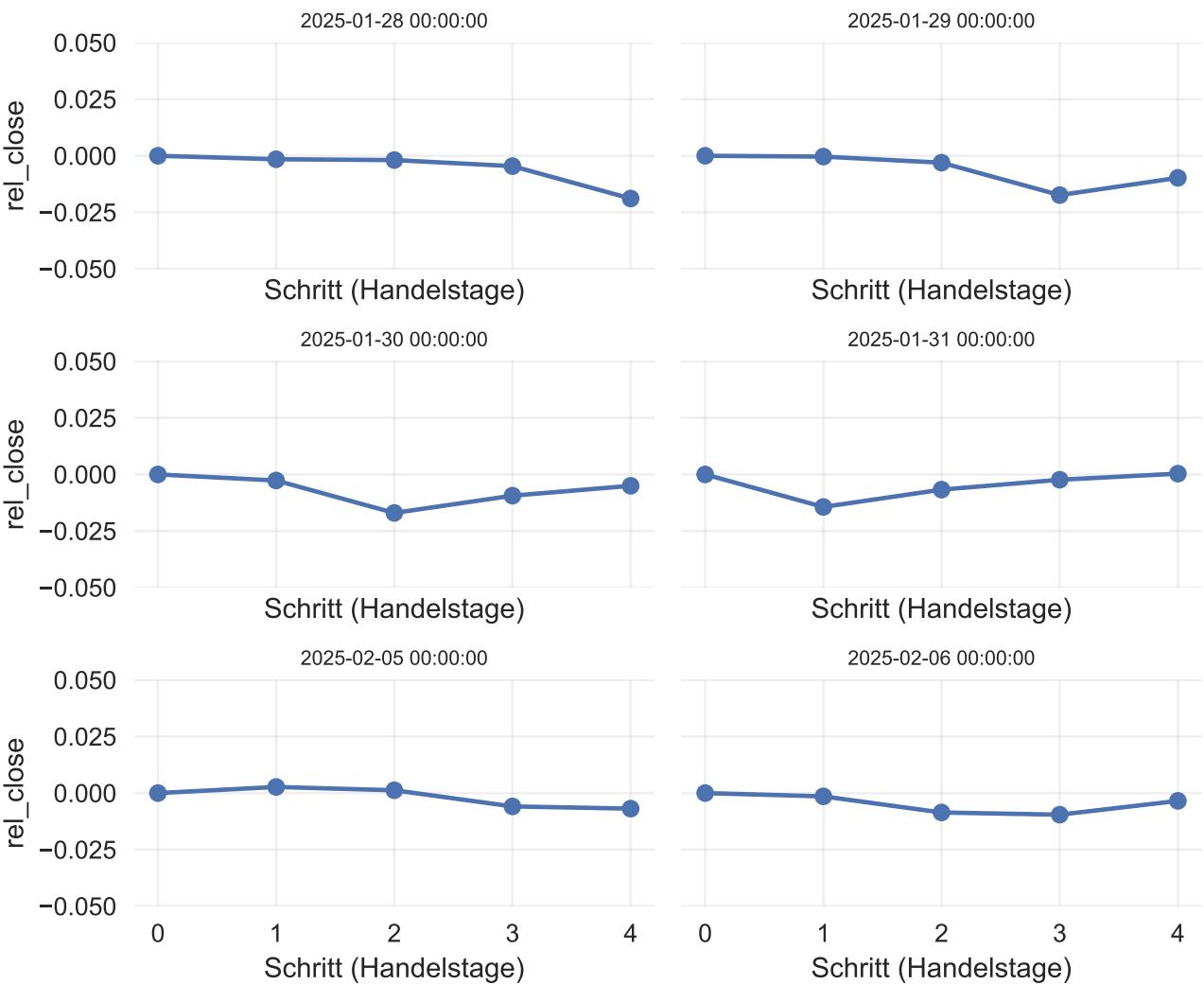


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

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

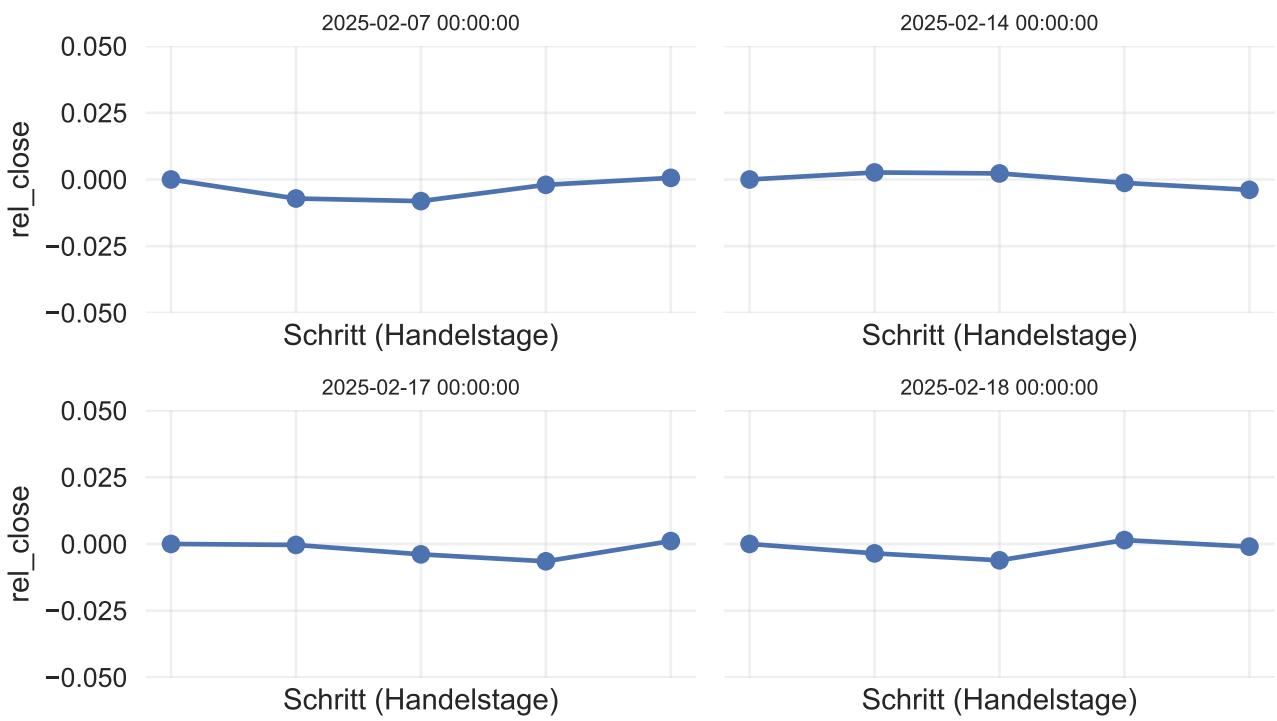


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'down'.

Signal-Modell – Kennzahlen für Klasse 'move' (train/val/test, thr=0.55)

