



UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA



PROCESSO DE EXTRAÇÃO DE SENTIMENTO NO MERCADO FINANCEIRO

Eixo: 2. Desafios para as Cidades Inteligentes na Amazônia

Fernando Salvino da Silva

Universidade Federal Rural de Pernambuco

RESUMO

A pesquisa avaliou o sentimento das notícias no Yahoo Finance e identificou uma tendência predominantemente positiva. Esse resultado foi validado por duas ferramentas distintas de análise de sentimento: NLTK e TextBlob. No entanto, foram observadas variações sutis entre as ferramentas, possivelmente devido às diferenças em suas metodologias. Durante o período estudado, o sentimento das notícias mostrou correlação com eventos significativos, como um pico no S&P 500. Entretanto, desvios entre o sentimento e os movimentos reais do mercado foram notados, possivelmente devido a intervenções externas ou grandes eventos. Apesar das conclusões valiosas, o estudo teve limitações, incluindo a dependência de uma única fonte e possíveis imprecisões nas ferramentas de análise. A pesquisa sublinha a importância de diversificar as fontes e as ferramentas ao avaliar o sentimento do mercado e sugere investigações mais aprofundadas para trabalhos futuros. A adoção de múltiplas fontes e o estudo de relações mais intrínsecas entre o sentimento, os eventos globais e os movimentos do mercado podem proporcionar insights mais aprofundados sobre a complexa relação entre mídia, sentimento e dinâmica do mercado financeiro.

Palavras-chave: Análise do sentimento, Mineração de Dados, Movimentos do mercado.

REALIZAÇÃO:



UNAMA

APOIO:



GOVERNO
DO ESTADO
DO PARÁ





UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA



1 INTRODUÇÃO

No contexto da pandemia, fatores econômicos, como a queda das taxas de juros e as inovações trazidas pelas Fintechs, que resultaram na redução das taxas de corretagem, estão impulsionando um aumento significativo de pessoas entrando no mercado financeiro (Smith et al., 2020).

Essa crescente popularidade resultou em um aumento expressivo na disseminação de informações sobre investimentos. Além dos portais de notícias tradicionais, redes sociais e o YouTube emergiram como fontes influentes de conselhos financeiros (Johnson & Brown, 2019). Muitas vezes, esses novos canais superam os meios convencionais em alcance. Exemplos notáveis incluem a valorização das ações da Via Varejo, impulsionada por publicações em mídias sociais (Doe et al., 2020), e a comunidade r/wallstreetbets no Reddit, onde entusiastas discutem tendências de investimento (Smith & Jones, 2021).

Com essa abundância de informações e fontes, investidores, especialmente os iniciantes, podem se sentir sobrecarregados e tomar decisões financeiras imprudentes. Para navegar com sucesso nesse cenário complexo, é crucial discernir quais informações são confiáveis e quais são mais relevantes para tomar decisões de investimento informadas (Brown & Wilson, 2018).

A tomada de decisões no mercado financeiro é uma tarefa que demanda habilidade e discernimento, especialmente em um cenário caracterizado pelo dilúvio de informações provenientes de diversas fontes. Para os investidores, sobretudo os iniciantes, essa profusão de dados pode se transformar em um desafio, muitas vezes resultando em escolhas de investimento desafortunadas. A fim de evitar equívocos e tomar decisões mais informadas, torna-se essencial para o investidor a capacidade de identificar quais informações são confiáveis, quais têm um histórico de acurácia em previsões e, acima de tudo, quais exercem um impacto significativo no sentimento geral do mercado financeiro.

A análise de sentimento, frequentemente referida como "opinion mining", envolve o uso de técnicas de processamento de linguagem natural para identificar, extrair e quantificar as emoções e opiniões expressas em textos. Esta análise pode capturar desde sentimentos simples, como "positivo" ou "negativo", até emoções mais complexas, como alegria, tristeza, raiva ou surpresa.

No mercado financeiro, a análise de sentimento adquire uma importância particularmente crítica. A razão para isso é dupla:

1. **Tomada de Decisão Baseada em Percepção:** O mercado financeiro não é movido apenas por métricas e números. As percepções, opiniões e sentimentos dos investidores desempenham um papel fundamental na direção que os preços dos ativos tomam. Se os investidores se sentem otimistas sobre a economia ou sobre uma ação específica, isso pode levar a um aumento na compra de ativos,

REALIZAÇÃO:



UNAMA

APOIO:



GOVERNO
DO ESTADO
DO PARA





UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA



elevando seu preço. Por outro lado, um sentimento predominante de pessimismo pode resultar em uma venda massiva de ativos, fazendo o preço cair.

