

UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS

NIKLAS BRANCIFORTI WENDLAND

Modelagem quantitativa para a análise da curva de juros no Brasil

São Carlos

2024

NIKLAS BRANCIFORTI WENDLAND

Modelagem quantitativa para a análise da curva de juros no Brasil

Monografia apresentada ao Curso de Engenharia de Produção, da Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Engenheiro de Produção

Orientador: Prof. Dr. André Carlos Ponce de Leon Ferreira de Carvalho

São Carlos

2024

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da
EESC/USP com os dados inseridos pelo(a) autor(a).

Wendland, Niklas Branciforti
W471m Modelagem quantitativa para a análise da curva de juros no Brasil / Niklas Branciforti Wendland; orientador André Carlos Ponce de Leon Ferreira de Carvalho. São Carlos, 2024.

Monografia (Graduação em Engenharia de Produção) -- Escola de Engenharia de São Carlos da Universidade de São Paulo, 2024.

1. Curva de juros brasileira. 2. Modelos transformers. 3. Aprendizado de máquina. 4. Mercado financeiro. I. Título.

Aos meus pais, Prof. Edson Cezar Wendland e Profa. Marcia Cristina Branciforti, professores da EESC, que foram minha maior inspiração e exemplo ao longo desta jornada.

Dedico este trabalho a vocês, que não apenas me apresentaram a USP, mas também plantaram em mim desde cedo a sede pelo conhecimento, a dedicação à excelência e o constante desafio de superar meus limites.

Vocês me ensinaram, com suas trajetórias e valores, que a educação é a base de todas as conquistas e que o esforço e a paixão pelo que fazemos nos levam a alcançar sonhos que antes pareciam distantes.

Agradeço por cada conversa, cada conselho, cada momento de apoio e, acima de tudo, por me ensinarem, através do exemplo, o significado de perseverança e comprometimento. Este trabalho não seria possível sem o que aprendi com vocês.

*Com amor e eterna gratidão,
Niklas*

AGRADECIMENTOS

Concluir esta etapa da minha vida acadêmica é um marco que não teria sido possível sem o apoio, o conhecimento e a inspiração de pessoas que me acompanharam durante esta jornada.

Agradeço, primeiramente, ao corpo docente do curso de Engenharia de Produção da Escola de Engenharia de São Carlos (EESC), por proporcionar uma formação de excelência e por transmitir não apenas conhecimento técnico, mas também valores que levarei para toda a vida.

Ao meu orientador, Prof. Dr. André Carlos Ponce de Leon Ferreira de Carvalho, expresso minha profunda gratidão por sua paciência, dedicação e pelos valiosos conselhos durante a elaboração deste trabalho. Sua orientação foi essencial para que este projeto atingisse a maturidade e a profundidade necessárias.

À minha namorada, Eugenia, pelo apoio, companheirismo e motivação durante os momentos mais desafiadores deste projeto.

Aos meus colegas de trabalho, Mauricio Pereira e Douglas Micena, devo imensa gratidão por compartilharem comigo o conhecimento prático e as experiências que enriqueceram este projeto de maneira significativa.

Por fim, agradeço à minha república, Poltergeist, por transformar meus anos na faculdade nos melhores de minha vida. O convívio, as amizades e os momentos vividos lá foram fundamentais para que essa jornada fosse não apenas produtiva, mas também memorável.

A cada uma dessas pessoas, meu mais sincero obrigado. Este trabalho é, em parte, fruto de tudo o que aprendi, vivi e compartilhei com vocês.

FOLHA DE APROVAÇÃO

Candidato: Niklas Branciforti Wendland
Título do TCC: Modelagem Quantitativa para a Análise da Curva de Juros no Brasil
Data de defesa: 13/12/2024

Comissão Julgadora	Resultado
Professor Doutor André Carlos Ponce de Leon Ferreira de Carvalho (orientador)	Aprovado.
Instituição: ICMC - Institute of Mathematics and Computer Sciences	
Professor Doutor Maicon Gouvêa de Oliveira	Aprovado
Instituição: EESC - SEP	
Professor Doutor Lucas Gabriel Zanon	Aprovado.
Instituição: EESC - SEP	

Presidente da Banca: Professor Doutor André Carlos Ponce de Leon Ferreira de Carvalho

RESUMO

WENDLAND, N. B. **Modelagem quantitativa para a análise da curva de juros no Brasil.** 2024. Trabalho de Conclusão de Curso (Bacharelado em Engenharia de Produção) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2024.

O presente trabalho explora a aplicação de um modelo transformer para a previsão da curva de juros brasileira, utilizando dados históricos de títulos públicos, a taxa SELIC e variáveis auxiliares, como datas de reuniões do COPOM. O estudo parte da fundamentação teórica sobre títulos públicos, curva de juros e aprendizado de máquina, com ênfase nos modelos transformers, para propor e avaliar a eficácia de uma metodologia inovadora na modelagem de séries temporais financeiras. A metodologia adotada incluiu a construção de um dataset robusto, a segmentação das variáveis em tenores e auxiliares, e a implementação de um modelo transformer adaptado para lidar com as especificidades da curva de juros brasileira. Foram realizadas diversas iterações e ajustes progressivos, como a introdução de penalização por volatilidade e o aumento da janela temporal de dados, culminando em análises dos resultados e limitações observadas. Os resultados indicaram que, embora o modelo tenha mostrado boa capacidade de captar padrões em períodos controlados, ele enfrentou desafios ao lidar com eventos extraordinários e a complexidade do mercado financeiro. Este estudo contribui para o campo da aplicação de aprendizado de máquina em finanças, destacando a importância de variáveis contextuais e da combinação de abordagens híbridas para aumentar a precisão das previsões.

