Ferramenta Web para Análise de Sentimento do Mercado Financeiro Brasileiro: uma evolução do modelo SentFinBERT-PT-BR

Murilo Anjos Leite Novais, Leonardo Barreto Campos

Coordenação de Sistemas de Informação – Instituto Federal de Educação, Ciência e Tecnologia da Bahia (IFBA) – Vitória da Conquista – BA – Brasil

muuriloanjos.dev@gmail.com, leonardobcampos@ifba.edu.br

Resumo. *Este trabalho apresenta o desenvolvimento de uma extensão de navegador para análise de sentimento de notícias financeiras em português brasileiro, integrando um modelo da família FinBERT-PT-BR diretamente ao ambiente de leitura. A solução adota arquitetura cliente-servidor: uma extensão para navegadores Chromium (Manifest V3) e um back-end local em FastAPI que executa o modelo. São propostos dois modos de análise: um modo baseline, com aplicação direta do modelo, e um modo híbrido, que combina segmentação contrastiva, heurísticas de domínio, score contínuo e explicabilidade por oclusão para mitigar erros em textos de sentimento misto. Uma avaliação sobre um conjunto de notícias rotuladas manualmente compara os dois modos, indicando que o modo híbrido aumenta a coerência da classificação em casos de jargão e polaridade mista, ampliando a acessibilidade da análise de sentimento financeiro.*

Abstract. *This work presents a browser extension for sentiment analysis of financial news in Brazilian Portuguese, embedding a FinBERT-PT-BR family model directly into the user's reading environment. The solution follows a client-server architecture: a Chromium extension (Manifest V3) and a local FastAPI back-end that runs the model. Two analysis modes are proposed: a baseline mode, applying the model directly, and a hybrid mode that combines contrastive chunking, domain heuristics, a continuous score and occlusion-based explainability to mitigate errors in mixed-sentiment texts. An evaluation over a manually labeled news set compares both modes, indicating that the hybrid mode improves classification coherence in cases of jargon and mixed polarity, broadening the accessibility of financial sentiment analysis.*

1. Introdução

Na economia contemporânea, o mercado financeiro é um dos ambientes mais dinâmicos e sensíveis a informações. Embora seja influenciado por aspectos objetivos, como indicadores macroeconômicos e dados quantitativos, ele também sofre forte influência de aspectos subjetivos, como opiniões, expectativas e emoções dos agentes econômicos (Lo, 2004). Esses fatores subjetivos manifestam-se sobretudo em notícias, artigos de opinião e relatórios de mercado, consumidos diariamente por investidores, analistas e por grande parte da população economicamente ativa.

Nesse cenário, a abundante quantidade de informações textuais geradas diariamente representa, simultaneamente, um desafio e uma oportunidade. Realizar a

curadoria desse grande volume de dados não estruturados, extraindo percepções relevantes em tempo hábil, é uma tarefa distante da capacidade humana, o que impulsiona a demanda por soluções automatizadas (Ashtiani; Raahemi, 2023). Ganha força, assim, a Análise de Sentimentos, subárea do Processamento de Linguagem Natural (PLN) (Caseli; Nunes; Pagano, 2023).

A partir dos textos, é possível extrair e classificar automaticamente a polaridade de opiniões e emoções, em positivo, negativo ou neutro, no domínio financeiro (Birjali; Kasri; Beni-Hssane, 2021; Medhat; Hassan; Korashy, 2014). Diversos estudos evidenciam a correlação entre o sentimento extraído de fontes textuais e os movimentos do mercado, reforçando seu potencial como ferramenta de apoio à decisão (Kraaijeveld; De Smedt, 2020; Galphade et al., 2024). As inovações recentes em PLN, especialmente a arquitetura Transformer (Vaswani et al., 2017) e modelos pré-treinados como o BERT (Devlin et al., 2019), ampliaram substancialmente a capacidade das ferramentas de inteligência artificial (IA) de perceber nuances da linguagem.

Contudo, a aplicação a textos financeiros é mais complexa, devido a jargões e estruturas de sentido próprias do domínio. Para atender essa demanda surgiram modelos especializados, como o FinBERT (Araci, 2019; Yang et al., 2020) para o inglês. No Brasil, Santos, Bianchi e Costa (2023) desenvolveram o FinBERT-PT-BR, ajustando um modelo BERT ao português do mercado financeiro, e derivaram dele o SentFinBERT-PT-BR, classificador que se mostrou superior a soluções genéricas na análise de notícias econômicas nacionais.

