

# LICEU *on-line*

**FINBERT BRASILEIRO**

**BRAZILIAN FINBERT**

**Paulo de Souza Knupp**

Universidade Federal de Santa Catarina.

[paulo\\_059@hotmail.com](mailto:paulo_059@hotmail.com)

Brasil

**Wilson Toshio Nakamura**

Universidade Presbiteriana Mackenzie

[wtnakamura@uol.com.br](mailto:wtnakamura@uol.com.br)

Brasil

**Carmen Campos Pereira**

[carmemcpereira@gmail.com](mailto:carmemcpereira@gmail.com)

Brasil

Recebido: 25/09/2023 – Aprovado: 13/02/2024. Publicado Julho/2024.

Processo de Avaliação: Double Blind Review.

## **RESUMO**

O presente artigo investigou o uso da inteligência artificial para, por meio de análise de notícias de jornais, informar qual foi a movimentação da bolsa de valores, B3. Para tanto, a ferramenta de inteligência artificial chamada BERTimbau foi treinada especificamente para esse propósito. Ao final, a ferramenta desenvolvida foi comparada com outras formas de análise e classificação de textos, também chamado de classificação de textos. O resultado aponta que o modelo baseado no BERTimbau teve maior sucesso na classificação no movimento da B3, com base nas notícias disponíveis.

**Palavras-chave: inteligência artificial; classificação de texto; análise do mercado**

## **ABSTRACT**

The present article investigated the use of artificial intelligence to inform about the movement of the stock market, B3, through the analysis of newspaper news. For this purpose, the artificial intelligence tool called BERTimbau was specifically trained. In the end, the developed tool was compared with other forms of text analysis and classification, also known as text classification. The result indicates that the model based on BERTimbau achieved greater success in classifying the movement of B3, based on the available news.

**Keywords: artificial intelligence; text classification; market analysis**

## **1. INTRODUÇÃO**

Não é incomum se deparar com a máxima de que o mercado reage às notícias. Aliás estar em posse de uma informação que poucos possuem é uma posição privilegiada no mercado de ações. Tão privilegiada que existem mecanismos que visam diminuir ao mínimo a disparidade de informações entre os participantes do mercado.

A ideia desses meios de controle é que todos tenham acesso as mesmas informações simultaneamente. Uma ferramenta importante para isso é a internet, que possibilita a divulgação dos fatos quase que instantaneamente. Dessa forma, um participante do mercado de ações deve estar atento às notícias. Pelo menos é isso o que pressupõem a tão conhecida teoria dos mercados eficientes.

A presente pesquisa buscou verificar essa premissa no mercado de ações do Brasil, por meio da observação das movimentações do seu principal índice, o Ibovespa. Para tanto, focou-se, em um primeiro momento, em ferramentas de inteligência artificial. O ponto de partida foi: Se os mercados são eficientes, ou apresentam algum nível de eficiência, será

possível treinar um algoritmo para ler as notícias e indicar ao investidor qual é o humor do mercado naquele exato momento, ou seja, se o mercado está subindo ou diminuindo.

A principal ferramenta utilizada durante esse projeto foi o modelo conhecido como Bert, que é uma abreviação para *Bidirectional Transformers for Language Understanding* (DELVIN et al 2019). Essa ferramenta, desenvolvida por pesquisadores do Google, tem uma capacidade inovadora para entender textos. Dentre as suas características que garantem esse título está o fato de ser bidirecional, ou seja, consegue associar e prever, não só a próxima palavra do texto, mas também a anterior, o que deixa o algoritmo mais preciso.

Dado o fato desse modelo ser aberto, pesquisadores brasileiros desenvolveram o BERTimbau (SOUZA, et al. 2020), ou seja, adaptaram o já existente modelo Bert para a língua portuguesa. Dessa forma, tornou-se possível desenvolver ferramentas sofisticadas para análise de textos em português, tais como a definição do humor do texto.

E foi essa oportunidade que este trabalho quis alcançar, utilizar uma ferramenta sofisticada e especializada na leitura de textos da área financeira, como o conhecido modelo FINBERT (ARACI, 2019).

Dessa forma, elaborou-se a seguinte pergunta de pesquisa:

- É possível treinar um algoritmo baseado no modelo BERTimbau para a interpretação de notícias financeiras associadas com a movimentação do mercado (Ibovespa)?

Para responder a essa pergunta desenvolvemos como objetivo principal:

- Treinar um algoritmo baseado no modelo BERTimbau para a interpretação de notícias financeiras associadas com a movimentação do mercado (Ibovespa)

E os objetivos específicos foram:

- Desenvolver um corpus, na língua portuguesa, de textos da área financeira;
- Desenvolver programas para a coleta de manchetes da área financeira de diferentes jornais online;
- Desenvolver códigos para o treinamento do algoritmo;
- Comparar os resultados obtidos por meio do BERTimbau com outras ferramentas; e
- Definir o nível de eficiência do mercado acionário nacional.