Palavras-chave: Curva de juros brasileira. Modelos transformers. Aprendizado de máquina. Mercado financeiro.

ABSTRACT

WENDLAND, N. B. **Quantitative modeling of the Brazilian yield curve.** 2024. Monography Bachelor Final Thesis (Undergraduate in Production Engineering) – São Carlos School of Engineering, University of São Paulo, São Carlos, 2024.

This study explores the application of a transformer model for forecasting the Brazilian yield curve, using historical data from public bonds, the SELIC rate, and auxiliary variables such as COPOM meeting dates. The research begins with a theoretical foundation on public bonds, yield curve, and machine learning, with an emphasis on transformer models, to propose and evaluate the effectiveness of an innovative methodology in modeling financial time series. The adopted methodology included building a robust dataset, segmenting variables into tenors and auxiliaries, and implementing a transformer model tailored to the specificities of the Brazilian yield curve. Several iterations and progressive adjustments were conducted, including the introduction of a volatility penalty and extending the data window, culminating in an analysis of the results and observed limitations. The results indicated that, although the model demonstrated a good ability to capture patterns in controlled periods, it faced challenges when dealing with extraordinary events and the complexity of the financial market. This study contributes to the field of machine learning in finance, emphasizing the importance of contextual variables and hybrid approaches to enhance forecasting accuracy.

Keywords: Brazilian yield curve. Transformers models. Machine learning. Financial market.

LISTA DE GRÁFICOS

Gráfico 1: Curva de juros brasileira ascendente, em 19/01/2018.....	51
Gráfico 2: Curva de juros brasileira descendente, em 19/01/2024.....	52
Gráfico 3: Curva de juros brasileira plana, em 19/01/2016	53
Gráfico 4: Previsão inicial da curva de juros sem suavização	74
Gráfico 5: Previsão com dados auxiliares e Huber Loss	77
Gráfico 6: Previsão da Curva de Juros com Taxa SELIC e Janela Reduzida	80
Gráfico 7: Previsão da curva de juros com penalização de volatilidade.....	83
Gráfico 8: Previsão da curva de juros com janela expandida.....	86

SUMÁRIO

1.	Introdução.....	19
1.1	O que são títulos públicos e qual foi sua evolução histórica.....	19
1.2	Como são os títulos públicos no Brasil, como surgiram e como foram importantes para diversos governos	20
1.3	Plano Real, Banco Central, Taxa SELIC, COPOM e impacto na economia brasileira	22
1.4	Instrumentos de Controle do Mercado, Combate à Inflação e o Papel da Taxa SELIC no Estímulo Econômico Brasileiro.....	24
1.5	A Relevância da Curva de Juros e Sua Relação com as Expectativas de Mercado	26
1.6	Alta Variabilidade da Curva de Juros e a Necessidade de Instrumentos Preditivos....	28
1.7	Avanços Tecnológicos e o Papel do Aprendizado de Máquina na Análise de Dados Complexos.....	30
1.8	O Papel do Aprendizado de Máquina na Análise Econômica e Financeira	32
1.9	Proposta de Utilização de um Modelo Transformer para Prever a Curva de Juros Brasileira.....	34
2.	Fundamentação Teórica.....	37
2.1	Revisão Bibliográfica sobre Trabalhos Semelhantes	37
2.1.1	Big Data Analytics and Machine Learning in Fixed Income Investment Strategies	38
2.1.1.1	Contextualização e Objetivo	38
2.1.1.2	Fontes de Dados Utilizadas.....	38
2.1.1.3	Metodologias de Machine Learning	39
2.1.1.4	Resultados e Contribuições.....	40
2.1.1.5	Limitações e Desafios	40
2.1.2	Big Data Analytics and Machine Learning in Fixed Income Investment Strategies Error! Bookmark not defined.	
2.1.2.1	Contextualização e Objetivo	41
2.1.2.2	Metodologia	42
2.1.2.3	Resultados e Contribuições.....	42
2.1.2.4	Limitações e Desafios	43
2.1.3	Machine Learning in Fixed Income Markets: Forecasting and Portfolio Management	43
2.1.3.1	Contextualização e Objetivo	44

2.1.3.2	Metodologia	44
2.1.3.3	Resultados	45
2.1.3.4	Limitações e Desafios	45
2.1.4	Machine Learning in Fixed Income Markets: Forecasting and Portfolio Management	Error! Bookmark not defined.
2.1.4.1	Contextualização e Objetivo	46
2.1.4.2	Metodologia	46
2.1.4.3	Resultados	47
2.1.4.4	Limitações e Desafios	48
2.2	Cálculo da Curva de Juros.....	49
2.2.1	Conceito e importância da Curva de Juros	49
2.2.1.1	Definição.....	49
2.2.1.2	Importância	50
2.2.2	Principais Tipos de Curvas de Juros	50
2.2.3	Ferramentas Matemáticas e Financeiras para Construção da Curva de Juros	53
2.2.3.1	Fatores de Desconto.....	54
2.2.3.2	Taxas Spot.....	54
2.2.3.3	Taxas Foward	54
2.2.3.4	Métodos de Ajuste e Suavização.....	55
2.2.3.5	Exemplo Prático.....	55
2.3	Modelos Transformer: Arquitetura e Justificativa para o Uso	56
2.3.1	Estrutura do Modelo Transformer	56
2.3.2	Componentes Técnicos do Transformer	57
2.3.3	Vantagens do Transformer	58
2.3.4	Justificativa para o Uso no Contexto do TCC	59
3.	Metodologia	61
3.1	Coleta e Pré-processamento dos Dados	61
3.1.1	Fontes de Dados	61
3.1.1.1	Títulos Públicos Brasileiros	61
3.1.1.2	Conversão de Taxas de Juros	62
3.1.1.3	Tenores utilizados	62