O objetivo geral deste trabalho é desenvolver e avaliar uma extensão para navegadores Chromium que aplique um modelo da família FinBERT-PT-BR à análise de sentimento de notícias financeiras brasileiras, fornecendo classificação contextualizada durante a leitura. Como contribuição adicional, são propostos e comparados dois modos de análise, um *baseline* e um modo híbrido com etapas de pós-processamento, discutindo-se seus efeitos sobre a qualidade da classificação.

O restante deste artigo está organizado em cinco seções. A Seção 2 apresenta o referencial teórico e os trabalhos relacionados que fundamentam a proposta. A Seção 3 descreve a arquitetura da ferramenta web, detalhando os componentes cliente e servidor e o fluxo de comunicação entre eles. A Seção 4 aprofunda o motor de análise, contrastando o modo baseline com o modo híbrido e o mecanismo de explicabilidade. A Seção 5 relata a avaliação experimental, comparando quantitativamente os dois modos. Por fim, a Seção 6 traz as considerações finais e aponta direções para trabalhos futuros.

2. Referencial Teórico

A hipótese dos mercados adaptativos, proposta por Lo (2004), sustenta que os preços não refletem apenas informações objetivas, mas também percepções, heurísticas e emoções coletivas dos agentes. Esse entendimento justifica a necessidade de ferramentas capazes de capturar o sentimento presente em comunicações e notícias econômicas, dado que tais percepções impactam diretamente o comportamento do mercado.

No campo técnico, Medhat, Hassan e Korashy (2014) revisam as principais técnicas de análise de sentimentos, das abordagens baseadas em regras aos algoritmos supervisionados, destacando seu papel em contextos em que opiniões humanas afetam decisões complexas. Birjali, Kasri e Beni-Hssane (2021) aprofundam os desafios contemporâneos da área (ironia, ambiguidade e informalidade) reforçando a necessidade

de modelos robustos e sensíveis ao contexto. A introdução da arquitetura Transformer (Vaswani et al., 2017) representou uma inflexão ao permitir o processamento paralelo e contextualizado do texto por meio do mecanismo de atenção, base sobre a qual foi construído o BERT (Devlin et al., 2019), capaz de compreender o contexto de uma palavra de forma bidirecional.

A partir do BERT surgiram variantes de domínio, como o FinBERT (Araci, 2019; Yang et al., 2020), voltado à linguagem financeira em inglês. Essa especialização garante leitura mais precisa de jargões técnicos, evitando erros comuns de modelos genéricos. No contexto brasileiro, Santos (2022) e, posteriormente, Santos, Bianchi e Costa (2023) adaptaram a tecnologia ao português, originando o FinBERT-PT-BR e o classificador de emoções SentFinBERT-PT-BR. O presente trabalho reutiliza diretamente esse esforço: o motor de análise emprega o modelo público FinBERT-PT-BR como componente central de classificação.

Em Galphade et al. (2024), os autores empregam aprendizado profundo sobre análise semântica para prever tendências de ações, e Yadav (2024) combina a análise de sentimento baseada em BERT com técnicas de aprendizado de máquina para aprimorar a previsão de tendências, relatando ganhos expressivos de desempenho. No plano metodológico, Shah et al. (2022) propõem arcabouços de negociação apoiados em arquiteturas de aprendizado profundo, enquanto Kraaijeveld e De Smedt (2020) demonstram que o sentimento expresso em redes sociais pode antecipar movimentos de preço, evidenciando o valor preditivo de percepções coletivas. Costa e Cozman (2024) contextualizam a aceleração da pesquisa em IA e seu impacto sobre aplicações práticas, cenário no qual a presente proposta se insere.

Apesar desses avanços, a aplicação dos modelos ainda esbarra em desafios de usabilidade. Ashtiani e Raahemi (2023) observam que muitos sistemas baseados em análise textual falham por não integrarem a tecnologia de forma intuitiva ao fluxo de trabalho do usuário. É justamente nessa lacuna de integração e acessibilidade, não na criação de um novo modelo, que este trabalho se posiciona, articulando um modelo já existente (o FinBERT-PT-BR) a uma interface de uso cotidiano e a uma camada de pós-processamento orientada ao domínio.

3. Arquitetura da Ferramenta Web

A solução foi concebida segundo uma arquitetura cliente-servidor composta por dois artefatos principais: uma extensão de navegador (cliente) e um serviço de back-end local (servidor) responsável por executar o modelo de IA. A Figura 1 apresenta a visão geral dessa organização e o fluxo de informação entre os componentes.

