There exists a collection of risky assets in the portfolio $\mathcal{P}$. An oracle provides the current price $p_i \in \mathbb{R}_+$ for each asset $i \in \mathcal{P}$, and a binary action vector $a \in \{0,1\}^{|\mathcal{P}|}$ indicating whether each asset is available for investment ($a_i = 1$) or not ($a_i = 0$). The goal of the investment agent is to allocate a fixed budget B across these assets to **maximize the utility** of the portfolio.

Utility Function: Cobb-Douglas utility measures investor satisfaction:

$$U(n_1, \ldots, n_P) = \kappa(\gamma) \prod_{i \in \mathcal{P}} n_i^{\gamma_i}$$

where γ_i is the preference for asset i (output of a preference model), $\kappa(\gamma)$ scales the function, and n_i is the number of shares (what we compute).

Investor preference model: The γ_i preference coefficients can reflect market conditions, sentiment, and other asset-specific information through a feature vector $\mathbf{x}_i \in \mathbb{R}^m$:

$$\gamma_i = \sigma\left(\mathbf{x}_i^\top \theta_i\right) \quad \forall i \in \mathcal{P}$$

where $\sigma : \mathbb{R} \to \mathbb{R}$ is an activation function $\sigma_\theta(x) \in [-1, 1]$, and $\theta_i \in \mathbb{R}^p$ $(p = m + 1)$ denotes the feature weights (and bias), learned from data or set based on subjective beliefs.