

JULIANA MARQUES DA CRUZ

**TEMPORAL FUSION TRANSFORMERS PARA A
PRECIFICAÇÃO DE ATIVOS FINANCEIROS**

São Paulo
2022

JULIANA MARQUES DA CRUZ

**TEMPORAL FUSSION TRANSFORMERS PARA A
PRECIFICAÇÃO DE ATIVOS FINANCEIROS**

Trabalho apresentado à Escola Politécnica
da Universidade de São Paulo para ob-
tenção do Título de Engenheiro Eletricista
com ênfase em Computação.

São Paulo
2022

JULIANA MARQUES DA CRUZ

**TEMPORAL FUSSION TRANSFORMERS PARA A
PRECIFICAÇÃO DE ATIVOS FINANCEIROS**

Trabalho apresentado à Escola Politécnica da Universidade de São Paulo para obtenção do Título de Engenheiro Eletricista com ênfase em Computação.

Orientador:

Anna Helena Reali Costa

São Paulo
2022

Prof. X

Prof. Y

Prof. Z

À minha mãe

AGRADECIMENTOS

Gostaria de agradecer primeiramente à minha mãe por todo o amor, incentivo, apoio incondicional e por ser a pessoa sem a qual nada disso seria possível.

Agradeço também à minha orientadora, Professora Doutora Anna Helena Reali Costa, pela disposição, suporte e por ser ter sido um exemplo durante toda a graduação.

Por fim, sou grata também a todos que direta ou indiretamente me encorajaram e me deram força nesse processo.

“Predicting the future isn’t magic, it’s artificial intelligence.”

-- Dave Waters

RESUMO

A bolsa de valores movimenta muito dinheiro todos os dias com a compra e venda de ações e recentemente está havendo um crescente interesse da população brasileira em investimentos, como mostra o aumento no número de pessoas físicas investidoras na B3. Neste contexto, uma forma de precificação assertiva dos ativos seria de enorme interesse, possibilitando um aumento da rentabilidade de diversos investimentos realizados diariamente. Atualmente, a previsão dos preços é realizada majoritariamente utilizando as abordagens de análise técnica e análise fundamentalista, sendo que a primeira busca identificar tendências futuras principalmente por gráficos e a segunda visa determinar o valor intrínseco das ações, entretanto, observa-se também uma crescente utilização de técnicas de aprendizado de máquina. Diante disso, este projeto propôs utilizar o *Temporal Fusion Transformer*, um modelo híbrido capaz de suportar diversos tipos de entradas e que tem como saídas não somente previsões pontuais, mas também intervalos de previsão, para prever o preço do ativo PETR4. Como entradas foram utilizados a série temporal de preços, indicadores econômicos, indicadores fundamentalistas e um índice de sentimentos gerado por um sistema de análise de sentimento de textos financeiros. Os resultados mostram que os indicadores econômicos são de grande importância para a precificação em comparação com os indicadores fundamentalistas e com o índice de sentimentos, pois resultam em um desempenho melhor do modelo.

Palavras-Chave – Mercado Financeiro, Ativos Financeiros, Ações, Aprendizado de Máquina, Aprendizado Profundo, *Transformer*, Previsão de Séries Temporais, *Temporal Fusion Transformer*.

ABSTRACT

The stock exchange moves a lot of money every day with the purchase and sale of shares and recently there has been a growing interest of the Brazilian population in investments, as shown by the growing number of new individuals investing in B3. In this context, a form of assertive pricing of assets would be of great interest, enabling an increase in the profitability of various investments made every day. Nowadays, price forecasting is carried out mostly using technical analysis and fundamental analysis approaches, the first one seeks to identify future trends mainly through charts and the second aims to determine the intrinsic value of stocks, however, it is observed that there is also an increasing use of machine learning. Therefore, this project proposes to use the Temporal Fusion Transformer, a hybrid model capable of supporting different types of inputs and which has not only point forecasts but also forecast intervals as outputs, to predict the price of the PETR4 asset. The inputs were the price time series, economic indicators, fundamental indicators and a sentiment index from a system of sentiment analysis of financial texts. The results show that economic indicators are of great importance for pricing compared to the others inputs, as they result in a better performing model.

Keywords – Financial Market, Financial Assets, Stocks, Machine Learning, Deep Learning, Transformer, Time Series Forecasting, Temporal Fusion Transformer.

LISTA DE FIGURAS

1	Uso de diferentes modelos de precificação no mercado financeiro	16
2	Diferentes entradas e saídas do <i>Temporal Fusion Transformer</i>	17
3	Operações de Atenção	28
4	Arquitetura do <i>Transformer</i>	29
5	Ilustração extraída de [1] da arquitetura do <i>Temporal Fusion Transformer</i>	31
6	Processo de Análise de Sentimento. Ilustração extraída de [2].	32
7	Fluxograma do Modelo Preditivo com o TFT.	34
8	Arquitetura geral do projeto, composta pelo <i>Temporal Fusion Transformer</i> e as entradas definidas	37
9	Proporção de cada setor da B3 em relação ao volume de negociações.	46
10	Arquitetura geral do projeto, composta pelo <i>Temporal Fusion Transformer</i> e as entradas definidas	47
11	Resultado das métricas de desempenho no Treinamento 1	51
12	Treinamento do modelo utilizando apenas a série temporal de preços	51
13	Resultado das métricas de desempenho no Treinamento 2	53
14	Treinamento do modelo utilizando a série temporal de preços e os indica- dores econômicos	53
15	Resultado das métricas de desempenho no Treinamento 3	55
16	Treinamento do modelo utilizando a série temporal de preços e os indica- dores econômicos e fundamentalistas	56
17	Resultado das métricas de desempenho no Treinamento 4	57
18	Treinamento do modelo utilizando a série temporal de preços e os indica- dores fundamentalistas	58
19	Resultado das métricas de desempenho no Treinamento 5	59

20	Treinamento do modelo utilizando a série temporal de preços e os indicadores econômicos e o índice de sentimentos	60
21	Comparação entre as métricas de desempenho de todos os treinamentos . .	61
22	Arquitetura final do projeto, composta pelo <i>Temporal Fusion Transformer</i> e as entradas definidas	62

LISTA DE TABELAS

1	Tabela com os treinamentos do TFT realizados no projeto.	35
2	Tabela com os indicadores fundamentalistas selecionados.	42
3	Tabela com os indicadores econômicos selecionados.	42
4	Tabela com as ações mais negociadas na bolsa em 2021.	45
5	Tabela com as variáveis de entrada e saída comuns a todos os treinamentos.	47
6	Tabela com as variáveis de entrada utilizadas dependendo do treinamento.	48
7	Tabela com os hiper-parâmetros utilizados.	49
8	Variáveis de entrada do primeiro treinamento.	50
9	Tabela com resultados do teste do <i>baseline</i> e do TFT no 1o. treinamento. .	50
10	Variáveis de entrada do segundo treinamento.	52
11	Tabela com resultados do teste do <i>baseline</i> e do TFT no 2o. treinamento. .	52
12	Variáveis de entrada do terceiro treinamento.	54
13	Tabela com resultados do teste do <i>baseline</i> e do TFT no 3o. treinamento. .	54
14	Variáveis de entrada do quarto treinamento.	57
15	Tabela com resultados do teste do <i>baseline</i> e do TFT no 4o. treinamento. .	57
16	Variáveis de entrada do quinto treinamento.	58
17	Tabela com resultados do teste do <i>baseline</i> e do TFT com as entradas do primeiro treinamento.	59
18	Tabela com todos os resultados do teste <i>baseline</i> e do TFT.	60
19	Tabela com todos os valores de perda dos treinamentos do TFT.	60
20	Variáveis de entrada do melhor treinamento.	63
21	Tabela com os hiper-parâmetros utilizados no melhor treino.	63

SUMÁRIO

1	Introdução	14
1.1	Motivação	15
1.2	Objetivo	17
1.3	Justificativa	17
1.4	Organização do trabalho	18
2	Aspectos Conceituais	19
2.1	Mercado de capitais	19
2.2	Métodos de análise financeira	20
2.2.1	Análise técnica	20
2.2.2	Análise fundamentalista	20
2.3	Previsão de séries temporais	25
2.4	Aprendizado de máquina	26
2.4.1	<i>Transformers</i>	27
2.4.2	<i>Temporal fusion transformers</i>	30
2.5	Análise de sentimento	32
3	Metodologia do Trabalho	34
3.1	Aquisição de dados	34
3.2	Treinamento do modelo	35
3.3	Validação do modelo	35
4	Especificação de Requisitos	37
4.1	Arquitetura do modelo	37
4.2	Requisitos funcionais	37

4.3	Requisitos não funcionais	38
5	Desenvolvimento do Trabalho	39
5.1	Tecnologias utilizadas	39
5.1.1	Python	39
5.1.2	Google Colaboratory	40
5.1.3	GitHub	40
5.2	Decisões de projeto	41
5.2.1	Definição das entradas	41
5.2.1.1	Indicadores fundamentalistas	41
5.2.1.2	Indicadores econômicos	41
5.2.1.3	Análise de sentimento de notícias	44
5.2.2	Definição do ativo	44
5.2.2.1	Volume de negociação por ação	45
5.2.2.2	Liquidez por setor da B3	45
5.2.2.3	Indicadores fundamentalistas disponíveis por ação	46
5.3	Implementação da arquitetura	47
5.3.1	Taxa de aprendizado	48
5.3.2	Prevenção de sobreajuste	48
5.3.3	Registro de resultados	48
5.3.4	Hiper-parâmetros	49
5.3.5	Modelo baseline	49
5.4	Treinos e avaliação detalhada	49
5.4.1	Primeiro treinamento	49
5.4.2	Segundo treinamento	52
5.4.3	Terceiro treinamento	54
5.4.4	Quarto treinamento	56

5.4.5	Quinto treinamento	58
5.5	Comparação entre os treinamentos	60
6	Considerações Finais	62
6.1	Conclusões do projeto de formatura	62
6.2	Perspectivas de continuidade	64
	Referências	65

1 INTRODUÇÃO

Segundo o Relatório Anual 2021 da bolsa de valores brasileira (B3 – Brasil, Bolsa, Balcão) [3], o ano de 2021 teve um recorde no número de novos investidores pessoa física, atingindo um total de 4.8 milhões de contas de pessoa física em renda variável, demonstrando um enorme interesse da população brasileira em investimentos.

Ao investir no mercado de capitais, o objetivo é comprar ações quando seu preço estiver baixo e vendê-las quando o preço estiver alto. Neste cenário, a capacidade de precificar ativos assertivamente resultaria em um aumento da rentabilidade dos investimentos.

Atualmente, a precificação de ativos é feita utilizando majoritariamente as análises técnica e/ou fundamentalista. A análise técnica é definida por [4] como o estudo de um ativo, principalmente com o uso de gráficos, visando de prever tendências futuras de preço; já a análise fundamentalista diz respeito à determinação do valor intrínseco de ações na bolsa de valores, através do estudo de previsões econômicas, a fim de medir a força das empresas, a eficiência da gestão e as oportunidades de negócios com base em demonstrações históricas e em condições atuais, permitindo determinar um valor justo para as ações e, então, compará-lo aos valores de mercado resultantes de interações de oferta e demanda e, assim, identificar oportunidades de investimento, conforme afirmado por [5].

Apesar disso, ambos os métodos de análise têm limitações. A análise técnica, ao focar apenas nos gráficos, deixa de considerar outras variáveis com impacto sobre o preço das ações, como indicadores fundamentalistas ou até notícias do mercado. Já a análise fundamentalista avalia o balanço patrimonial, fluxo de caixa, conjuntura econômica, entre outros fatores, deixando de fora padrões gráficos, informações dinâmicas, entre outras, que podem ter impacto sobre o preço também.

Outra forma de precificação de ativos é utilizando a inteligência artificial. Uma área de grande sucesso da inteligência artificial é o aprendizado de máquina, que oferece, entre outras, técnicas de regressão e de classificação para diversos problemas. Segundo [6],

técnicas de regressão são comumente utilizadas para a previsão de preços, uma vez que fornecem resultados numéricos ou contínuos, enquanto técnicas de classificação, que fornecem resultados categóricos, são frequentemente utilizadas para prever tendências na mudança de preços de ativos. Todavia, por mais promissoras que essas técnicas sejam, há problemas e limitações atrelados a elas: o impacto do horizonte de previsão na precisão, a dificuldade de processamento de uma grande quantidade de dados e a falta de interpretabilidade dos modelos (também conhecidos como modelos caixa preta cuja complexidade é tão alta que não são interpretáveis por humanos).

1.1 Motivação

O crescimento da utilização de aprendizado de máquina em finanças é explicado por [7] como sendo decorrente do crescimento da quantidade de dados para monitorar, registrar e comunicar através do sistema financeiro. Neste cenário, as máquinas conseguem modelar processos complexos e de alta dimensão ao varrer milhões de configurações e, em seguida, avaliar e corrigir esses modelos em resposta a novas informações.

Do ponto de vista de [8], a vantagem da utilização de aprendizado de máquina no mercado financeiro reside na capacidade do algoritmo aprender padrões complexos em um espaço de alta dimensão com pouca orientação humana na especificação do modelo, além de ajudar na formação de teorias baseadas em evidências empíricas.

Mais especificamente, o aprendizado profundo conta com as habilidades de lidar com grandes volumes de dados e de aprender as relações não lineares entre dados de entrada e a previsão esperada, obtendo um desempenho melhor do que os modelos lineares e os de aprendizado de máquina tradicional na previsão do mercado de ações, de acordo com [9].

Como é possível ver na Figura 1a, a utilização de modelos de aprendizado profundo é mais presente em vista de modelos lineares e outros modelos de aprendizado de máquina tradicionais. Dentre esses modelos, o mais utilizados são as Redes Neurais Recorrentes, doravante RNN (*Recurrent Neural Network*), entretanto, de 2018 para 2019, houve uma queda na utilização e o crescimento na utilização de modelos híbridos, como é possível ver na Figura 1b.

RNNs são um tipo de modelo de aprendizado profundo que processa dados sequencialmente de forma linear e suas vantagens residem no fato de que ela é capaz de se lembrar de informações ao longo do tempo, conforme descrito por [10]. Entretanto, esse tipo de modelo não é indicado para analisar longos períodos, pois ele tende a esquecer

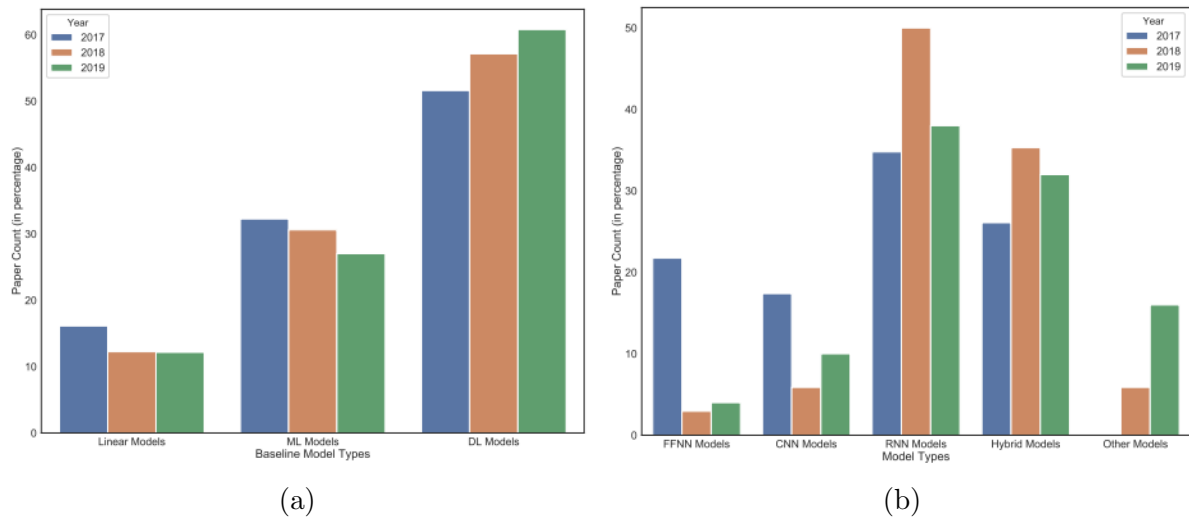


Figura 1: O uso de diferentes modelos de aprendizado de máquina para precificação no mercado financeiro: (a) modelos lineares (esquerda), tradicionais (meio) e de aprendizado profundo (direita); (b) diferentes modelos de aprendizado profundo. Ilustração extraída de [9].

informações de posições distintas; além disso, tendo em vista sua natureza sequencial, é difícil aproveitar ao máximo os dispositivos modernos de computação, como TPUs e GPUs.

Em comparação com a RNN, podemos citar o *Transformer* que, segundo [11], também é um modelo de aprendizado profundo de processamento sequencial, entretanto o processamento não é ordenado (graças ao Mecanismo de Atenção, *Self Attention*, empregado), o que faz com que ele não tenha problemas em lembrar de informações relevantes e que ele permita mais paralelização.

Por fim, o *Temporal Fusion Transformer* (TFT), como explica [1], é uma arquitetura baseada na combinação de redes *Long Short Term Memory* (uma classe de RNN) com *Transformers*, tirando o melhor proveito de ambos. Como pode ser visto na Figura 2, esta arquitetura permite diferentes tipos de entrada para trabalhar com previsão multi-horizonte, como, por exemplo, informações conhecidas sobre o futuro, dados estáticos e séries de variáveis exógenas. Além disso, as saídas são, além das previsões pontuais, intervalos de precisão, ou seja, não é um modelo caixa preta, o que o torna um modelo mais vantajoso para a previsão do mercado de ações.

