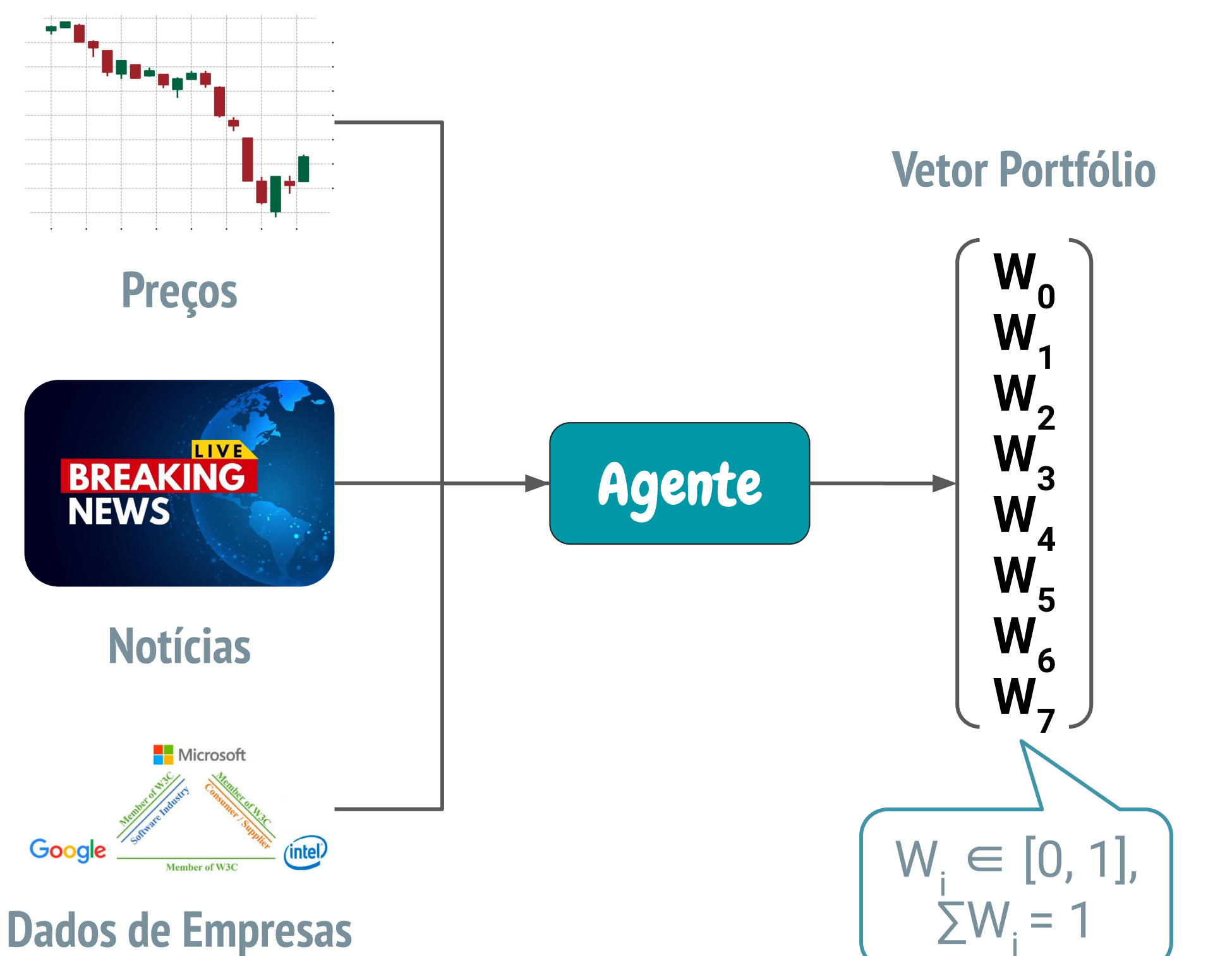


Tema:

## Otimização de Carteiras de Ações com Aprendizado por Reforço

### Otimização de Carteiras



A cada instante de tempo, um agente otimizador de carteiras define um vetor chamado **vetor portfólio** com base em dados do mercado.