3.1.1.4	Taxa SELIC e Reuniões do COPOM.....	63
3.1.2	Limpeza e Transformação dos Dados.....	63
3.1.3	Estruturação dos Dados para o Modelo	64
3.2	Configuração do Modelo Transformer.....	64
3.2.1	Estrutura do Modelo	64
3.2.2	Penalização por Volatilidade.....	65
3.2.3	Configurações Globais do Modelo	66
3.2.4	Configuração Dinâmica.....	66
3.3	Treinamento do Modelo	67
3.3.1	Divisão dos Dados	67
3.3.2	Configurações do Treinamento.....	67
3.3.3	Procedimentos de Treinamento	68
3.3.4	Avaliação no Conjunto de Teste	69
3.3.5	Visualização dos Resultados.....	69
3.4	Ferramentas e Ambiente de Desenvolvimento.....	70
3.4.1	Linguagens e Bibliotecas Utilizadas.....	70
3.4.2	Ambiente de Execução	71
4.	Resultados e Discussões.....	73
4.1	Progresso Inicial: Configuração Básica do Modelo	73
4.1.1	Gráfico 4: Previsão da Curva de Juros Sem Suavização	74
4.1.2	Análise do Gráfico 4.....	75
4.1.3	Conclusões e Ajustes Necessários	75
4.2	Ajustes Progressivos na Arquitetura e Função de Perda	76
4.2.1	Gráfico 5: Previsão com Dados Auxiliares e Huber Loss	77
4.2.2	Análise do Gráfico 5:.....	78
4.2.3	Conclusão e Ajustes Necessários.....	78
4.3	Adição de Variáveis Auxiliares e Redução da Janela.....	79
4.3.1	Gráfico 6: Previsão com Taxa SELIC e Janela Reduzida.....	80
4.3.2	Análise do Gráfico 6.....	80
4.3.3	Conclusão e Ajustes Necessários.....	81
4.4	Refinamento da Arquitetura com Penalização por Volatilidade.....	81

4.4.1	Gráfico 7: Refinamento com Penalização de Volatilidade	83
4.4.2	Análise do Gráfico 7:.....	83
4.4.3	Conclusões e Preparação para a Análise Final	84
4.5	Análise Final com Janela Expandida	84
4.5.1	Gráfico 8: Previsão com Janela Expandida	86
4.5.2	Análise do Gráfico 8.....	86
4.5.3	Conclusões Preliminares	87
5.	Conclusões e Sugestões para Trabalhos Futuros	89
5.1	Conclusões Gerais	89
5.2	Aprendizados e Reflexões do Estudo.....	91
5.3	Sugestões para Trabalhos Futuros.....	92
5.3.1	Ampliação das Variáveis Auxiliares	93
5.3.2	Segmentação Temporal.....	93
5.3.3	Integração de Modelos Híbridos	94
5.3.4	Expansão do Horizonte de Dados.....	94
5.3.5	Abordagem Contextualizada de Crises.....	95
5.4	Limitações do Estudo	95
5.5	Reflexão Final	96
6.	Referências	99

1. Introdução

A seção introdutória tem como objetivo contextualizar o tema dos títulos públicos, sua importância histórica e papel no mercado financeiro global e brasileiro. Iniciaremos explicando como esses instrumentos emergiram como ferramentas fundamentais para o financiamento de governos e desenvolvimento econômico, evoluindo ao longo dos séculos. Será destacado o contexto brasileiro, onde os títulos públicos desempenham uma função essencial tanto na gestão da dívida pública quanto como instrumentos de política monetária. Essa introdução prepara o leitor para entender o pano de fundo econômico e histórico que sustenta a relevância da curva de juros como objeto de estudo neste trabalho.

1.1 O que são títulos públicos e qual foi sua evolução histórica

Os títulos públicos são instrumentos financeiros emitidos por governos com o objetivo de captar recursos no mercado para financiar seus gastos e implementar políticas públicas. Eles funcionam como uma espécie de "emprestimo coletivo" em que investidores adquirem os títulos e, em troca, recebem a promessa de pagamento de juros e do valor principal em uma data futura. Esses instrumentos representam um elo entre a administração pública e o mercado financeiro, sendo fundamentais tanto para a gestão da dívida pública quanto para a condução de políticas econômicas.

A origem dos títulos públicos remonta à Idade Média, quando cidades-estados italianas, como Veneza e Florença, começaram a emitir instrumentos de dívida para financiar guerras e projetos públicos. Segundo Reinhart e Rogoff (2009), Veneza foi pioneira ao criar o *prestito*, um título perpétuo que oferecia retornos fixos anuais e era negociado no mercado secundário. Esse mecanismo não apenas ajudava o governo a financiar despesas de guerra, mas também criava um mercado financeiro rudimentar, que posteriormente inspiraria sistemas mais sofisticados em outras partes da Europa.

Durante o século XVII, a emissão de títulos públicos tornou-se mais sistemática, especialmente no Reino Unido e na Holanda, que desenvolveram mercados financeiros estruturados. O Reino Unido, em particular, destacou-se ao criar os primeiros títulos de longo prazo, como os *consols*, que ajudaram a financiar a expansão militar e comercial do império britânico. De acordo com Ferguson (2008), o uso de títulos públicos foi crucial para financiar a

Revolução Industrial, pois permitiu aos governos evitar aumentos excessivos de impostos enquanto impulsionavam investimentos em infraestrutura e inovação tecnológica.