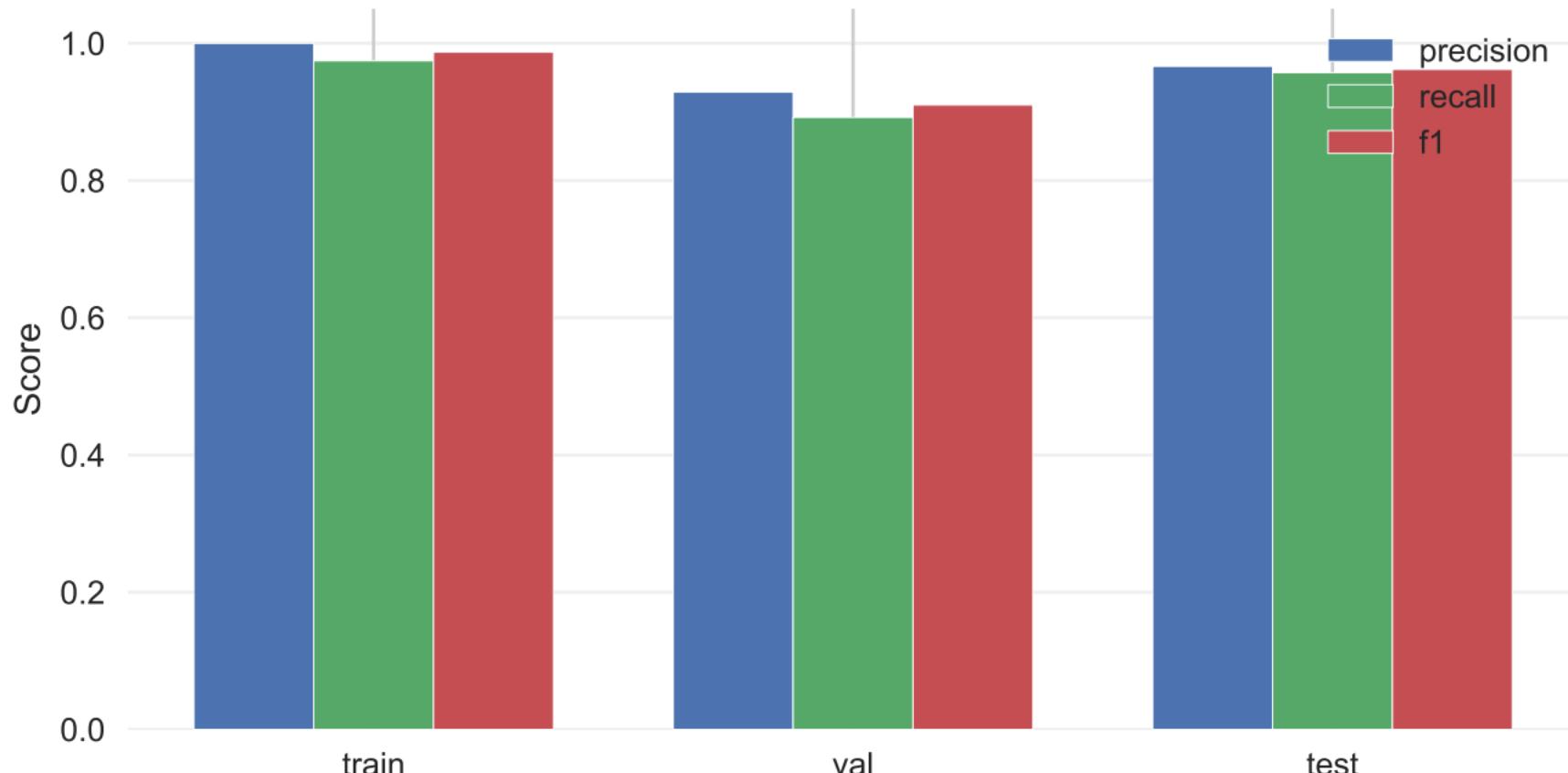


Abbildung: Precision, Recall und F1 der positiven Klasse je Split (train/val/test).

Signal-Modell – Tabelle (Klasse 'move', thr=0.55)

split	precision	recall	f1	support
train	1.000	0.974	0.987	703.000
val	0.929	0.892	0.910	176.000
test	0.967	0.957	0.962	211.000