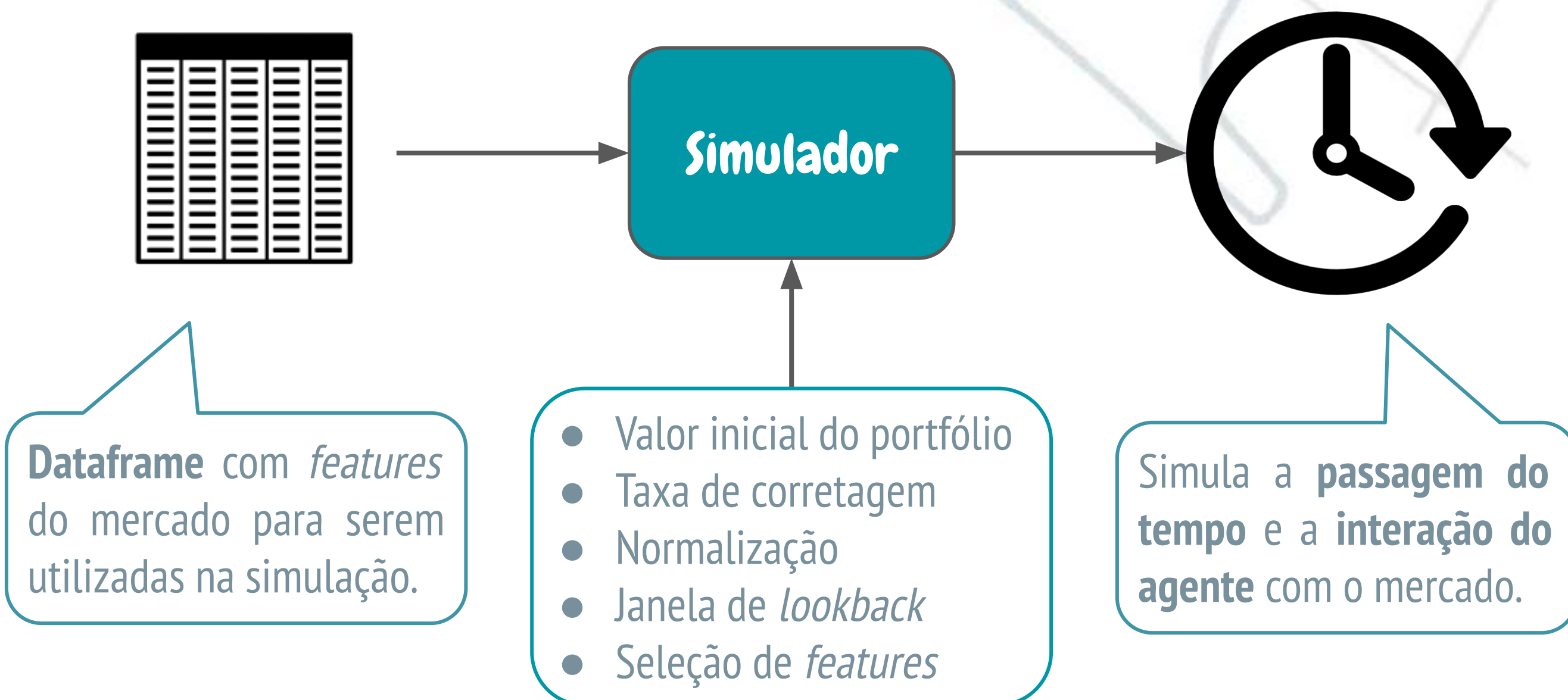
### Ações do Agente

A ação que o agente aplica sobre o ambiente é o **vetor portfólio**.