Nos séculos XVIII e XIX, a prática de emissão de títulos públicos espalhou-se para outros países, acompanhando a expansão do capitalismo e a necessidade de financiar Estados em crescimento. Nos Estados Unidos, por exemplo, os títulos do Tesouro foram amplamente utilizados para financiar a Guerra Civil e, mais tarde, a Segunda Guerra Mundial. Esses períodos de emissão intensa de dívida pública marcaram a consolidação dos mercados de capitais e a introdução de conceitos como a confiança na solvência do emissor e o papel do governo como garantidor de última instância.

Com o avanço do século XX, os títulos públicos ganharam uma nova dimensão: além de sua função primária de financiamento, eles passaram a ser instrumentos centrais na condução da política monetária e na estabilização econômica. Nos Estados Unidos, o Federal Reserve utilizou a compra e venda de títulos como uma ferramenta para controlar a oferta monetária e influenciar as taxas de juros, estabelecendo um precedente que seria seguido por bancos centrais em todo o mundo (Friedman, 1968). A partir desse ponto, os títulos públicos passaram a desempenhar um papel duplo: viabilizar o financiamento dos Estados e servir como instrumento de regulação econômica.

Atualmente, os títulos públicos são elementos essenciais para a saúde financeira de qualquer país, compondo a base dos sistemas de dívida soberana e dos mercados de capitais. Sua evolução reflete não apenas as mudanças nas necessidades fiscais dos governos, mas também a crescente sofisticação do mercado financeiro global. Em um mundo cada vez mais integrado, os títulos públicos têm sido utilizados como reservas de valor por investidores institucionais e bancos centrais, além de funcionarem como uma referência de risco soberano para outros ativos financeiros.

1.2 Como são os títulos públicos no Brasil, como surgiram e como foram importantes para diversos governos

No Brasil, os títulos públicos desempenharam um papel central na organização das finanças nacionais e no desenvolvimento da economia desde o período imperial. Embora rudimentares em sua concepção inicial, esses instrumentos evoluíram ao longo dos séculos para atender às demandas

fiscais e de política econômica, assumindo múltiplas funções, desde o financiamento de infraestrutura até o combate à inflação. A história dos títulos públicos brasileiros reflete as transformações econômicas e políticas do país, sendo marcada por períodos de crise fiscal, reformas econômicas e avanços institucionais.

Os primeiros títulos públicos emitidos no Brasil datam do período imperial, no início do século XIX, como parte do esforço para financiar a infraestrutura e a independência econômica do país. Segundo Lobo (2006), durante o reinado de Dom Pedro II, os títulos públicos foram amplamente utilizados para financiar projetos de infraestrutura, como a construção de ferrovias e portos, essenciais para a integração territorial e o desenvolvimento econômico. Entretanto, a falta de um mercado financeiro estruturado dificultava a negociação desses instrumentos, limitando sua eficácia como ferramenta de gestão fiscal.

Com a proclamação da República, em 1889, o uso de títulos públicos tornou-se mais sistemático, particularmente para cobrir déficits fiscais e financiar as demandas de um Estado em expansão. No entanto, a política de emissão descontrolada de títulos em certos períodos gerou inflação e crises de confiança, como no caso do Encilhamento, no final do século XIX. Essa experiência evidenciou a necessidade de maior disciplina fiscal e de um mercado financeiro mais robusto para sustentar a dívida pública. Segundo Villela e Suzigan (1973), a criação do Banco do Brasil e de outros mecanismos financeiros foi um marco importante para estabilizar o uso de títulos públicos e fomentar o desenvolvimento do mercado de capitais no país.

Durante a primeira metade do século XX, os títulos públicos continuaram a desempenhar um papel central no financiamento governamental, especialmente em tempos de crise, como durante a Grande Depressão e a Segunda Guerra Mundial. Nesses períodos, o governo utilizou esses instrumentos para sustentar a economia e financiar projetos estratégicos. No entanto, foi somente na década de 1960, com a criação do Banco Central do Brasil, que os títulos públicos começaram a ser utilizados de forma mais sofisticada como ferramenta de política monetária. De acordo com Giambiagi e Além (2001), o Banco Central passou a emitir títulos para regular a oferta monetária e controlar a inflação, marcando uma nova era para a gestão da dívida pública no Brasil.

Nos anos seguintes, especialmente durante o regime militar (1964-1985), os títulos públicos ganharam ainda mais relevância. Durante esse período, o governo utilizou a dívida pública interna para financiar grandes projetos de infraestrutura, como a construção de usinas hidrelétricas e rodovias. Esses investimentos foram fundamentais para o crescimento econômico do país, mas

também contribuíram para o aumento da dívida pública, que se tornou um dos principais desafios econômicos nas décadas seguintes. Segundo Pastore (1994), a emissão de títulos públicos foi amplamente utilizada como forma de financiar o déficit fiscal, especialmente em um contexto de alta inflação e desequilíbrios macroeconômicos.

A partir da década de 1990, com o advento do Plano Real e a estabilização da economia brasileira, o mercado de títulos públicos passou por uma transformação significativa. A criação do Tesouro Direto, em 2002, democratizou o acesso a esses instrumentos, permitindo que investidores individuais comprassem títulos diretamente do governo, sem intermediários. Essa iniciativa, aliada a reformas institucionais, contribuiu para aumentar a transparência e a eficiência do mercado de dívida pública no Brasil. Além disso, a modernização do sistema de gestão da dívida, sob a coordenação do Tesouro Nacional e do Banco Central, permitiu ao governo melhorar o perfil da dívida pública, reduzindo sua vulnerabilidade a choques externos.