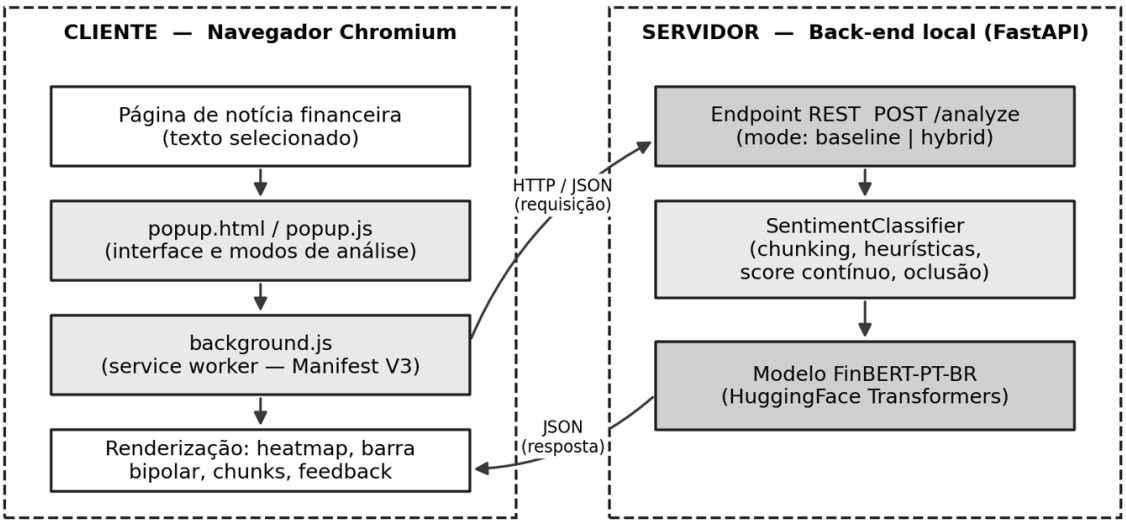


Figura 1. Arquitetura cliente-servidor da solução.

O cliente é uma extensão para navegadores baseados em Chromium (Google Chrome, Microsoft Edge e Brave), construída sobre o Manifest V3. Ela é composta por uma interface em janela (popup.html/popup.js), na qual o usuário cola ou digita o trecho a ser analisado e seleciona o modo de análise, e por um service worker em segundo plano (background.js), que registra um item de menu de contexto, permitindo analisar qualquer texto selecionado na página diretamente pela ação de clique. A renderização dos resultados ocorre na própria interface da extensão e, no caso do menu de contexto, por meio de um aviso (toast) injetado na página visitada.

O servidor é um serviço REST desenvolvido em Python com o framework FastAPI, executado localmente. Ele expõe o endpoint POST /analyze, que recebe um objeto JSON contendo o texto e o modo desejado (baseline ou hybrid), e um endpoint auxiliar GET /health para verificação de disponibilidade. O carregamento e a inferência do modelo FinBERT-PT-BR são realizados pela biblioteca HuggingFace Transformers (WOLF et al., 2020), que oferece uma interface de alto nível para modelos baseados em Transformer. A escolha de um back-end local privilegia a privacidade — os textos não trafegam para serviços de terceiros – e elimina custos de hospedagem, ao preço de exigir a execução do serviço na máquina do usuário.

Para executar o protótipo, deve-se instalar as dependências indicadas em requirements.txt, iniciar o servidor, a partir da pasta backend, com uvicorn app.main:app --reload, e carregar a pasta extension como extensão descompactada em um navegador Chromium. A extensão então acessa a API local em http://127.0.0.1:8000. A comunicação entre os componentes ocorre por requisições HTTP assíncronas: a extensão envia o texto extraído da página e recebe, em resposta, um objeto JSON com o rótulo de sentimento, a confiança, o score contínuo e, no modo híbrido, a análise por sentença e as importâncias por palavra.

4. O Motor de Análise: do Baseline ao Modo Híbrido

O núcleo da solução é a classe SentimentClassifier, que encapsula duas estratégias de classificação, atendendo a uma orientação metodológica de comparar o modelo aplicado de forma direta a uma versão aprimorada com etapas adicionais de pós-processamento.