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

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

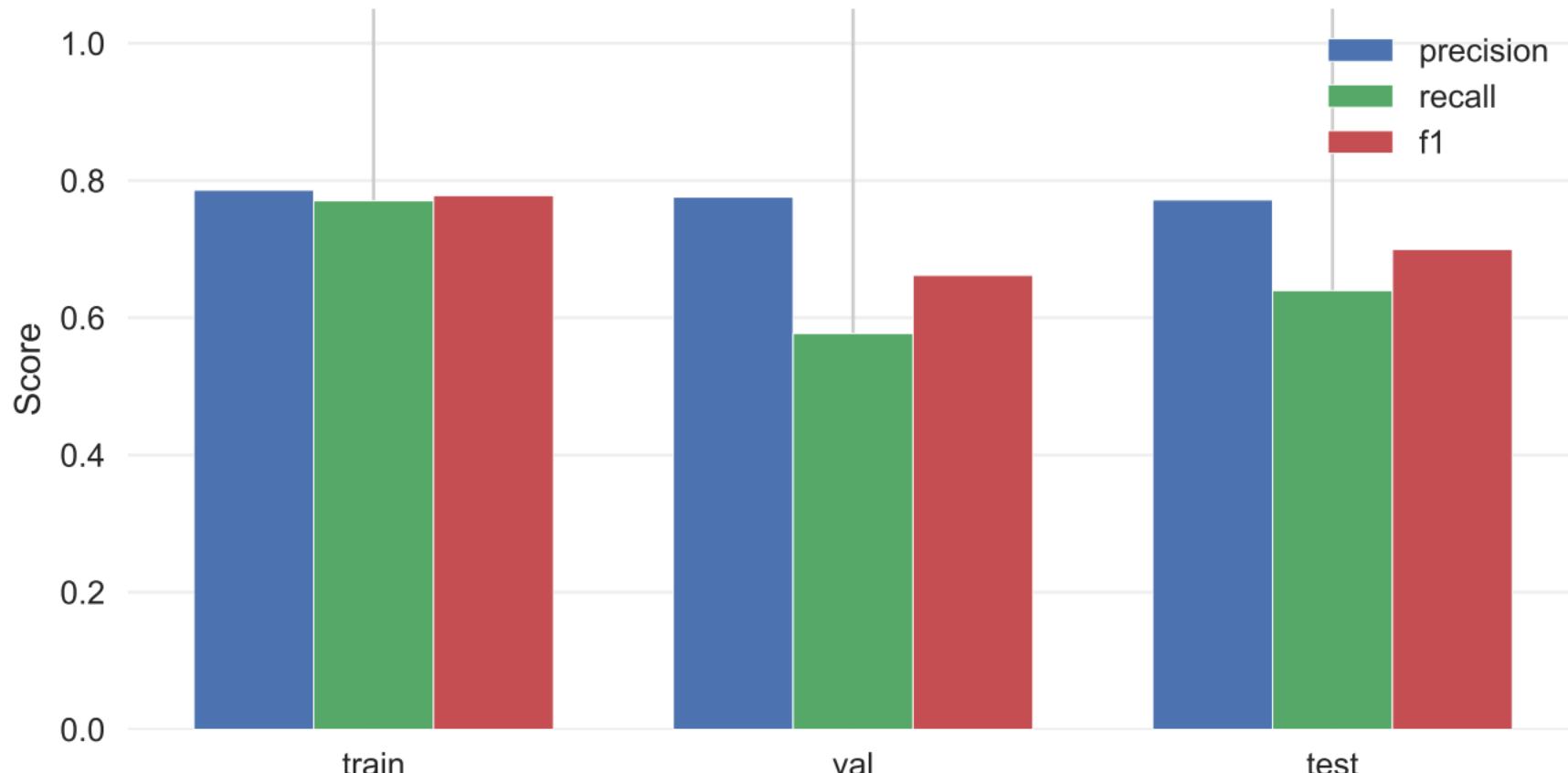


Abbildung: Precision, Recall und F1 der positiven Klasse je Split (train/val/test).

Richtungs-Modell – Tabelle (Klasse 'up')

split	precision	recall	f1	support
train	0.786	0.770	0.778	305.000