Hoje, os títulos públicos no Brasil são amplamente utilizados não apenas para financiar o déficit público, mas também como ferramentas de política monetária e de controle macroeconômico. Eles desempenham um papel crucial na formação da curva de juros, que reflete as expectativas do mercado sobre a evolução das taxas de juros e da economia. Além disso, os títulos brasileiros, como as Letras do Tesouro Nacional (LTN) e as Notas do Tesouro Nacional (NTN), são considerados ativos de referência no mercado financeiro, atraindo investidores nacionais e internacionais devido à sua liquidez e previsibilidade.

1.3 Plano Real, Banco Central, Taxa SELIC, COPOM e impacto na economia brasileira

O Plano Real, implementado em 1994, representou um divisor de águas na economia brasileira, marcando o fim de um longo período de hiperinflação que corroía o poder de compra da população e desestabilizava a economia nacional. Este plano não apenas introduziu uma nova moeda, mas também instituiu uma série de reformas macroeconômicas que tiveram impactos profundos e duradouros na política monetária e no papel das instituições financeiras do país. De acordo com Franco (2000), o sucesso do Plano Real dependia da adoção de um regime de metas para a inflação e do fortalecimento do Banco Central do Brasil como principal agente na condução

da política monetária. Nesse contexto, a taxa SELIC (Sistema Especial de Liquidação e Custódia) tornou-se um instrumento central para controlar a inflação e estabilizar a economia.

A taxa SELIC é a taxa básica de juros da economia brasileira e serve como referência para as demais taxas de juros praticadas no mercado. Ela é determinada pelo Banco Central por meio das reuniões do Comitê de Política Monetária (COPOM), criado em 1996. O COPOM é composto por membros da diretoria do Banco Central e se reúne regularmente para avaliar a conjuntura econômica e decidir o nível da taxa SELIC, com base em metas inflacionárias predefinidas. Segundo Bogdanski, Tombini e Werlang (2000), as decisões do COPOM são fundamentais para alinhar as expectativas dos agentes econômicos e sinalizar a direção da política monetária. A adoção de uma taxa básica de juros clara e transparente aumentou a previsibilidade das decisões do Banco Central, contribuindo para a credibilidade da política econômica brasileira.

A independência do Banco Central, formalizada em 2021 com a Lei Complementar nº 179, consolidou uma prática que vinha sendo gradualmente implementada desde o Plano Real. Essa independência conferiu ao Banco Central maior autonomia para implementar políticas monetárias sem interferências políticas diretas, permitindo que o COPOM ajustasse a taxa SELIC de acordo com as condições econômicas e as metas inflacionárias, sem pressões externas. Segundo Carvalho e Castro (2021), a autonomia do Banco Central foi um avanço institucional significativo, fortalecendo a confiança do mercado financeiro na capacidade do Brasil de conduzir políticas econômicas eficazes.

O papel do COPOM e das reuniões periódicas para decidir a taxa SELIC vai além da simples determinação da taxa básica de juros. Essas decisões influenciam diretamente o custo do crédito, a dinâmica da dívida pública e o comportamento dos investidores no mercado financeiro. Quando o COPOM eleva a taxa SELIC, os custos de financiamento aumentam, reduzindo a demanda agregada e ajudando a conter a inflação. Por outro lado, uma redução na SELIC tende a estimular o consumo e os investimentos, promovendo o crescimento econômico. Essa dinâmica faz da taxa SELIC um dos principais instrumentos de controle da economia brasileira, especialmente em um contexto de alta sensibilidade às flutuações inflacionárias.

Além disso, a SELIC exerce influência direta sobre o mercado de títulos públicos, já que as taxas de juros associadas a esses títulos geralmente seguem a trajetória da SELIC. Segundo Oreiro (2017), a ligação entre a taxa SELIC e os títulos públicos é uma das razões pelas quais a política monetária no Brasil tem um impacto tão significativo sobre a curva de juros. As decisões

do COPOM moldam as expectativas dos investidores quanto às condições econômicas futuras, o que, por sua vez, influencia o comportamento da curva de juros e, consequentemente, o custo de financiamento do governo.

A operacionalização da taxa SELIC e sua relação com o mercado financeiro dependem de um sistema financeiro altamente integrado e de instituições robustas. O Sistema Especial de Liquidação e Custódia, criado em 1983, foi um marco na modernização do mercado financeiro brasileiro. Ele permitiu a negociação, registro e liquidação eletrônica de títulos públicos, aumentando a eficiência e a transparência do sistema. Segundo Silva *et al.* (2011), o avanço tecnológico representado pela SELIC contribuiu para reduzir os custos de transação e aumentar a liquidez no mercado de títulos públicos, consolidando a importância desses instrumentos na condução da política monetária.

A história recente da economia brasileira mostra como o Banco Central, o COPOM e a taxa SELIC foram fundamentais para estabilizar a economia e criar um ambiente mais previsível para os agentes econômicos. Embora desafios persistam, especialmente em um cenário de volatilidade global, a consolidação dessas instituições e instrumentos posiciona o Brasil de forma competitiva no cenário econômico internacional.

1.4 Instrumentos de Controle do Mercado, Combate à Inflação e o Papel da Taxa SELIC no Estímulo Econômico Brasileiro

Antes da consolidação da taxa SELIC como principal instrumento de política monetária no Brasil, o país enfrentou décadas de instabilidade econômica marcada por uma inflação crônica que corroía o poder de compra e minava a confiança no sistema financeiro. Entre as décadas de 1980 e 1990, conhecidas como a "década perdida", o Brasil conviveu com taxas de inflação anuais que chegavam a superar 1.000%, um fenômeno que dificultava a previsibilidade econômica, desestimulava investimentos e penalizava principalmente a população de menor renda (GIAMBIAGI; ALÉM, 2001). A introdução de um regime de metas inflacionárias no final da década de 1990 e a consolidação da taxa SELIC como instrumento central da política monetária foram fundamentais para a estabilização econômica e o controle do mercado.