### Recompensas do Ambiente

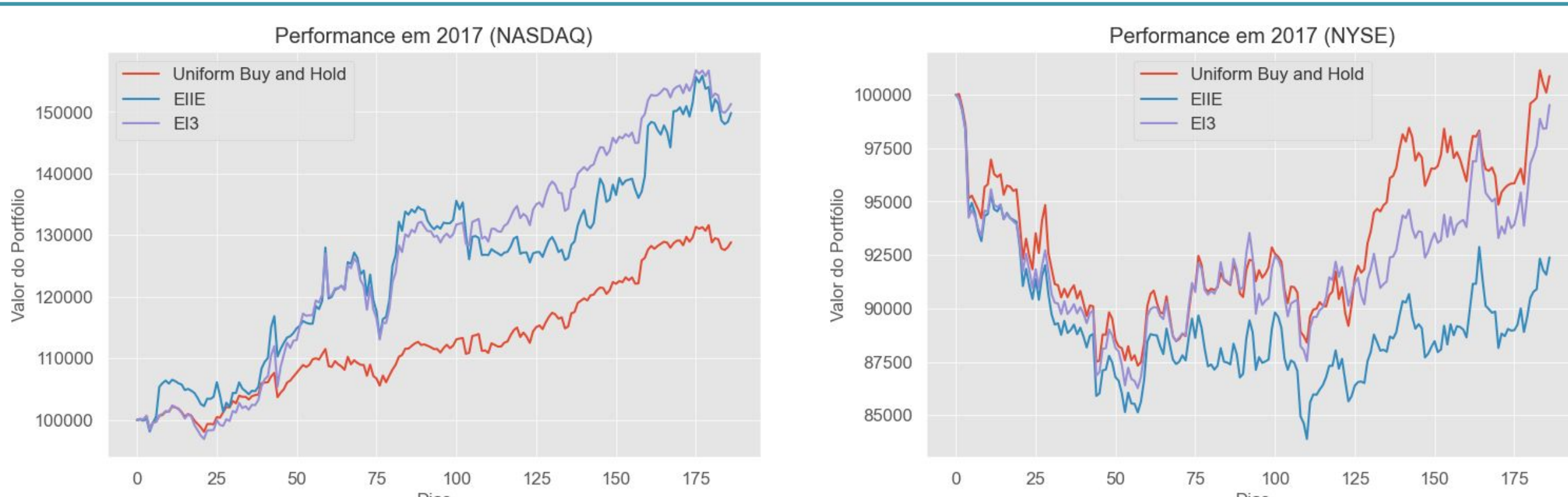
A recompensa do ambiente pode ser qualquer **métrica de lucro ou risco** da carteira no período de simulação.