val	0.776	0.577	0.662	78.000
test	0.772	0.640	0.700	111.000

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

Kombinierte Test-Auswertung – neutral / up / down

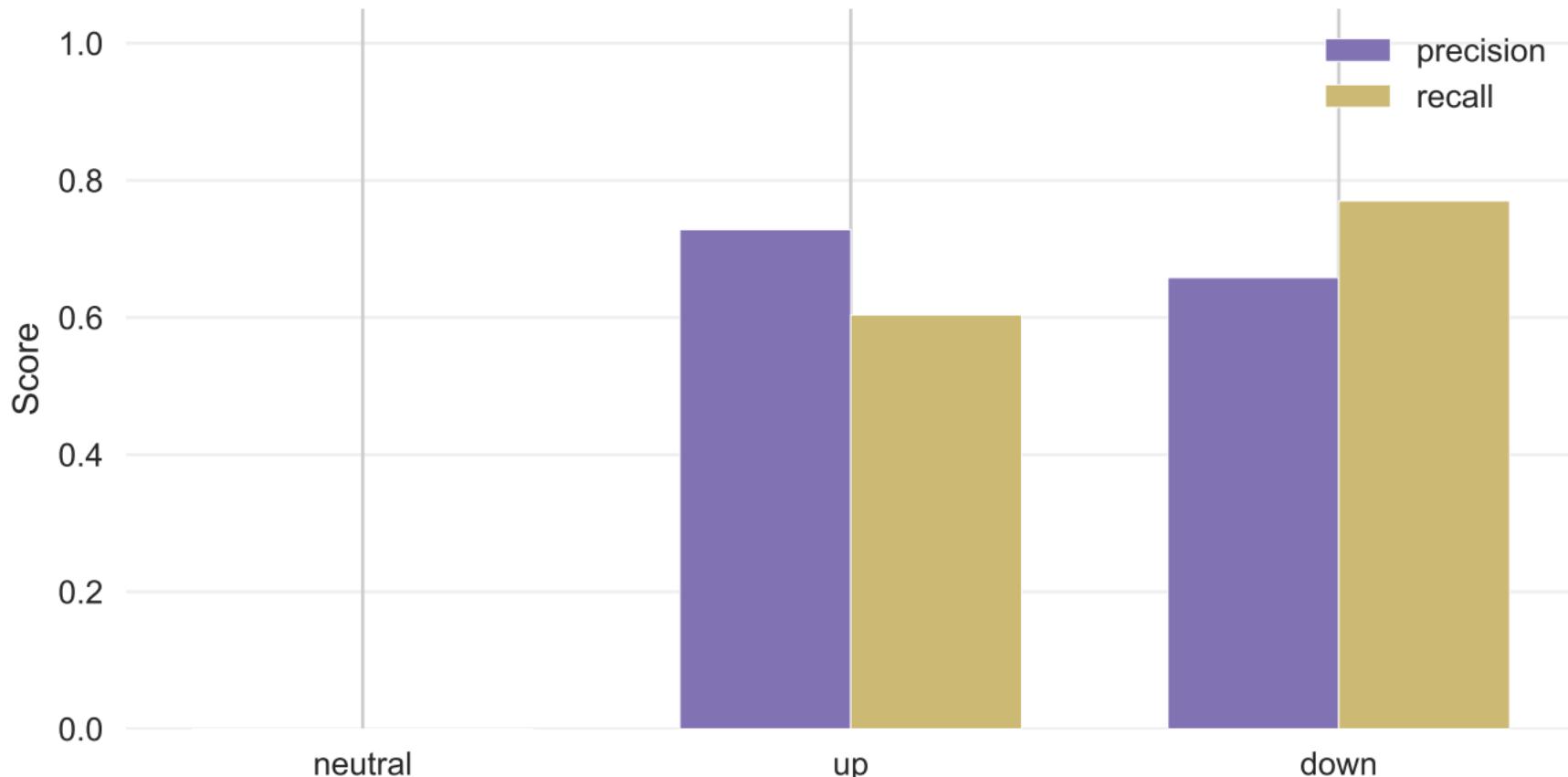


Abbildung: Precision und Recall der kombinierten 3-Klassen-Vorhersage (neutral/up/down) auf dem Test-Split.