No modo baseline, o texto é submetido integralmente ao pipeline de análise de sentimentos do FinBERT-PT-BR, e o rótulo de maior probabilidade (positivo, negativo ou neutro) é retornado junto à respectiva confiança. Essa abordagem é simples e rápida, servindo como linha de base de comparação. Contudo, ao tratar o texto como um bloco único, ela é suscetível a erros em manchetes e parágrafos que misturam polaridades, por exemplo, quando a queda de um indicador é positiva para o mercado local e não oferece qualquer explicação sobre a decisão. A Figura 2 ilustra concretamente essa limitação: uma mesma notícia de polaridade mista recebe um rótulo equivocado no modo baseline, ao passo que o modo híbrido chega à classificação correta.

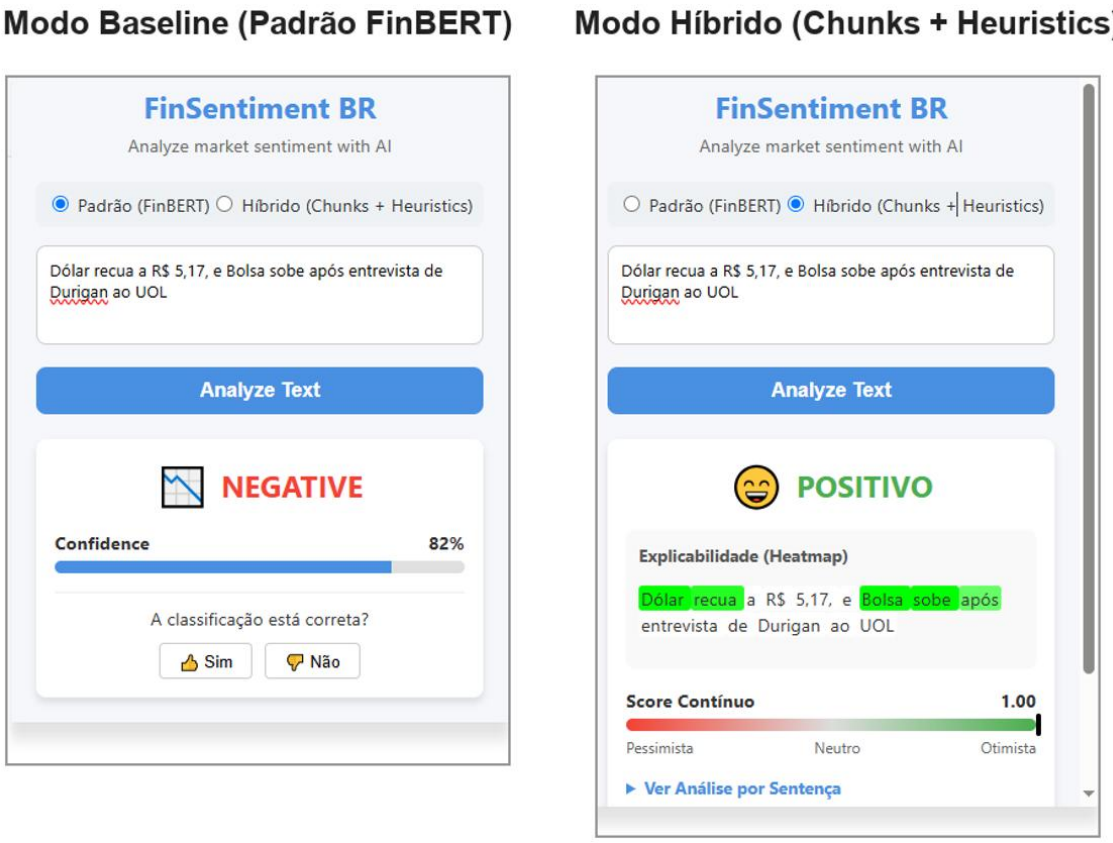


Figura 2. Classificações distintas para a mesma notícia (Baseline × Híbrido).

O modo híbrido foi proposto para mitigar essas limitações por meio de uma camada de pós-processamento orientada ao domínio financeiro. Seu funcionamento ocorre em etapas encadeadas, são elas: (i) pré-processamento; (ii) chucking; (iii) classificação; (iv) heurística; (v) score; (vi) agregação ponderada; (vii) sentiment geral.

A primeira etapa é um pré-processamento que normaliza símbolos monetários (R$, US$, €) para suas formas por extenso, melhorando a interpretação do modelo. Em seguida, o texto passa por uma segmentação (*chunking*) que o divide não apenas por sentenças, mas também por conectivos contrastivos (mas, porém, contudo, no entanto, apesar de). A cada segmento é atribuído um peso: orações introduzidas por adversativas recebem peso maior, por concentrarem a informação dominante, ao passo que construções de concessão recebem peso menor, refletindo sua função subordinada no sentido global.