### Ambiente de Simulação



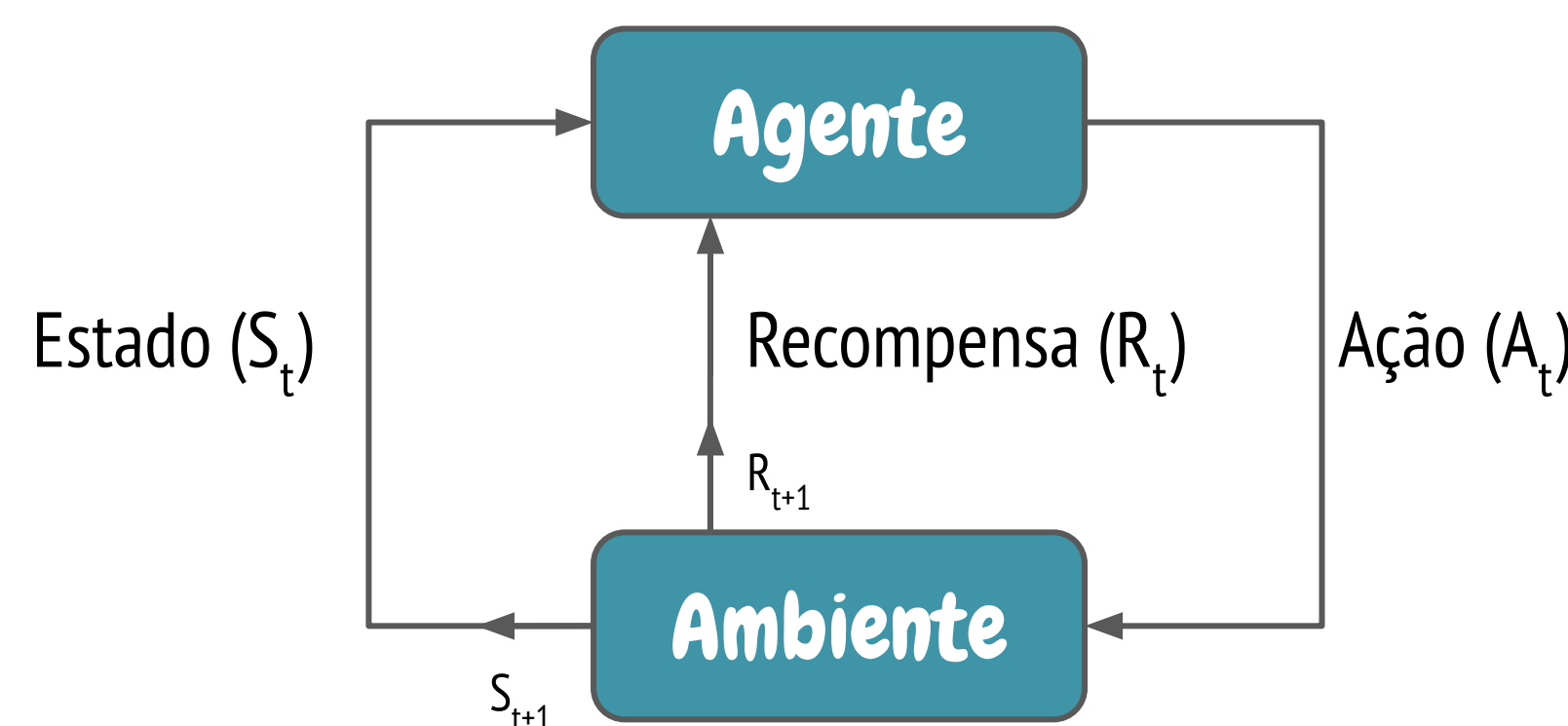
### Resultados

Este teste comparativo entre EIIE, EI<sup>3</sup> e a estratégia *buy and hold* (comprar ações igualmente e nunca vendê-las) foi feito em **dois portfólios de dez e onze ações de alto volume da NASDAQ e da NYSE, respectivamente.**



Resultados mostram que, ao treinar um agente para maximizar o lucro, ele tem bom desempenho em algumas situações mas pode perder dinheiro em momentos voláteis.