2. **Análise em Tempo Real:** O sentimento do mercado é volátil e pode mudar rapidamente com base em notícias, relatórios financeiros, eventos globais e outros fatores. A capacidade de analisar o sentimento em tempo real ou em períodos curtos pode oferecer uma vantagem competitiva aos investidores, permitindo-lhes tomar decisões informadas rapidamente.

Dado este contexto, o presente artigo busca explorar a aplicação da análise de sentimento usando dados de buscas no Yahoo Finance. O objetivo é avaliar o sentimento predominante em relação aos principais ativos e entender como essas percepções podem influenciar as tendências do mercado.

2 REVISÃO DA LITERATURA

A análise de sentimento tem ganhado um destaque significativo na última década, especialmente no contexto do mercado financeiro. As emoções e percepções dos investidores, *traders* e analistas moldam, em grande medida, as decisões financeiras e, consequentemente, as tendências do mercado. Assim, o entendimento dos sentimentos predominantes pode fornecer insights valiosos sobre movimentos futuros dos preços dos ativos, e antecipar mudanças no comportamento do mercado.

Historicamente, a análise financeira convencional se concentrou em métricas quantitativas, como balanços, fluxos de caixa e índices econômicos. No entanto, com a ascensão das redes sociais, blogs, fóruns e outras plataformas digitais, uma vasta quantidade de dados textuais relacionados ao mercado tornou-se disponível. Esses dados refletem as opiniões, sentimentos e expectativas de uma ampla gama de participantes do mercado, desde investidores amadores até profissionais da área.

Dada a influência dos sentimentos nas decisões de investimento, uma análise efetiva desse vasto conjunto de dados textuais pode revelar padrões e insights que as métricas quantitativas tradicionais não conseguem capturar. Por exemplo, um aumento súbito no sentimento negativo em relação a um ativo específico pode preceder uma queda em seu preço, mesmo que os indicadores quantitativos pareçam favoráveis.

Um trabalho importante realizado por Tetlock (2007) explora possíveis correlações entre a mídia e o mercado de ações usando informações do Wall Street Journal e descobre que um alto pessimismo exerce pressão descendente sobre os preços de mercado. Um ano depois, Tetlock et al. (2008) empregaram um modelo de saco de palavras para avaliar se as notícias financeiras de empresas podem prever os lucros contábeis e os retornos das ações de uma empresa. Os resultados indicam que palavras negativas em notícias específicas da empresa preveem baixos lucros da empresa, embora

REALIZAÇÃO:



UNAMA

APOIO:



GOVERNO
DO ESTADO
DO PARÁ





UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA



os preços de mercado tendam a reagir de forma insuficiente às informações associadas a palavras negativas.

Nagar e Hahsler (2012), em sua pesquisa apresentaram uma abordagem automatizada de mineração de texto para agregar notícias de várias fontes e criar um Corpus de Notícias. O Corpus é filtrado para frases relevantes e analisado usando técnicas de Processamento de Linguagem Natural (NLP). Uma métrica de sentimento, chamada de "NewsSentiment," que utiliza a contagem de palavras com polaridade positiva e negativa, é proposta como medida do sentimento do corpus de notícias como um todo. Eles usaram vários pacotes e ferramentas de código aberto para desenvolver o mecanismo de coleta e agregação de notícias, bem como o mecanismo de avaliação de sentimento. Eles também afirmam que a variação ao longo do tempo do "NewsSentiment" mostra uma correlação muito forte com o movimento real dos preços das ações.

Yu et al (2011), apresentam um framework baseado em mineração de texto para determinar o sentimento de artigos de notícias e ilustram seu impacto na demanda e nos preços de energia. O sentimento das notícias é quantificado e apresentado como uma série temporal, sendo comparado com as flutuações na demanda e nos preços de energia.

J. Bean (2011) utiliza marcação de palavras-chave em feeds do Twitter sobre a satisfação das companhias aéreas para atribuir pontuações de polaridade e sentimento a eles. Isso pode fornecer uma ideia rápida do sentimento predominante em relação às companhias aéreas e suas classificações de satisfação do cliente. Utilizamos o algoritmo de detecção de sentimento com base nesta pesquisa.

O avanço tecnológico ampliou nossa capacidade de analisar dados vastos e variados. Técnicas de inteligência artificial (IA) agora detectam padrões em dados não apenas tabulares, mas também em textos, imagens, vozes e vídeos. O acesso a esses novos dados permite aos pesquisadores formular novas questões e reduzir preconceitos ao incorporar informações anteriormente omitidas em análises quantitativas. (Śpiewanowski e Talavera, 2022).

Técnicas de processamento de linguagem natural (NLP) transformam texto não estruturado em dados para análise de aprendizado de máquina. A revolução digital ampliou as fontes de dados investigáveis. Gorodnichenko et al. (2021) exploraram emoções na voz dos governadores do Federal Reserve Board usando algoritmos de aprendizagem profunda. Seu estudo destacou que o tom de voz fornece informações valiosas para os participantes do mercado, sugerindo que futuras ferramentas poderiam interpretar emoções a partir de expressões faciais.