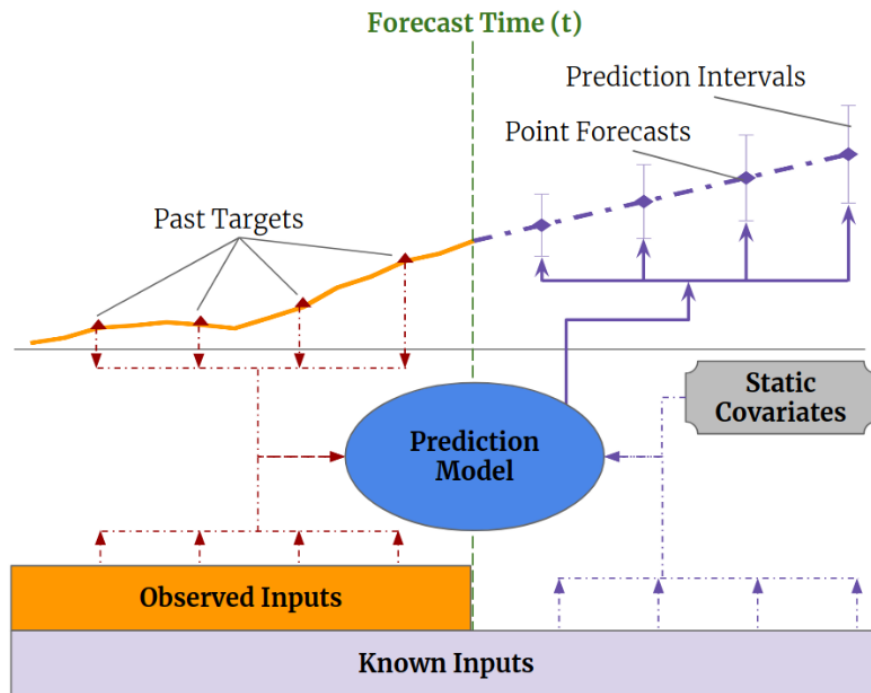


Figura 2: Ilustração extraída de [1] dos diferentes tipos de entradas (informações conhecidas sobre o futuro, dados estáticos e séries de variáveis exógenas) e saídas (previsões pontuais e intervalos de precisão) do *Temporal Fusion Transformer*.

1.2 Objetivo

Este projeto visa construir um modelo de precificação de ativos financeiros utilizando *Temporal Fusion Transformers*.

As entradas mais comuns para modelos de precificação de ativos financeiros, segundo [6], são a própria série histórica de preços, indicadores econômicos e indicadores fundamentalistas. Neste projeto, é ainda incluída um índice de sentimentos de notícias do mercado financeiro como entrada.

Portanto, o ativo a ser precificado foi selecionado considerando três critérios: volume de negociação por ação, liquidez por setor da B3 e volume de indicadores fundamentalistas disponíveis por ação.

1.3 Justificativa

O desenvolvimento do modelo de precificação de ativos financeiros é justificado pelo crescente número de novos investidores pessoa física na bolsa de valores brasileira que se

beneficiariam de um modelo capaz de gerar lucros e reduzir riscos e prejuízos por meio da antecipação do comportamento de ativos.

O modelo utilizando *Temporal Fusion Transformer* se justifica pela tendência de desaceleração na utilização de RNN em vista de modelos híbridos. Os benefícios deste modelo são os diferentes tipos de entrada que são permitidos, além de permitir o processamento não ordenado dos dados, possibilitando mais paralelização e evitando o problema de lembrar de informações no horizonte de tempo. Outro fator relevante consiste em não somente possibilitar o acesso a previsões pontuais, mas também a intervalos de previsão.

1.4 Organização do trabalho

No segundo capítulo há uma revisão da literatura no tocante a mercado financeiro (mercado de capitais e métodos de análise financeira), previsão de séries temporais, aprendizado de máquina (*transformers* e *temporal fusion transformers*) e análise de sentimento.

Em seguida, no terceiro capítulo, é apresentada a metodologia do trabalho, em que é explicada a aquisição de dados e o treinamento e validação do modelo.

Já o quarto capítulo trata da especificação dos requisitos do projeto, definindo a arquitetura do modelo e os requisitos funcionais e não funcionais.

O quinto capítulo traz o desenvolvimento do trabalho, detalhando as tecnologias utilizadas, as decisões de projeto, a implementação da arquitetura e os treinos e avaliação detalhada.

Finalmente, no sexto capítulo ocorrem as considerações finais, passando pelos tópicos de conclusões do projeto e perspectivas de continuidade.

2 ASPECTOS CONCEITUAIS

Este capítulo contém os conceitos teóricos fundamentais para o desenvolvimento do projeto. São eles conceitos de mercado de capitais, de métodos de análise financeira, de previsão de séries temporais, de aprendizado de máquina, passando por *Transformers* e por *Temporal Fusion Transformers*, e, por fim, de análise de sentimento.

2.1 Mercado de capitais

O mercado de capitais, segundo [12], é um dos quatro grandes mercados em que o mercado financeiro é segmentado, sendo os outros os mercados monetário, de crédito e de câmbio.

Diante disso, o mercado de capitais é responsável por viabilizar a captação de recursos por empresas de investidores diretamente por meio da emissão de instrumentos financeiros, possibilitando seu processo de capitalização, conforme descrito em [13], constituído por bolsas, corretoras e outras instituições financeiras autorizadas, cujos principais títulos negociados são representativos do capital de empresas (ações) e empréstimos tomados, via mercado, por empresas.

Ações – Ações são valores mobiliários que representam uma fração do capital social da empresa emissora [13] sendo consideradas títulos de renda variável tendo em vista que seu valor depende do comportamento do mercado acionário [12].

Bolsa de valores – Bolsas de valores (ou apenas bolsas) são definidas como locais que oferecem condições e sistemas necessários para a negociação de compra e venda de ativos, além de orientar e fiscalizar os serviços prestados pelos seus intermediários e divulgar constantemente informações sobre os ativos sob seu controle, conforme explicado em [13].

2.2 Métodos de análise financeira

De acordo com [6], existem dois métodos de análise financeira que consideram que os preços do mercado variam de acordo com uma tendência: análise técnica, que defende haver propensão no movimento dos preços, portanto, utiliza a série histórica de preços para a previsão; e a análise fundamentalista, que argumenta que o contexto socioeconômico da empresa em questão interfere no preço futuro da ação, logo, utiliza esse tipo de informação para prever os preços.

2.2.1 Análise técnica

A análise técnica é descrita por [4] como o estudo da movimentação do mercado, principalmente por meio de gráficos, com o propósito de prever o futuro das tendências de preço. Diante disso, há três premissas das quais a análise técnica parte: a movimentação de mercado desconta tudo, os preços se movimentam segundo tendências e a história se repete.

Segundo [4], o primeiro princípio reflete a hipótese de que tudo que pode afetar o preço do ativo (fundamental, política e psicologicamente ou outros) está refletido no preço do mercado. Já o segundo possui um corolário que diz que um movimento de tendência é provável de continuar do que reverter: basicamente, essa premissa sustenta a utilização de gráficos na análise técnica, pois o objetivo deles é identificar tendências no começo. Por fim, o terceiro diz respeito a padrões de comportamento identificados no decorrer dos anos que são relevantes e continuarão sendo.

2.2.2 Análise fundamentalista

A análise fundamentalista é definida por [5] como sendo o conhecimento das regras e etapas necessárias para conseguir determinar o valor intrínseco de ações do mercado financeiro. Em outras palavras, é o estudo da situação financeira e das perspectivas da empresa visando identificar o potencial de crescimento no futuro.

Na estrutura proposta por [5], é primeiramente feito o estudo das previsões econômicas para obter uma seleção de setores que geram um aumento nas vendas e lucros e, por fim, é feita a medição da solidez das empresas financeiras, a eficiência da gestão e as oportunidades de negócios com base em demonstrações financeiras históricas e condições atuais. Portanto, uma vez determinado o valor justo para a ação, é feita uma comparação

com o preço real para identificar possíveis oportunidades de investimento.

Assim sendo, há três etapas bem definidas, como foi detalhado em [14]: análise macroeconômica, em que são analisados indicadores econômicos como PIB (Produto Interno Bruto) do país, índices de inflação, taxa de câmbio, taxa de juros e nível de renda, entre outros; análise setorial, onde é considerado que a análise macroeconômica não é uniforme em todos os setores (setores com correlação negativa podem se comportar de maneira contrária diante diferentes cenários) e, além disso, são inclusas questões específicas a cada setor, como, por exemplo, incentivos econômicos, regulamentações e mudanças estruturais; e, por fim, a análise empresarial, em que serão analisados os indicadores fundamentalistas de cada empresa individualmente.

Para o estudo da situação atual e das perspectivas futuras, a análise fundamentalista lança mão de indicadores. Os indicadores fundamentalistas são métricas para avaliar a saúde financeira da empresa e constam nos balanços financeiros que as empresas divulgam trimestralmente, de acordo com [15].

Normalmente, eles são utilizados como método de comparação entre empresas do mesmo setor ou internamente, conforme explicado por [16] sendo segmentados em quatro tipos: indicadores de *valuation*, de rentabilidade, de eficiência e de endividamento.

Os indicadores de *valuation*, também conhecidos como indicadores de mercado, relacionam valores contábeis com o valor de mercado da empresa. Existem diversos indicadores desse tipo, porém neste projeto serão abordados treze deles, descritos a seguir.

O $D.Y$ é responsável por relacionar os proventos (ou dividendos) pagos por uma companhia e o preço atual da ação,

$$D.Y = \frac{\text{Dividendos Pagos no Período}}{\text{Preço da Ação}}. \quad (2.1)$$

O P/L indica o quanto o mercado está disposto a pagar pelos lucros da empresa,

$$P/L = \frac{\text{Preço Atual}}{\text{Lucro Por Ação (LPA)}}. \quad (2.2)$$

O P/VP viabiliza a comparação do preço de negociação de um ativo com seu valor patrimonial por ação,

$$P/VP = \frac{\text{Preço Atual}}{\text{Valor Patrimonial por Ação (VPA)}}. \quad (2.3)$$

O $EV/EBITDA$ e o $EV/EBIT$ calculam quanto tempo seria necessário para o valor do $EBITDA$ e do $EBIT$, respectivamente, pagar o investimento feito para comprar a companhia em questão,

$$EV/EBITDA = \frac{Enterprise\ Value}{EBITDA}, \quad (2.4)$$

$$EV/EBIT = \frac{Enterprise\ Value}{EBIT}. \quad (2.5)$$

O $P/EBITDA$ e o $P/EBIT$ (Equações 2.6 e 2.7) indica o preço da ação em relação ao valor do $EBITDA$ e do $EBIT$, respectivamente,

$$P/EBITDA = \frac{Preço\ Atual}{EBITDA}, \quad (2.6)$$

$$P/EBIT = \frac{Preço\ Atual}{EBIT}. \quad (2.7)$$

O VPA , que também pode ser chamado de VP , indica o real valor contábil de cada ação

$$VPA = \frac{Patrimônio\ Líquido}{Número\ de\ Ações}. \quad (2.8)$$

O $P/ATIVO$ é o preço da ação dividido pelos ativos totais

$$P/ATIVO = \frac{Preço\ Atual}{Ativos}. \quad (2.9)$$

O LPA é uma forma de normalizar o lucro e serve como um indicador de quão lucrativa uma empresa é

$$LPA = \frac{Lucro\ Líquido}{Número\ de\ Ações}. \quad (2.10)$$

O P/SR é o valor de mercado da empresa dividido pela receita operacional líquida

$$P/SR = \frac{Preço\ Atual}{Receita\ Líquida\ por\ Ação}. \quad (2.11)$$

O $P/CAP.GIRO$ é o quociente entre o preço da ação e o capital de giro da empresa,

que é dado pela diferença entre ativo e passivo circulantes

$$P/CAP.GIRO = \frac{\textit{Preço Atual}}{\textit{Ativo Circulante} - \textit{Passivo Circulante}}. \quad (2.12)$$

Por fim, o $P/ATIVOCIRC.LIQ.$ é a fração do preço atual da ação pelos ativos circulantes líquidos por ação

$$P/ATIVOCIRC.LIQ. = \frac{\textit{Preço Atual}}{\textit{Ativos Circulantes Líquidos por Ação}}. \quad (2.13)$$

Os indicadores de rentabilidade exibem os retornos obtidos sobre a receita líquida, sobre o investimento, entre outros. Neste projeto, serão estudados quatro tipos: ROE, ROA, ROIC e Giro Ativos, descritos a seguir.

O ROE ,

$$ROE = \frac{\textit{Lucro Líquido}}{\textit{Patrimônio Líquido}}, \quad (2.14)$$

reflete o retorno sobre o patrimônio líquido - ou seja, mede a capacidade da empresa de agregar valor com seus próprios recursos e dinheiro dos investidores - e, por essa razão, é um indicador da saúde financeira da empresa: quanto maior esta porcentagem, mais saudável ela está.

O ROA ,

$$ROA = \frac{\textit{Lucro Líquido}}{\textit{Ativo Total}}, \quad (2.15)$$

é o retorno sobre os ativos, portanto calcula a capacidade que a empresa tem de gerar lucro a partir de seus ativos.

O $ROIC$, por sua vez, mede o que uma empresa é capaz de gerar tendo em vista todo o capital investido, incluindo dívidas

$$ROIC = \frac{\textit{EBIT} - \textit{Impostos}}{\textit{Patrimônio Líquido} + \textit{Endividamento}}. \quad (2.16)$$

Por fim, o $GiroAtivos$ relaciona o quanto a empresa vendeu em um ano, após abatimentos, em relação ao seu total médio de ativos.

$$\textit{GiroAtivos} = \frac{\textit{Receita Líquida}}{\textit{Total Médio de Ativos}} \quad (2.17)$$

Os indicadores de eficiência relacionam a receita líquida com valores contábeis e são *M.BRUTA*, *M.EBITDA*, *M.EBIT* e *M.LIQ.*

O *M.BRUTA* mede o quanto a empresa ganha com a venda de seus produtos

$$M.BRUTA = \frac{LucroBruto}{ReceitaLíquida}. \quad (2.18)$$

Por sua vez, o *M.EBITDA* é utilizado para analisar a lucratividade operacional da empresa

$$M.EBITDA = \frac{EBITDA}{ReceitaLíquida}. \quad (2.19)$$

Já o *M.EBIT* é uma forma de avaliar o crescimento da eficiência produtiva da empresa ao longo do tempo

$$M.EBIT = \frac{EBIT}{ReceitaLíquida}. \quad (2.20)$$

Por fim, o *M.LIQ* é a porcentagem de lucro em relação às receitas da empresa

$$M.LIQ. = \frac{LucroLíquido}{ReceitaLíquida}. \quad (2.21)$$

Os indicadores de endividamento são uma maneira de calcular o nível de alavancagem da empresa. Este projeto irá abordar seis desses indicadores:

$$\begin{aligned} & DIV.LIQUIDA/PL, \\ & DIV.LIQUIDA/EBITDA, \\ & DIV.LIQUIDA/EBIT, \\ & PL/ATIVOS, \\ & PASSIVOS/ATIVOS \text{ e} \\ & LIQ.CORRENTE. \end{aligned}$$

O *DIV.LIQUIDA/PL* é um indicador do quanto de dívida a empresa utiliza para financiar seus ativos em relação ao patrimônio dos acionistas

$$DIV.LÍQUIDA/PL = \frac{Dívida Líquida}{Patrimônio Líquido}. \quad (2.22)$$

O *DIV.LIQUIDA/EBITDA* e o *DIV.LIQUIDA/EBIT* calculam o tempo necessário para pagar a dívida líquida da empresa considerando o *EBITDA* e o *EBIT* atual, respectivamente,

$$DIV.LÍQUIDA/EBITDA = \frac{Dívida Líquida}{EBITDA}, \quad (2.23)$$

$$DIV.LÍQUIDA/EBIT = \frac{Dívida Líquida}{EBIT}. \quad (2.24)$$

O $PL/ATIVOS$ mostra a relação dos ativos no patrimônio da empresa,

$$PL/ATIVOS = \frac{Patrimônio Líquido}{Ativos}. \quad (2.25)$$

O $PASSIVOS/ATIVOS$ calcula uma relação entre os ativos e os passivos da empresa

$$PASSIVOS/ATIVOS = \frac{Passivos}{Ativos}. \quad (2.26)$$

Por fim, o $LIQ.CORRENTE$ indica a capacidade de pagamento da empresa no curto prazo

$$LIQ.CORRENTE = \frac{Ativo Circulante}{Passivo Circulante}. \quad (2.27)$$

2.3 Previsão de séries temporais

Série temporal é definida como sendo um conjunto de observações x_t , cada uma sendo registrada em um tempo específico t , diante disso há duas classificações: série temporal discreta e série temporal contínua, de acordo com [17].

Segundo [17], a série temporal discreta é o caso em que o conjunto de observações é um conjunto discreto, ou seja, são feitas em intervalos de tempo fixados. A série temporal contínua, por sua vez, é caracterizada pelas observações serem registradas de forma contínua dentro de um intervalo de tempo.

Além dessa classificação, as séries temporais também podem ser classificadas em univariante e multivariante: o primeiro é caracterizado por utilizar apenas informações contidas nos valores passados e valores atuais e passados de um erro, como é descrito em [18]; em contrapartida, o segundo é definido por ter mais de uma variável dependente do tempo e, cada variável dependente não apenas de seus valores passados, mas também com dependência em outras variáveis, de acordo com [17].