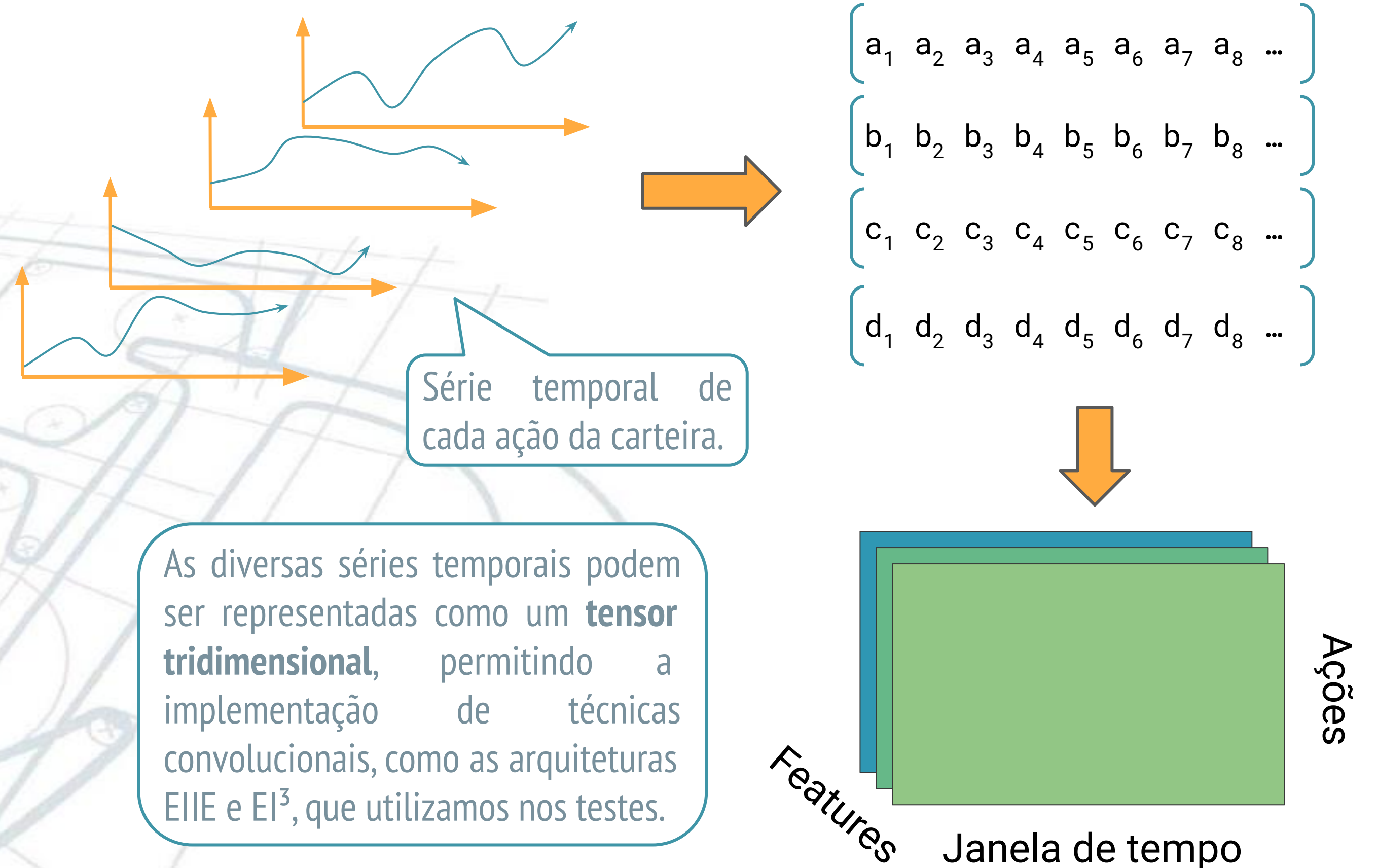
### Aprendizado por Reforço



O agente interage continuamente com o ambiente e recebe recompensas associadas às ações praticadas.

Para tentar **maximizar essa recompensa a longo prazo**, o agente aprende uma política de ações ótima.