A seguir será apresentado uma revisão bibliográfica, seguida da descrição do método e dos resultados observados.

## **2. REVISÃO BIBLIOGRÁFICA**

Nesta seção serão apresentados trabalhos sobre os aspectos que nortearam a pesquisa. Em primeiro lugar, será apresentado a hipótese dos mercados eficientes, que é a base adotada nesta pesquisa para explicar o comportamento dos tomadores de decisão no mercado. Em seguida será discutido aspectos das pesquisas que utilizam análise de sentimentos em trabalhos da área financeira.

### **2.1 Hipótese dos Mercados Eficientes**

A hipótese dos mercados eficientes é, talvez, um dos conceitos mais relevantes da teoria moderna de finanças. De acordo com essa hipótese o preço de um ativo irá refletir todas as informações disponíveis para os investidores, dessa forma tanto as pessoas físicas quanto as jurídicas envolvidas nas transações devem esperar que o preço do ativo reflita de forma justa o valor presente do fluxo de caixa futuro da empresa (ROSS, WESTERFIELD e JAFFE, 2010).

Ainda de acordo com essa proposição, o mercado pode apresentar uma de três formas de eficiência, sendo elas a i) fraca; ii) semiforte; e iii) forte. Como é possível deduzir pelos nomes, cada uma das formas aponta para um grau de como as informações são incorporadas nos preços dos ativos, no sentido de que na forma fraca as informações disponíveis no mercado têm pouco impacto na formação dos preços dos ativos, enquanto que na forma forte todas as informações disponíveis são incorporadas nesses preços (ROSS, WESTERFIELD e JEFF, 2010).

Assim, diversos estudos se aprofundam sobre o tema buscando testar essa hipótese, ou verificar em qual forma o mercado estudado se enquadra, como é o caso de Sanvicente e Monteiro (2005) que observam no mercado o “efeito pressão sobre preços” das ações, levando-os a afirmar que a formação de preço desses ativos não é inteiramente eficiente. Nesse mesmo sentido, Antunes, Lamounier e Bressan (2006) também apontam que o mercado de ações brasileira tem características de um mercado semiforte, dado que informações contábeis da empresa relacionados ao tamanho, possuem coeficientes de correlações baixos, porém significativos em relação ao valor das ações.

Outro trabalho mais recente e que corrobora com as conclusões do parágrafo anterior foi realizado por Evard e Cruz (2016). Esses autores observaram, porém, maior

ineficiência no mercado de ações brasileiro. Os pesquisadores não encontraram correlação negativa entre risco e retorno o que, de acordo com os autores indicaria a ineficiência.

Contudo outros pesquisadores tais como Fioresi e Galdi (2018) e Ely (2011) observaram o mercado brasileiro como mais eficiente. Assim observamos que ainda existem contradições sobre a eficiência do mercado acionário brasileiro.

Por fim, o trabalho que tratou esse assunto e motivou a elaboração desse projeto foi desenvolvido por Galdi e Gonçalves (2018). Esses analisaram como o tom das notícias do jornal Valor Econômico está correlacionado com o preço das ações no dia da publicação do jornal.

Esses pesquisadores não encontraram uma correlação entre o tom das notícias e o valor das ações, contudo isso pode ter ocorrido pela metodologia adotada por esses autores, que abrangeu apenas as notícias gratuitas de um jornal em um período relativamente curto.

No próximo capítulo será discutido melhor os aspectos relacionados a um tom de um texto e como isso é utilizado nos estudos em finanças.

## **2.2 Análise de Sentimento**

Pesquisas na área de Finanças e da Contabilidade-Financeira passaram a utilizar os textos como uma fonte de informações para se entender melhor a tomada de decisões dos investidores, tal como na situação de compra ou venda de uma ação. Dessa forma, os pesquisadores que se interessam por essa linha de análise, buscam quantificar e classificar os textos de acordo com o seu “Tom”. Essa análise pode ser denominada como “análise do sentimento”. Ou seja, os pesquisadores buscam classificar os textos como tendo um tom (sentimento) positivo, neutro ou negativo, por exemplo. Nas palavras de Oliveira et al. (2019; p. 235) a análise de sentimentos é uma técnica que busca “revelar a opinião das pessoas sobre temas específicos”.

De acordo com a revisão de literatura realizada por Ye e Li (2017) a análise de sentimento é utilizada para mensurar o humor dos investidores. Esses autores observaram que as principais técnicas utilizadas pelos pesquisadores dessa área são a “*machine learning*” e a elaboração de dicionários que buscam classificar as palavras. Ainda de acordo com essa pesquisa os trabalhos usam a análise de sentimento para prever a movimentação do mercado acionário, contudo os resultados encontrados não são conclusivos devido, provavelmente, as limitações da amostra e da análise dos métodos utilizados.