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

klasse	precision	recall	f1	support
neutral	0.0	0.0	0.0	7
up	0.728	0.604	0.66	111
down	0.658	0.77	0.71	100

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

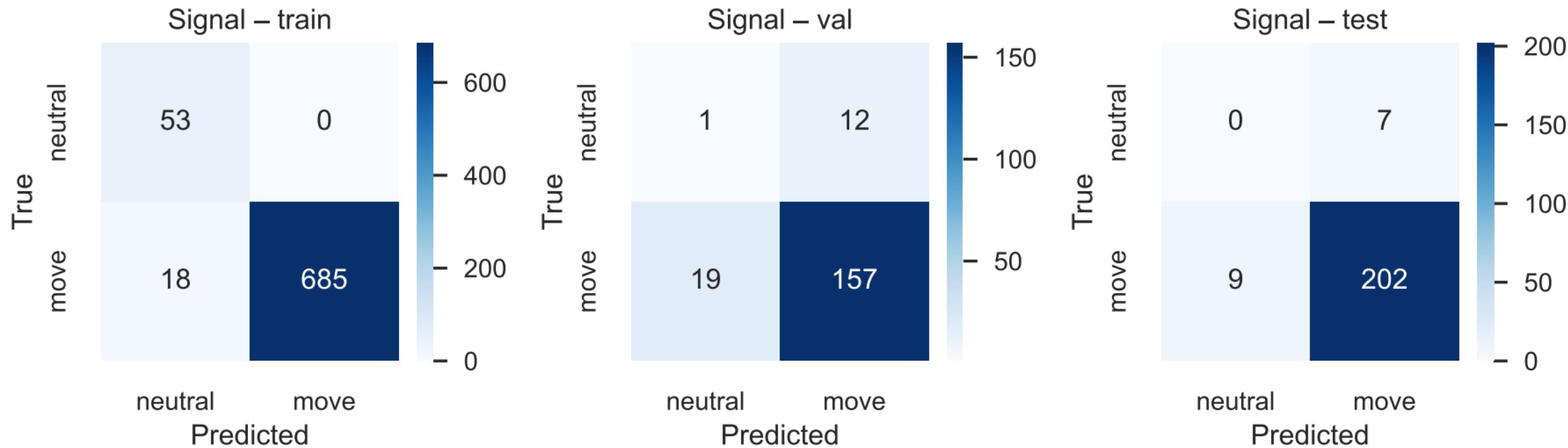


Abbildung: Confusion-Matrizen des Signal-Modells (neutral vs move) für Train-, Validierungs- und Test-Split.

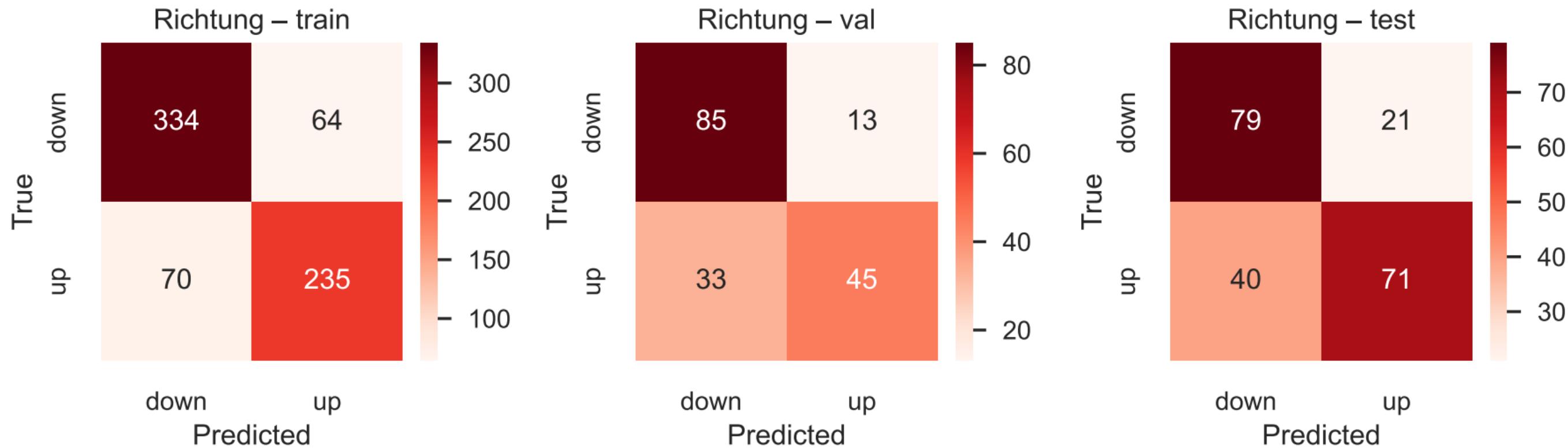


Abbildung: Confusion-Matrizen des Richtungs-Modells (down vs up) für Train-, Validierungs- und Test-Split.