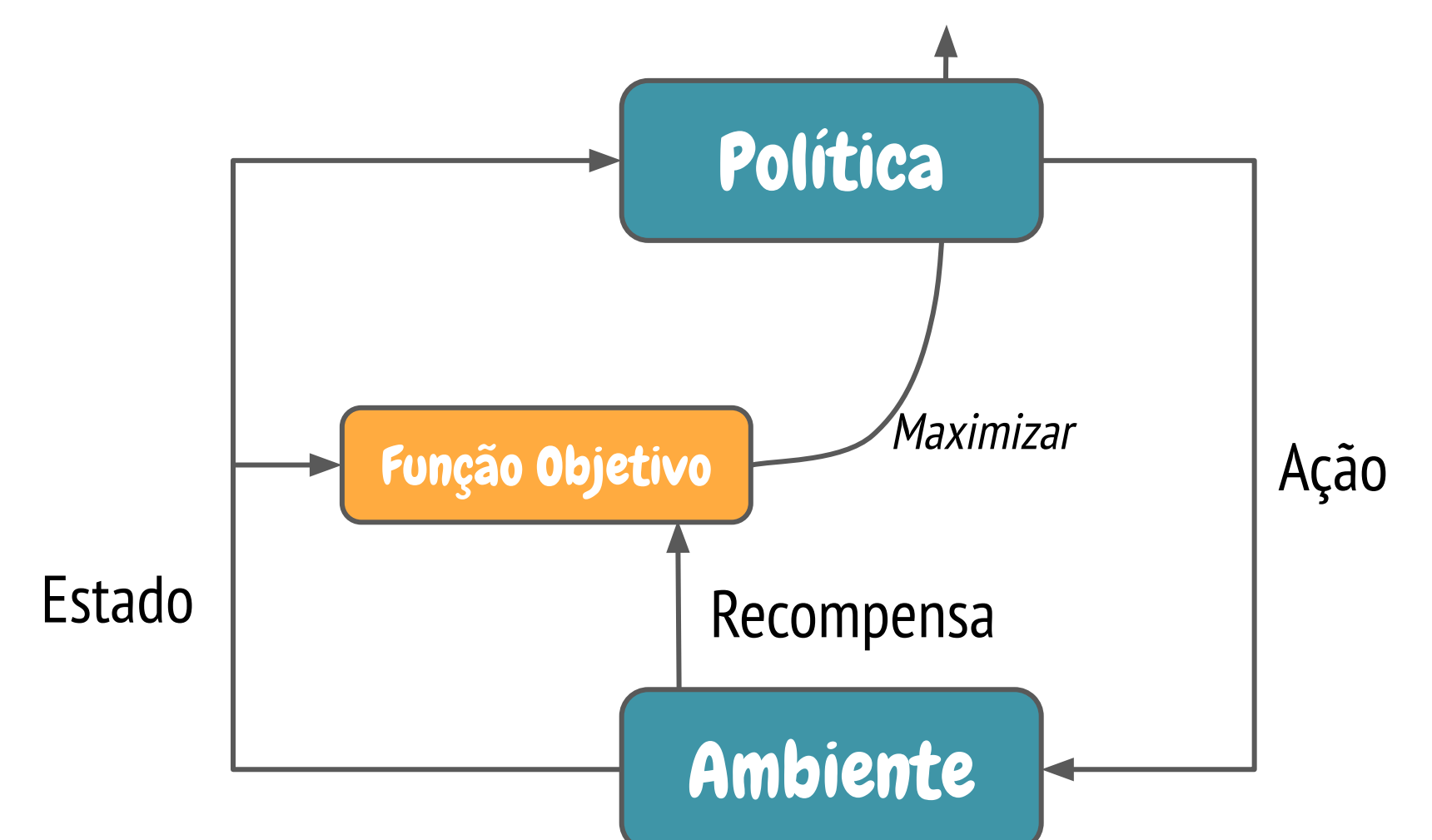
### Representação de Estados do Agente



As diversas séries temporais podem ser representadas como um **tensor tridimensional**, permitindo a implementação de técnicas convolucionais, como as arquiteturas EIIE e EI<sup>3</sup>, que utilizamos nos testes.

### Treinamento

O treinamento é feito usando um algoritmo de **policy gradient** em uma política definida por uma rede neural convolucional.



Essa técnica gera uma política por meio da maximização de uma função objetivo (que pode levar em consideração o lucro, o risco, entre outros fatores) utilizando **gradiente ascendente**.

### Agradecimentos



< **c<sup>2</sup>d** \_  
centro de ciência de dados