A pesquisa realizada por Fernandez (2019) está entre os trabalhos dessa linha que encontrou uma relação entre o sentimento analisado e a variação no preço das ações. Esse pesquisador analisou os sentimentos dos textos no Twitter para verificar se esses refletiam a junção de duas empresas na Inglaterra. Os resultados indicaram que o Twitter poderia ser utilizado para observar eventos financeiros relevantes.

Nessa mesma direção Junior et al (2019) também observaram que a variação nos preços das ações poderia ser explicada pelos sentimentos detectados em textos, sendo que nesse trabalho os textos analisados foram os relativos às *Conference Calls* trimestrais das empresas. Resultado semelhante foi alcançado por Machado e Silva (2017) que observaram que os sentimentos expressos nos relatórios trimestrais das empresas se refletem nas movimentações do mercado.

Outro trabalho que também verificou que o tom dos textos impacta um aspecto financeiro das empresas foi o realizado por Song, Wang e Zhu (2018), contudo esses autores analisaram o impacto dos sentimentos dos relatórios financeiros sobre o índice CSR (*corporate social responsibility*). O resultado dos pesquisadores apontou que a análise dos sentimentos tinha um nível de 86% de acerto na previsão do CSR, ou seja, é uma ferramenta viável para esse fim.

Contudo, como alertaram Ye e Li (2017), não há consenso nos resultados das pesquisas sobre esse tema. A exemplo disso podemos citar os resultados de Lee et al (2018) que apontam que o tom dos relatórios financeiros não tem relação com a performance de vendas das empresas. Ainda nessa linha contrária está a pesquisa de Galdi e Gonçalves (2018) que não observaram uma relação significativa entre o tom das notícias de jornal e a movimentação da bolsa de valores.

Para explicar esse distanciamento entre os resultados, voltamos a recorrer a Ye e Li (2017) que expõe que essa divergência pode ocorrer devido às limitações metodológicas. Isso pode ser exemplificado pelas afirmações de Souma, Vodenska e Aoyama (2019) que melhoram o desempenho do seu modelo com um pequeno ajuste na metodologia, mais especificamente na maneira em que consideravam uma notícia positiva ou negativa. O mesmo pode se dizer sobre o trabalho de Renault (2019), que aponta uma melhoria nas previsões do modelo quando fazem alterações no tamanho da base de dados.

Visto que pequenas alterações na metodologia podem trazer diferenças relevantes nos resultados, os resultados desta pesquisa podem ser somados aos alcançados pelos artigos produzidos por Carosia, Coelho e Silva publicados em 2020 e 2021. Em ambos os artigos os pesquisadores utilizam inteligência artificial para verificar se o humor de textos pode ser

utilizados para a elaboração de estratégias de investimentos, sendo que em ambos eles utilizam banco de dados que contem textos do Twitter, sendo que no segundo artigo eles acrescentam a esses textos, frases traduzidas e classificadas no trabalho de Martin et al (2015), assim como textos de jornais. Os resultados dos artigos mostram grande potencial do uso da Inteligência artificial para o mercado brasileiro.

Contudo, de acordo com o último artigo desses autores, o modelo BERT deve ser testado, pois há a expectativa de resultados melhores com ele (Carosia, Coelho e Silva, 2021). E é exatamente isso o que essa pesquisa traz.

### **3. MÉTODO**

Para a realização desta pesquisa optou-se pela aplicação da técnica conhecida como análise de sentimentos. Essa técnica permite, por diferentes meios, a análise e classificação dos textos. O mais comum é encontrar análises que buscam uma classificação binária, ou seja, entre dois valores, como é o caso da presente pesquisa. Contudo, também é possível encontrar pesquisas que analisam uma gama maior de humor.

O grande objetivo desta pesquisa foi o de treinar um algoritmo a partir do modelo Bert para identificar, a partir de textos de jornais, o humor do mercado. As etapas para se alcançar esse objetivo serão apresentadas nos subtópicos a seguir.

#### **3.1 Coleta dos Dados**

A primeira etapa foi a coleta de dados para o desenvolvimento de um corpus (conjunto de textos). Além disso também se buscou coletar as movimentações do índice Bovespa. A coleta desses dois itens, assim como a associação deles, foi crucial para o desenvolvimento de um banco de dados que permitisse um aprendizado supervisionado, ou seja, o sistema sabia, ao final do processo, como cada texto deveria ser classificado.