Confusion Matrix – Test (neutral / up / down)

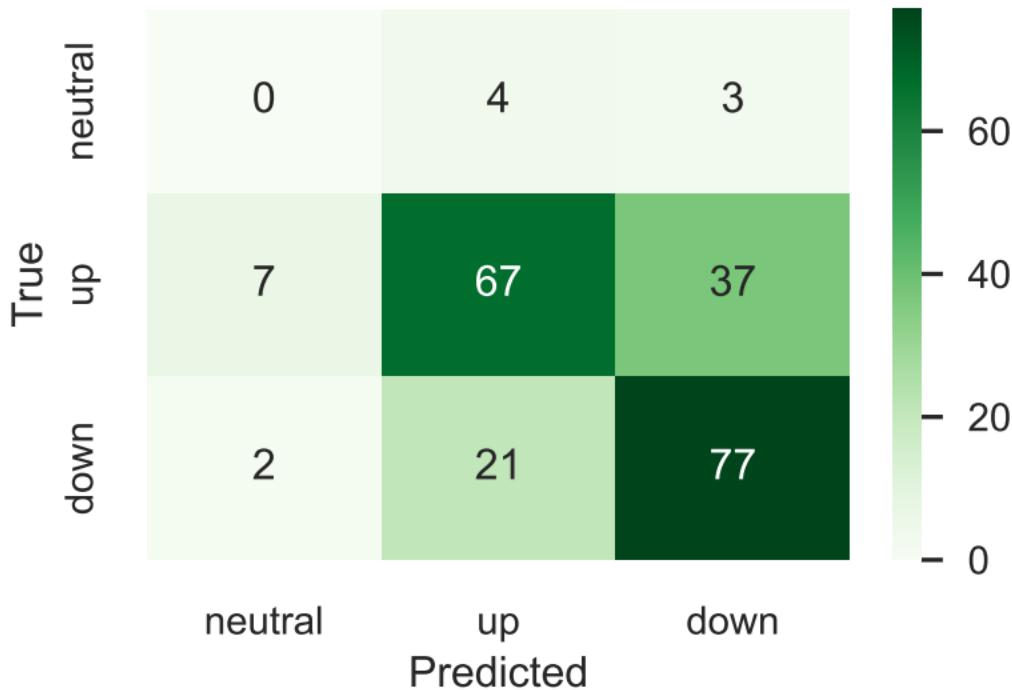


Abbildung: Confusion-Matrix des kombinierten Modells (neutral/up/down) auf dem Test-Split.

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

modell	split	TN	FP	FN	TP
signal	train	53	0	18	685
signal	val	1	12	19	157
signal	test	0	7	9	202
direction	train	334	64	70	235
direction	val	85	13	33	45
direction	test	79	21	40	71

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	25	neutral:4, down:21
combined	down	40	neutral:3, up:37
signal	move	7	neutral:7

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')