Por fim, a maneira que são realizadas as previsões também são classificatórias, existindo três tipos: previsão de um passo no futuro, previsão do k -ésimo passo no futuro e previsão multi-horizonte.

Previsão de um passo no futuro – Neste caso, também conhecido como previsão de um horizonte, t instantes são utilizadas para prever o instante $t+1$ da série temporal [19].

Previsão do k -ésimo passo no futuro – Semelhante à previsão de um passo no futuro, a diferença é que aqui as n observações são utilizadas para prever o k valor da série temporal, sendo $k > n + 1$.

Previsão multi-horizonte – Por fim, esta previsão tem mais de um horizonte, portanto são previstos múltiplos pontos no futuro e é equivalente a diversas previsões do k -ésimo passo no futuro em paralelo [19].

2.4 Aprendizado de máquina

Aprendizado de máquina é definido por [20] como uma forma de inteligência artificial que permite que o sistema aprenda dos dados ao invés de aprender através de programação explícita, ou seja, são utilizados dados de exemplo ou experiências passadas para otimizar um critério de desempenho.

Diante disso, há dois objetivos: primeiro, no tocante ao treinamento do modelo, é necessário que o algoritmo seja eficiente para solucionar o problema de otimização, assim como em armazenar e processar os dados; segundo, com o modelo já treinado, sua representação e solução algorítmica também deve ser eficiente (complexidade de espaço e tempo), como é descrito por [21].

Inicialmente, a maioria das técnicas de aprendizado de máquina exploravam em especial o que é conhecido como *Shallow Neural Networks*. Segundo [22], essas arquiteturas de redes neurais contam com no máximo uma ou duas camadas de transformações não lineares. Apesar de ter se mostrado efetiva na solução de diversos problemas simples e restritos, a capacidade de modelagem e representação deixa a desejar ao lidar com aplicações mais complexas. Isso acontece pois o desempenho desses algoritmos simples de aprendizado de máquina depende fortemente da representação utilizada para os dados, como foi pontuado por [23], portanto os problemas que eles podem resolver devem ter fácil representação e o processamento de informações mais complexas requerem representações não triviais.

Uma solução para esse problema é utilizar o aprendizado de máquina para desenvolver a própria representação dos dados – isso é conhecido como aprendizagem de características. É nesse cenário que o aprendizado profunda se destaca: segundo [23], essa técnica introduz a ideia de expressar representações em função de outras representações mais simples, ou seja, o aprendizado profunda permite a construção de conceitos complexos a partir de representações simples.

Diante disso, o aprendizado profunda é definido por [22] como uma série de algoritmos

de aprendizado de máquina cujo funcionamento consiste em aprender a representar dados por meio da modelagem de relacionamentos complexos entre eles. Basicamente, recursos e conceitos de nível superior são definidos em termos de recursos de nível inferior, formando uma hierarquia de recursos denominada arquitetura profunda. *Transformers* constituem uma classe de aprendizado profundo.

2.4.1 *Transformers*

Transformer é definido por [11] como um modelo de arquitetura que evita recorrência, dependendo totalmente de um mecanismo de atenção, a Auto-atenção (em inglês, *Self-attention*), para capturar dependências globais entre entrada e saída.

O mecanismo de atenção é definido em [11] como o mapeamento de uma consulta (em inglês, *query*) e um conjunto de pares chave-valor (em inglês, *key-value*) em uma saída, sendo todos esses elementos vetores.

A saída é calculada como sendo uma soma ponderada de todos esses valores, sendo que o peso de cada valor é computado de acordo com uma função de compatibilidade que relaciona a consulta com a chave, como é explicado em [11].

Diante disso, o mecanismo de atenção deve receber três entradas que podem ser representadas, respectivamente, como sendo X_Q , $X_K \in \mathbb{R}^{d_k \times m}$ e $X_V \in \mathbb{R}^{d_v \times m}$. Cada uma dessas entradas passa por uma camada linear distinta, resultando nos vetores Q, K e V que são a consulta, chave e valor, respectivamente:

$$Q = W_Q^T \cdot X_Q + B_Q, \quad (2.28)$$

$$K = W_K^T \cdot X_K + B_K, \quad (2.29)$$

$$V = W_V^T \cdot X_V + B_V. \quad (2.30)$$

Desta forma, há duas formas de realizar a operação de atenção: Atenção de Produto Escalar Escalonado (em inglês, *Scaled Dot-Product Attention*) – em que há apenas uma camada de atenção, como pode ser visto na Figura 3a – e Atenção Multi-cabeça (em inglês, *Multi-Head Attention*), caracterizada por muitas camadas de atenção operando em paralelo, cuja representação pode ser vista na Figura 3b.

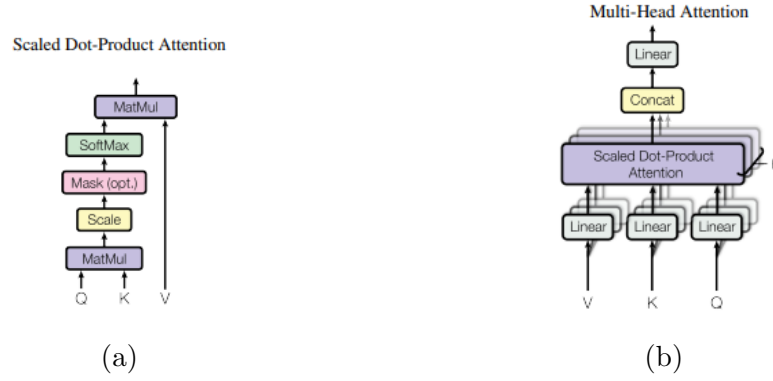


Figura 3: Formas de realizar a operação de atenção: (a) *Atenção de Produto Escalar Escalonado*; (b) *Atenção Multi-cabeça*. Ilustração extraída de [11].

A atenção de produto escalar escalonado é dada por

$$Atenção(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (2.31)$$

sendo que as entradas consulta Q e chave K têm dimensão d_k e a entrada valor V tem dimensão d_v .

A função *softmax* determinará a similaridade entre consultas e chaves, o que será multiplicada pelos valores, realizando uma ponderação.

A atenção multi-cabeça utiliza mais de um mecanismo de atenção para o mesmo conjunto de entradas, sendo que cada um desses mecanismos é denominado cabeça (em inglês, *head*).

Isso é descrito em [11] como a projeção linear de consultas, chaves e valores h vezes com projeções lineares diferentes para as dimensões d_k , d_k e d_v , respectivamente,

$$MultiCabeça(Q, K, V) = Concatenação(cabeça_1, \dots, cabeça_h)W^O, \quad (2.32)$$

em que

$$cabeça_i = Atenção(QW_i^Q, KW_i^K, VW_i^V). \quad (2.33)$$

Em cada uma dessas versões projetadas é executada uma função de atenção em paralelo, resultando em valores de saída de dimensão d_v , que são concatenadas e projetadas novamente, a fim de obter os valores finais, como é possível ver na Equação 2.32.

Esta é a abordagem utilizada nos *Transformers*, como é explicado em [11] e o número de cabeças é um dos hiper-parâmetros destes tipos de modelos. Mais especificamente, é utilizada a auto-atenção: trata-se de um caso particular do mecanismo de atenção multi-cabeça em que $d_k = d_v$, logo, o mesmo vetor pode ser fornecido como as três entradas.

Isso consiste em comparar o vetor de entrada consigo mesmo e, assim, são determinados quais valores devem ser passados adiante.

O *Transformer* possui uma arquitetura codificador-decodificador, como pode ser visto na Figura 4.

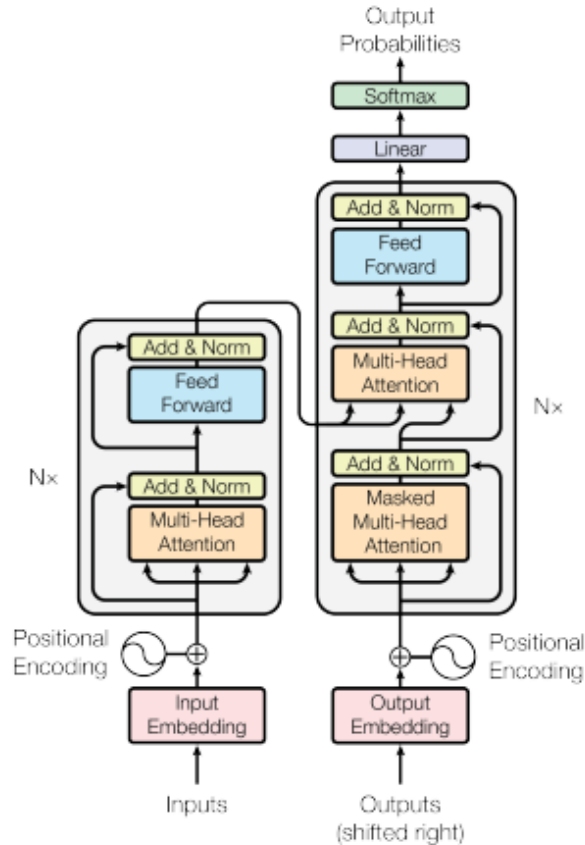


Figura 4: Arquitetura do *Transformer*. Ilustração extraída de [11].

Como é explicado em [11], inicialmente o modelo mapeia a sequência de entrada de símbolos (x_1, \dots, x_n) em uma sequência de representação contínua $z = (z_1, \dots, z_n)$, denominada *embeddings*, porém neste processo a ordem dos elementos na sequência é perdida. Diante disso, os dados passam por um codificador posicional responsável por gerar uma sequência de saída (y_1, \dots, y_m) com informações sobre a posição de cada elemento.

Posteriormente, essa sequência de saída (*embeddings* somados à codificação posicional) são recebidas pelo decodificador que é composto por um bloco de N camadas idênticas. Cada camada é composta por duas subcamadas: a primeira é um mecanismo de atenção multi-cabeça de auto-atenção e a segunda é rede com alimentação direta (do inglês, *feedforward*) posicional. Ao final, é aplicada uma conexão residual e uma normalização de camada [11].

O decodificador também é composto por um bloco de N camadas idênticas e também

possui as duas subcamadas do codificador, como é descrito em [11], entretanto a subcamada de auto-atenção é modificada visando que previsões para uma posição i dependam apenas das saídas conhecidas de entradas em posições anteriores a i . Há também o acréscimo de uma terceira camada que executa o mecanismo de atenção multi-cabeça na saída dos blocos de camadas do decodificador. Por fim, novamente é aplicada uma conexão residual e uma normalização de camada. A cada passo, o modelo é auto-regressivo, consumindo os símbolos gerados anteriormente como entrada adicional para gerar o próximo.

O modelo explorado neste trabalho é um tipo especial de *Transformer*, denominado *Temporal Fusion Transformer*, descrito na próxima seção.

2.4.2 *Temporal fusion transformers*

O *Temporal Fusion Transformer*, doravante TFT, é uma rede neural profunda baseada em atenção que combina previsão multi-horizonte de alto desempenho com informações sobre a dinâmica temporal, como é descrito em [1]. Basicamente, esta arquitetura integra diversos mecanismos de outras arquiteturas neurais, principalmente LSTM (*Long Short-Term Memory*) e *Transformers*.

A arquitetura pode ser vista na Figura 5. Segundo [1], os componentes mais relevantes são cinco: mecanismos de bloqueio (*gating mechanisms*), redes de seleção de variável (*variable selection networks*), codificadores de variáveis estáticas (*static covariate encoders*), processamento temporal (*Temporal processing*) e intervalos de previsão (*Prediction intervals*).

Os mecanismos de bloqueio têm a função de pular componentes não utilizados da arquitetura, permitindo uma profundidade e complexidade de rede adaptativas, tornando o modelo adaptável para uma grande variedade de base de dados e cenários. Já as redes de seleção de variáveis será responsável por selecionar as variáveis de entrada que são relevantes a cada passo de tempo. Os codificadores de variáveis estáticas irão integrar as entradas estáticas na rede e o processamento temporal irá aprender relações temporais de longo e curto prazo. Por fim, os intervalos de previsão irão determinar o intervalo de prováveis saídas em cada horizonte de previsão.

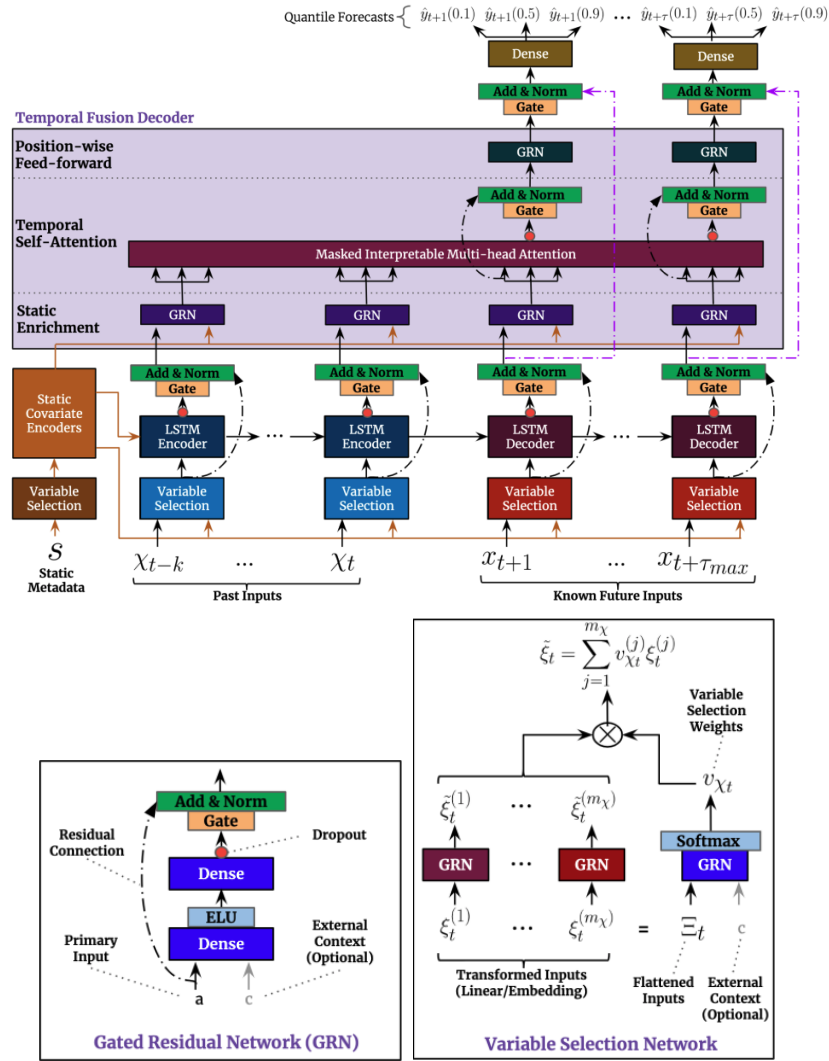


Figura 5: Ilustração extraída de [1] da arquitetura do *Temporal Fusion Transformer*.

O modelo é composto por um bloco temporal com atenção multi-cabeça responsável por identificar os padrões de longo prazo que a série temporal pode conter e priorizar aquelas que forem mais relevantes; por blocos LSTM que identificam padrões de curto prazo; por blocos GRNs (*Gated Residual Network*) que eliminam entradas que não são importantes ou que não são utilizadas, além de poderem também descartar nós aleatoriamente para evitar sobreajuste¹ (em inglês, *overfitting*).

Os tipos de entrada permitidas pelo TFT são: variáveis estáticas (constantes por série temporal), variáveis conhecidas no futuro e séries de variáveis exógenas (variáveis cujos instantes futuros não serão previstos).

Por fim, o *Temporal Fusion Transformer* utiliza como função de custo a perda por

¹Em ciência de dados, esse termo descreve quando um modelo tem um desempenho excelente no conjunto de treino, porém é ineficaz para prever novos resultados.

quantil (em inglês, *Quantile Loss*), em que Ω é o domínio dos dados de treinamento, \mathbf{W} são os pesos do TFT e Q é o conjunto de quantis de saída.

$$\mathcal{L}(\Omega, \mathbf{W}) = \sum_{y_t \in \Omega} \sum_{q \in Q} \sum_{\tau=1}^{\tau_{max}} \frac{QL(y_t, \hat{y}(q, t - \tau, \tau), q)}{M\tau_{max}} \quad (2.34)$$

$$QL(y, \hat{y}, q) = q(y - \hat{y})_+ + (1 - q)(\hat{y} - y)_+ \quad (2.35)$$

Desta forma, o *Quantile Loss* permite que o modelo realize uma previsão pontual - que é uma média condicional da variável que deve ser prevista - e também intervalos de precisão em torno da resposta, que são os percentis. Portanto, as saídas do modelo são previsões pontuais e estimativas de intervalos de precisão.

2.5 Análise de sentimento

Análise de sentimento (em inglês *Sentiment Analysis* ou *Opinion Mining*), de acordo com [2], é o estudo computacional das opiniões, atitudes e emoções das pessoas em relação a uma entidade, sendo que esta entidade pode representar indivíduos, eventos ou tópicos.

Esse estudo é dito como um processo de classificação como na Figura 6, em que o objetivo é encontrar opiniões, identificar os sentimentos que elas expressam e, por fim, classificar a polaridade desses sentimentos, segundo [2]. Além disso, há três níveis de classificação: nível de documento, nível de frase e nível de aspectos.