Dessa forma os dados foram organizados em dias e horas, o que permitiu o desenvolvimento de dois bancos de dados. No primeiro as manchetes publicadas ao decorrer de todo o dia foram agrupadas em um único texto, e esse recebia uma única classificação que retratava o movimento do Ibovespa durante o dia de negociações, sendo ‘1’ para quando esse índice subisse e ‘0’ no caso contrário. Chamamos esse banco de dados de ‘Dia’

Já no segundo banco de dados foram utilizadas apenas as manchetes publicadas durante o período das negociações da bolsa, ou seja, ao contrário do banco de dados Dia, nesse novo banco de dados as informações não englobam manchetes publicadas antes ou depois do funcionamento da bolsa. Os dados dessa segunda coleta também foram agrupados, contudo, formaram dois grupos para cada dia. No primeiro foram adicionados apenas as notícias relacionadas a movimentos positivos na bolsa e no segundo apenas notícias negativas. Esse segundo banco de dados chamamos de ‘IntraDia’.

O uso desses dois bancos de dados se fez necessário, devido aos alertas apresentados na revisão bibliográfica, os quais indicam que simples mudanças na metodologia podem influenciar no sucesso ou fracasso da pesquisa.

### **3.2 Fonte dos Dados**

As manchetes foram coletadas a partir de dois sites de grandes jornais brasileiros: Estadão e Folha de São Paulo. Foram coletadas manchetes ligadas aos editoriais de economia e finanças de cada um dos jornais. Observa-se que as duas coletas foram realizadas de forma automatizada a partir de um programa escrito pelo pesquisador utilizando, principalmente, as bibliotecas *Beautiful Soup*, *Requests* e *selenium*.

Destaca-se que os dados coletados do jornal Estadão foram destinados para o treinamento dos modelos, enquanto que os dados coletados do jornal Folha de São Paulo foram destinados para testar os modelos, com exceção, do Modelo I, que usou apenas os dados obtidos a partir do Estadão, Sendo tendo 15% do seu conteúdo retirado da etapa de treinamento e destinado para o teste de acurácia do modelo.

Já em relação às cotações foram usadas duas fontes para os dados. Para as movimentações diárias foi utilizado os dados do portal *Yahoo Finanace*, enquanto que para as movimentações *intraday* foi utilizado o sistema ValorPro. Os dados *intraday* foram coletados a cada 30 minutos de forma que as notícias publicadas neste intervalo eram juntadas e associadas a movimentação do mercado desse mesmo período

Esses dados foram organizados e tabulados por meio das ferramentas *Pandas* e *Numpy*, tudo isso por meio do *Python*. Em certos momentos os dados também foram trabalhados no Excel, mas apenas para facilitar a observação dos textos.

Com os dados organizados, passou-se para a etapa de treinamento dos modelos.



### **3.3 Treinamento dos Modelos**

Como já exposto, o foco desta pesquisa foi o treinamento de um modelo baseado na tecnologia Bert, porém, aproveitou-se a oportunidade para o estudo de outros dois modelos a partir dos mesmos bancos de dados. Dessa forma, foram realizados 6 treinamentos distintos.

O primeiro modelo é também o mais simples e aquele que consome menos tempo para ser treinado. Trata-se de um modelo baseado na regressão logística. Para o treinamento deste modelo de regressão, o texto deve passar por etapas de pré-tratamento, visando retirar palavras que agregam pouca informação para a inferência a ser feita pelo modelo. Essas são chamadas de *stopwords*. Além disso também são retirados acentos, pontuações e outros caracteres especiais. Após essa primeira etapa de preparação os dados são transformados em números e guardados e uma variável chamada de *bag of words*, que nada mais é do que uma “sacola” contendo todas as palavras presentes nas manchetes coletadas.

Realizado essa preparação, aplicou-se a técnica *TF IDF* que dá um valor para cada palavra de acordo com a frequência em que ela aparece no texto, de forma que palavras muito frequentes têm menor relevância do que as palavras usadas muitas vezes.

Em adição a isso também foi aplicado o método *Ngrams* o qual permite que as palavras sejam analisadas em duplas, no caso do nosso trabalho. Assim a palavra é observada pelo sistema junto com a sua vizinha.

O segundo modelo estudado, já é um modelo mais sofisticado e que exige um maior poder de processamento. Utilizou-se um pacote chamado *spacy*, que se baseia na técnica da rede neural convolucional. Ao se utilizar essa rede o pesquisador também deve preparar o texto, retirando palavras pouco relevantes, assim como pontuações e outros caracteres especiais. Após esse processo os dados estão prontos para serem enviados ao sistema que irá “treinar” buscando diminuir os erros das suas previsões.

Por fim, o último modelo a ser treinado é o Bert que é uma sigla para *Bidirectional Transformers for Language Understanding*. Esse modelo, apesar de ser pré-treinado é o que mais exige capacidade de processamento computacional.

Como dito o Bert é um modelo pré-treinado, cabendo ao usuário final, treinar as últimas camadas da sua rede. Observa-se que é possível treinar o modelo a partir do zero, contudo, isso exigiria um corpus muito grande, além de uma capacidade de processamento extravagante.