Cada segmento é então classificado individualmente pelo FinBERT-PT-BR, e sua saída categórica é convertida em um score contínuo no intervalo de -1,0 a +1,0. Sobre esse score atua uma camada de heurísticas de domínio, implementada por expressões regulares, que ajusta a polaridade conforme regras específicas do mercado: a relação entre moeda e direção do movimento (a queda do dólar tende a ser positiva para o mercado local, ao passo que a queda do IBOVESPA é negativa), além de jargões macroeconômicos e corporativos (como fechamento da curva de juros, EBITDA recorde, abertura da curva ou compressão de margens). O sentimento geral do texto resulta da média ponderada dos scores dos segmentos pelos respectivos pesos, sendo convertido novamente em um rótulo final por limiares simétricos.

4.1. Explicabilidade por Oclusão

Para tornar a decisão do modelo interpretável, o modo híbrido incorpora um mecanismo de explicabilidade baseado em oclusão (occlusion), técnica de atribuição por perturbação alinhada à literatura de IA explicável (Ribeiro; Singh; Guestrin, 2016). Conforme a Figura 4, calcula-se o *score* do texto completo e, a seguir, remove-se uma palavra de cada vez, recalculando o *score* do texto perturbado. A diferença entre o score original e o score sem a palavra estima o impacto daquele termo na classificação. Esses impactos são enviados à extensão e renderizados como um mapa de calor (heatmap), que destaca visualmente as palavras que mais empurraram a decisão para o polo positivo ou negativo. Por exigir uma inferência por palavra, o procedimento é restrito a textos de até cinquenta palavras, preservando uma latência aceitável para uso interativo.

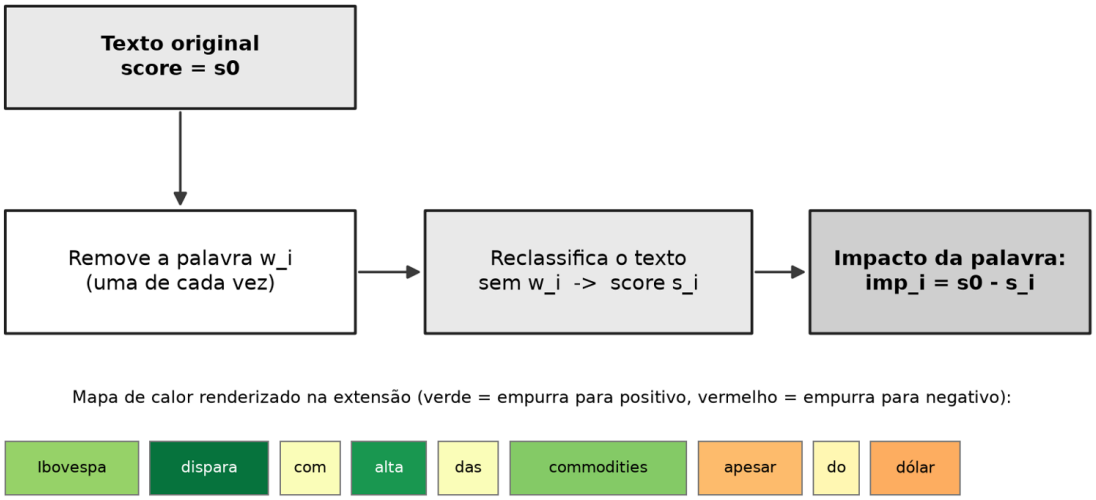


Figura 4. Explicabilidade por oclusão e geração do mapa de calor.

Além do mapa de calor, a interface apresenta uma barra bipolar que posiciona o *score* contínuo entre os polos pessimista e otimista, um detalhamento da análise por sentença e um instrumento de *feedback* (“A classificação está correta?”), que permite coletar avaliações dos usuários para análises futuras. Esse conjunto de recursos transforma a saída, antes um rótulo opaco, em uma explicação acionável e contextualizada.

5. Avaliação Experimental

A avaliação teve como objetivo comparar quantitativamente o modo baseline e o modo híbrido, verificando o efeito das etapas de pós-processamento sobre a qualidade da

classificação. Para isso, foi construído um conjunto diagnóstico (*challenge set*) com 200 amostras de notícias financeiras em português. Optou-se por um conjunto independente do corpus de treinamento do FinBERT-PT-BR, evitando o viés de memorização (data leakage) que ocorreria ao reutilizar a base original do modelo, e construído de modo a estressar sistematicamente os fenômenos linguísticos de interesse.