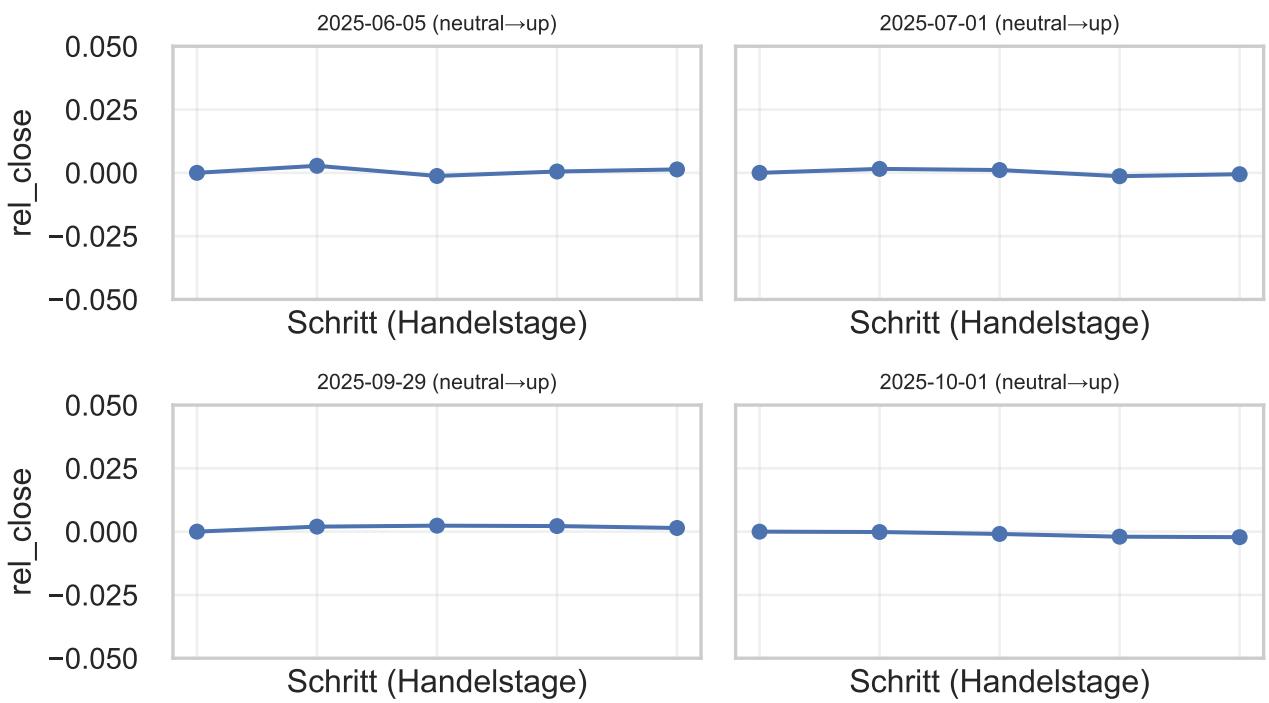


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'neutral', die im kombinierten Test als 'up' klassifiziert wurden.

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

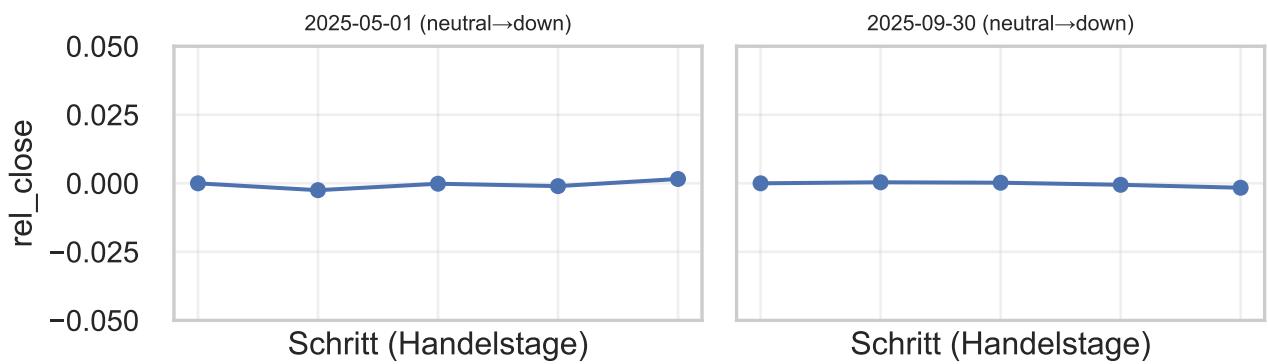


Abbildung: Relativer Verlauf der Close-Preise für alle Testtage mit true label 'neutral', die im kombinierten Test als 'down' klassifiziert wurden.

