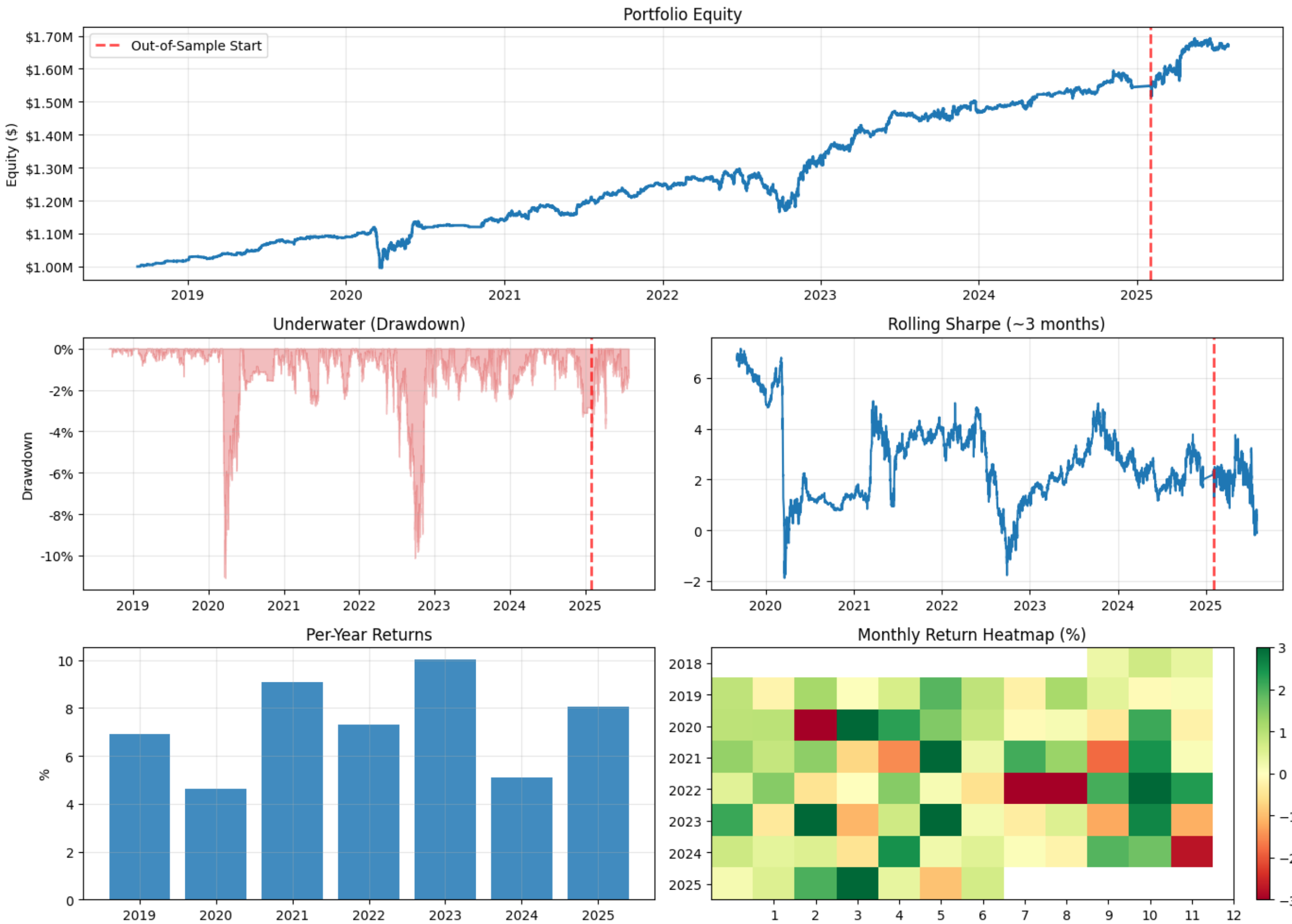
Systematic FX Alpha from State-Space Neural Networks

Levente Szabo — Quant Research & AI Engineering — levbszabo@gmail.com • levbszabo.github.io

Executive Summary. We built a model that learns the "state" of the FX market from hourly data across seven major pairs. Instead of making one-point predictions, it forecasts the full distribution of returns at two horizons: one day ahead (24h) and one week ahead (168h). We only trade when the forecast looks strong relative to its own uncertainty (z-score gating), and we size our positions so risk stays balanced. The result is a USD-neutral portfolio that remains profitable out-of-sample, even after accounting for realistic trading costs.



Backtest equity across EURUSD, USDCHF and other majors remains profitable in- and out-of-sample, with controlled drawdowns

Performance (net, costs included)

	IS (2019–2024)	OOS (2025 YTD)
CAGR	7.16%	15.87%
Ann. Vol	12.84%	8.47%
Sharpe	2.20	1.75
Sortino	2.58	2.36
Max DD	-11.09%	-3.85%
Calmar	0.65	4.12
Trades	1,071	129
Win Rate	51.3%	52.7%