É relevante informar que o ambiente utilizado para o treinamento desses modelos foi o Google Colab Pro. Isso foi necessário para agilizar o treinamento do modelo Bert, por meio

de uma GPU com maior capacidade de processamento do que as de computadores comuns. Entretanto há desvantagens no uso dessa plataforma. A primeira delas é que a empresa requer que o sistema receba sinais de tempos em tempos de que o usuário está presente. Dentre esses sinais estão a movimentação de mouse, digitação de palavras, etc. Caso o sistema fique muito tempo sem um sinal, ele para de processar. Além disso, só é possível utilizar o sistema por 24 horas seguidas. Ou seja, sistemas muito grandes com grande necessidade de treinamento podem ser cancelados ao fim desse prazo.

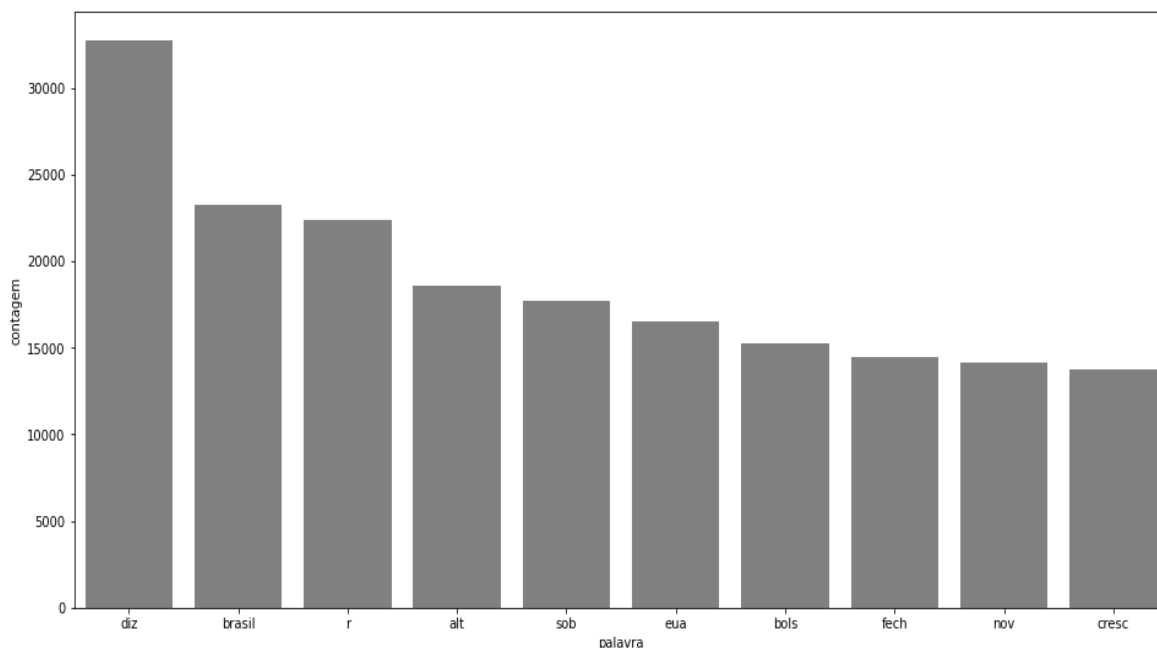
Destaca-se também que o sistema exige que o usuário esteja conectado a internet. A instabilidade dessa ou sua ausência podem interromper a execução dos treinamentos.

#### **4. RESULTADOS OBSERVADOS**

Foram coletadas, no total 33.918 manchetes do site do Estadão de um período que se inicia no mês de março de 2000 e termina no mês de setembro de 2021. Observa-se que essas manchetes são aquelas publicadas em dias úteis, ou seja, são as manchetes publicadas nos dias em que houveram negociações de ações na bolsa de valores de São Paulo. As manchetes publicadas durante os finais de semana e feriados foram ignoradas da amostra. Em relação aos dados de validação do modelo, foram coletadas 13.151 manchetes publicadas online pelo Jornal Folha de São Paulo no mesmo período.

A figura 1 apresenta as palavras mais comuns no corpus usado para o treinamento dos modelos. Nesse gráfico estão as palavras que se mantiveram no corpus após a retirada de *stopwords*, caracteres especiais e números. É relevante informar que nessa figura estão apenas os radicais das palavras. Por exemplo, o ‘diz’, engloba diferentes palavras como o próprio ‘diz’ e o ‘dizem’. Outro ponto de destaque é que o “r” se refere ao “R\$”. O “\$” não aparece na figura, pois foi retirado na etapa de preparação do texto, uma vez que é um caractere especial.