O conjunto combina dois subconjuntos. O primeiro, um núcleo de 20 amostras escritas e revisadas manualmente, concentra casos difíceis de polaridade mista, paradoxos macroeconômicos e jargão de mercado. O segundo, com 180 amostras, foi gerado por uma técnica de geração por templates (inspirada em metodologias de behavioral testing, como o CheckList), na qual cada frase é instanciada a partir de moldes cujo rótulo de sentimento é conhecido por construção, o que torna o padrão-ouro confiável sem depender de anotação manual posterior. Cada amostra registra sua origem (curada ou sintética) e a categoria do fenômeno avaliado, permitindo recortes analíticos. As categorias foram agrupadas em casos fáceis (manchetes de polaridade única) e difíceis (direção contraintuitiva de moeda, orações de concessão e múltiplos eventos contrastantes). O conjunto é reprodutível: o gerador (build_validation_dataset.py) emprega semente fixa.

Cada amostra é submetida aos dois modos, então são calculados, para cada um dos modos, a acurácia geral, a precisão, a revocação e a medida F1 por classe, além das médias macro e ponderada da F1 e da matriz de confusão. A Figura 5 consolida o desempenho global comparativo dos dois modos sobre as 200 amostras, e a Figura 6 detalha o desempenho segmentado entre casos fáceis e difíceis.

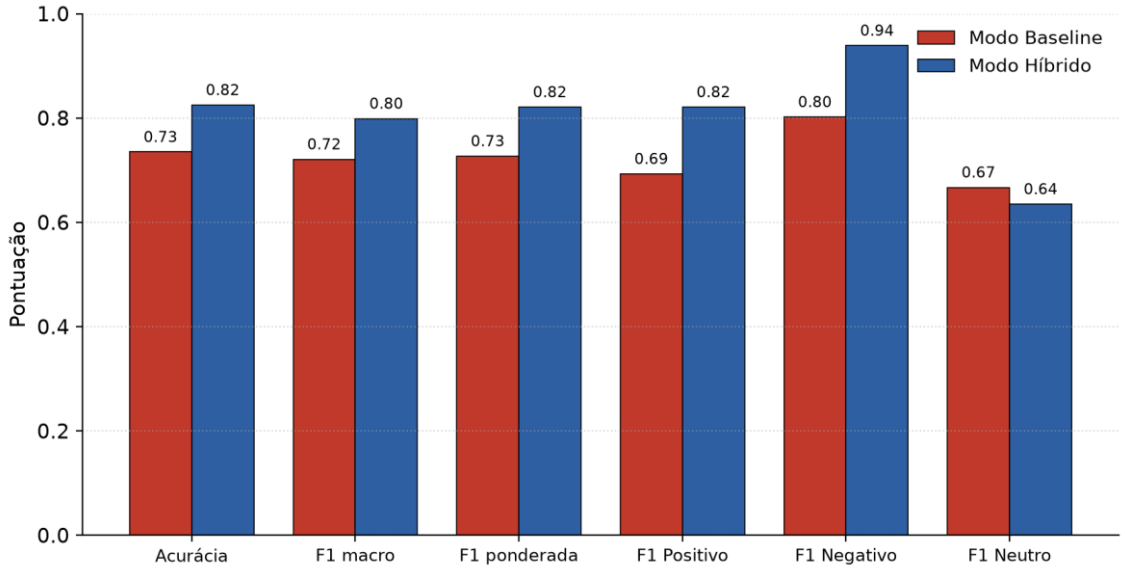


Figura 5. Desempenho global dos modos baseline e híbrido por métrica (200 amostras).

Enquanto a Figura 5 sintetiza o desempenho agregado por métrica, a Figura 6 isola o comportamento dos dois modos nos subconjuntos de casos fáceis e difíceis. Esse recorte é essencial para compreender onde efetivamente reside o ganho do modo híbrido.

No conjunto completo, o modo híbrido superou o baseline em todos os indicadores agregados: a acurácia passou de 0,74 para 0,83 e a F1 macro de 0,72 para 0,80, ganhos de 9 e 8 pontos percentuais, respectivamente. Por classe, as melhoras mais expressivas ocorreram em Positivo (F1: 0,69 → 0,82) e Negativo (F1: 0,80 → 0,94).