Technical Outline. Our core model is a recurrent state-space neural network implemented in `PyTorch`, trained on GPUs using hourly data for the seven major FX pairs from 2018–2025 (with February 2025 onwards held out as true out-of-sample). Data is structured into 512h sequences, letting the network learn momentum, volatility regimes, and cross-currency interactions. At each step t, raw features x_t are encoded, and a deterministic recurrent update

$$h_t = f_\theta(h_{t-1}, \mathrm{Enc}_\phi(x_t))$$

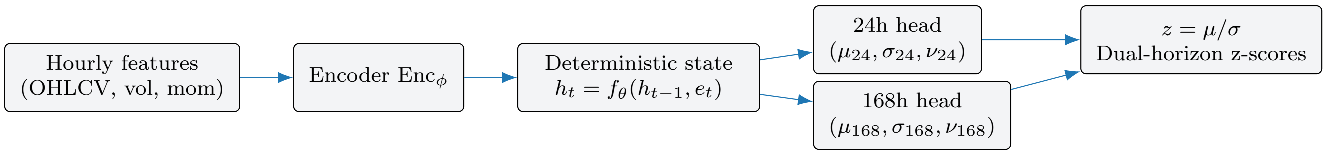
propagates into a hidden market state h_t. From h_t the model predicts distributional forward returns at 24h and 168h horizons for each asset a:

$$r^{(a)}_{t+H} \sim \text{Student-}t\big(\mu_{H,a}, \sigma_{H,a}, \nu_{H,a}\big).$$

Where μ is the expected log return, σ is the scale (uncertainty), and v the degrees of freedom capturing tail risk. We convert each forecast into a standardized confidence score

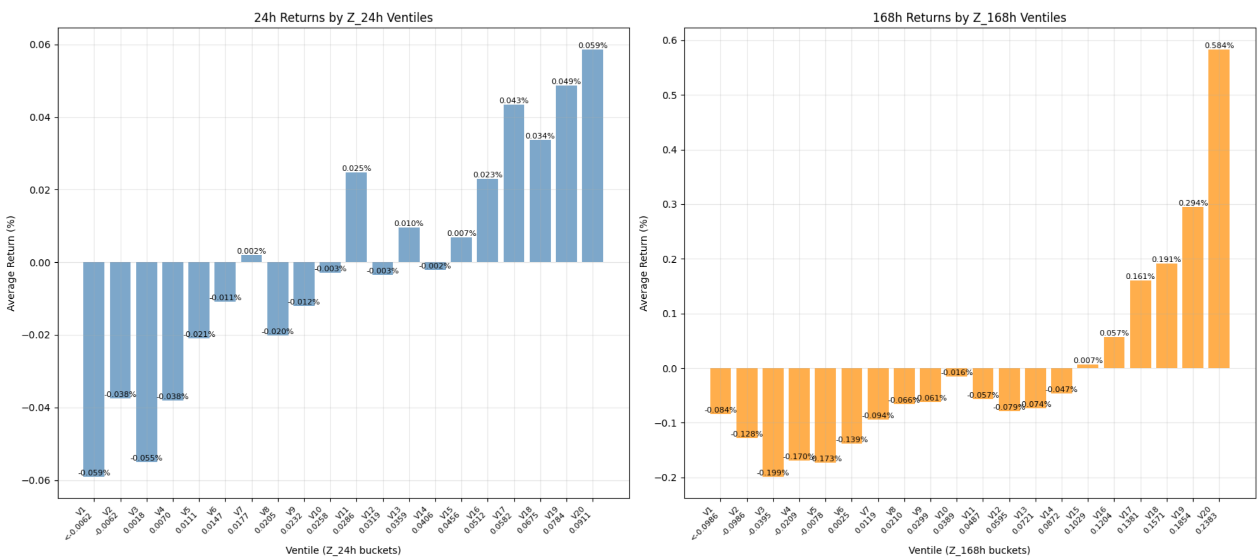
$$z_{H,a} = \mu_{H,a}/\sigma_{H,a},$$

which measures the predicted return relative to its uncertainty. These z-scores drive the trading system, which gates entries, sizes positions, and controls risk accordingly.



Model pipeline. The model processes hourly features (prices, volatility, momentum, etc.) through an encoder and maintains a hidden "market state" that evolves over time. From this state, it predicts the likely distribution of returns for each pair at 24h and 168h horizons, including both expected return and its uncertainty (fat-tailed, via a Student-t distribution).

Signal Calibration. To make sure the forecasts actually line up with reality, we bucketed predictions by their confidence level (z-scores) and checked average outcomes. For the weekly horizon, the relationship is clean: high z-scores lead to higher realized returns, and negative z-scores line up with losses. The daily horizon is noisier, but still adds value as a timing filter. This ventile analysis gave us the confidence to set cutoffs around the top/bottom deciles for the 168h forecasts.



24h and 168h ventile plots. Indicates a clear monotonic relationship between z-score and realized returns

Cutoff Selection. We don't trade on every forecast. Instead, we require the weekly model to be in the strongest 10–20% of signals before entering. When the daily and weekly forecasts agree at an extreme, we size up; when they disagree, we cut the trade size. This keeps turnover low and avoids over-trading in noisy regimes.

Risk Management Summary.

Positions are sized dynamically using volatility targeting, so higher-confidence signals (large $|z|$) scale exposure while keeping portfolio risk within bounds. The book is always USD-neutral, balancing long and short USD exposures across pairs, with both per-pair and portfolio-level caps to prevent concentration.

Exits are not fixed percentages; they scale with the model's predicted uncertainty. Profit-taking, stop-loss, and trailing levels are expressed as multiples of σ from the 168h forecast, which allows the system to adapt naturally between calm and volatile regimes. During stress periods, a *drawdown brake* automatically halves gross exposure after a 7% portfolio drawdown, only restoring risk once losses have partially recovered.

Disclaimer: Results are based solely on backtested simulations. This research demonstrates a systematic edge that we are preparing to validate through walk-forward and live paper trading