

# 1. Introdução e Contexto

O modelo Black-Litterman tradicional combina equilíbrio de mercado com opiniões de investidores. Sua abordagem inovadora integra três camadas de informação[1]:

1. **Equilíbrio de mercado** (via Yahoo Finance): retornos implícitos no portfólio de market cap
2. **Análise de sentimento** (análise de notícias): score de sentimento público sobre o ativo
3. **Previsão de preços** (modelo preditivo): estimativa de retorno esperado baseada em dados históricos e padrões

Essa combinação transforma o Black-Litterman em um sistema híbrido bayesiano-preditivo, onde as views não são apenas opiniões, mas previsões informadas por NLP e séries temporais[2][3].

## 2. Análise de Sentimento Financeiro (FSA)

### 2.1 Papel do Sentimento no Modelo

O sentimento não é um modelo de retorno. Ele é tratado como:

- uma fonte de opinião (view),
- sujeita a erro,
- com grau de confiança variável.

Ou seja: sentimento  $\neq$  previsão direta  
sentimento  $\rightarrow$  expectativa condicional  $\rightarrow$  view probabilística

### 2.2 Abordagens Avaliadas

#### 2.2.1 Modelos Especializados (BERT-like)

- FinBERT
  - Prós: léxico financeiro, baixo custo, fácil implementação.
  - Contras:
    - insensível a valores numéricos,
    - dificuldade com ironia, causalidade e contexto curto.
- BloombergGPT
  - Prós: estado da arte.
  - Contras: inviável computacionalmente (custo e acesso).

Conclusão: bons como baseline, insuficientes como núcleo do sistema.

### **2.2.2 LLMs Generalistas com Engenharia de Prompt**

Resultados empíricos mostram que:

- LLMs bem instruídos superam FinBERT em F1-score,
- especialmente em notícias curtas e ambíguas.

Técnicas-chave:

- Role-playing (“aja como analista de commodities”),
- prompts estruturados,
- cadeias de raciocínio implícitas.

Limitação estrutural:

- fraqueza em raciocínio numérico e comparativo se não guiados.

### **2.2.3 Frameworks Cognitivos**

- FAP (Financial Attribute Prompting)  
Força o modelo a analisar:
  - semântica,
  - temporalidade,
  - causalidade,
  - risco,
  - comparação implícita.
- HAD (Multi-agentes)  
Debate interno entre agentes especializados reduz erros sutis.

Trade-off: melhora qualidade, aumenta custo de inferência.

### **2.2.4 Recuperação de Contexto (RAG)**

Notícias financeiras frequentemente:

- omitem contexto histórico,
- usam linguagem ambígua.

RAG resolve isso ao:

- recuperar eventos passados,
- relatórios oficiais (USDA, CONAB),
- contexto setorial.

Resultado: redução significativa de classificações erradas.

### 2.3 Escolha Metodológica

Para aplicações práticas e acadêmicas:

- LLM via API + engenharia de prompt + agregação temporal é o melhor trade-off entre:
- desempenho,
- custo,
- simplicidade,
- interpretabilidade.

### 3. Componentes Completos da Fórmula Black-Litterman com Origem e Justificativa

A fórmula central do modelo Black-Litterman é uma atualização bayesiana que combina equilibrium returns (prior do mercado) com suas views (evidência):

$$E(R) = [(\tau\Sigma)^{-1} + P^T\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\Pi + P^T\Omega^{-1}Q]$$

#### 3.1 Detalhamento de Cada Componente

Símbolo	Significado	Dimensão	Origem/Justificativa
$\Pi$	Retornos implícitos de equilíbrio (prior)	$N \times 1$	<b>Reverse optimization:</b> Assume que o portfólio de market cap é ótimo em média-variância. Calculado como $\Pi = \lambda \Sigma w_{mkt}$ . Representa "o que o mercado já precificou".
$\Sigma$	Matriz de covariância dos retornos	$N \times N$	<b>Histórico ou modelo:</b> Covariância empírica (retornos passados) ou modelo fator (PCA/GARCH). Captura risco individual + correlações entre ativos. Essencial para propagar views via covariâncias.
$\tau$	Escalar que mede incerteza sobre o prior (tipicamente 0,01–0,05)	$1 \times 1$	<b>Escalar definido:</b> Diagonal com variâncias $\varepsilon$ das views; comum: $\omega_k = \left(\frac{\pi_k}{\rho}\right)^2 \frac{1}{K}$ , onde $\rho$ é confiança. Reflete subjetividade das views, evitando dominância.

$P$	Matriz picking (como views combinam ativos)	$K \times N$	<b>Estrutura das opiniões:</b> Cada linha define um `portfólio de view". View absoluta → linha com 1; view relativa → [1, --1]; grupo → pesos proporcionais. Para sentimento → identidade (1 view por ativo).
$Q$	Retornos esperados das views	$K \times 1$	<b>Opiniões do investidor:</b> Para sentimento → $Q_i = \Pi_i + \alpha \cdot s_i$ . Sua previsão de preços → $Q_i =$ $\frac{\hat{P}_{t+1} - P_t}{P_t}$ . Fusão das duas fontes.
$\Omega$	Incerteza das views (matriz diagonal)	$K \times K$	<b>Confiança nas opiniões:</b> $\omega_{ii}$ pequeno = view confiável (puxa muito); grande = view incerta (pouco impacto). Para sentimento → função de $ s_i $ ; para previsão → inverso da variância do modelo.
$E(R)$	Retornos esperados posteriores	$N \times 1$	<b>Resultado bayesiano:</b> Média ponderada ótima entre prior (mercado) e views (você). Usado na otimização final de portfólio Markowitz.