A leitura mais informativa, porém, emerge do recorte entre casos fáceis e difíceis (Figura 6). Nas manchetes de polaridade única, ambos os modos têm desempenho próximo (0,80 contra 0,82), o que confirma que o modelo generalista já é competente quando o sinal de sentimento é direto e não há ruído contextual. É nos casos difíceis que a contribuição do trabalho se materializa: a acurácia do baseline cai para 0,58, enquanto o modo híbrido sustenta 0,85, uma diferença de 27 pontos percentuais. São justamente as situações que motivaram a arquitetura proposta: a direção contraintuitiva de moedas (a queda do dólar como sinal positivo para o mercado local), as orações de concessão ("apesar de", "mas") e os múltiplos eventos contrastantes na mesma sentença, em que o baseline se confunde com a coexistência de termos como "queda" e "dispara", ao passo que a segmentação contrastiva e as heurísticas de moeda e mercado conduzem o modo híbrido ao rótulo correto.

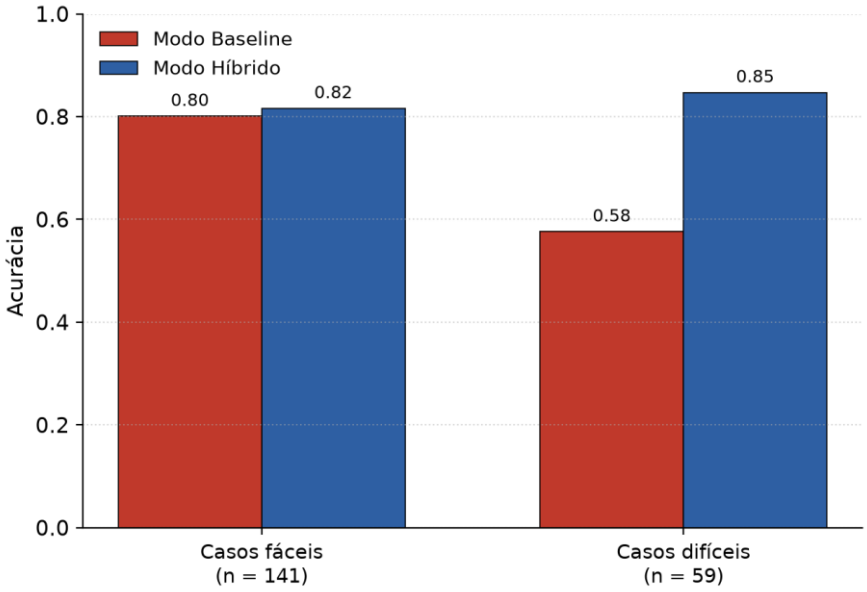


Figura 6. Acurácia dos modos em casos fáceis e difíceis.

Em contrapartida, dois limites devem ser registrados. Primeiro, a classe Neutro teve leve regressão (F1: 0,67 → 0,64): por serem orientadas à polaridade, as heurísticas ocasionalmente deslocam para positivo ou negativo textos que deveriam permanecer neutros. Segundo, as regras têm caráter dependente de domínio e exigem manutenção, pois novos jargões ou contextos não previstos podem não ser capturados — o que delimita o alcance da abordagem e aponta direções para trabalhos futuros.

6. Considerações Finais

Este trabalho apresentou o projeto, a implementação e a avaliação de uma extensão de navegador para análise de sentimento financeiro, integrando o checkpoint FinBERT-PTBR ao ambiente de leitura por meio de uma arquitetura cliente-servidor. A principal contribuição não está na criação ou evolução de um modelo, mas na integração do classificador e na proposta de uma camada híbrida de pós-processamento com segmentação contrastiva, heurísticas de domínio, score contínuo e explicabilidade por oclusão. No conjunto diagnóstico, o híbrido melhorou significativamente o desempenho, sobretudo nos casos difíceis; a ablação atribuiu o ganho agregado principalmente às regras de moeda e direção. Entre as limitações, destacam-se a predominância de dados

sintéticos, a ausência de um conjunto independente de notícias publicadas, a regressão da classe Neutro, a falta de avaliação da fidelidade das explicações e de estudo de usabilidade, a execução de um back-end local e a manutenção contínua das heurísticas.

Como trabalhos futuros, prevêse ampliar o conjunto rotulado com notícias reais e anotadores independentes, validar os resultados em teste held-out, avaliar explicações e usabilidade, investigar ABSA aprendida, viabilizar inferência no navegador e automatizar a extração do corpo das notícias. O protótipo constitui uma base de engenharia para essas extensões, sem substituir análise financeira profissional.