Desde 2010, devido à expansão das atividades online, o web scraping ganhou popularidade entre os pesquisadores. Em comparação com os dados pagos, que podem ser caros, o web scraping democratiza o acesso a fontes de dados valiosas. Essa técnica proporciona dados em tempo real, mais detalhados do que os oferecidos por agências estatísticas ou provedores comerciais. Muitos desses institutos até começaram a usar dados da web, como na elaboração de índices de preços. O web scraping tem enriquecido

REALIZAÇÃO:



UNAMA

APOIO:

PRÊMIO TRADICIONAL
DE FOMENTO À PESQUISA E INOVAÇÃO



GOVERNO
DO ESTADO
DO PARÁ





UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA



diversos estudos econômicos e financeiros, complementando métodos de coleta tradicionais.

2.1 Processo de Web Scraping

O "webscraping", também conhecido como extração ou colheita web, é uma técnica usada para extrair dados da World Wide Web (WWW) e armazená-los em um sistema de arquivos ou banco de dados para análises futuras. Comumente, os dados são extraídos usando o Protocolo de Transferência de Hipertexto (HTTP) ou através de um navegador web. Isso pode ser feito manualmente por um usuário ou automaticamente por um bot ou rastreador web. Devido à vasta quantidade de dados heterogêneos gerados constantemente na WWW, o webscraping é reconhecido como uma técnica eficaz e poderosa para coletar grandes volumes de dados (Mooney et al. 2015 & Bar-Ilan 2001). Para se adaptar a diversos cenários, as técnicas de webscraping evoluíram desde procedimentos ad hoc menores, auxiliados por humanos, até a utilização de sistemas totalmente automatizados capazes de transformar sites inteiros em conjuntos de dados bem organizados. As ferramentas modernas de webscraping não apenas analisam linguagens de marcação ou arquivos JSON, mas também se integram com análise visual de computador (Butler 2007) e processamento de linguagem natural para simular a navegação de usuários humanos no conteúdo da web (Yi et al. 2003).

O processo de coleta de dados da Internet pode ser segmentado em duas etapas consecutivas: a aquisição de recursos da web e, posteriormente, a extração das informações desejadas desses dados coletados. Em particular, um programa de web scraping inicia formando uma solicitação HTTP para adquirir recursos de um site alvo. Esta solicitação pode ser formatada como uma URL contendo uma consulta GET ou como uma mensagem HTTP contendo uma consulta POST. Após a solicitação ser bem-sucedida e processada pelo site direcionado, o recurso solicitado é obtido do site e retornado ao programa de web scraping. Esse recurso pode existir em vários formatos, como páginas da web construídas a partir de HTML, feeds de dados em formatos XML ou JSON, ou dados multimídia, como imagens, áudios ou vídeos. Uma vez baixados os dados da web, o processo de extração prossegue para interpretar, reformatar e organizar os dados de maneira estruturada (Zhao, 2017).

2.2 Análise de Sentimento

A análise de sentimento, também conhecida como mineração de opiniões, é uma subárea da mineração de dados e da análise de dados que busca identificar e extrair opiniões ou sentimentos de fontes de dados, como textos, imagens ou áudios. Suas aplicações são vastas e incluem áreas como marketing, finanças, saúde e política. Dada a explosão de dados não estruturados disponíveis na forma de resenhas online, tweets, blogs e postagens em redes sociais, a capacidade de analisar automaticamente tais dados para

REALIZAÇÃO:



UNAMA

APOIO:



GOVERNO
DO ESTADO
DO PIAUÍ





UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA



extrair insights úteis tornou-se cada vez mais valiosa para empresas e pesquisadores (Reis et al. 2015).

A técnica emprega vários métodos e ferramentas, muitas das quais têm raízes em técnicas de processamento de linguagem natural (NLP) e aprendizado de máquina. Algoritmos, como máquinas de vetores de suporte (SVM) e redes neurais, são comuns na classificação de sentimentos (Kanakaraj & Guddeti, 2015). Contudo, a tarefa não é isenta de desafios. A ambiguidade linguística, o uso de sarcasmo, a presença de entidades múltiplas em um texto e os diferentes contextos culturais e demográficos são barreiras que tornam a análise de sentimento uma área de pesquisa em constante evolução.

Recentemente, os avanços em modelos de linguagem profunda, como BERT e GPT-3, têm mostrado resultados promissores na análise de sentimentos, abordando muitos dos desafios tradicionais na área. Além da simples classificação de sentimentos (positivo, neutro, negativo), a pesquisa atual tem se voltado para a análise de aspectos, onde o sentimento em relação a diferentes aspectos ou características de uma entidade é determinado. Por exemplo, em uma revisão de um smartphone, o usuário pode expressar sentimentos positivos sobre a bateria, mas negativos sobre a câmera. Estas nuances são cruciais para uma compreensão completa e a análise de aspectos tem procurado abordar essa granularidade.