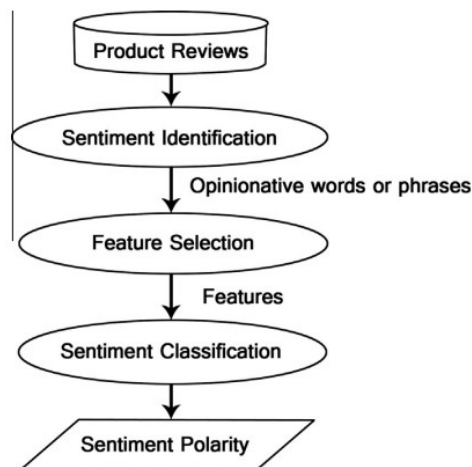


Figura 6: Processo de Análise de Sentimento. Ilustração extraída de [2].

Nível de documento – O nível de documento tem como objetivo classificar a opinião

de um documento como expressão de um sentimento negativo ou positivo. Para isso, todo o documento é considerado uma unidade básica de informação, como se tratasse de apenas um tópico, como é explicado por [2].

Nível de frase – O nível de frase deve classificar o sentimento expresso em cada sentença. Segundo [2], o primeiro passo para isso é identificar a frase como subjetiva ou objetiva e, caso seja a primeira opção, a frase é determinada como expressando sentimentos positivos ou negativos.

Nível de aspectos – O nível de aspectos é responsável por classificar o sentimento de cada frase (entidade) de acordo com as opiniões diferentes (aspectos) que são expressadas. Para isso, o primeiro passo é identificar as entidades e seus respectivos aspectos [2].

3 METODOLOGIA DO TRABALHO

A metodologia do trabalho é composta de três fases: aquisição de dados, treinamento e validação do modelo, como pode ser visto na Figura 7.

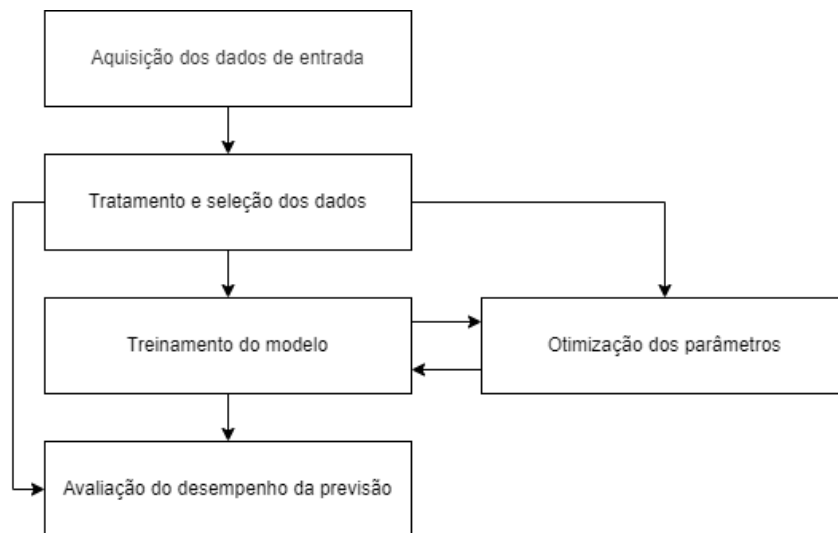


Figura 7: Fluxograma básico de um modelo preditivo com o TFT. Ilustração inspirada em [6].

O primeiro passo é a aquisição dos dados necessários para treinar e testar o modelo preditivo. Em seguida, esses dados podem ser tratados, transformados ou reduzidos a fim de remover informações ruidosas e destacar as informações importantes. Então, o preditor usa o dado tratado para treinar o modelo, neste caso, o TFT, cujos hiper-parâmetros podem ser otimizados em uma etapa de validação do treino. Por fim, o desempenho do modelo treinado com hiper-parâmetros ajustados é avaliada.

3.1 Aquisição de dados

A aquisição de dados diz respeito à série histórica de preços, aos indicadores fundamentalistas e aos indicadores econômicos, pois o índice de sentimentos foi fornecido em um arquivo formato *csv*.

Primeiramente, a série histórica de preços das ações foi obtida por meio da *Yahoo Finance API* utilizando o método *download* tendo como argumento o *ticker*¹ da ação escolhida.

Em seguida, os indicadores econômicos foram obtidos através do IPEA Data em arquivos no formato *csv*. Esta plataforma é uma base de dados macroeconômicos, financeiros e regionais do Brasil mantida pelo IPEA (Instituto de Pesquisa Econômica Aplicada).

Por fim, os indicadores fundamentalistas foram obtidos no *Status Invest* que fornece uma planilha com todos os indicadores do ano atual até o 2008. Esta planilha foi transformada em arquivo *csv* para ser utilizada no projeto.

3.2 Treinamento do modelo

O modelo utilizado foi o *Temporal Fusion Transformer* e foram realizados cinco treinamentos mudando as variáveis de entrada de acordo com a Tabela 1.

Tabela 1: Tabela com os treinamentos do TFT realizados no projeto.

Treinamento	Variáveis de Entrada
1o.	Série temporal de preços
2o.	Série temporal de preços e indicadores econômicos
3o.	Série temporal de preços, indicadores econômicos e indicadores fundamentalistas
4o.	Série temporal de preços e indicadores fundamentalistas
5o.	Série temporal de preços, indicadores econômicos e índice de sentimentos

O objetivo destes treinamentos é analisar quais entradas têm maior importância e encontrar uma configuração que resulta no melhor desempenho do modelo.

3.3 Validação do modelo

A validação do modelo utilizou três métricas de desempenho: *MASE*, *RMSE* e *MAPE*, que são comumente utilizadas em trabalhos envolvendo o TFT, como é o caso de [24].

O valor de saída das métricas *RMSE* e *MAE* tem a mesma escala dos dados utilizados para a previsão, portanto não podem ser utilizadas para comparar modelos operando em

¹Código utilizado na bolsa de valores para identificar e negociar um determinado ativo.

séries com escalas diferentes. Em contrapartida, o valor de saída da métrica *MAPE* é em porcentagem e, portanto, pode ser utilizada para comparar modelos utilizando séries em escalas diferentes.

O *MAE* (*Mean Absolute Error*, em português Erro Médio Absoluto) mensura a magnitude média do erro em uma série de previsões em n amostras, sem considerar sua direção, ao calcular a diferença entre o valor predito (y_i) e o valor real (\hat{y}_i), como é explicado em [25],

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (3.1)$$

Esta métrica não é tão afetada por valores atípicos, em inglês *outliers*, dando mais importância para grandes números de erros médios, de acordo com [26].

O *MAPE* (*Mean Absolute Percentual Error*, em português Erro Percentual Absoluto Médio) funciona da mesma forma que o *MAE*, porém há uma divisão pelo valor predito a fim da resposta ser em porcentagem,

$$MAPE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \quad (3.2)$$

sendo n o número de número de amostras, y_i o valor real e \hat{y}_i o valor predito.

Por fim, o *RMSE* (Raiz do Erro Quadrático Médio, em inglês *Root Mean Squared Error*) tem a mesma ideia do *MAE* de calcular a diferença entre o valor predito (y_i) e o valor real (\hat{y}_i); entretanto, neste caso essa diferença é elevada ao quadrado com o objetivo de penalizar mais os valores atípicos [25],

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (3.3)$$

4 ESPECIFICAÇÃO DE REQUISITOS

Este projeto deve respeitar a arquitetura proposta e uma série de requisitos funcionais e não funcionais.

4.1 Arquitetura do modelo

A arquitetura proposta neste projeto (Figura 8) é composta pelo *Temporal Fusion Transformer* e tem como entradas os índices de tempo e de tempo relativo (provenientes do TFT), a série temporal de preços, os indicadores econômicos os indicadores fundamentalistas, o índice de sentimentos e o mês. A saída é o preço previsto.

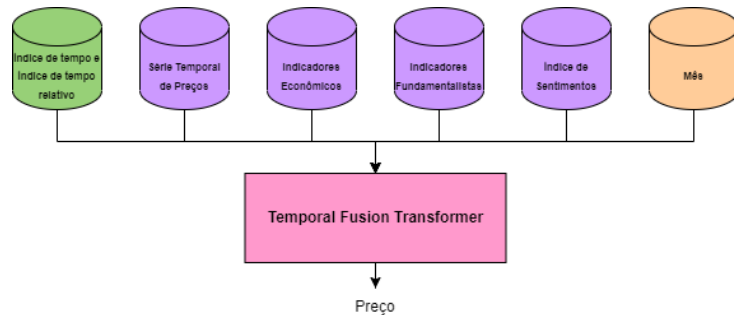


Figura 8: Arquitetura geral do projeto, composta pelo *Temporal Fusion Transformer* e as entradas definidas. Em verde estão as entradas inerentes ao TFT, em roxo as entradas do tipo variável contínua não conhecida no futuro e em laranja estão as entradas do tipo variável categórica conhecida no futuro

4.2 Requisitos funcionais

Os requisitos funcionais do projeto devem definir o que o projeto faz. Diante disso, eles são a previsão da precificação de ações, o cálculo do erro entre os preços previsto e observado e o registro dos resultados.

Previsão da precificação de ações: O principal objetivo do projeto é previsão da pre-

cificação de ativos financeiros.

Cálculo do erro entre previsto e observado: O sistema deve ser capaz de comparar a previsão com o preço observado no instante em questão, gerando métricas de erro que indiquem possíveis diferenças entre os preços previsto e observado.

Registro dos resultados: O sistema deve ser capaz de registrar os resultados obtidos, como por exemplo hiper-parâmetros do modelo testado, as especificações dos dados que entram no sistema e estatísticas de erro.

4.3 Requisitos não funcionais

Os requisitos não funcionais são responsáveis por dizer como o projeto vai realizar os requisitos funcionais. Portanto, eles são as métricas de erro baixas e a facilidade de realização, documentação e acompanhamento.

Métricas de erro baixas: As métricas de erro de validação e de teste devem se manter baixas, com a finalidade de que a precificação dos ativos em questão seja assertiva.

Facilidade para realização, documentação e acompanhamento: O projeto deve permitir a realização de experimentos, por meio da alteração de variáveis, que devem ser registradas de forma automática com métricas de desempenho, de modo que os experimentos possam ser reproduzidos e comparados.

5 DESENVOLVIMENTO DO TRABALHO

Neste capítulo são descritos todos os passos para o desenvolvimento do projeto, passando por tecnologias utilizadas, decisões de projeto, implementação da arquitetura do *software*, treinos e avaliação detalhada e, por fim, comparação entre os treinamentos.

5.1 Tecnologias utilizadas

Esta seção descreve as principais tecnologias utilizadas durante o desenvolvimento do projeto.

5.1.1 Python

Python é definido por [27] como uma linguagem de programação de código aberto¹ de alto nível² interpretada que suporta *scripts*³, é imperativa⁴ e orientada a objetos⁵, além de funcional⁶ e de tipagem dinâmica e forte.

Essa linguagem é normalmente utilizada na ciência de dados por diversos motivos, dos quais se destacam a grande comunidade de programadores, facilitando encontrar soluções para problemas recorrentes, o crescente número de bibliotecas de análise de dados, a escalabilidade e a grande quantidade de bibliotecas de visualização de dados.

Biblioteca YFinance: Segundo [28], a biblioteca *YFinance* oferece uma maneira automatizada de baixar dados de mercado do *Yahoo Finance*. Neste projeto, foi utilizado o método *download* para obter a série temporal de preços da ação que foi precificada.

¹Também conhecido como *Open Source*, é um *software* cujo código fonte pode ser visto, adaptado e utilizado por qualquer um.

²Linguagem de programação com forte abstração dos detalhes do computador.

³O termo *script* se refere a pequenos programas (pouco mais de alguns milhares de linhas de código).

⁴Possui ações, enunciados ou comandos que mudam o estado (variáveis) de um programa.

⁵Paradigma de análise, projeto e programação de sistemas de software baseado na composição e interação entre diversas unidades de software chamadas de objetos.

⁶Paradigma de programação que trata a computação como uma avaliação de funções matemáticas e que evita estados ou dados mutáveis.

Biblioteca Pandas: A biblioteca *Pandas*, segundo [29], é uma ferramenta *open source* de análise e manipulação de dados que se destaca por ser rápida, poderosa, flexível e fácil de utilizar. No projeto, ela foi utilizada para organizar as séries temporais e as entradas do modelo em *dataframes*.

Biblioteca Matplotlib: A biblioteca *Matplotlib*, segundo [30], é uma biblioteca para a criação de gráficos 2D de *arrays* em *Python*. No projeto, ela foi utilizada para construir os gráficos para melhor visualização dos resultados.

Biblioteca PyTorch Forecasting: A biblioteca *PyTorch Forecasting*, segundo [31], visa facilitar a previsão de séries temporais de última geração com redes neurais para casos do mundo real e pesquisas, por meio de uma aplicação de alto nível. Há diversos modelos de previsão implementados nesta biblioteca, destacando-se o *Temporal Fusion Transformer*, que é o modelo utilizado neste projeto. Além disso, ela também forneceu o modelo *baseline*, o *time series dataset*, o *group normalizer* e o *quantile loss*.

Biblioteca PyTorch Lightning: A biblioteca *PyTorch Lightning*, segundo [31], fornece uma interface de alto nível para o *PyTorch*. Neste projeto, ela foi responsável por fornecer o *TensorBoard Logger*, o *Early Stopping* e o *Learning Rate Monitor*.

Biblioteca Scikit-Learn: A biblioteca de código aberto *Scikit-Learn* provê ferramentas para aprendizado de máquina, como por exemplo algoritmos de classificação e regressão, além de diversas métricas de análise de desempenho e outras implementações [32]. Neste projeto, ela foi utilizada para o cálculo das métricas de avaliação de desempenho.

5.1.2 Google Colaboratory

O *Google Colaboratory* é um serviço de nuvem gratuito hospedado pelo *Google* que permite programar em *Python* e utilizar GPUs sem custo financeiro, dentro outras vantagens, como é descrito em [33]. Neste projeto, ele foi utilizado para desenvolver e treinar a arquitetura proposta com GPU.

5.1.3 GitHub

O *GitHub* é descrito por [34] como uma plataforma de hospedagem de código-fonte e arquivos com controle de versão, sendo, portanto, uma plataforma completa para cons-

truir, escalar e entregar *softwares* seguros. Neste projeto, ele foi utilizado para controlar as versões e para armazenar o conteúdo desenvolvido.

5.2 Decisões de projeto

Nesta seção são detalhados os processos de decisões de projeto que é dividido em duas partes: definição das entradas e definição do ativo.

5.2.1 Definição das entradas

As variáveis de entrada para o modelo são as séries históricas de preço e de indicadores fundamentalistas dos ativos selecionadas, além da série histórica de indicadores econômicos e da análise de sentimento de notícias dos setores selecionados.

5.2.1.1 Indicadores fundamentalistas

Os indicadores fundamentalistas que são utilizados no projeto, conforme a Tabela 2, são indicadores de *valuation* (treze indicadores, cujas equações podem ser vistas em 2.1 a 2.13), indicadores de endividamento (seis indicadores, explicitados nas equações 2.22 a 2.27), indicadores de eficiência (quatro indicadores, com as equações em 2.18 e 2.21) e indicadores de rentabilidade (quatro indicadores, que podem ser vistos em 2.14 a 2.17).

5.2.1.2 Indicadores econômicos

Os indicadores econômicos selecionados foram INPC, IGP-M, IPCA, Taxa de Juros Selic (acumulada no mês), Taxa de Juros CDI (acumulada no mês) e IBOVESPA (índice mensal), conforme a Tabela 3.

O Índice Nacional de Preços ao Consumidor (INPC) visa orientar os reajustes dos salários dos trabalhadores por meio da "mensuração das variações de preços da cesta de consumo da população assalariada com mais baixo rendimento" [35]. O cálculo é feito em três partes, como está pontuado em [35]:

Primeira parte – Cálculo da média aritmética dos preços dos produtos pesquisados no mês corrente em diferentes estabelecimentos, comparando o resultado final obtido no mês anterior.

Tabela 2: Tabela com os indicadores fundamentalistas selecionados como entradas do projeto.

Nome	Classe	Equação
<i>M.BRUTA</i>	Eficiência	2.18
<i>M.EBITDA</i>	Eficiência	2.19
<i>M.EBIT</i>	Eficiência	2.20
<i>M.LIQ</i>	Eficiência	2.21
<i>DIV.LIQUIDA/PL</i>	Endividamento	2.22
<i>DIV.LIQUIDA/EBITDA</i>	Endividamento	2.23
<i>DIV.LIQUIDA/EBIT</i>	Endividamento	2.24
<i>PL/ATIVOS</i>	Endividamento	2.25
<i>PASSIVOS/ATIVOS</i>	Endividamento	2.26
<i>LIQ.CORRENTE</i>	Endividamento	2.27
<i>ROE</i>	Rentabilidade	2.14
<i>ROA</i>	Rentabilidade	2.15
<i>ROIC</i>	Rentabilidade	2.16
<i>GiroAtivos</i>	Rentabilidade	2.17
<i>D.Y</i>	<i>Valuation</i>	2.1
<i>P/L</i>	<i>Valuation</i>	2.2
<i>P/VP</i>	<i>Valuation</i>	2.3
<i>EV/EBITDA</i>	<i>Valuation</i>	2.4
<i>EV/EBIT</i>	<i>Valuation</i>	2.5
<i>P/EBITDA</i>	<i>Valuation</i>	2.6
<i>P/EBIT</i>	<i>Valuation</i>	2.7
<i>VPA</i>	<i>Valuation</i>	2.8
<i>P/ATIVO</i>	<i>Valuation</i>	2.9
<i>LPA</i>	<i>Valuation</i>	2.10
<i>P/SR</i>	<i>Valuation</i>	2.11
<i>P/CAP.GIRO</i>	<i>Valuation</i>	2.12
<i>P/ATIVOCIRC.LIQ</i>	<i>Valuation</i>	2.13

Tabela 3: Tabela com os indicadores econômicos selecionados como entradas do projeto.