Declaração de Uso de Inteligência Artificial

Na elaboração deste trabalho, ferramentas de inteligência artificial generativa foram utilizadas como apoio à revisão da redação, à organização do código e à estruturação dos experimentos. Uma IA generativa atuou como segundo anotador auxiliar nas 20 amostras curadas. O autor verificou as fontes, o código e os resultados quantitativos; as decisões de projeto, interpretações e conclusões permanecem sob sua responsabilidade.

Referências

ARACI, Dogu. FinBERT: financial sentiment analysis with pre-trained language models. 2019. Disponível em: https://arxiv.org/abs/1908.10063. Acesso em: 21 abr. 2025.

ASHTIANI, Matin N.; RAAHEMI, Bijan. News-based intelligent prediction of financial markets using text mining and machine learning: a systematic literature review. Expert Systems with Applications, v. 217, 119509, 2023.

BIRJALI, Mohammed; KASRI, Mouad; BENI-HSSANE, Abderrahim. A comprehensive survey on sentiment analysis: approaches, challenges, and trends. Knowledge-Based Systems, v. 226, 107134, 2021.

CASELI, Helena de Medeiros; NUNES, Maria das Graças Volpe; PAGANO, Adriana Silvina. Processamento de linguagem natural: conceitos, técnicas e aplicações em português. São Carlos: BPLN, 2023. Disponível em: https://brasileiraspln.com/livro-pln/. Acesso em: 16 set. 2025.

COSTA, Anna Helena Reali; COZMAN, Fabio Gagliardi. O futuro da pesquisa em inteligência artificial. Revista USP, São Paulo, n. 141, p. 133-146, abr./jun. 2024.

DEVLIN, Jacob et al. BERT: pre-training of deep bidirectional transformers for language understanding. In: NAACL-HLT, 2019, Minneapolis. Proceedings... Minneapolis: ACL, 2019. p. 4171-4186.

GALPHADE, Manisha et al. Semantic analysis using deep learning for predicting stock trends. Procedia Computer Science, v. 235, p. 820-829, 2024.

KRAAIJEVELD, Otto; DE SMEDT, Jelle. The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. Journal of International Financial Markets, Institutions and Money, v. 65, 101188, 2020.

LO, Andrew W. The adaptive markets hypothesis. The Journal of Portfolio Management, v. 30, n. 5, p. 15-29, 2004.

MEDHAT, Walaa; HASSAN, Ahmed; KORASHY, Hoda. Sentiment analysis algorithms and applications: a survey. Ain Shams Engineering Journal, v. 5, n. 4, p. 1093-1113, 2014.

RIBEIRO, Marco Tulio; SINGH, Sameer; GUESTRIN, Carlos. “Why should I trust you?”: explaining the predictions of any classifier. In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 22., 2016, San Francisco. Proceedings... New York: ACM, 2016. p. 1135-1144.

SANTOS, Lucas Leme. FinBERT-PT-BR: análise de sentimentos de textos em português referentes ao mercado financeiro. 2022. Trabalho de Conclusão de Curso (Engenharia de Computação) — Escola Politécnica, Universidade de São Paulo, São Paulo, 2022.

SANTOS, Lucas L.; BIANCHI, Reinaldo A. C.; COSTA, Anna H. R. FinBERT-PT-BR: análise de sentimentos de textos em português do mercado financeiro. In: Simpósio Brasileiro de Sistemas de Informação (SBSI), 19., 2023, Passo Fundo. Anais... Porto Alegre: SBC, 2023. p. 825-832.

SHAH, Anshul et al. A stock market trading framework based on deep learning architectures. Multimedia Tools and Applications, v. 81, n. 10, p. 14153-14171, 2022.

VASWANI, Ashish et al. Attention is all you need. Advances in Neural Information Processing Systems, v. 30, 2017.

WOLF, Thomas et al. Transformers: state-of-the-art natural language processing. In: Conference on Empirical Methods in Natural Language Processing (EMNLP): System Demonstrations, 2020. p. 38-45.

YADAV, Nikesh. Enhancing stock trend prediction using BERT-based sentiment analysis and machine learning techniques. International Journal of Quantitative Research and Modeling, v. 5, n. 1, p. 1-11, 2024.

YANG, Yaping et al. FinBERT: a pretrained language model for financial communications. 2020. Disponível em: https://arxiv.org/abs/2006.08097. Acesso em: 21 abr. 2025.