Feature Importance – Signal-Modell

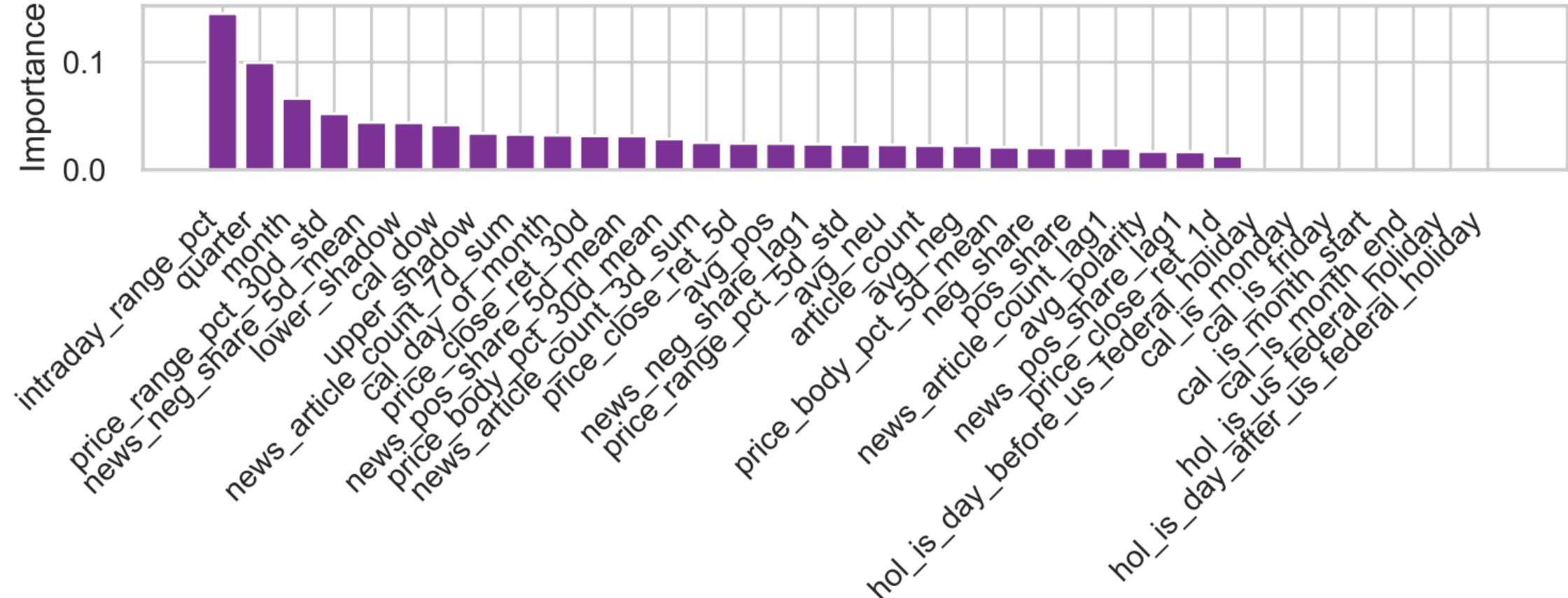


Abbildung: Wichtigkeit der Features für das Signal-Modell (neutral vs move).

Feature Importance – Richtungs-Modell

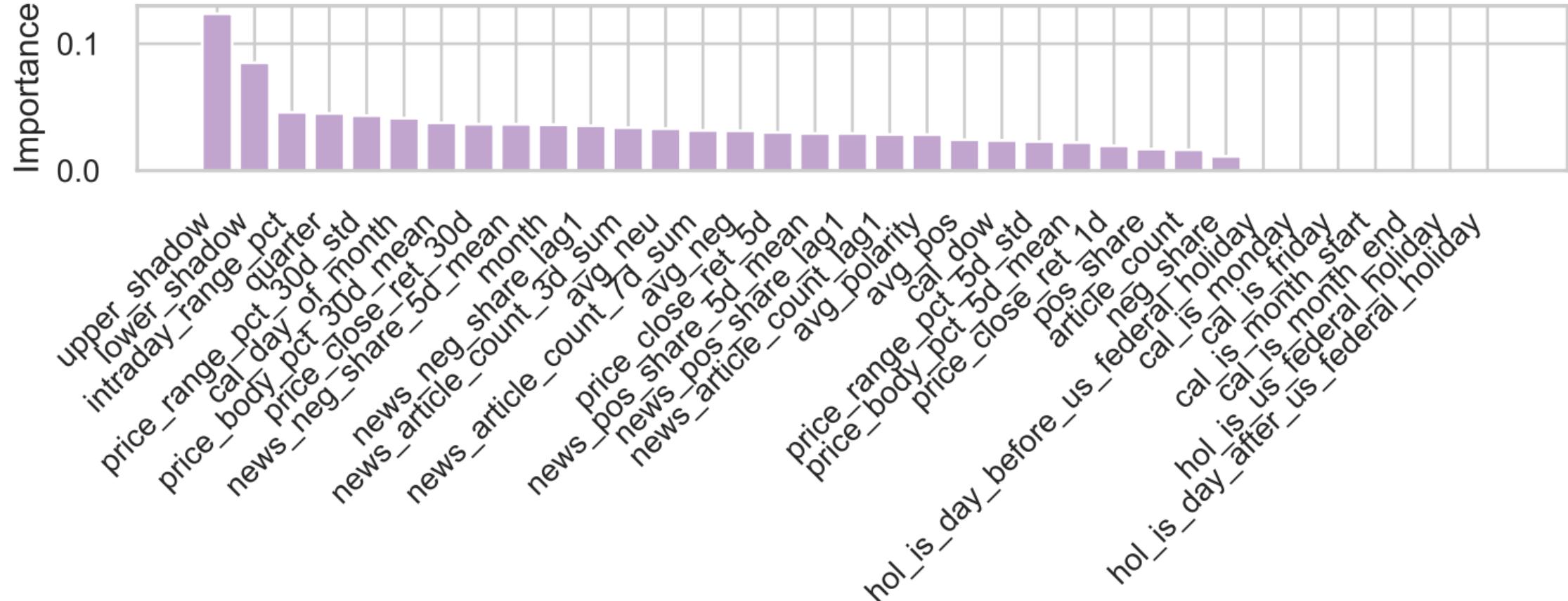


Abbildung: Wichtigkeit der Features für das Richtungs-Modell (down vs up).