Indicador Econômico
Índice Nacional de Preços ao Consumidor (INPC)
Índice Geral de Preços - Mercado (IGP-M)
Índice Nacional de Preços ao Consumidor-Amplo (IPCA)
Taxa de Juros Selic
Taxa de Juros CDI
IBOVESPA

Segunda parte – Cálculo do índice do subitem por meio da média geométrica simples para agregação dos resultados dos produtos pertencentes ao subitem.

Terceira parte – Aplicação da fórmula de Laspeyres da Equação 5.1 (definida em [36]) em todos os níveis superiores de agregação.

$$L_{T/0} = \sum_{k=1}^K W_{0,k} R_{T,k} = \sum_{k=1}^K W_{0,k} \left(\frac{P_{T,k}}{P_{0,k}} \right) \quad (5.1)$$

, onde

$P_{T,k}$ – Preço do subitem k no período T ;

$P_{0,k}$ – Preço do subitem k no período base 0;

$R_{T,k}$ – Relativo de preços do subitem k no período T comparado com o período base 0;

K – Número total de subitens que compõem o item considerado;

$W_{0,k}$ – Peso do subitem k ;

O Índice Geral de Preços - Mercado (IGP-M) é uma medida abrangente da movimentação de preços, indicando mensalmente o nível de atividade econômica do país, como é explicado em [37]. O cálculo é uma média aritmética ponderada do IPA (Índice de Preços ao Produtor Amplo), do IPC (Índice de Preços ao Consumidor) e do INCC (Índice Nacional de Custo da Construção), como mostra a Equação 5.2.

$$IGP - M = 0.6 \cdot IPA + 0.3 \cdot IPC + 0.1 \cdot INCC \quad (5.2)$$

O Índice Nacional de Preços ao Consumidor-Amplo (IPCA) mede "a inflação de um conjunto de produtos e serviços comercializados no varejo" [38]. O cálculo deste índice é descrito da mesma forma que o cálculo do INPC.

A diferença entre INPC e IPCA, portanto, é que o primeiro abrange famílias assalariadas com rendimentos entre 1 a 5 salários mínimos, fazendo uma cobertura populacional de 50% das famílias em área de cobertura do SNIPC (Sistema Nacional de Índices de Preços ao Consumidor), já o segundo reúne famílias cujo rendimento varia de 1 a 40 salários mínimos independentemente da fonte de renda, com cobertura populacional de 90% das famílias em área de cobertura do SNIPC.

A taxa Selic indica a taxa de juros básica do Brasil e seu cálculo é feito diariamente por meio de uma "média ponderada do volume financeiro das taxas de juros de operações compromissadas¹ com títulos públicos, no prazo de um dia útil" [39].

¹Operações de compra e a venda de títulos com compromisso de revenda assumido pelo comprador e compromisso de recompra assumido pelo vendedor.

A taxa de Juros CDI (Certificado de Depósito Interbancário) é uma média diária referente aos juros cobrados nas operações de empréstimos entre bancos.

Por fim, o IBOVESPA (índice mensal) é ”o principal indicador de desempenho das ações negociadas na B3” [40] e é resultado de uma carteira teórica de investimentos que corresponde a cerca de 80% do número de negócios e do volume financeiro do mercado de capitais brasileiro aplicando a Equação 5.3.

$$IBOVESPA = \sum \text{Preço} \cdot \text{Peso da Empresa} \quad (5.3)$$

5.2.1.3 Análise de sentimento de notícias

Os dados utilizados para o treinamento do modelo de análise de sentimento são de notícias extraídas dos veículos de comunicação Exame, *Info Money* e Valor Econômico.

O modelo utilizado é um modelo de linguagem baseado na arquitetura BERT¹, denominado *FinBERTPTBR*, que constrói um índice de sentimentos do mercado baseado nas classificações das notícias ao longo do tempo de acordo com a Equação 5.4 explicada em [42].

$$BSI_t^i = \frac{Pos_t^i - Neg_t^i}{Pos_t^i + Neu_t^i + Neg_t^i} \quad (5.4)$$

, onde

BSI_t^i – Índice de sentimentos para a ação i no tempo t ;

Pos_t^i – Número de notícias positivas para a ação i no tempo t ;

Neu_t^i – Número de notícias neutras para a ação i no tempo t ;

Neg_t^i – Número de notícias negativas para a ação i no tempo t ;

5.2.2 Definição do ativo

A definição do ativo precificado neste projeto levou em conta três fatores: o volume de negociação por ação, a liquidez por setor e a quantidade de indicadores fundamentalistas disponíveis para a ação.

¹*Bidirectional Encoder Representations from Transformers* (BERT): modelo de aprendizado de máquina para o processamento de linguagem natural. [41]

5.2.2.1 Volume de negociação por ação

De acordo com [43], as três ações com maior volume médio diário de negociação em 2021 (janeiro a novembro) foram Vale (VALE3), Petrobras (PETR4), Itaú Unibanco (ITUB4) (Tabela 4).

Tabela 4: Tabela com as ações mais negociadas na bolsa em 2021. Tabela extraída de [43]

Ação	Volume médio diário em 2021
Vale (VALE3)	R\$ 2.652.006
Petrobras (PETR4)	R\$ 2.107.543
Itaú Unibanco (ITUB4)	R\$ 1.072.465

5.2.2.2 Liquidez por setor da B3

A classificação setorial da B3, de acordo com [44], agrupa as empresas listadas de acordo com os tipos e os usos dos produtos ou serviços desenvolvidos pelas empresas, com os seguintes propósitos:

- fornecer uma identificação mais objetiva dos setores de atuação das empresas, já a partir do primeiro nível da estrutura;
- permitir uma visão sobre empresas que, embora com atividades diferentes, atuem em estágios similares da cadeia produtiva ou com produtos/serviços relacionados e tendam a responder de forma semelhante às condições econômicas;
- facilitar a localização dos setores de atuação das empresas negociadas; e
- aproximar-se de critérios utilizados pelo mercado financeiro nacional e internacional.

Diante disso, há um total de dez setores e alguns exemplos são Petróleo, Gás e Biocombustíveis, Financeiro e Saúde.

A liquidez de um ativo, segundo [45], é uma medida da facilidade e velocidade com que ele pode ser convertido em dinheiro, sem perda significativa de seu valor.

A análise da liquidez foi feita construindo carteiras igualmente ponderadas (também conhecidas como *equal weight*) de cada setor no período de um ano, ou seja, todas as ações têm a mesma importância dentro de seu setor.

Desta forma, obtém-se o volume médio de negociação anual por setor, como pode ser visto no gráfico de setores da Figura 9. Os setores com maior volume de negociação são Petróleo, Gás e Biocombustíveis, Financeiro e Consumo Cíclico.

Proporção de cada setor da B3 em relação ao volume de negociações

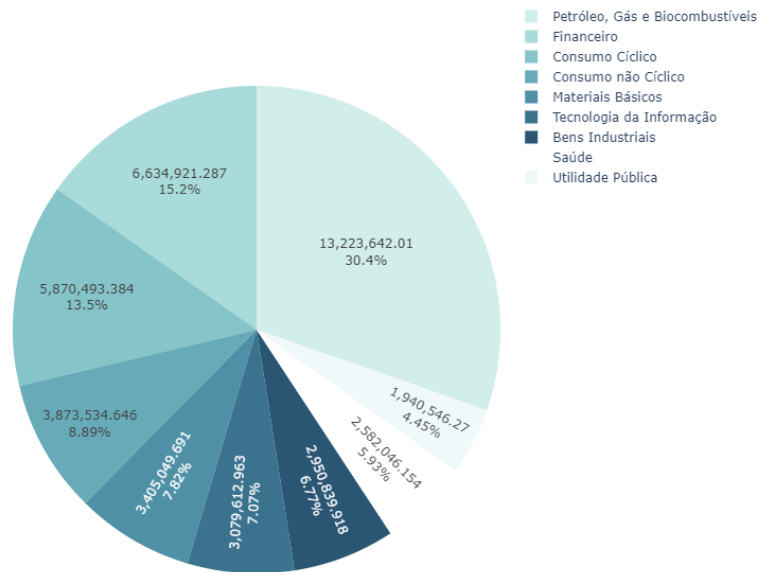


Figura 9: Proporção de cada setor da B3 em relação ao volume de negociações. Figura do autor.

Portanto, como a ação Vale (VALE3) é do setor Materiais Básicos e não está entre os setores com maior volume de negociação, ela não foi escolhida.

5.2.2.3 Indicadores fundamentalistas disponíveis por ação

De acordo com os resultados das seções anteriores, restaram apenas as ações Petrobras (PETR4) e Itaú Unibanco (ITUB4).

A escolha entre os dois foi feita levando em conta o fato de que as ações Petrobras (PETR4) têm indicadores fundamentalistas por um período maior de tempo (desde 2008) [46] do que as ações Itaú Unibanco (desde 2009) [47].

Portanto, dentre as opções, o código das ações que serão precificadas neste projeto será PETR4, ações preferenciais da Petrobras que são negociadas na B3.

5.3 Implementação da arquitetura

A implementação da arquitetura do *software* foi feita de acordo com a Figura 10 utilizando a implementação do *Temporal Fusion Transformer* da biblioteca *PyTorch Forecasting*.

A variável de saída é o preço de fechamento da ação (*close*, da série temporal de preços) e as variáveis de entrada fixas são o mês, o índice de tempo e o índice de tempo relativo, detalhados na Tabela 5.

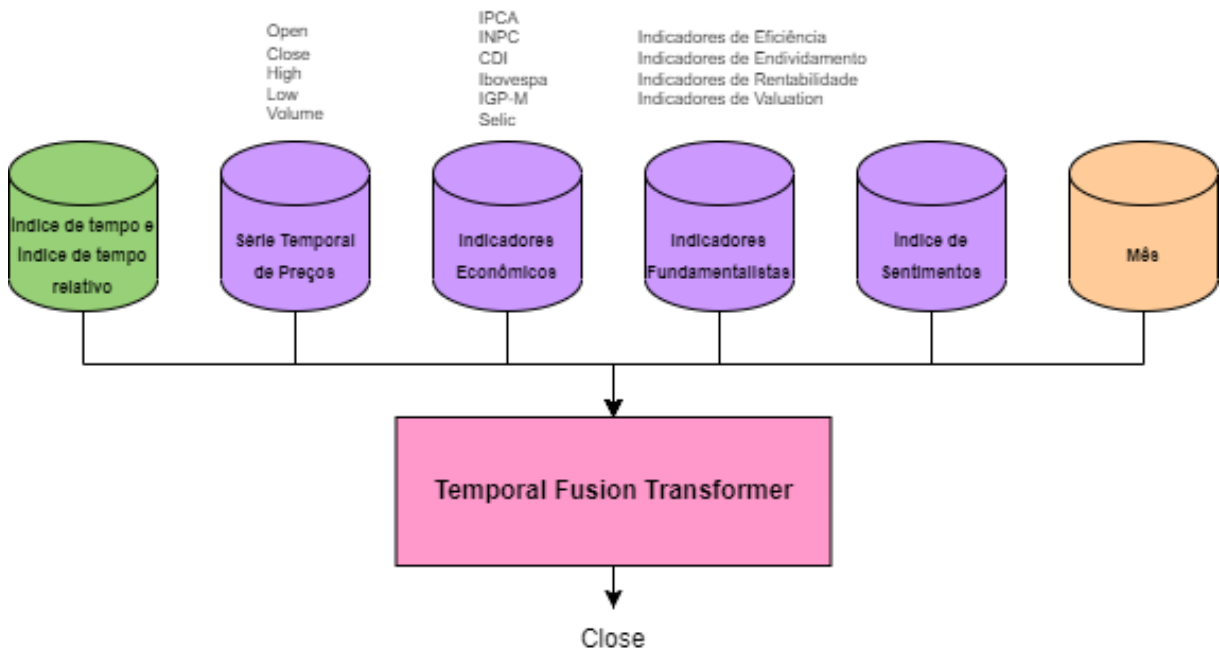


Figura 10: Arquitetura geral do projeto, composta pelo *Temporal Fusion Transformer* e as entradas definidas

Tabela 5: Tabela com as variáveis de entrada e saída comuns a todos os treinamentos.

Variáveis de entrada fixas
<i>Mês</i> : variável categórica conhecida no futuro
<i>Índice de tempo relativo</i> : <i>relative time index</i> do TFT
<i>Índice de tempo</i> : <i>time index</i> do TFT
Variável de saída
<i>Preço de fechamento</i> : <i>close</i> (da série temporal de preços)

As variáveis de entrada que são utilizadas ou não dependendo do treinamento são a série temporal de preços, indicadores econômicos, indicadores fundamentalistas e índice de sentimentos e estão detalhadas na Tabela 6.

Tabela 6: Tabela com as variáveis de entrada utilizadas dependendo do treinamento.

Variáveis de entrada
<i>Série temporal de preços</i> : variável contínua desconhecida no futuro
<i>Indicadores econômicos</i> : variável contínua desconhecida no futuro
<i>Indicadores fundamentalistas</i> : variável contínua desconhecida no futuro
<i>Índice de sentimentos</i> : variável contínua desconhecida no futuro

5.3.1 Taxa de aprendizado

A taxa de aprendizado ideal do modelo foi definida utilizando o método *Tuner* do *PyTorch Lightning* que, dado um intervalo de possíveis taxas de aprendizado, realiza testes e define um bom valor inicial.

Para isso é necessário passar como parâmetros o modelo utilizado, os conjuntos de teste e validação e os valores mínimo e máximo de taxa de aprendizado a serem investigados (*min lr* e *max lr*).

5.3.2 Prevenção de sobreajuste

A prevenção de sobreajuste nos treinamentos foi feita utilizando o *Early Stopping* da biblioteca *PyTorch Lightning* que têm como parâmetros principais:

Parâmetros:

- *Monitor*: Métrica a ser monitorada
- *Min delta*: Alteração mínima em *monitor* para se qualificar como uma melhoria
- *Patience*: Número de épocas sem melhora após o qual o treinamento será interrompido
- *Mode*: Pode ser *min* ou *max*, no primeiro o *min delta* deve ser negativo e no segundo, positivo

5.3.3 Registro de resultados

O registro dos resultados de todos os treinos foi feito utilizando o *Tensor Board Logger* da biblioteca *PyTorch Lightning* que salva o melhor modelo de cada treinamento, possibilitando reavê-lo a qualquer momento sem precisar retreinar o modelo.

5.3.4 Hiper-parâmetros

Os hiper-parâmetros utilizados no projeto estão na Tabela 7.

Tabela 7: Tabela com os hiper-parâmetros utilizados.

Hiper-parâmetros
<i>Max prediction length: 54 meses</i>
<i>Max encoder length: 542</i>
<i>Min encoder length: 271</i>
<i>Batch size: 128</i>
<i>Hidden size: 16</i>
<i>Attention head size: 1</i>
<i>Drop out: 0.1</i>
<i>Hidden continuous size: 8</i>
<i>Output size: 7 quantis</i>
<i>Log interval: 10 batches</i>
<i>Função de perda: perda por quantil</i>
<i>Épocas: 500</i>
<i>Patience: 10 épocas</i>
<i>Monitor: val_loss</i>
<i>Mode: min</i>
<i>Min lr: 10</i>
<i>Max lr: 1^{-6}</i>

5.3.5 Modelo baseline

O modelo *baseline* utilizado para comparação de desempenho em cada treinamento foi o *Baseline* do *PyTorch Forecasting*. Ele utiliza o último valor conhecido para fazer previsões.

5.4 Treinos e avaliação detalhada

Este projeto teve cinco treinamentos de acordo com o que está descrito na Tabela 1, utilizando os hiper-parâmetros da Tabela 7 e as variáveis de entrada fixas e variável de saída descrita em Tabela 5.

5.4.1 Primeiro treinamento

As variáveis de entrada contínuas e não conhecidas no futuro do primeiro treinamento são *open* (preço de abertura), *close* (preço de fechamento), *high* (preço máximo), *low*

(preço mínimo) e volume (de negociação), que são informações que compõem a série temporal de preços da PETR4 (vide Tabela 8).

Tabela 8: Variáveis de entrada contínuas não conhecidas no futuro do primeiro treinamento.

Variáveis de entrada do 1o. treinamento
<i>open</i> – preço de abertura
<i>close</i> – preço de fechamento
<i>high</i> – preço máximo
<i>low</i> – preço mínimo
volume de negociação

Os resultados da Tabela 9 mostram que o TFT teve um desempenho melhor do que o modelo *baseline*, tendo em vista que todas as métricas reduziram, sendo reduções de \$2.1312 e \$21.2380 no erro médio de MAE e RMSE, respectivamente e de 6.4442% no erro médio do MAPE. A comparação também pode ser feita pela Figura 11 que mostra na Figura 11a a comparação entre as métricas RMSE e MAE e na Figura 11b a comparação entre as métricas MAPE para os modelos *baseline* e TFT.

