1) Shiller/Campbell–Shiller backbone (valuation → expected returns)

Let $p_t = \log P_t$ and $d_t = \log D_t$ where D_t is a "cash-flow" proxy. In equities, D_t is dividends/earnings; in FX we stand in a constant daily "carry" (interest differential) $D_t \equiv \bar{D}$ to emulate a dividend-price ratio. Define the log dividend-price ratio

$$dp_t \;=\; d_t - p_t \;=\; \log \frac{D_t}{P_t}.$$

Campbell–Shiller's log-linear identity (first-order) connects valuation to future returns:

$$p_t - d_t \;\approx\; \kappa + \rho\,(p_{t+1} - d_{t+1}) \;+\; \Delta d_{t+1} \;-\; r_{t+1},$$

hence in expectations,

$$\mathrm{E}_t[r_{t+1}] \;\approx\; \kappa' \;-\; (p_t - d_t) \;+\; \rho\,\mathrm{E}_t[p_{t+1} - d_{t+1}] \;+\; \mathrm{E}_t[\Delta d_{t+1}].$$

Holding the last two expectation terms "slow" (or absorbed into features), the key monotone implication is:

$$\mathrm{E}_t[r_{t\to t+K}] \text{ increases with } dp_t$$

i.e., higher dp_t (cheap price vs. "cash-flow") predicts higher future returns on horizon K.

In the code, D_t is a constant daily carry $\bar{D} = 0.015/252$, so $dp_t = \log(\bar{D}/P_t)$. To stabilize scales and mimic stationarity, the script z-scores dp_t over a rolling window W_z:

$$zdp_t \;=\; \frac{dp_t - \mu_t(dp)}{\sigma_t(dp)}.$$

2) Feature construction (signals driving the learner)

The model ingests a compact vector

$$\mathbf{x}_t \;=\; \left[zdp_t,\; m_{5,t},\; m_{20,t},\; \sigma_{20,t},\; \mathrm{RSI}_{14,t}/100,\; \frac{\mathrm{ATR}_{20,t}}{P_t},\; r_{t,t-1},\; r_{t,t-10}\right],$$

with

- $m_{k,t} = \log(P_t/P_{t-k})$ (5/20-bar momentum),
- $\sigma_{20,t} = \mathrm{StdDev}\big(\log(P_\tau/P_{\tau-1})\big)_{\tau=t-19}^{t}$ (realized vol),
- ATR_{20}/P_t (scale-free range proxy),
- $r_{t,t-1}$ and $r_{t,t-10}$ (1- and 10-bar log returns).

The **z-scored valuation** zdp_t is the Shiller anchor; other terms capture short- to medium-term dynamics to disambiguate regimes where valuation alone is slow-moving.

3) Learner, scores, and "edge"

Zorro's built-in **PERCEPTRON + BALANCED** is a linear neural unit (a one-layer NN). It produces two directional scores:

$$s_L(t) \ = \ \mathbf{w}_L^\top \mathbf{x}_t, \qquad s_S(t) \ = \ \mathbf{w}_S^\top \mathbf{x}_t,$$

trained internally on class balance (the exact internal target is Zorro-specific, but conceptually aligned with directional profitability).

Define the **edge**

$$e_t \ = \ s_L(t) - s_S(t).$$

A practical probabilistic reading is $p_t \approx \sigma(\alpha\, e_t)$ with $\sigma(u) = 1/(1 + e^{-u})$, so larger e_t tilts odds to long and vice-versa.

Decision rule used in the script:

$$\text{go long if } e_t > \tau, \quad \text{go short if } e_t < -\tau, \quad \text{else flat},$$

with a small $\tau = \mathrm{EdgeMin}$ to ensure activity. The code also carries a minimal fallback using RSI_{14} and $m_{5,t}$ when e_t is close to zero.

4) How Shiller enters the NN decisively

- The raw valuation **level** $dp_t = \log(\bar{D}/P_t)$ rises when price P_t falls (asset becomes "cheaper").
- The **standardized** zdp_t places that cheap/expensive signal on a stable scale, preventing it from being dwarfed by momentum/volatility magnitudes.
- During training, the perceptron estimates $\mathbf{w}_L, \mathbf{w}_S$ so that positive weight on zdp_t rewards states where elevated dp_t historically preceded higher forward returns, in line with Campbell–Shiller.
- Momentum/volatility features let the linear NN modulate the valuation effect across regimes (for instance, discounting zdp_t during sharp negative momentum and amplifying it when volatility/RSI states are supportive).

5) Optional links to the variance-bound idea

Shiller's variance-bound logic compares the variance of actual prices P_t with that of discounted cash-flow values $P_t^* = \sum_k \beta^k D_{t+k}$. In the script's didactic setting, with a near-constant "cash-flow" proxy D_t, the discounted sum becomes smooth, illustrating that **price swings can exceed changes justified by fundamentals**. Using $dp_t = \log(D_t/P_t)$ as a feature is a compact way to inject that valuation gap into the learner, letting the NN learn when those gaps tended to close via future returns.