Historicamente, o combate à inflação no Brasil foi marcado por medidas heterodoxas e de eficácia limitada, como congelamento de preços, tabelamentos e planos econômicos que buscavam,

muitas vezes, resultados de curto prazo sem atacar as causas estruturais da inflação. O Plano Cruzado (1986) e o Plano Collor (1990) são exemplos de tentativas frustradas de estabilizar a economia, que acabaram contribuindo para o agravamento da crise inflacionária. De acordo com Barbosa (2010), essas políticas falharam em grande parte porque negligenciaram a importância de um arcabouço monetário consistente e de uma política fiscal responsável, dois pilares que só seriam consolidados com a adoção do Plano Real e das reformas subsequentes.

Com a introdução do Plano Real, em 1994, e a estabilização inicial dos preços, o Banco Central do Brasil ganhou maior protagonismo na formulação de políticas econômicas voltadas para o controle da inflação. Foi nesse contexto que a taxa SELIC começou a ser utilizada como o principal instrumento para regular a liquidez do mercado e influenciar as expectativas inflacionárias. Segundo Bogdanski, Tombini e Werlang (2000), a SELIC funciona como uma âncora para a economia brasileira, servindo como referência para o custo do crédito e para a remuneração de títulos públicos, ao mesmo tempo em que orienta as decisões de consumo e investimento.

O Brasil adotou formalmente o regime de metas para a inflação em 1999, após a crise cambial que marcou o início daquele ano. Esse regime foi um marco na condução da política monetária, pois estabeleceu um objetivo claro para a inflação anual, permitindo ao Banco Central alinhar suas ações com as expectativas do mercado e da sociedade. Nesse modelo, o COPOM ajusta a taxa SELIC com base na avaliação de fatores como a trajetória da inflação, o desempenho do PIB, e as condições externas. Segundo Carvalho (2021), o regime de metas trouxe maior previsibilidade à política econômica brasileira e reforçou o papel da SELIC como principal instrumento para manter a inflação dentro dos limites estabelecidos.

Além de seu papel no controle da inflação, a taxa SELIC também desempenha uma função crucial no estímulo à atividade econômica, especialmente em períodos de desaceleração. Quando a economia enfrenta um crescimento abaixo do esperado, o COPOM pode reduzir a taxa SELIC para baratear o crédito e incentivar o consumo e o investimento. Por outro lado, em momentos de superaquecimento econômico ou de pressão inflacionária, o aumento da SELIC tem como objetivo restringir a liquidez e conter a demanda agregada. Essa flexibilidade faz da SELIC uma ferramenta poderosa para equilibrar os ciclos econômicos, ainda que seus efeitos não sejam imediatos e dependam de uma série de variáveis externas e internas.

Um exemplo recente da importância da SELIC no estímulo econômico ocorreu durante a pandemia de COVID-19, quando o Banco Central reduziu a taxa para seu menor nível histórico, 2% ao ano, em um esforço para amortecer os impactos econômicos da crise sanitária. Segundo dados do Banco Central (2020), essa medida foi crucial para evitar uma recessão ainda mais profunda, embora tenha gerado desafios adicionais, como a alta do dólar e pressões inflacionárias nos meses seguintes. Esse episódio evidencia a complexidade da gestão da política monetária no Brasil, onde a SELIC precisa equilibrar múltiplos objetivos, incluindo o controle da inflação, a estabilidade do câmbio e o estímulo ao crescimento econômico.

Além de influenciar diretamente o mercado financeiro e o custo do crédito, a taxa SELIC tem impactos indiretos sobre diversos setores da economia. Por exemplo, taxas de juros mais altas tendem a atrair investidores estrangeiros para o mercado de renda fixa, valorizando a moeda local, enquanto juros mais baixos podem incentivar investimentos produtivos e o consumo interno. Esses efeitos tornam a SELIC um dos pilares da política econômica brasileira, com implicações de longo alcance para o desenvolvimento do país.

Além de influenciar diretamente o mercado financeiro e o custo do crédito, a taxa SELIC desempenha um papel fundamental na formação da curva de juros, servindo como um ponto de ancoragem para os títulos de curto prazo. No entanto, a dinâmica da curva de juros brasileira é marcada por uma alta sensibilidade a fatores como mudanças na política monetária, riscos fiscais e oscilações no mercado internacional. Esse comportamento reflete a volatilidade da economia brasileira, onde variações nas expectativas do mercado podem causar alterações significativas na inclinação e no formato da curva. Essa alta variabilidade diária da curva representa um desafio para investidores e empresas financeiras, que dependem de previsões confiáveis para tomar decisões estratégicas.

1.5 A Relevância da Curva de Juros e Sua Relação com as Expectativas de Mercado

Dado o contexto de constante variabilidade da curva de juros no Brasil, compreender e prever seus movimentos torna-se uma tarefa crítica para o mercado financeiro. Instrumentos que possibilitem a análise preditiva da curva de juros são indispensáveis para melhorar a eficiência das decisões de investimento e gestão de risco. Tais instrumentos são particularmente valiosos em um

mercado onde a volatilidade diária pode afetar diretamente o custo do crédito, a precificação de ativos financeiros e a estratégia de alocação de recursos.

A interpretação da curva de juros é fundamental para entender as dinâmicas da economia. Normalmente, a curva apresenta uma inclinação positiva, indicando que títulos de longo prazo oferecem taxas de juros mais altas do que os de curto prazo. Essa configuração reflete a expectativa de que o risco e a incerteza aumentem com o tempo, levando os investidores a demandar uma remuneração maior. No entanto, em determinados momentos, a curva pode se inverter, com títulos de curto prazo apresentando taxas superiores às de longo prazo. Esse fenômeno, frequentemente associado à antecipação de recessões, ocorre quando os investidores esperam uma queda nas taxas futuras devido a um enfraquecimento da economia (Duffee, 2018).

