



Stacked Market Ensemble

A classifier that learns, match by match, how much of your model and how much of the market to quote.

Multinomial logistic regression over the model's and the market's log-odds and their disagreement, fitted walk-forward on an expanding window. It closes almost the whole gap between a public model and a sharp market without ever seeing a result before it prices it.

SPORT

Football

FAMILY

The machine layer

LICENCE

£599

one-off, perpetual

RUNNING IN THE PLATFORM

<https://edgewortharena.com/arena/api/benchmark/providers>

Inputs

NAME	TYPE	UNITS	WHERE IT COMES FROM
model_probabilities	dict per match	probability	any match model, including this platform's
market_probabilities	dict per match	probability	quoted prices with the overround removed
results	list	-	only used for matches already played
warmup, refit_every, l2 (optional)	ints and float	-	your choice; 1,500 / 250 / 1.0 are what the validation used

Outputs

NAME	TYPE	UNITS	NOTES
stacked probabilities	dict per match	probability	
attention	dict	share of weight	how much of the answer came from the model, the market and their gap
walk-forward scores	dict	Brier, nats	beside the model's and the market's on the same matches

Method

- 1 Turn each side's probabilities into log-odds, and add their difference as a third feature per outcome: the model, the market, and the disagreement.

THIS IS THE SAMPLE

The method runs to 4 steps; one is shown. The assumptions, the limits, the module layout and the reference implementation come with the licence. Everything on the page below this -- what the model was measured at -- is the same in both versions, because that is the part you should not have to take on trust.



What it was measured at

Walk-forward on an expanding window: every match priced by a fit that had not seen it, refitted every 250 matches.

MEASURE	RESULT	READING
Matches priced out of sample	7,726	
Brier, stacked	0.5845	lower is better
Brier, the model alone	0.5968	
Brier, the market	0.5841	the benchmark
Gain over the model	+2.07%	Brier skill
Gain over the market	-0.07%	negative: it does not beat the market, and is not sold as doing so

Assumptions

What has to be true for the numbers to mean what they say.

- Both inputs are honest probabilities for the same match at the same moment; a stale market price makes the disagreement feature meaningless.

Limits

Where it is known to be wrong. Published because a model without stated limits is one nobody has pushed on.

- It does not beat the market, and the comparison that matters is the strict one. Fed only prices that exist before the close, on 7,097 walk-forward matches it scored 0.5872 Brier against our own model's 0.5972 and the closing line's 0.5850: it recovers about 1.7% of the model's error and stays 0.4% behind the close.

LICENCE

Issued to Evaluation copy. One perpetual licence to use and modify this model inside your own organisation. Redistribution of the specification or the reference implementation is not included. No warranty is given: the validation above is the whole of the claim being made.