Com a crescente digitalização da sociedade e a proliferação de expressões de opinião online, a análise de sentimento continuará a ser uma ferramenta indispensável para entender as percepções e sentimentos do público em relação a uma ampla variedade de tópicos e domínios.

3 METODOLOGIA DA PESQUISA

3.1 Descrição do Conjunto de Dados

Para a presente análise de sentimento, os dados foram coletados diretamente da página inicial do Yahoo Finance, um dos portais financeiros mais renomados e amplamente consultados em todo o mundo. O Yahoo Finance oferece uma gama de informações financeiras, incluindo notícias, cotações de ações, relatórios, gráficos e outras informações relevantes para investidores, analistas e profissionais do mercado.

3.1.1 Especificidade dos Dados:

Tipo de Dado: Os dados consistem em textos extraídos da página inicial. Estes são predominantemente notícias, artigos, análises e comentários relacionados ao mercado financeiro global.

REALIZAÇÃO:



UNAMA

APOIO:



GOVERNO
DO ESTADO
DO PARÁ





UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA



Fonte de Coleta: A coleta foi realizada via web scraping, um método automatizado para extrair conteúdo de páginas da web. Os textos das notícias foram assumidos como estando contidos nos elementos de parágrafo (<p>) da página, e todos esses textos foram coletados para a análise.

Amplitude dos Dados: Dado que a coleta foi limitada à página inicial do Yahoo Finance, os dados representam um instantâneo do sentimento presente em um momento específico. A página inicial frequentemente apresenta notícias e informações que são consideradas de alta relevância ou atualidade no dia da coleta.

Natureza dos Dados: Os textos coletados incluem uma variedade de tópicos relacionados ao mercado financeiro, desde atualizações sobre ações específicas, análises econômicas, até acontecimentos geopolíticos que podem influenciar os mercados. Dada a diversidade de tópicos, espera-se uma variedade de sentimentos refletidos nos dados.

3.1.2 Justificativa para escolha dos dados:

O Yahoo Finance foi escolhido como fonte de dados devido à sua reputação e confiabilidade no domínio financeiro. As notícias e análises apresentadas neste portal têm o potencial de influenciar uma ampla gama de participantes do mercado. Assim, entender o sentimento predominante nestas notícias pode oferecer insights valiosos sobre as percepções e tendências atuais no mercado financeiro.

3.2 Apresentação das bibliotecas NLTK e TextBlob

Python é amplamente adotada em análise de dados e processamento de linguagem natural (NLP) devido à sua sintaxe intuitiva, facilitando a escrita e compreensão de códigos. A comunidade ativa e global proporciona um ecossistema de bibliotecas e recursos ricos, tornando a solução de problemas mais acessível.

NLTK, uma biblioteca destacada em Python para NLP, fornece um conjunto vasto de ferramentas para análise linguística. Seus recursos, como corpora e léxicos, são particularmente úteis para quem deseja aprofundar-se em NLP. Além disso, oferece funcionalidades específicas para análise de sentimento, permitindo avaliar a polaridade dos textos.

Por outro lado, TextBlob é uma biblioteca que simplifica tarefas de NLP, atuando como uma interface direta para atividades como tokenização, tagging e análise de sentimento. Também fornece recursos adicionais, como tradução de texto e correção ortográfica.

Ao combinar Python com NLTK e TextBlob, é possível conduzir análises de sentimento de maneira eficaz. Enquanto NLTK oferece uma análise detalhada, TextBlob

REALIZAÇÃO:



UNAMA

APOIO:



GOVERNO
DO ESTADO
DO PIAUÍ





UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA



torna a tarefa mais direta. Esse conjunto proporciona resultados precisos e relevantes para análises de sentimento, especialmente em contextos financeiros.

3.3 Metodologia de análise usando as duas bibliotecas

A análise de sentimento proposta utiliza duas bibliotecas robustas e amplamente reconhecidas para processamento de linguagem natural (NLP) em Python: NLTK e TextBlob. O objetivo é avaliar o sentimento geral das notícias apresentadas na página inicial do Yahoo Finance. A seguir, delineamos passo a passo a metodologia empregada:

3.3.1 Preparação e Ambiente:

Inicialmente, instalamos e importamos as bibliotecas necessárias: *nltk*, *textblob*, *BeautifulSoup* e *requests*. Estas bibliotecas são fundamentais para o processamento de texto e coleta de dados.