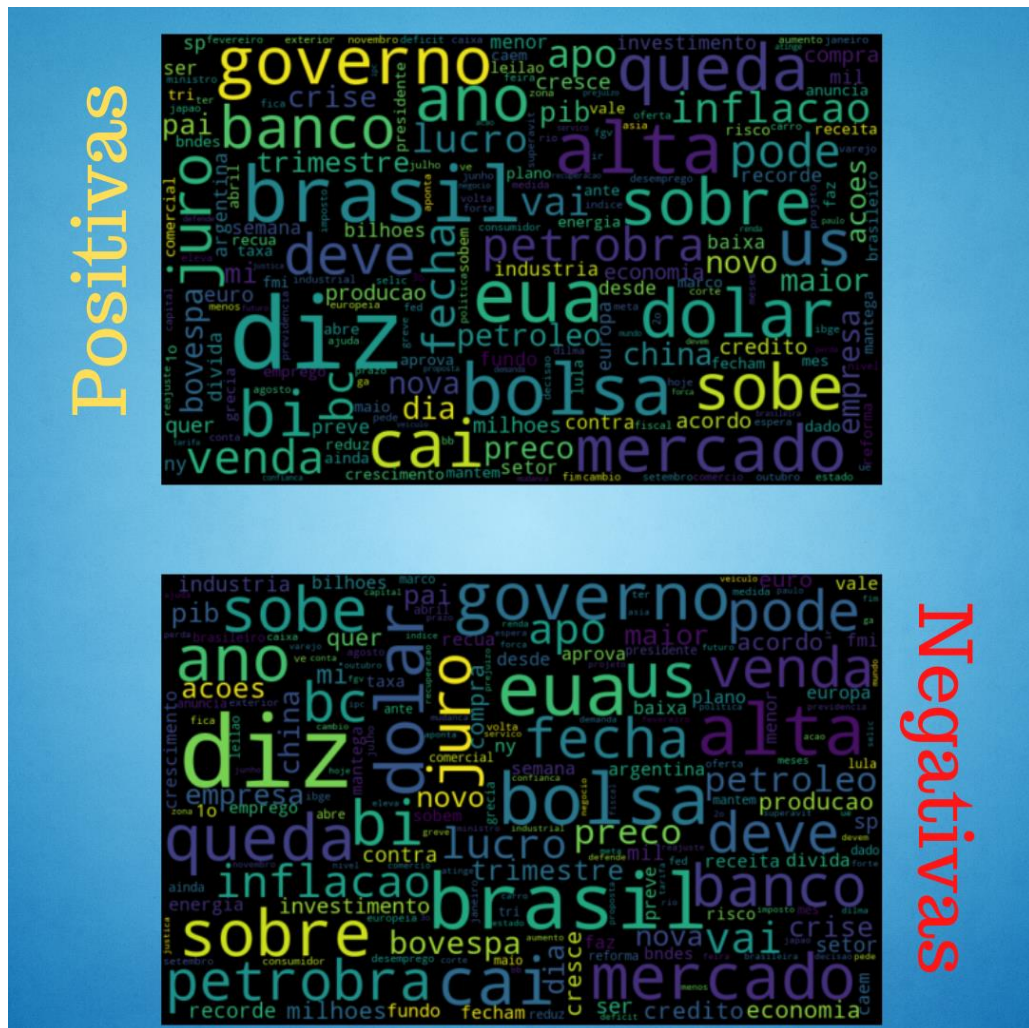
**Figura 1:** Palavras mais comuns no corpus de treinamento



Por meio desse gráfico podemos deduzir que as notícias comumente tratam da fala de uma pessoa, já que o ‘diz’ é a palavra mais comum no corpus. Outro ponto de destaque é o foco nas quantidades monetárias em reais, visível pela presença do ‘r’ que se refere ao ‘R\$’. Também notamos a importância dada pela imprensa aos EUA, uma vez que esse país aparece na sexta posição, antes mesmo termo referente à Bolsa.

A figura 2, adicionada a seguir, nos permite observar quais são as palavras mais comuns classificadas nos textos positivos e negativos do corpus de treinamento. É importante frisar que, apesar das figuras 1 e 2 apresentarem informações relevantes sobre o que é tratado nas manchetes jornalísticas sobre economia e finanças, as palavras mais comuns são aquelas com menores pesos para o nosso modelo de regressão logística, que utilizou a técnica *TF IDF*.

Figura 2: Palavras mais comuns nos textos classificados como positivos e negativos



#### 4.1 Resultados dos Modelos

A tabela 1 apresenta o resultado de cada modelo para cada tipo banco de dados. O que se destaca nesta tabela é o fato de que o banco de dados IntraDia apresentou uma acurácia próxima a 50% em todos os modelos. Ou seja, agrupar as notícias publicadas durante o pregão em boas e ruins para o treinamento dos modelos não apresentou um resultado muito melhor do que o de se jogar uma moeda para cima e sortear cara ou coroa para se saber o humor do mercado.

**Tabela 1:** Acurácia dos modelos para cada base de dados.

Acurácia dos modelos para cada base de dados		
Modelos	Base de dados	
	Dia	IntraDia
Regressão Logística	64,89%	51,03%
Rede neural convolucional	64,64%	50,65%
Bert	86,49%	50,82%

A tabela 2 joga uma luz sobre o motivo da diferença entre os dois bancos de dados usados nesta pesquisa.