Os valores de RMSE e MAE para o TFT indicam erros médios de aproximadamente \$10.9210 e \$2.8069, respectivamente, ou ainda 8.6042% do preço de fechamento de acordo com o MAPE. Como o RMSE está muito maior que o MAE (comparação feita na Figura 11a), há um indicativo que o modelo está errando muito em valores atípicos (*outliers*).

Tabela 9: Tabela com resultados do teste do *baseline* e do TFT no 1o. treinamento. Os melhores resultados estão em negrito.

	MAE (\$)	RMSE (\$)	MAPE (%)
<i>Baseline</i>	4.9381	32.1590	15.0484
TFT	2.8069	10.9210	8.6042

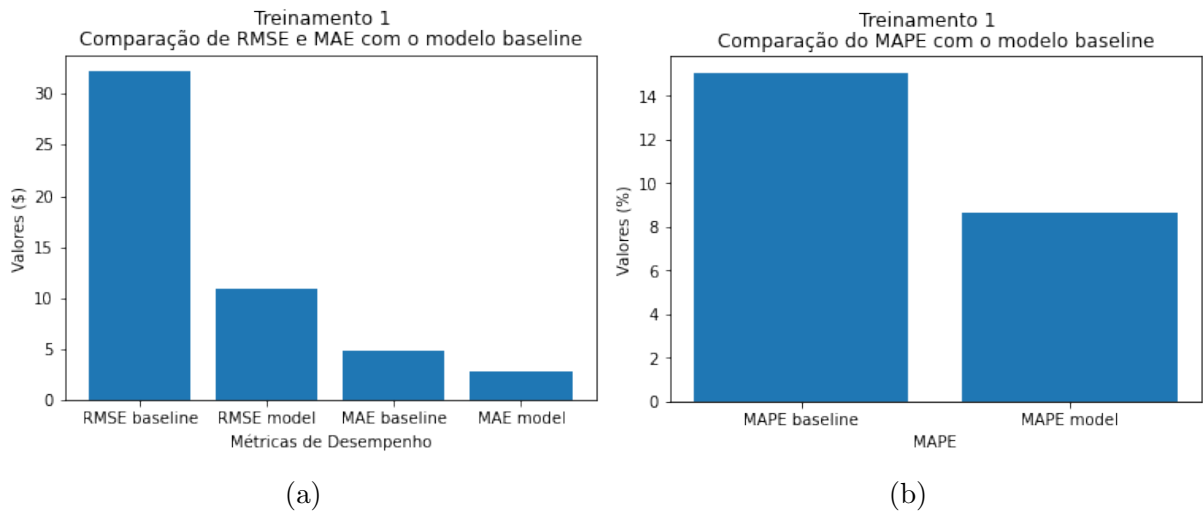


Figura 11: Resultado das métricas de desempenho no treinamento do modelo utilizando a série temporal de preços: (a) RMSE e MAE; (b) MAPE

A Figura 12a mostra como foram as previsões (*predicted*) em relação ao valor real (*observed*) com uma perda (em inglês, *Loss*) de 0.7007 e, ao fundo em cinza, é possível ver em quais momentos o modelo dedicou mais atenção - quanto mais antigas eram as informações, menor atenção foi prestada.

Por fim, na Figura 12b há uma lista das variáveis mais importantes para o treinamento do modelo, sendo que as cinco mais importantes o *high*, o índice de tempo relativo, o *open*, o mês e o índice de tempo.

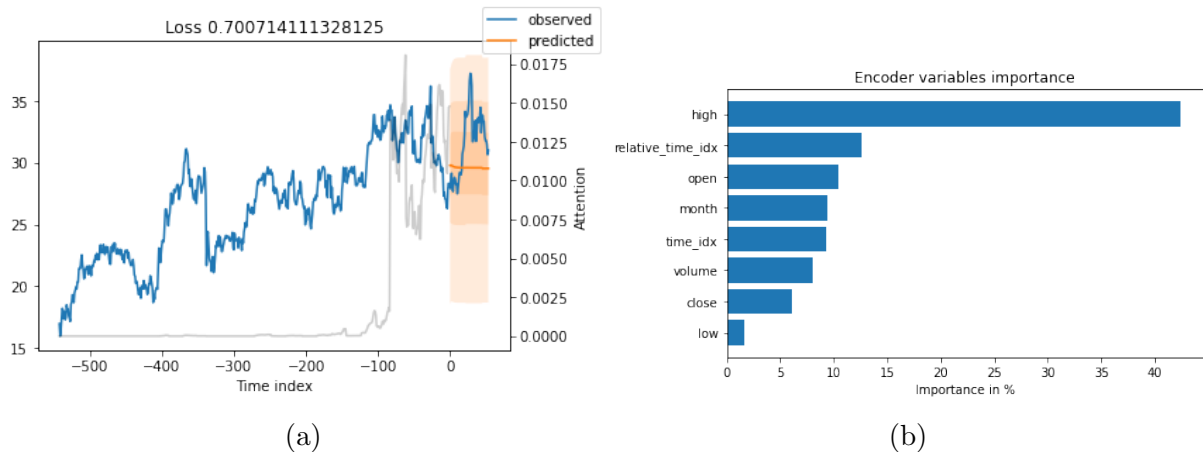


Figura 12: Treinamento do modelo utilizando apenas a série temporal de preços: (a) Desempenho do modelo: a curva em cinza é a quantidade de atenção que o modelo presta a cada momento, a curva em azul é o valor real e a curva em laranja escuro (que corresponde a 54 meses, de acordo com o hiper-parâmetro *max prediction length*) são as previsões pontuais com os intervalos de previsão em laranja claro; (b) Variáveis com maior importância

5.4.2 Segundo treinamento

O segundo treinamento apresenta as mesmas entradas do primeiro treinamento com o acréscimo dos indicadores econômicos (IPCA, INPC, CDI, IBOVESPA, IGP-M, Selic) como variáveis de entrada contínuas não conhecidas no futuro.

Tabela 10: Variáveis de entrada contínuas não conhecidas no futuro do segundo treinamento.

Variáveis de entrada do 2o. treinamento
<i>open</i> – preço de abertura
<i>close</i> – preço de fechamento
<i>high</i> – preço máximo
<i>low</i> – preço mínimo
volume de negociação
indicadores econômicos

Novamente, o TFT teve um desempenho melhor do que o modelo *baseline*, como é possível notar na Tabela 11 e na Figura 13, sendo as diferenças de \$3.2125, \$21.5418 e 11.4188% para MAE, RMSE e MAPE, respectivamente.

Neste caso também há uma grande diferença entre RMSE e MAE para o TFT, então o modelo segue errando em valores atípicos (*outliers*).

Tabela 11: Tabela com resultados do teste do *baseline* e do TFT no 2o. treinamento. Os melhores resultados estão em negrito.

	MAE (\$)	RMSE (\$)	MAPE (%)
<i>Baseline</i>	4.4983	24.4978	15.6320
TFT	1.2858	2.956	4.2132

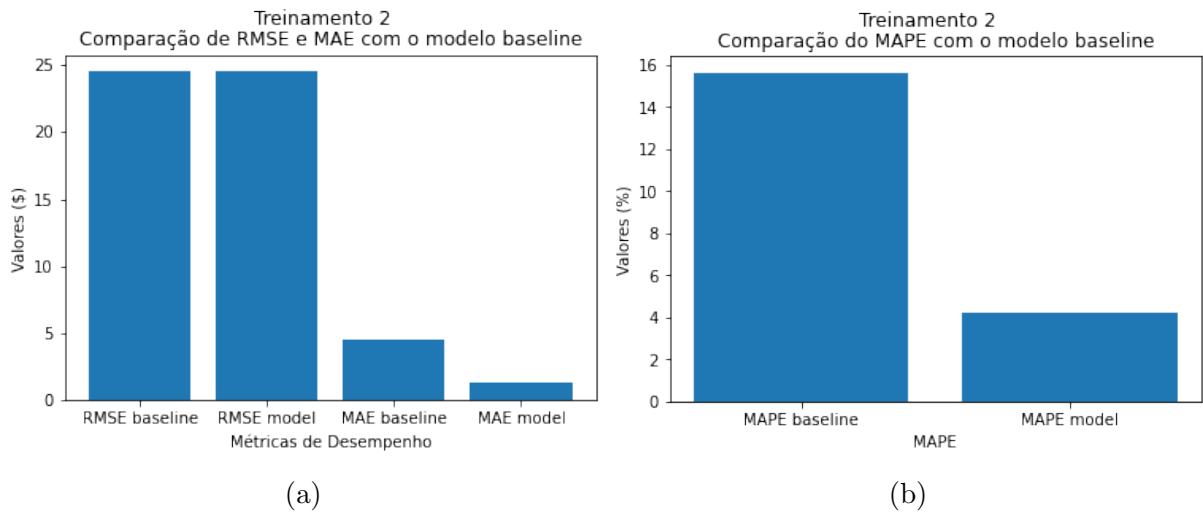


Figura 13: Resultado das métricas de desempenho no treinamento do modelo utilizando a série temporal de preços e indicadores econômicos: (a) RMSE e MAE; (b) MAPE

Na Figura 14a é possível ver a melhora significativa no desempenho do modelo com uma perda (*loss*) de 0.3633. Além disso, como é possível ver na Figura 14b, as variáveis de entrada com maior importância para este resultado foram índice de tempo relativo, *close*, IBOVESPA, *high* e mês.

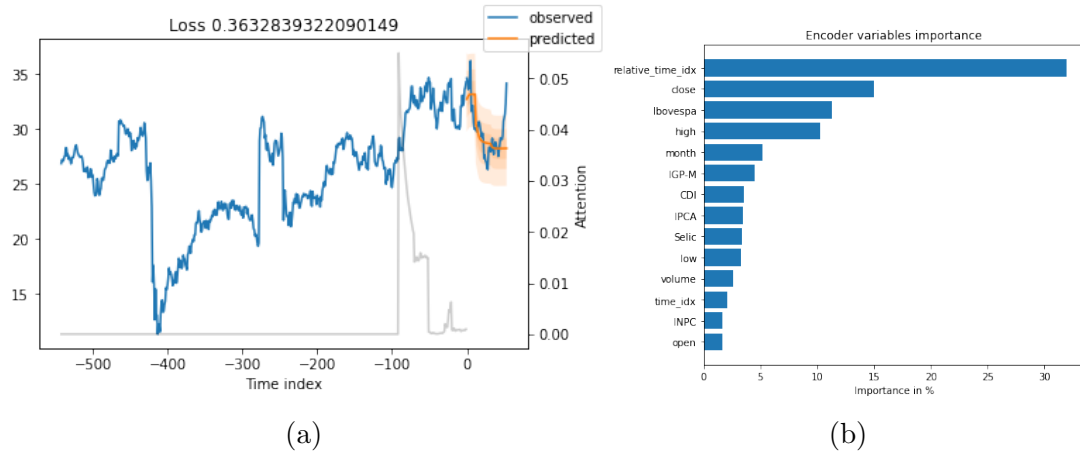


Figura 14: Treinamento do modelo utilizando a série temporal de preços e os indicadores econômicos: (a) Desempenho do modelo: a curva em cinza é a quantidade de atenção que o modelo presta a cada momento, a curva em azul é o valor real e a curva em laranja escuro (que corresponde a 54 meses, de acordo com o hiper-parâmetro *max prediction length*) são as previsões pontuais com os intervalos de previsão em laranja claro; (b) Variáveis com maior importância

5.4.3 Terceiro treinamento

Neste terceiro treinamento, foram adicionados os indicadores fundamentalistas como variáveis de entrada contínuas não conhecidas no futuro.

Tabela 12: Variáveis de entrada contínuas não conhecidas no futuro do terceiro treinamento.

Variáveis de entrada do 3o. treinamento
<i>open</i> – preço de abertura
<i>close</i> – preço de fechamento
<i>high</i> – preço máximo
<i>low</i> – preço mínimo
volume de negociação
indicadores econômicos
indicadores fundamentalistas

Como é possível ver na Tabela 13, as métricas de desempenho do modelo *baseline* se mantiveram iguais às métricas do Treinamento 2, entretanto, o desempenho do TFT foi pior com diferenças de \$2.7348, \$19.7704 e 9.7562% para o MAE, RMSE e MAPE, respectivamente.

Tabela 13: Tabela com resultados do teste do *baseline* e do TFT no 3o. treinamento. Os melhores resultados estão em negrito.

	MAE (\$)	RMSE (\$)	MAPE (%)
<i>Baseline</i>	4.4983	24.4978	15.6320
TFT	1.7635	4.7274	5.8758

Apesar do desempenho ter sido pior do que o do treinamento anterior, como é possível ver nas Figuras 15a e 15b, o desempenho do TFT ainda foi melhor do que a performance do *baseline*.

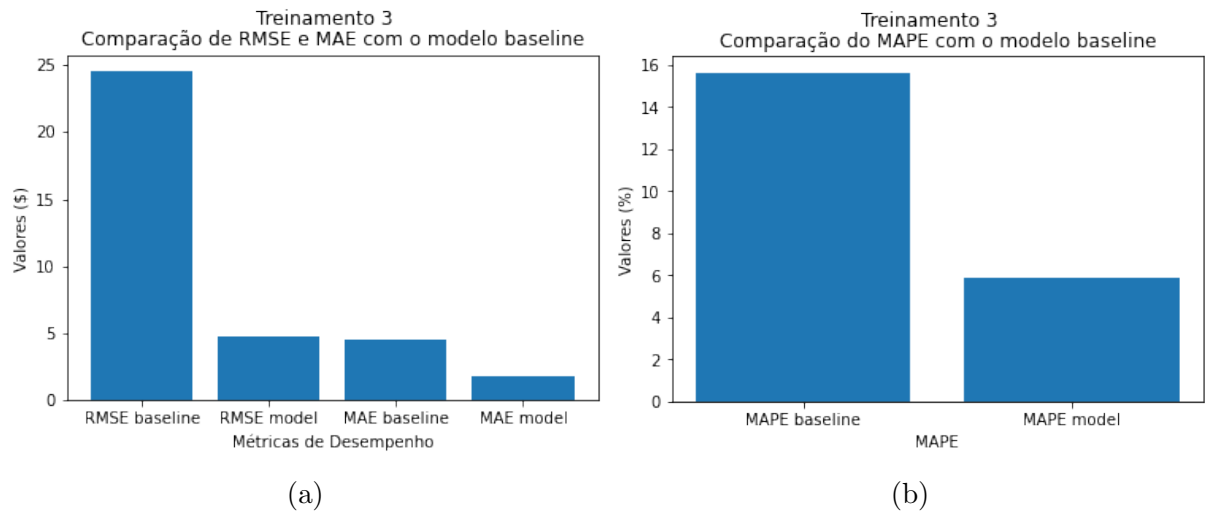


Figura 15: Resultado das métricas de desempenho no treinamento do modelo utilizando a série temporal de preços e indicadores econômicos e fundamentalistas: (a) RMSE e MAE; (b) MAPE

Na Figura 16a é possível ver como foram as previsões (*predicted*) e o valor real (*observed*) do TFT com uma perda (*Loss*) de 0.6611. Já na Figura 16b, a lista de variáveis mais importantes indica que os indicadores fundamentalistas estão nas três primeiras posições com P/VP (Equação 2.3), $M.BRUTA$ (Equação 2.18) e $EV/EBIT$ (Equação 2.5), seguido dos indicadores econômicos IPCA e Selic nas quarta e quinta posições.

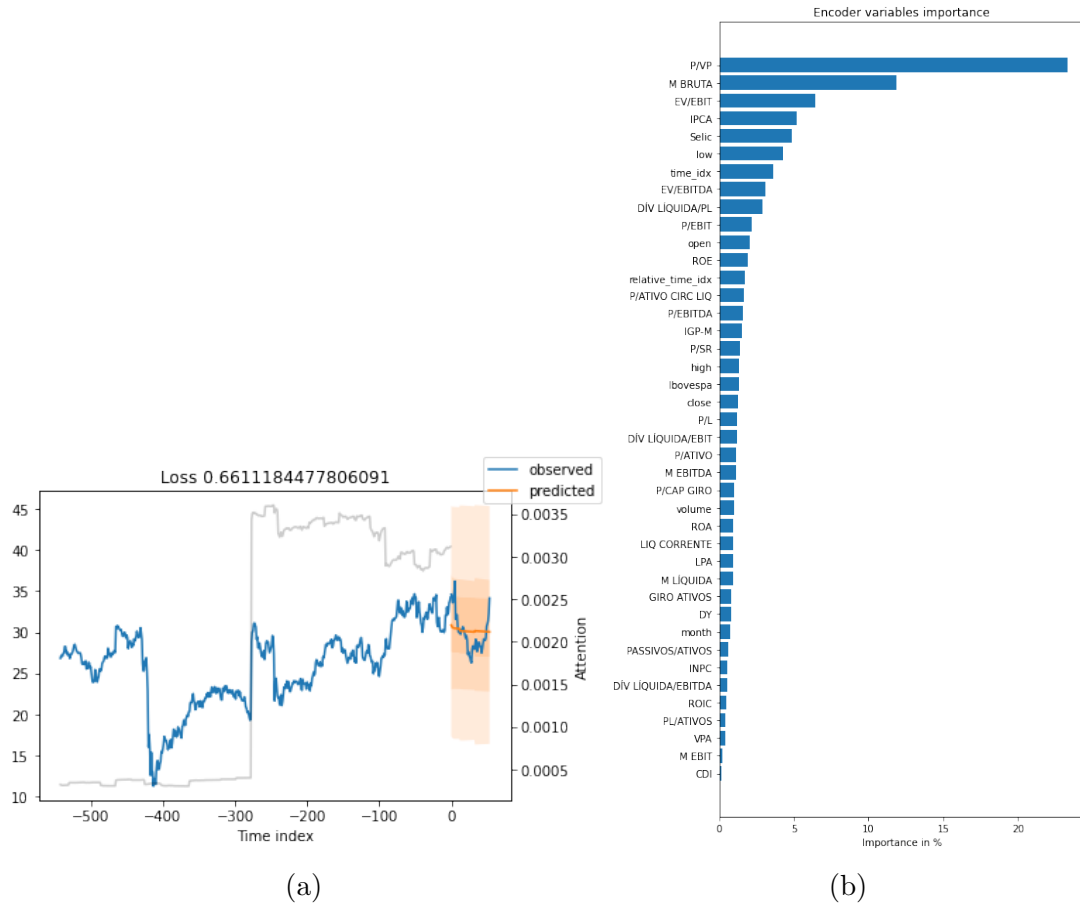


Figura 16: Treinamento do modelo utilizando a série temporal de preços e os indicadores econômicos e fundamentalistas: (a) Desempenho do modelo: a curva em cinza é a quantidade de atenção que o modelo presta a cada momento, a curva em azul é o valor real e a curva em laranja escuro (que corresponde a 54 meses, de acordo com o hiper-parâmetro *max prediction length*) são as previsões pontuais com os intervalos de previsão em laranja claro; (b) Variáveis com maior importância

5.4.4 Quarto treinamento

Neste treinamento, foram utilizadas como variáveis de entrada contínuas não conhecidas no futuro a série temporal de preços (composta por *open*, *close*, *high*, *low* e volume) e os indicadores fundamentalistas, ou seja, sem os indicadores econômicos.