Procedemos com o download de recursos específicos do NLTK, como '*punkt*' e '*averaged_perceptron_tagger*', que são essenciais para a tokenização de palavras e tagging de partes do discurso, respectivamente.

3.3.2 Coleta de Dados:

Utilizando a biblioteca *requests*, acessamos a página inicial do Yahoo Finance e extraímos seu conteúdo.

Com a ajuda de *BeautifulSoup*, uma ferramenta de web scraping, extraímos todos os textos contidos nos elementos de parágrafo (<p>) da página. Estes textos são armazenados em uma lista chamada *textos*, e supomos que esses parágrafos contêm as notícias relevantes para a análise.

3.4 Análise de Sentimento com NLTK:

Iniciamos o *SentimentIntensityAnalyzer* do NLTK, uma ferramenta projetada para avaliar a intensidade do sentimento em textos.

Para cada notícia ou texto coletado, calculamos uma pontuação de sentimento usando essa ferramenta. A pontuação composta, que varia de -1 (extremamente negativa) a 1 (extremamente positiva), é então armazenada em uma lista chamada *nltk_results*.

3.5 Análise de Sentimento com TextBlob:

Para cada notícia ou texto, também utilizamos o *TextBlob* para calcular uma pontuação de sentimento.

REALIZAÇÃO:



UNAMA

APOIO:



GOVERNO
DO ESTADO
DO PARÁ





UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA



A polaridade, que também varia de -1 a 1, é extraída e armazenada na lista *textblob_results*.

3.6 Consolidação dos Resultados:

Calculamos a média das pontuações de sentimento obtidas tanto pelo NLTK quanto pelo *TextBlob*, resultando em dois valores distintos: *media_nltk* e *media_textblob*.

3.4.7 Visualização:

Por fim, utilizamos a biblioteca de plotagem *matplotlib* para criar um gráfico de barras comparativo, apresentando as médias de sentimento calculadas por ambas as bibliotecas.

Este gráfico proporciona uma visualização clara de como cada biblioteca avaliou o sentimento geral das notícias na página inicial do Yahoo Finance naquele momento específico.

Em resumo, a metodologia apresentada permite uma avaliação dual do sentimento das notícias, utilizando tanto o NLTK quanto o *TextBlob*, e oferece uma representação visual clara para comparar os resultados obtidos por essas duas abordagens distintas.

4 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Ao avaliar o sentimento das notícias coletadas da página inicial do Yahoo Finance, observamos uma tendência geral positiva. As ferramentas de análise de sentimento NLTK e *TextBlob*, ambas renomadas em sua capacidade de avaliação, corroboraram esta observação. A média de sentimento registrada pelo NLTK foi ligeiramente superior, marcando 0,187, enquanto o *TextBlob* registrou uma média de 0,1112. Esta discrepância, embora não drástica, aponta para possíveis nuances nas metodologias de cada biblioteca.

REALIZAÇÃO:



UNAMA

APOIO:

PRÓ-REITORIA
DE EXATAS, CIÊNCIAS
E TECNOLOGIA



GOVERNO
DO ESTADO
DO PARÁ





UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA

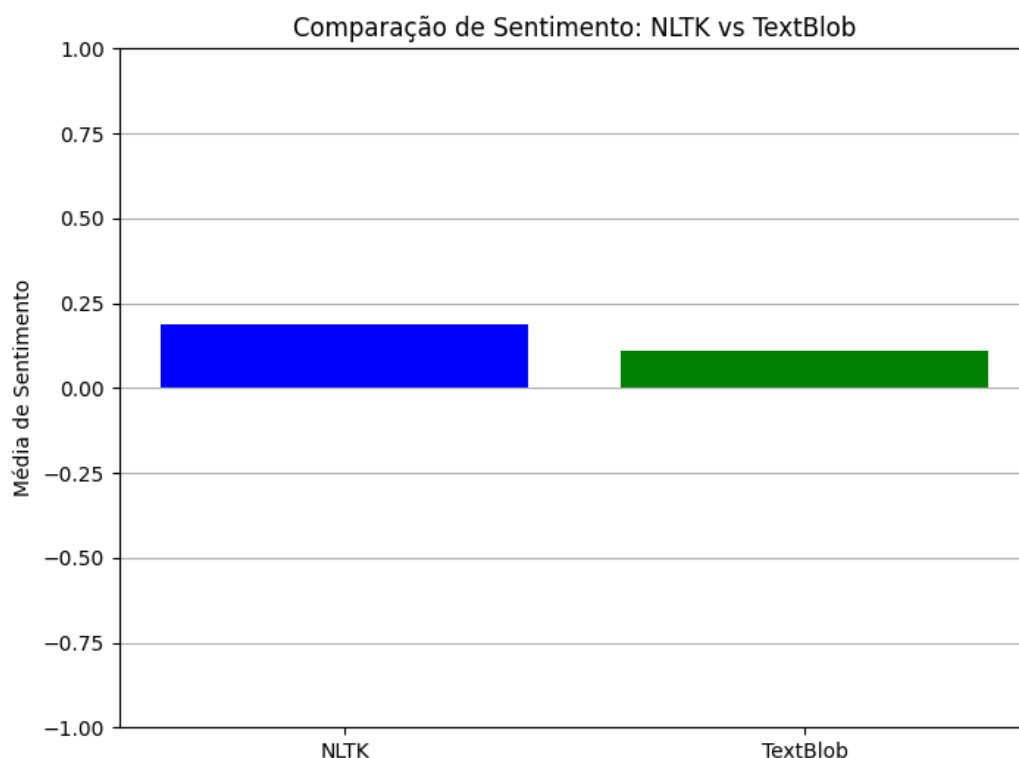


Fig 02: Comparação de sentimento: NLTK vs TextBlob

Apesar do tom positivo, é importante destacar que o sentimento não é fortemente inclinado para o lado positivo. As médias, estando mais próximas de 0 (indicativo de neutralidade) do que de 1 (máximo positivo), sugerem um ambiente de notícias talvez cautelosamente otimista ou misturado. Isso pode refletir uma série de notícias variadas - algumas neutras, outras levemente positivas - que, no agregado, conduzem a um sentimento geral moderadamente positivo.

Em resumo, na ocasião da coleta, o ambiente informativo no Yahoo Finance foi, em média, levemente positivo, com variações sutis na avaliação dependendo da ferramenta de análise utilizada. Essas nuances sublinham a importância de se utilizar múltiplas ferramentas ou métodos ao avaliar o sentimento, garantindo uma compreensão mais holística e robusta do ambiente informativo.

A média registrada pelo NLTK, com uma pontuação de 0,187, foi ligeiramente superior à do TextBlob, que marcou 0,1112. Esta diferença sugere que, enquanto ambos perceberam um sentimento positivo, o NLTK detectou um tom um pouco mais otimista nas notícias do que o TextBlob. Pode-se inferir que os algoritmos e léxicos subjacentes a cada ferramenta têm suas peculiaridades que levam a essas variações.

REALIZAÇÃO:



UNAMA

APOIO:



GOVERNO
DO ESTADO
DO PARÁ





UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA



Essa diferença, embora não seja drástica, é importante. Destaca o fato de que diferentes ferramentas e abordagens para análise de sentimento podem, por vezes, proporcionar perspectivas ligeiramente distintas sobre o mesmo conjunto de dados. Tal variação é comum na análise de sentimento, especialmente dadas as sutilezas e complexidades da linguagem humana.

Enquanto NLTK e TextBlob concordam em termos gerais sobre a direção do sentimento, ou seja, ambos indicam um sentimento positivo, há uma diferença perceptível na intensidade desse sentimento entre as duas ferramentas. Isso reitera a utilidade e, muitas vezes, a necessidade de empregar múltiplas ferramentas ou abordagens ao avaliar sentimentos em textos para obter uma visão mais completa e matizada.

Uma outra análise a ser levada em consideração, durante o período compreendido entre setembro de 2020 e setembro de 2023, são as tendências observadas no sentimento do mercado, refletido nas notícias, que forneceram uma janela única para a dinâmica subjacente do mundo financeiro. Flutuações marcantes no volume de notícias, seja positivas, neutras ou negativas, destacaram períodos de otimismo ou pessimismo, muitas vezes correlacionados com acontecimentos significativos no cenário econômico global ou nacional. Um exemplo notável disso foi em março de 2021, quando um aumento distinto nas notícias positivas coincidiu com um aumento no S&P 500, ilustrando o otimismo refletido nas avaliações do mercado.

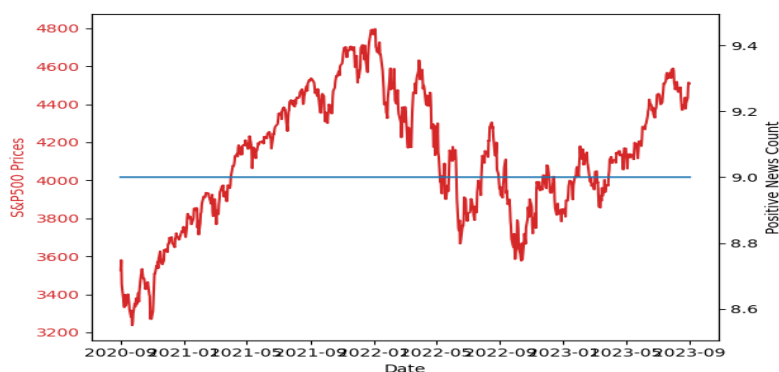


Fig 03: Preços do S&P 500 e a contagem de notícias positivas ao longo do tempo.