No Brasil, a curva de juros tem particular importância devido ao impacto direto da taxa SELIC nos títulos públicos. Como a SELIC é a taxa básica de juros da economia, ela serve como âncora para os títulos de curto prazo, influenciando a formação de preços ao longo da curva. Segundo Oreiro (2017), a estrutura da curva de juros brasileira é profundamente influenciada pela política monetária conduzida pelo Banco Central, bem como pelas expectativas do mercado em relação à inflação, à dívida pública e ao desempenho econômico. Em um ambiente de alta volatilidade como o Brasil, a curva de juros se torna um reflexo direto das percepções de risco e das incertezas econômicas.

Além de ser um termômetro econômico, a curva de juros é amplamente utilizada como referência para precificação de ativos financeiros. Fundos de investimento, bancos e empresas utilizam as taxas da curva para calcular o custo de captação, o valor presente de fluxos de caixa futuros e a viabilidade de projetos de longo prazo. Conforme apontado por Campbell (1995), a curva de juros é um indicador-chave para a gestão de risco financeiro, permitindo que investidores antecipem movimentos no mercado e ajustem suas estratégias conforme necessário.

Outro aspecto relevante da curva de juros é sua capacidade de sintetizar as expectativas do mercado em relação às decisões de política monetária e ao comportamento econômico futuro. Quando o Banco Central sinaliza uma postura mais rígida no controle da inflação, por exemplo, as taxas de curto prazo podem subir rapidamente, alterando a inclinação da curva. Por outro lado, políticas de estímulo econômico tendem a reduzir as taxas de curto prazo, incentivando o consumo e os investimentos. Essa sensibilidade da curva às expectativas de mercado faz dela uma ferramenta

poderosa tanto para análise econômica quanto para a formulação de políticas públicas (Mishkin, 2019).

Historicamente, a curva de juros no Brasil tem sido caracterizada por sua alta volatilidade, reflexo de um ambiente econômico marcado por ciclos de inflação elevada, crises cambiais e instabilidade fiscal. Durante a crise econômica de 2015-2016, por exemplo, a inclinação da curva mudou drasticamente à medida que as expectativas de mercado se ajustaram às perspectivas de recessão e ao aumento dos riscos fiscais. Situações como essa demonstram como a curva de juros pode servir como um termômetro da confiança dos investidores na capacidade do governo de honrar suas obrigações e manter a estabilidade macroeconômica.

A compreensão da curva de juros é particularmente relevante em um país como o Brasil, onde o mercado financeiro é altamente integrado às decisões de política monetária e às condições fiscais. Para investidores, a análise da curva permite identificar oportunidades de arbitragem e prever movimentos de mercado, enquanto para os formuladores de políticas públicas, ela oferece uma visão consolidada das expectativas econômicas. Nesse sentido, a curva de juros se consolida como um instrumento indispensável para a gestão da economia e para a mitigação de riscos financeiros em um ambiente de alta complexidade e incerteza.

1.6 Alta Variabilidade da Curva de Juros e a Necessidade de Instrumentos Preditivos

A dinâmica da curva de juros brasileira é amplamente reconhecida por sua elevada variabilidade, reflexo das condições econômicas internas e externas, da política monetária e de fatores macroeconômicos que afetam diretamente a percepção de risco no país. Essa variabilidade se manifesta tanto em movimentos graduais quanto em mudanças abruptas, muitas vezes impulsionadas por decisões do Banco Central, alterações na política fiscal ou choques externos. De acordo com Oreiro (2017), a curva de juros é um indicador sensível, reagindo rapidamente às expectativas dos agentes econômicos quanto à inflação, à política monetária e à solvência fiscal do governo. Isso torna sua análise e projeção um elemento essencial para a gestão de riscos e estratégias de investimento.

No Brasil, as condições econômicas e políticas historicamente instáveis amplificam essa variabilidade. Durante períodos de maior incerteza, como crises fiscais ou econômicas, a curva

pode experimentar oscilações significativas em um curto intervalo de tempo, afetando negativamente os custos de financiamento e a precificação de ativos financeiros. Por exemplo, crises como as de 2015-2016, marcadas pelo aumento da percepção de risco fiscal, resultaram em uma curva de juros com inclinação mais acentuada e altas taxas para prazos curtos e longos. Segundo Duffee (2018), essas mudanças abruptas criam desafios significativos para investidores e empresas que dependem de estabilidade para planejar suas operações financeiras.

Diante desse cenário, a capacidade de prever a trajetória futura da curva de juros adquire uma importância estratégica. Modelos que permitem antecipar seus movimentos, mesmo com um grau razoável de incerteza, fornecem informações valiosas para investidores, gestores de portfólios e formuladores de políticas públicas. Segundo Mishkin (2019), a análise preditiva da curva de juros é uma ferramenta indispensável para entender o comportamento futuro das taxas de juros e suas implicações para o custo do crédito, a alavancagem empresarial e a dinâmica dos mercados financeiros.

Essa necessidade se torna ainda mais evidente para grandes instituições financeiras e empresas que gerenciam volumes substanciais de dívida. Essas entidades precisam de previsões confiáveis para ajustar suas estratégias de captação de recursos, refinanciamento e alocação de investimentos. Por exemplo, em momentos de expectativa de alta das taxas de juros, bancos e empresas podem optar por antecipar emissões de dívida de longo prazo para travar custos menores, enquanto em cenários de queda esperada, podem postergar essas operações para se beneficiar de condições mais favoráveis.