Neste caso, novamente as métricas para o modelo *baseline* se mantiveram como mostra a Tabela 15, entretanto agora a performance do modelo TFT foi diferente, com uma diferença de \$1.9415, \$16.2383 e 6.8524% para o MAE, RMSE e MAPE, como pode ser visto na Figura 17.

Apesar do desempenho não ter sido melhor do que o do segundo treinamento, como é possível ver nas Figuras 17a e 17b, o desempenho do TFT ainda foi melhor do que a

Tabela 14: Variáveis de entrada contínuas não conhecidas no futuro do quarto treinamento.

Variáveis de entrada do 4o. treinamento
<i>open</i> – preço de abertura
<i>close</i> – preço de fechamento
<i>high</i> – preço máximo
<i>low</i> – preço mínimo
volume de negociação
indicadores fundamentalistas

Tabela 15: Tabela com resultados do teste do *baseline* e do TFT no 4o. treinamento. Os melhores resultados estão em negrito.

	MAE (\$)	RMSE (\$)	MAPE (%)
<i>Baseline</i>	4.4983	24.4978	15.6320
TFT	1.7635	4.7274	5.8758

performance do *baseline*.

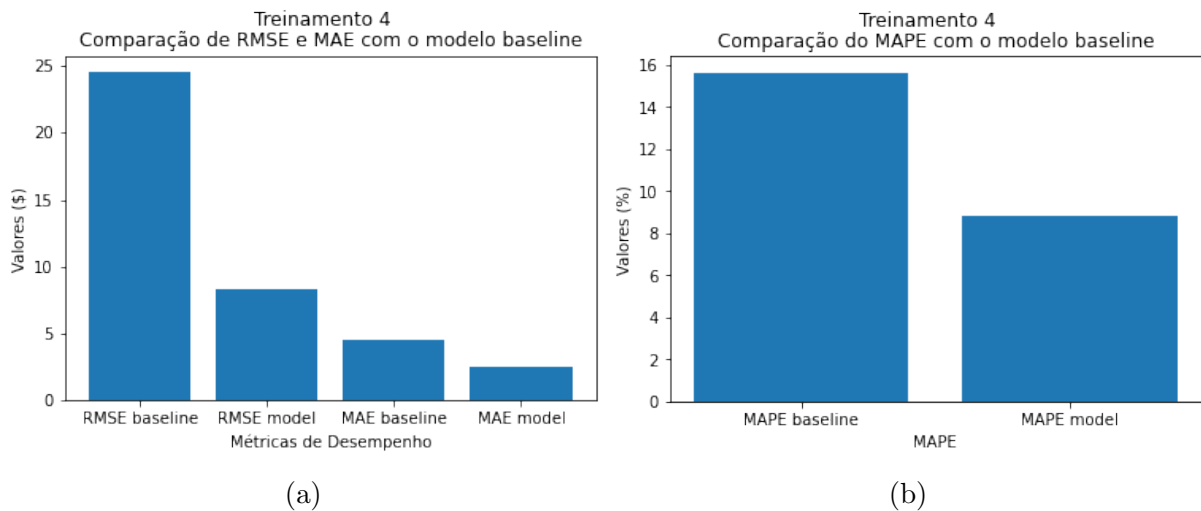


Figura 17: Resultado das métricas de desempenho no treinamento do modelo utilizando a série temporal de preços e indicadores fundamentalistas: (a) RMSE e MAE; (b) MAPE

Por fim, na Figura 18a é possível ver como o TFT se comportou no conjunto de validação com uma perda (*Loss*) de 0.6915 e na Figura 18b é possível ver que as variáveis mais importantes no treinamento foram os indicadores fundamentalistas, destacando-se *ROE* (Equação 2.14), *PASSIVOS/ATIVOS* (Equação 2.26), *P/EBIT* (Equação 2.7), *ROIC* (Equação 2.16) e *P/ATIVOCIRC.LIQ.* (Equação 2.13).

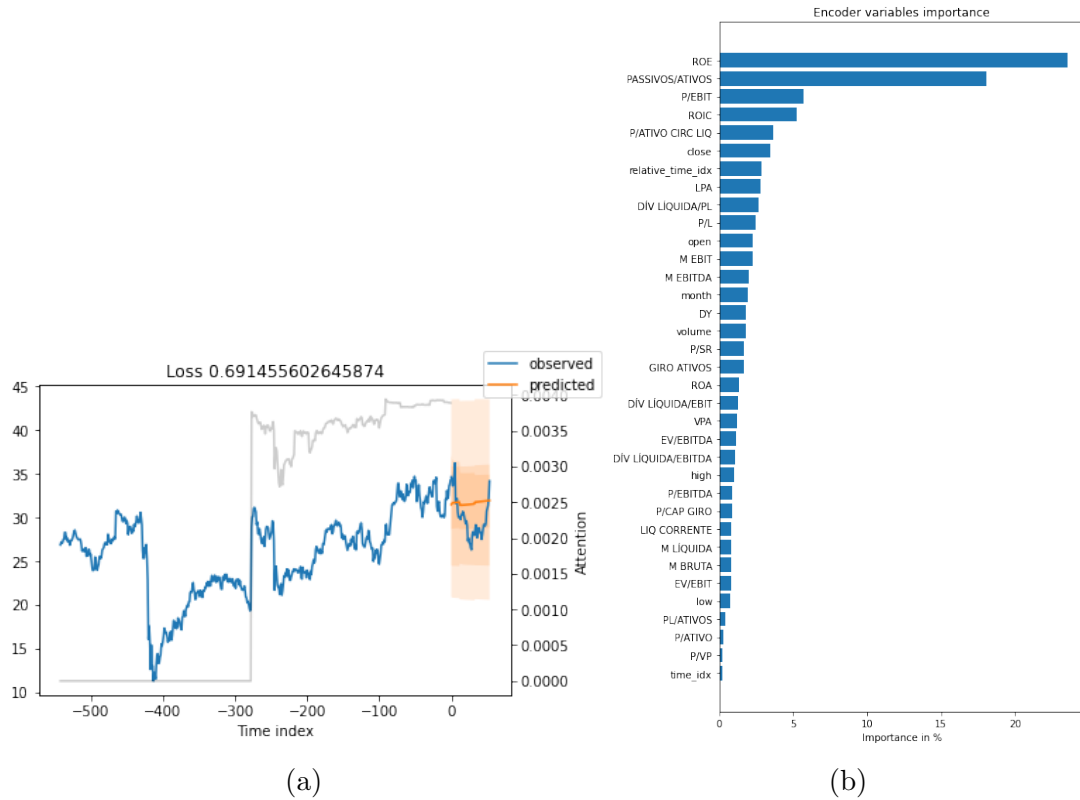


Figura 18: Treinamento do modelo utilizando a série temporal de preços e os indicadores fundamentalistas: (a) Desempenho do modelo: a curva em cinza é a quantidade de atenção que o modelo presta a cada momento, a curva em azul é o valor real e a curva em laranja escuro (que corresponde a 54 meses, de acordo com o hiper-parâmetro *max prediction length*) são as previsões pontuais com os intervalos de previsão em laranja claro; (b) Variáveis com maior importância

5.4.5 Quinto treinamento

Neste treinamento, foram utilizadas como variáveis de entrada contínuas não conhecidas no futuro a série temporal de preços (composta por *open*, *close*, *high*, *low* e volume) os indicadores econômicos e o índice de sentimentos.

Tabela 16: Variáveis de entrada contínuas não conhecidas no futuro do quinto treinamento.

Variáveis de entrada do 5o. treinamento

open – preço de abertura
close – preço de fechamento
high – preço máximo
low – preço mínimo
 volume de negociação
 indicadores econômicos
 índice de sentimentos

Neste caso, novamente as métricas para o modelo *baseline* se mantiveram como mostra a Tabela 17, entretanto agora a performance do modelo TFT teve diferenças de \$2.3765, \$18.4300 e 8.4361% para o MAE, RMSE e MAPE, como pode ser visto na Figura 19.

Tabela 17: Tabela com resultados do teste do *baseline* e do TFT com as entradas do primeiro treinamento. Os melhores resultados estão em negrito.

	MAE (\$)	RMSE (\$)	MAPE (%)
<i>Baseline</i>	4.4983	24.4978	15.6320
TFT	2.1218	6.0678	7,1958

Apesar do desempenho não ter sido melhor do que o do segundo treinamento novamente, como é possível ver nas Figuras 19a e 19b, o desempenho do TFT ainda foi melhor do que a performance do *baseline*.

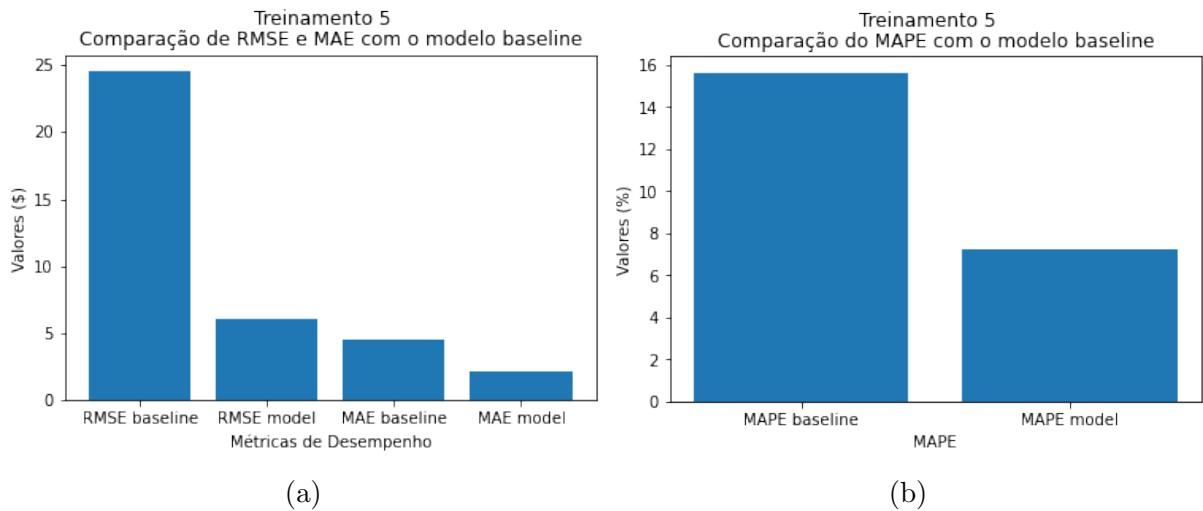


Figura 19: Treinamento do modelo utilizando a série temporal de preços, os indicadores econômicos e o índice de sentimento: (a) RMSE e MAE; (b) MAPE

Por fim, na Figura 20a é possível ver como o TFT se comportou no conjunto de validação com uma perda (*Loss*) de 0.6159 e na Figura 20b é possível ver que as variáveis mais importantes no treinamento foram *close*, índice de tempo relativo, índice de tempo, Selic e *open*. O índice de sentimentos (*sentiment_metrics*) está como uma das variáveis menos importantes, atrás apenas da variável mês.

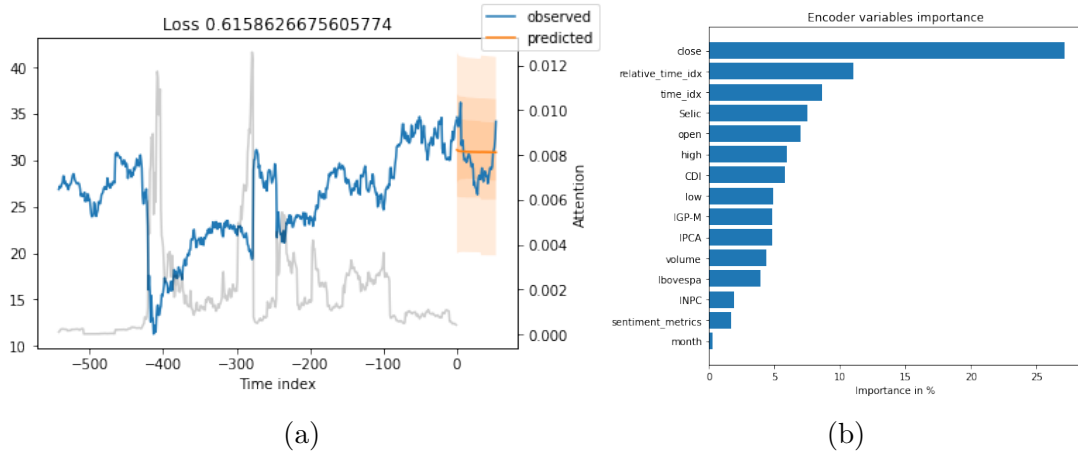


Figura 20: Treinamento do modelo utilizando a série temporal de preços e os indicadores econômicos e o índice de sentimentos: (a) Desempenho do modelo: a curva em cinza é a quantidade de atenção que o modelo presta a cada momento, a curva em azul é o valor real e a curva em laranja escuro (que corresponde a 54 meses, de acordo com o hiperparâmetro *max prediction length*) são as previsões pontuais com os intervalos de previsão em laranja claro; (b) Variáveis com maior importância

5.5 Comparação entre os treinamentos

Nas Tabelas 18 e 19 é possível ver uma comparação das métricas de desempenho do baseline e do TFT e perda (*loss*) para todos os treinamentos.

Tabela 18: Tabela com todos os resultados do teste *baseline* e do TFT.

Treinamento	1	2	3	4	5
$MAE_{baseline}$ (\$)	4.9381	4.4983	4.4983	4.4983	4.4983
MAE_{modelo} (\$)	2.8069	1.2858	1.7635	2.5568	2.1218
$RMSE_{baseline}$ (\$)	32.1590	24.4978	24.4978	24.4978	24.4978
$RMSE_{modelo}$ (\$)	10.9210	2.9560	4.7274	8.2595	6.0678
$MAPE_{baseline}$ (%)	15.0484	15.6320	15.6320	15.6320	15.6320
$MAPE_{modelo}$ (%)	8.6042	4.2132	5.8758	8.7796	7.1958

Tabela 19: Tabela com todos os valores de perda dos treinamentos do TFT.

Treinamento	1	2	3	4	5
Perda	0.7007	0.3633	0.6611	0.6915	0.6159

Em todos os treinamentos, o modelo teve um desempenho melhor do que o modelo *baseline*, pois as métricas RMSE, MAE e MAPE reduziram e, entre os quatro treinamentos, o segundo teve a melhor performance (treinamento com a série temporal de preços e indicadores econômicos), como é mostrado na Figura 21.

Os treinamentos 2, 3 e 4 tiveram um desempenho melhor do que o primeiro treinamento, então os indicadores econômicos e os indicadores fundamentalistas têm relevância na determinação do preço da ação PETR4 e não apenas a série temporal de preços.

Já a comparação entre os treinamentos 2 e 3 mostra que o acréscimo dos indicadores fundamentalistas aos indicadores econômicos acrescenta ruído ao modelo ou faz com que ele estabeleça correlações que não são verdadeiras.

Por fim, a comparação entre os treinamentos 2 e 4 mostra que, provavelmente, a análise macroeconômica tem um impacto maior sobre o preço da ação PETR4.