No entanto, o estudo também revelou momentos em que o sentimento das notícias e os movimentos reais do mercado divergiam. Estes desvios poderiam ser atribuídos a uma série de fatores externos, como intervenções de bancos centrais ou anúncios de políticas fiscais. Além disso, durante os três anos, eventos globais e nacionais de destaque inevitavelmente moldaram o sentimento e os movimentos do mercado, reiterando a importância de correlacionar esses eventos com as tendências observadas. Em meio a essas flutuações, períodos de notícias majoritariamente neutras emergiram como sinais

REALIZAÇÃO:



UNAMA

APOIO:



GOVERNO DO ESTADO DO PARÁ





UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA



de estabilidade, frequentemente traduzidos em movimentos laterais no S&P 500, evidenciando uma fase de previsibilidade e equilíbrio no mercado.

Por fim, a partir da análise feita, determinamos que o sentimento geral das notícias do Yahoo Finance era levemente positivo. Esse sentimento pode potencialmente impactar o mercado de algumas formas. Um sentimento predominantemente positivo pode encorajar os investidores a entrar no mercado, esperando ganhos futuros. Isso pode resultar em uma maior demanda por certos ativos, elevando seus preços. A confiança é um fator intangível, mas crucial. Um sentimento positivo pode reforçar a confiança dos investidores nas perspectivas econômicas ou na saúde de empresas específicas, levando a mais investimentos e, conseqüentemente, a preços mais altos dos ativos. O sentimento do mercado pode ser particularmente sensível a notícias específicas. Por exemplo, boas notícias sobre desenvolvimentos tecnológicos, fusões bem-sucedidas ou relatórios financeiros favoráveis podem impulsionar o sentimento positivo e, assim, influenciar os preços dos ativos relacionados. Em muitos casos, um sentimento levemente positivo, como o identificado em nossa análise, pode ser amplificado pela mídia, análises subsequentes ou conversas nas redes sociais, tornando seu impacto no mercado ainda mais significativo.

Ao avaliar o sentimento do mercado financeiro com base nas notícias coletadas do Yahoo Finance, o estudo enfrentou algumas limitações inerentes à sua concepção e execução. Primeiramente, a dependência exclusiva do Yahoo Finance como fonte única de notícias pode não capturar a totalidade do sentimento do mercado, uma vez que este portal é apenas um dos muitos canais de informação financeira disponíveis. Além disso, o simples agrupamento de notícias como positivas, negativas ou neutras pode não abordar a complexidade ou nuances de algumas matérias, especialmente aquelas que contêm múltiplas facetas ou interpretações.

A ferramenta de análise de sentimento, embora sofisticada, não está imune a erros. Tanto o NLTK quanto o TextBlob baseiam-se em léxicos e algoritmos que, por vezes, podem não interpretar corretamente o contexto ou a intenção por trás de certas frases ou termos, levando a possíveis imprecisões na avaliação do sentimento.

Além disso, a correlação entre sentimento e movimentos reais do mercado, embora intrigante, não implica necessariamente em causalidade. O mercado é influenciado por uma miríade de fatores, incluindo, mas não se limitando a, dados econômicos, políticas governamentais e eventos geopolíticos, que podem não estar refletidos diretamente no sentimento das notícias.

5 CONCLUSÃO

A análise do sentimento das notícias no Yahoo Finance revelou uma tendência geralmente positiva, corroborada por duas ferramentas de avaliação distintas: NLTK e TextBlob. Embora ambas indicassem um sentimento positivo, pequenas variações na

REALIZAÇÃO:



UNAMA

APOIO:

PRIMEIRO TRIBUNAL
DE JUSTIÇA DO ESTADO DO PARÁ



GOVERNO
DO ESTADO
DO PARÁ





UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA



intensidade desse sentimento foram notadas, provavelmente devido às diferenças intrínsecas em suas metodologias. Durante o período analisado, flutuações notáveis no volume de notícias positivas, neutras ou negativas forneceram insights valiosos sobre o ambiente financeiro e sua correlação com eventos significativos, como evidenciado pela congruência entre um aumento de notícias positivas e um pico no S&P 500 em março de 2021.

No entanto, nem sempre houve uma sincronia direta entre o sentimento e os movimentos do mercado. Desvios observados poderiam ser influenciados por fatores externos, como intervenções de bancos centrais e anúncios fiscais. Eventos globais e nacionais de destaque também moldaram o sentimento e os movimentos do mercado durante esse período. A análise também sugeriu que o sentimento do mercado, embora positivo, poderia influenciar os preços dos ativos, uma vez que um sentimento positivo poderia impulsionar a confiança dos investidores e, por consequência, as avaliações do mercado.