Table 1: Componentes completos da fórmula Black-Litterman com origem/justificativa

### 3.2 Fluxo Bayesiano Integrado

A fórmula inteira representa uma **atualização bayesiana**:

Prior (mercado) + Evidência (suas views) → Posterior (portfólio ótimo)

**Black & Litterman (1992):** resolveram o problema do Markowitz (retornos estimados = fraco → portfólios ruim) partindo do mercado como ponto neutro e adicionando suas views de forma probabilística[1].

### 3.3 Aplicação em Seu Sistema

Para a abordagem com commodities brasileiras:

- $\Pi$ : Yahoo Finance → market caps → reverse optimization
- $Q$ : Fusão de sentimento (notícias via LLM) + previsão (LSTM/XGBoost)

- $\Omega$ : 1/confiança do LLM + 1/MSE da previsão (Passível de mudanças)
- $P$ : Identidade (views absolutas por ativo/commodity)
- $E(R)$ : Retornos posteriores para otimizar portfólio final

Essa estrutura garante que as views de sentimento/previsão **nunca dominem completamente** o mercado, produzindo sempre portfólios diversificados e estáveis.

## 4. Arquitetura da Solução

### 4.1 Diagrama

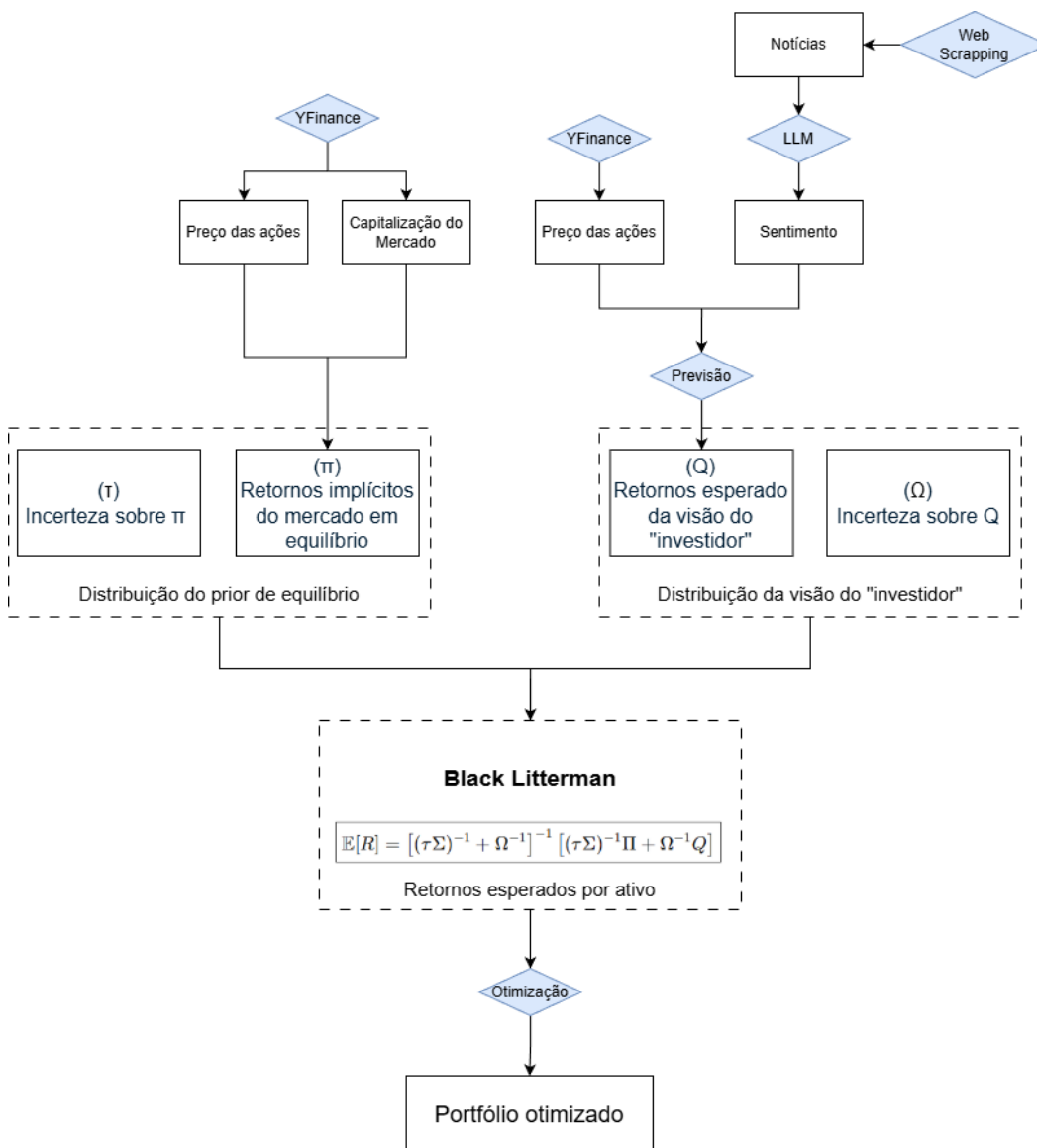


Imagem 1: Diagrama da solução proposta