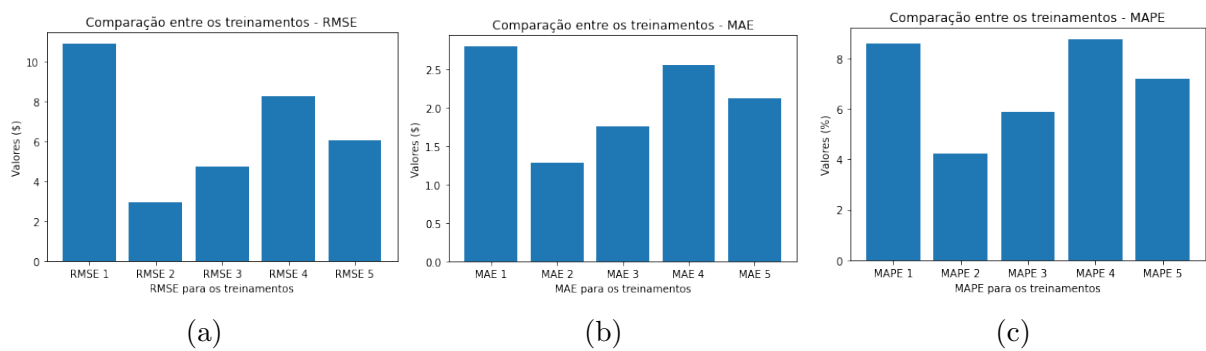


Figura 21: Comparação entre as métricas de desempenho de todos os treinamentos: (a) RMSE; (b) MAE; (c) MAPE

6 CONSIDERAÇÕES FINAIS

As considerações finais do projeto visam explicitar as conclusões obtidas e relatar possíveis perspectivas de continuidade.

6.1 Conclusões do projeto de formatura

Este projeto contou com cinco treinamentos de acordo com a Tabela 1 e o segundo treino foi o que teve melhor desempenho.

Diante disso, a arquitetura final proposta está na Figura 22 e é composta pelas entradas descritas na Tabela 20 (mês, índice de tempo, índice de tempo relativo, série temporal de preços e indicadores econômicos) e pelos hiper-parâmetros da Tabela 21.

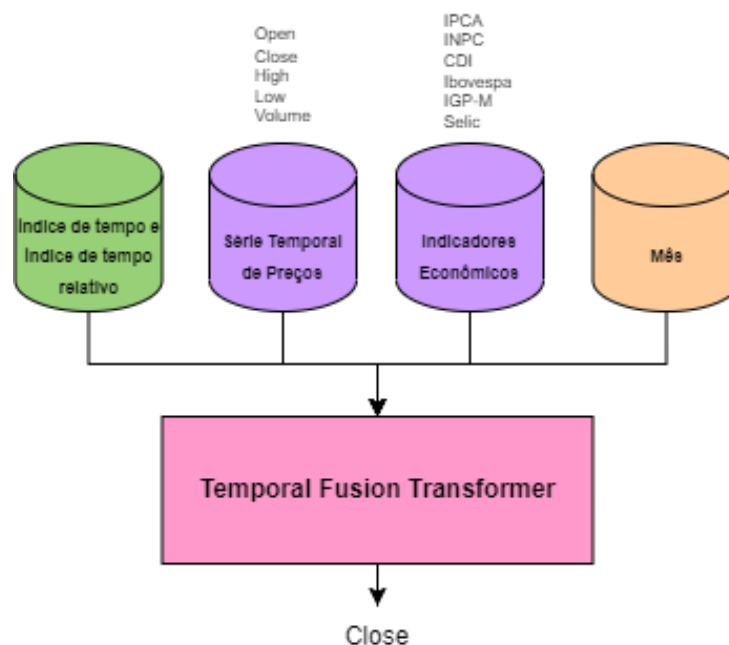


Figura 22: Arquitetura final do projeto, composta pelo *Temporal Fusion Transformer* e as entradas definidas.

Esta configuração teve a menor perda (em inglês, *Loss*), como pode ser visto na

Tabela 20: Variáveis de entrada contínuas do melhor treinamento.

Variáveis de entrada do melhor treinamento
<i>month</i> – mês
<i>relative time index</i> – índice de tempo relativo
<i>time index</i> : índice de tempo
<i>open</i> – preço de abertura
<i>close</i> – preço de fechamento
<i>high</i> – preço máximo
<i>low</i> – preço mínimo
volume de negociação
INPC
IGP-M
IPCA
Taxa de Juros Selic
Taxa de Juros CDI
IBOVESPA

Tabela 21: Tabela com os hiper-parâmetros utilizados no melhor treino.

Hiper-parâmetros
<i>Max prediction length</i> : 54 meses
<i>Max encoder length</i> : 542
<i>Min encoder length</i> : 271
<i>Batch size</i> : 128
<i>Hidden size</i> : 16
<i>Attention head size</i> : 1
<i>Drop out</i> : 0.1
<i>Hidden continuous size</i> : 8
<i>Output size</i> : 7 quantis
<i>Log interval</i> : 10 <i>batches</i>
<i>Função de perda</i> : perda por quantil
<i>Épocas</i> : 500
<i>Patience</i> : 10 épocas
<i>Monitor</i> : <i>val_loss</i>
<i>Mode</i> : <i>min</i>
<i>Min lr</i> : 10
<i>Max lr</i> : 1^{-6}

Tabela 19, e erros baixos de MAE, RMSE e MAPE, de acordo com a Tabela 18, entre todos os treinamentos.

Diante disso, há duas principais conclusões: a primeira é de que muitas entradas geram ruído e aceleram o sobre-ajuste do modelo, justificando porque a adição de indicadores fundamentalistas e índice de sentimentos aos indicadores econômicos e série temporal de preços não melhorou o desempenho; e a segunda é uma explicação para os indicadores

fundamentalistas terem o pior desempenho.

A primeira conclusão diz respeito ao fato de que muitas entradas podem levar o modelo a estabelecer relações não verdadeiras entre os dados no conjunto de treino, acelerando o sobre-ajuste e, por conseguinte, deteriorando o desempenho.

A respeito dos indicadores fundamentalistas, uma explicação para eles terem o pior desempenho (Treinamento 4, em comparação com os indicadores econômicos, e treinamento 3, em comparação com o índice de sentimentos) é que, segundo [5], eles são capazes de estimar o valor justo (intrínseco) da ação e, assim, permitir ganhos ao notar que algum preço está abaixo ou acima do valor real. Portanto, não necessariamente os indicadores fundamentalistas irão exprimir o valor de mercado da ação, que é o objetivo do modelo, diferentemente dos indicadores econômicos que, em sua maioria, têm relação com o preço das ações: por exemplo, o IBOVESPA (terceira variável de maior importância no treinamento 2, como mostra a Figura 14b) que indica desempenho médio das cotações das ações negociadas na B3.

6.2 Perspectivas de continuidade

Há diversas perspectivas de continuidade para o projeto e aqui serão descritas três: tornar possível treinar o modelo para qualquer ação escolhida pelo usuário, experimentar outros tipos de entrada que podem ser relevantes e experimentar outros hiper-parâmetros em busca de um desempenho maior.

O primeiro ponto precisaria de alterações para que o modelo buscasse na *Yahoo Finance API* a série temporal de preços da ação desejada e seria necessário - caso fosse utilizar os indicadores fundamentalistas - ter este arquivo *csv* para a ação escolhida.

A respeito de experimentar outros tipos de entrada, seria necessário obter essa nova entrada e acrescentá-la no *dataframe* do modelo de acordo com a categoria que ela se encaixa.

Por fim, para experimentar outros hiper-parâmetros uma sugestão seria utilizar a função *optimize_hyperparameters()* presente na implementação do *Temporal Fusion Transformer* da biblioteca *PyTorch Forecasting*.

REFERÊNCIAS

- [1] LIM, B. et al. *Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting*. arXiv, 2019. Disponível em: <<https://arxiv.org/abs/1912.09363>>.
- [2] MEDHAT, W.; HASSAN, A.; KORASHY, H. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, v. 5, n. 4, p. 1093–1113, 2014. ISSN 2090-4479. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2090447914000550>>.
- [3] B3. *Relatório Anual 2021*. 2021. Acesso em 2022-04-06. Disponível em: <<https://ri.b3.com.br/pt-br/informacoes-financeiras/relatorio-anual/>>.
- [4] MURPHY, J. J.; MURPHY, J. J. *Technical analysis of the financial markets*. Fishkill, N.Y.: New York Institute of Finance, 1999. ISBN 0735200661 9780735200661. Disponível em: <http://www.worldcat.org/search?qt=worldcat_org_all&q=0735200661>.
- [5] WAFI, A. S.; HASSAN, H.; MABROUK, A. Fundamental analysis models in financial markets – review study. *Procedia Economics and Finance*, v. 30, p. 939–947, 2015. ISSN 2212-5671. IISES 3rd and 4th Economics and Finance Conference. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2212567115013441>>.
- [6] FERREIRA, F. G. D. C.; GANDOMI, A. H.; CARDOSO, R. T. N. Artificial intelligence applied to stock market trading: A review. *IEEE Access*, v. 9, p. 30898–30917, 2021.
- [7] DIXON, M.; HALPERIN, I.; BILOKON, P. *Machine Learning in Finance: From Theory to Practice*. Springer International Publishing, 2020. ISBN 9783030410681. Disponível em: <<https://books.google.com.br/books?id=0pruDwAAQBAJ>>.
- [8] DIXON, M.; HALPERIN, I.; BILOKON, P. *Machine Learning in Finance: From Theory to Practice*. Springer International Publishing, 2020. ISBN 9783030410681. Disponível em: <<https://books.google.com.br/books?id=0pruDwAAQBAJ>>.
- [9] JIANG, W. Applications of deep learning in stock market prediction: Recent progress. *Expert Systems with Applications*, Elsevier BV, v. 184, p. 115537, dec 2021. Disponível em: <<https://doi.org/10.1016/j.eswa.2021.115537>>.
- [10] SHERSTINSKY, A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *CoRR*, abs/1808.03314, 2018. Disponível em: <<http://arxiv.org/abs/1808.03314>>.
- [11] VASWANI, A. et al. *Attention Is All You Need*. arXiv, 2017. Disponível em: <<https://arxiv.org/abs/1706.03762>>.

- [12] PESENTE, R. *Mercados Financeiros*. 2019. Acesso em 2022-06-07. Disponível em: <<http://educapes.capes.gov.br/handle/capes/553591>>.
- [13] BOVESPA, B. *Introdução ao Mercado de Capitais*. 2010. Acesso em 2022-06-07. Disponível em: <<https://web.archive.org/web/20100601204536/http://www.bmfbovespa.com.br/pt-br/a-bmfbovespa/download/merccap.pdf>>.
- [14] INFOMONEY. *Análise Fundamentalista de Ações: como identificar empresas sólidas e rentáveis a longo prazo*. 2022. Acesso em 2022-06-07. Disponível em: <<https://www.infomoney.com.br/guias/analise-fundamentalista/>>.
- [15] INFOMONEY. *Indicadores fundamentalistas: guia para analisar empresas listadas*. 2022. Acesso em 2022-06-07. Disponível em: <<https://www.infomoney.com.br/guias/indicadores-fundamentalistas/>>.
- [16] CAPITALIZO. *Análise Fundamentalista*. 2019. Acesso em 2022-06-07. Disponível em: <<https://capitalizo.com.br/wp-content/uploads/2019/10/ebook-analise-fundamentalista-capitalizo-1.pdf>>.
- [17] BROCKWELL, P. J.; DAVIS, R. A. *Introduction to time series and forecasting*. 2nd ed. ed. [S.l.]: Springer, 2002. (Springer texts in statistics). ISBN 978-0-387-95351-9.
- [18] BROOKS, C. Univariate time series modelling and forecasting. In: _____. *Introductory Econometrics for Finance*. 2. ed. [S.l.]: Cambridge University Press, 2008. p. 206–264.
- [19] SURADHANIWAR, S. et al. Time series forecasting of univariate agrometeorological data: A comparative performance evaluation via one-step and multi-step ahead forecasting strategies. *Sensors*, v. 21, 04 2021.
- [20] MUELLER, J. P.; MASSARON, L. *Machine Learning For Dummies*. 1st. ed. [S.l.]: For Dummies, 2016. ISBN 1119245516.
- [21] ALPAYDIN, E. *Introduction to Machine Learning*. 3. ed. Cambridge, MA: MIT Press, 2014. (Adaptive Computation and Machine Learning). ISBN 978-0-262-02818-9.
- [22] DENG, L.; YU, D. Deep learning: Methods and applications. *Foundations and Trends® in Signal Processing*, v. 7, n. 3–4, p. 197–387, 2014. ISSN 1932-8346. Disponível em: <<http://dx.doi.org/10.1561/20000000039>>.
- [23] GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. [S.l.]: MIT Press, 2016. <http://www.deeplearningbook.org>.
- [24] WU, B.; WANG, L.; ZENG, Y.-R. Interpretable wind speed prediction with multivariate time series and temporal fusion transformers. *Energy*, v. 252, p. 123990, 2022. ISSN 0360-5442. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0360544222008933>>.
- [25] HYNDMAN, R. J. Another Look at Forecast Accuracy Metrics for Intermittent Demand. *Foresight: The International Journal of Applied Forecasting*, n. 4, p. 43–46, June 2006. Disponível em: <<https://ideas.repec.org/a/for/ijafaa/y2006i4p43-46.html>>.

- [26] CHAI, T.; DRAXLER, R. Root mean square error (rmse) or mean absolute error (mae)?– arguments against avoiding rmse in the literature. *Geoscientific Model Development*, v. 7, p. 1247–1250, 06 2014.
- [27] PYTHON. *What is Python? Executive Summary*. 2022. Acesso em 2022-06-07. Disponível em: <<https://www.python.org/doc/essays/blurb/>>.
- [28] YFINANCE. *YFinance*. 2022. Acesso em 2022-07-07. Disponível em: <<https://pypi.org/project/yfinance/>>.
- [29] TEAM, T. pandas development. *pandas-dev/pandas: Pandas*. Zenodo, 2020. Disponível em: <<https://doi.org/10.5281/zenodo.3509134>>.
- [30] HUNTER, J. D. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, IEEE COMPUTER SOC, v. 9, n. 3, p. 90–95, 2007.
- [31] PASZKE, A. et al. Pytorch: An imperative style, high-performance deep learning library. In: *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 2019. p. 8024–8035. Disponível em: <<http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>>.
- [32] PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011.
- [33] GOOGLE. *Olá, este é o Colaboratory*. 2022. Acesso em 2022-11-10. Disponível em: <<https://colab.research.google.com/notebooks/intro.ipynb>>.
- [34] GITHUB. *About*. 2022. Acesso em 2022-11-10. Disponível em: <<https://github.com/about>>.
- [35] IBGE. *INPC - Índice Nacional de Preços ao Consumidor*. 2022. Acesso em 2022-11-11. Disponível em: <<https://www.ibge.gov.br/estatisticas/economicas/precos-e-custos/9258-indice-nacional-de-precos-ao-consumidor.html?&t=o-que-e>>.
- [36] IBGE. *Nota Metodológica 01/2018*. 2018. Acesso em 2022-11-11. Disponível em: <https://ftp.ibge.gov.br/Precos_Indices_de_Precos_ao_Consumidor/Sistema_de_Indices_de_Precos_ao_Consumidor/Notas_Metodologicas/Nota_metodologica_012018.pdf>.
- [37] FGV, P. *IGP-M: Resultados 2022*. 2022. Acesso em 2022-11-11. Disponível em: <<https://portal.fgv.br/noticias/igpm-resultados-2022>>.
- [38] IBGE. *IPCA - Índice Nacional de Preços ao Consumidor Amplo*. 2022. Acesso em 2022-11-11. Disponível em: <<https://www.ibge.gov.br/estatisticas/economicas/precos-e-custos/9256-indice-nacional-de-precos-ao-consumidor-amplo.html?&t=o-que-e>>.
- [39] CONCURSOS, M. *Taxa SELIC e operações compromissadas*. 2022. Acesso em 2022-11-11. Disponível em: <<https://www.metodoconcursos.com.br/2021/08/taxa-selic-e-operacoes-compromissadas.html#:~:text=A%20taxa%20Selic%2C%20que%20indica,prazo%20de%20um%20dia%20%C3%BAtil.>>>.

- [40] B3. *Índice Bovespa (IBOVESPA B3)*. 2022. Acesso em 2022-11-10. Disponível em: <https://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/ibovespa.htm>.
- [41] DEVLIN, J. et al. BERT: pre-training of deep bidirectional transformers for language understanding. *CoRR*, abs/1810.04805, 2018. Disponível em: <<http://arxiv.org/abs/1810.04805>>.
- [42] HIEW, J. Z. G. et al. *BERT-based Financial Sentiment Index and LSTM-based Stock Return Predictability*. arXiv, 2019. Disponível em: <<https://arxiv.org/abs/1906.09024>>.
- [43] INVEST, E. *VALE3 lidera entre as ações mais negociadas na bolsa em 2021*. 2021. Acesso em 2022-11-10. Disponível em: <<https://exame.com/invest/mercados/as-acoes-mais-negociadas-na-bolsa-em-2021/>>.
- [44] B3. *Critério de Classificação Setorial*. 2022. Acesso em 2022-11-10. Disponível em: <https://www.b3.com.br/pt_br/produtos-e-servicos/negociacao/renda-variavel/acoes/consultas/classificacao-setorial>.
- [45] ROSS, S. A.; JORDAN, B. D. *Princípios de Administração Financeira*. 2a. ed.. ed. São Paulo: Atlas, 2009. Tradução de: Essentials of Corporate Finance.
- [46] INVEST, S. *PETR4 - PETROBRAS PN: cotação e indicadores*. 2022. Acesso em 2022-11-10. Disponível em: <<https://statusinvest.com.br/acoes/petr4>>.
- [47] INVEST, S. *ITUB4 - BANCO ITAU UNIBANCO PN: cotação e indicadores*. 2022. Acesso em 2022-11-10. Disponível em: <<https://statusinvest.com.br/acoes/itub4>>.