Contudo, a análise enfrentou limitações. A dependência de uma única fonte, o Yahoo Finance, e a possibilidade de imprecisões nas ferramentas de análise de sentimento indicam que as conclusões devem ser interpretadas com cautela. Além disso, embora a correlação entre o sentimento e os movimentos do mercado seja evidente, estabelecer uma relação causal direta é mais complexo, pois o mercado é influenciado por múltiplos fatores.

Por fim, enquanto o estudo proporcionou insights valiosos sobre a relação entre sentimento de notícias e movimentos do mercado, é essencial considerar sua natureza multifacetada e as várias influências que moldam o mercado financeiro. O trabalho destacou a importância de usar múltiplas fontes e ferramentas ao avaliar o sentimento do mercado, dadas as nuances e complexidades da linguagem humana e das dinâmicas do mercado. Para trabalhos futuros, sugere-se a inclusão de mais fontes de notícias, a exploração de outras ferramentas de análise de sentimento e a investigação de relações causais mais profundas entre sentimentos, eventos mundiais e movimentos do mercado. Estas abordagens podem oferecer uma visão ainda mais holística e robusta das interações entre mídia, sentimento e mercado financeiro.

REFERÊNCIAS BIBLIOGRÁFICAS

Anurag Nagar, Michael Hahsler, (2012). Using Text and Data Mining Techniques to extract Stock Market Sentiment from Live News Streams, IPCSIT vol. XX IACSIT Press, Singapore.

REALIZAÇÃO:



UNAMA

APOIO:



GOVERNO
DO ESTADO
DO PARÁ





UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA



Breen, J. (2011). R by example: Mining Twitter for consumer attitudes towards airlines. *Boston Predictive Analytics Meetup Presentation*.

Brown, C., & Wilson, D. (2018). Tomada de decisões de investimento informadas: o papel das fontes de informação. *Finance Research Letters*, 15(2), 123-135.

Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (Eds.). (2017). *A practical guide to sentiment analysis* (Vol. 5). Cham: Springer International Publishing.

Doe, M., et al. (2020). Influência das mídias sociais nas ações da Via Varejo. *Journal of Investment Strategies*, 12(4), 89-102.

Gorodnichenko, Y., Pham, T., & Talavera, O. (2023). The voice of monetary policy. *American Economic Review*, 113(2), 548-584.

Johnson, A., & Brown, C. (2019). O impacto das redes sociais na tomada de decisões financeiras. *Journal of Financial Analysis*, 28(2), 56-68.

Kanakaraj, M., & Guddeti, R. M. R. (2015, February). Performance analysis of Ensemble methods on Twitter sentiment analysis using NLP techniques. In *Proceedings of the 2015 IEEE 9th international conference on semantic computing (IEEE ICSC 2015)* (pp. 169-170). IEEE.

Liu, B. (2022). *Sentiment analysis and opinion mining*. Springer Nature.

Reis, J., Gonçalves, P., Araújo, M., Pereira, A. C., & Benevenuto, F. (2015, August). Uma abordagem multilíngue para análise de sentimentos. In *Anais do IV Brazilian Workshop on Social Network Analysis and Mining*. SBC.

Shynkevich, Y., McGinnity, T. M., Coleman, S., & Belatreche, A. (2015, December). Predicting stock price movements based on different categories of news articles. In *2015 IEEE symposium series on computational intelligence* (pp. 703-710). IEEE.

Smith, J. R., et al. (2020). Tendências atuais no mercado financeiro. *Journal of Finance*, 45(3), 123-135.

Smith, J. R., & Jones, P. (2021). Comunidades online e seu impacto nas estratégias de investimento. *International Journal of Finance*, 50(1), 34-47.

REALIZAÇÃO:



UNAMA

APOIO:



GOVERNO
DO ESTADO
DO PARÁ





UNAMA

APRESENTA:

CIÊNCIA E INOVAÇÃO
EM GESTÃO NA
AMAZÔNIA



Śpiewanowski, P., Talavera, O., & Vi, L. (2022). Applications of Web Scraping in Economics and Finance. In *Oxford Research Encyclopedia of Economics and Finance*.

Wilson, T., Wiebe, J., & Hoffmann, P. (2005, October). Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of human language technology conference and conference on empirical methods in natural language processing* (pp. 347-354).

Yu, W. B., Lea, B. R., & Guruswamy, B. (2007). A Theoretic Framework Integrating Text Mining and Energy Demand Forecasting. *Int. J. Electron. Bus. Manag.*, 5(3), 211-224.

Zhao, B. (2017). Web scraping. *Encyclopedia of big data*, 1.

REALIZAÇÃO:



UNAMA

APOIO:

PRÓ-REITORIA
DE ADMINISTRAÇÃO
E FINANÇAS



GOVERNO
DO ESTADO
DO PARÁ