Peso dos termos no modelo de regressão logística			
Positivo		Negativo	
Palavras	Peso	Palavras	Peso
fech alt	2.82	bovesp cai	-2.39
bovesp sob	2.76	dol sob	-2.28
bols sob	2.5	fech qued	-2.2
dol cai	2.03	bols cai	-2.18
ibovesp sob	1.52	ibovesp cai	-1.39
sob dol	0.98	bols recu	-1.24
dol recu	0.9	fech baix	-0.94
mil pont	0.82	cai dol	-0.9
ny sob	0.67	bovesp recu	-0.86
oper alt	0.64	qued dol	-0.81

Nota-se pela tabela 2 que os termos mais relevantes para a classificação do humor do mercado são aqueles que descrevem o que já ocorreu, como o “fech alt” que representa “fechou em alta”.

Analisando conjuntamente os resultados das tabelas 1 e 2 podemos verificar que é difícil para um investidor saber qual é o movimento atual da bolsa de valores apenas pelas manchetes das notícias publicadas na internet pelos jornais.

Ainda sobre a tabela 2 é importante apontar a relevância do mercado dos EUA sobre o brasileiro, demonstrada pelo termo “ny sob”, “dol cai” e “dol sob” que tratam respectivamente de “bolsa de NY sobe”, “dólar cai” e “dólar sobe”. Isso reforça o que já foi apontado na figura 1 onde o termo EUA foi o sexto mais usado nas manchetes sobre a área econômica/financeira.

Voltando à tabela 1 é possível notar que não há uma diferença grande entre o modelo baseado na regressão logística e o desenvolvido a partir de uma rede neural convolucional. Ambos foram capazes de acertar, aproximadamente, 65% das vezes. Diante desse fato, indica-se o uso do modelo baseado na regressão logística, somado as técnicas de *TF IDF* e *Ngrams*. Esse modelo, apesar de exigir um pré-tratamento dos dados, também é o que necessitou um menor tempo de treinamento e de capacidade de processamento. Nesse sentido, o mais simples foi o mais adequado para atender aos objetivos desta pesquisa.

Por fim, o modelo de maior sucesso, como o esperado, é o modelo treinado a partir da técnica mais robusta. O nosso modelo, baseado no BERTimbau que por sua vez é adaptado do Bert, apresentou uma acurácia de 86,49%, aproximadamente 21 pontos percentuais acima dos seus concorrentes nesta pesquisa e muito próximo do melhor modelo desenvolvido por Carosia, Coelho e Silva (2021), que usaram uma base de dados muito mais extensa para treinar os seus modelos.

Confirmando o sucesso do modelo, ele também apresentou um bom valor para a correlação de Matheus (0,7294). Assim, mesmo considerando a grande exigência de tempo e recursos para o processamento e treinamento deste modelo, o seu resultado justifica o investimento nesses pontos.

Dessa forma, podemos afirmar que temos um FinBertMackenzista, que ainda pode ser aperfeiçoado, porém já apresenta resultados que nos permite utilizá-lo em nossas pesquisas.

## **5. CONCLUSÃO**

Essa pesquisa visou analisar a eficiência do mercado acionário brasileiro por meio do treinamento de um modelo de Inteligência Artificial. O preceito é simples, caso o mercado possua alguma eficiência é possível criar um modelo que retrate o seu humor. Contudo os resultados nesta pesquisa são desafiadores para a realização de uma interpretação definitiva pois, por um lado conseguimos bons resultados em modelos com uma acurácia que vão de 65% para os modelos mais simples para até 86% para o modelo mais sofisticado treinado

aqui. Mas, por outro lado, observamos que se esses modelos forem treinados apenas com notícias publicadas durante o pregão, o nível de acerto de todos os modelos cai para cerca de 50%.

Isso nos leva na direção de que os modelos são capazes de entender o mercado caso sejam alimentados com manchetes publicadas antes da abertura ou após o fechamento do pregão, como é possível ver na tabela 2. Dessa forma sugere-se que novas pesquisas sejam feitas sobre o assunto utilizando metodologias semelhantes, mas que comparem os resultados dos modelos treinados com manchetes publicadas antes e durante o pregão com modelos alimentados com manchetes do período que aporte o horário de funcionamento do pregão e as horas posteriores a ele.

De toda forma, o resultado principal desta pesquisa foi o desenvolvimento do FinBertMackenzista, que com uma taxa de acerto de 86% pode contribuir para o desenvolvimento de novas pesquisas na área financeira, inclusive na área de gerenciamento de impressão.

Ainda em tempo, destaca-se que esta pesquisa também contribui com o desenvolvimento de um corpus focado em textos financeiros. Além das notícias usadas nos dois bancos de dados, também foram coletadas manchetes dos editoriais da área financeira de outros jornais e revistas. Juntando todos esses bancos de dados há um corpus com 1.065.838 manchetes da área financeira. Esse banco extenso banco de dado é útil para a realização de treinamentos não supervisionados de I.A.s a serem aplicadas na área financeira.

## **REFERÊNCIAS**

ANTUNES, G. A.; LAMOUNIER, W. M.; BRESSAN, A. A. Análise do “efeito tamanho” nos retornos das ações de empresas listadas na Bovespa. *Revista Contabilidade & Finanças*, v. 17, n. 40, p. 87–101, 2006.

ARACI, Dogu. FinBERT: Financial Sentiment Analysis with Pre-trained Language Models. 20019. 11 f. Dissertação (Mestrado) - Curso de Computer Science, University Of Amsterdam, Amsterdam, 2019.

BAIARDO CAVALCANTE VIANA JUNIOR, D. *et al.* Do the Words Matter? an Analysis of the Linguistic Sentiment of Earnings Conference Calls and Abnormal Stock Returns in Brazilian Companies. *Advances in Scientific and Applied Accounting*, v. 12, n. 3, p. 144–158, 2019.

Carosia, A. E., Coelho, G. P., & Silva, A. E. A. Investment strategies applied to the Brazilian stock market: A methodology based on Sentiment Analysis with deep learning. *Expert Systems with Applications*, v. 184, 115470, 2021

Carosia, A. E. O., Coelho, G. P., & Silva, A. E. A. Analyzing the Brazilian Financial Market through Portuguese Sentiment Analysis in Social Media. *Applied Artificial Intelligence*, 34(1), 1–19. 2020.

ELY, R. A. Returns Predictability and Stock Market Efficiency in Brazil. *Revista Brasileira de Finanças*, v. 9, n. 4, p. 571–584, 2011.

EVARD, H. S.; CRUZ, J. A. W. A União Faz a Força? Um Teste Usando Fatores de Retorno. *Revista Brasileira de Finanças*, v. 15, n. 1, p. 59–92, 2017.

FERNÁNDEZ VILAS, A. *et al.* Twitter permeability to financial events: an experiment towards a model for sensing irregularities. *Multimedia Tools and Applications*, v. 78, n. 7, p. 9217–9245, 2019.

FIORISI, A.; GALDI, F. C. Estrutura de capital e custo de capital subsidiado no Brasil: influência no valor das ações das empresas. *Revista Contemporânea de Contabilidade*, v. 15, n. 34, p. 42–57, 2018.

GALDI, F. C.; GONÇALVES, A. M. PESSIMISMO E INCERTEZA DAS NOTÍCIAS E O COMPORTAMENTO DOS INVESTIDORES NO BRASIL. *RAE-Revista de Administração de Empresas*, v. 58, n. 2, p. 130–148, 2018.

JUNIOR, D. B. C. V. *et al.* Do the Words Matter? an Analysis of the Linguistic Sentiment of Earnings Conference Calls and Abnormal Stock Returns in Brazilian Companies. *Advances in Scientific and Applied Accounting*, v. 12, n. 3, p. 144–158, 2019.

LEE, B. R. *et al.* About relationship between business text patterns and financial performance in corporate data. *Journal of Open Innovation: Technology, Market, and Complexity*, v. 4, n. 1, 2018.

MACHADO, M. A. V.; SILVA, M. D. DE O. P. D. Análise do Sentimento Textual dos Relatórios de Desempenho Trimestral das Indústrias Brasileiras. *Sociedade, Contabilidade e Gestão*, v. 12, n. n.1, jan/abr, p. 1–25, 2017.

OLIVEIRA, D. J. S. *et al.* The application of the sentiment analysis technique in social media as a tool for social management practices at the governmental level. *Revista de Administração Pública*, v. 53, n. 1, p. 235–251, 2019.

ROSS, S. A.; WESTERFIELD, R. W.; JORDAN, B. D. Princípios de administração financeira. 2a ed. São Paulo : Atlas, 2000.

SANVICENTE, A. Z.; MONTEIRO, R. DA C. Eficiência do mercado de opções da Bolsa de Valores de São Paulo. *R. Adm.*, v. 40, n. 4, p. 34–43, 2005. Disponível em: <<http://www.rausp.usp.br/principal.asp?artigo=494>>.



SONG, Y.; WANG, H.; ZHU, M. Sustainable strategy for corporate governance based on the sentiment analysis of financial reports with CSR. *Financial Innovation*, v. 4, n. 1, 2018.

SOUZA F., NOGUEIRA R., LOTUFO R. (2020) BERTimbau: Pretrained BERT Models for Brazilian Portuguese. *Computer Science*, v 12319. Springer,

YE, M.; LI, G. Internet big data and capital markets: a literature review. *Financial Innovation*, v. 3, n. 1, 2017.