No entanto, as ferramentas tradicionais utilizadas para modelar a curva de juros apresentam limitações em contextos de alta variabilidade, como o mercado brasileiro. Métodos econométricos clássicos, como os modelos de Nelson-Siegel ou Svensson, são amplamente empregados para ajustar curvas em momentos de estabilidade, mas sua precisão tende a se deteriorar em cenários de volatilidade elevada. Nesse sentido, a aplicação de técnicas mais avançadas, como os modelos de machine learning, emerge como uma alternativa promissora. Segundo Gu *et al.* (2020), a incorporação de modelos baseados em aprendizado de máquina permite explorar padrões não lineares e interações complexas nos dados financeiros, resultando em previsões mais robustas e adaptáveis.

A possibilidade de empregar algoritmos de aprendizado de máquina para prever a curva de juros abre novos horizontes para a gestão de risco e tomada de decisão no mercado financeiro

brasileiro. Esses modelos, alimentados por dados históricos da curva de juros, da taxa SELIC e de outras variáveis macroeconômicas, podem identificar padrões que escapam às abordagens tradicionais, oferecendo um diferencial competitivo para investidores e gestores. Além disso, sua flexibilidade para lidar com grandes volumes de dados e integrar múltiplas fontes de informação os torna ferramentas particularmente úteis em um ambiente complexo como o mercado brasileiro.

Essa perspectiva de avanço tecnológico cria uma conexão direta com a próxima etapa deste trabalho, que explora o papel dos modelos de machine learning na sociedade contemporânea e suas capacidades para detectar padrões de comportamento em dados complexos. Ao discutir o impacto dessas tecnologias, será possível entender como elas podem transformar a análise e a previsão de variáveis econômicas, como a curva de juros, contribuindo para maior eficiência e segurança no mercado financeiro.

1.7 Avanços Tecnológicos e o Papel do Aprendizado de Máquina na Análise de Dados Complexos

O avanço tecnológico das últimas décadas transformou profundamente a maneira como analisamos dados e tomamos decisões, especialmente em cenários de alta complexidade e variabilidade. O aprendizado de máquina (*machine learning*), uma subárea da inteligência artificial, emergiu como uma das ferramentas mais poderosas nesse contexto, permitindo que sistemas computacionais processem grandes volumes de dados, identifiquem padrões ocultos e realizem previsões com alto grau de precisão. Segundo Bishop (2006), a principal força do aprendizado de máquina reside em sua capacidade de adaptar-se continuamente aos dados disponíveis, ajustando-se a novas informações para melhorar a precisão de suas análises.

Embora os fundamentos do aprendizado de máquina remontem à década de 1950, foi apenas com o advento de novas tecnologias, como processadores mais rápidos, armazenamento em nuvem e o crescimento exponencial de dados digitais, que ele passou a desempenhar um papel crucial em aplicações práticas. De acordo com Goodfellow *et al.* (2016), técnicas como o aprendizado profundo (*deep learning*), baseadas em redes neurais complexas, impulsionaram o desenvolvimento da área, tornando possível resolver problemas antes considerados intratáveis, como o reconhecimento de padrões em grandes conjuntos de dados não estruturados.

Essas inovações tornaram o aprendizado de máquina uma tecnologia indispensável em setores como saúde, transporte e segurança, mas seu impacto é particularmente evidente no mercado financeiro. Instituições financeiras e investidores utilizam algoritmos de aprendizado de máquina para analisar grandes volumes de dados históricos, prever comportamentos de mercado e otimizar estratégias de investimento. Segundo Gu *et al.* (2020), a flexibilidade desses modelos para lidar com a não linearidade e a alta dimensionalidade dos dados financeiros oferece vantagens significativas em comparação com métodos estatísticos tradicionais. Além disso, sua capacidade de processar dados em tempo real permite uma adaptação rápida às mudanças do mercado, criando oportunidades para decisões mais informadas.

No contexto econômico, o aprendizado de máquina se destaca por sua habilidade de modelar variáveis interdependentes em sistemas complexos. Por exemplo, a previsão da curva de juros — influenciada por múltiplos fatores, como política monetária, inflação e condições globais — requer modelos que sejam capazes de capturar essas interações e identificar padrões que escapam a abordagens mais simples. Segundo Hastie, Tibshirani e Friedman (2009), técnicas de aprendizado supervisionado, como regressões avançadas e árvores de decisão, têm mostrado grande potencial para superar as limitações dos métodos lineares em análises econômicas e financeiras.

Além disso, os avanços em aprendizado profundo trouxeram uma nova dimensão à análise de dados. Redes neurais profundas são especialmente úteis para detectar padrões em séries temporais complexas, como os movimentos da curva de juros, onde os dados possuem características dinâmicas e dependências não triviais. LeCun, Bengio e Hinton (2015) apontam que as redes neurais convolucionais e recorrentes têm se mostrado eficazes em problemas que envolvem análise sequencial, abrindo caminho para aplicações inovadoras no setor financeiro.

No Brasil, onde a variabilidade dos dados econômicos e financeiros é uma constante, o aprendizado de máquina oferece uma abordagem promissora para lidar com a complexidade. Modelos baseados nessa tecnologia podem integrar múltiplas variáveis, como a curva de juros, a taxa SELIC e indicadores macroeconômicos, para fornecer previsões mais robustas e adaptáveis às condições de mercado. Essa capacidade é particularmente relevante em cenários de incerteza, onde as ferramentas tradicionais podem falhar em capturar a dinâmica subjacente dos dados.

Os avanços no aprendizado de máquina não apenas permitem análises mais precisas, mas também ampliam as possibilidades de inovação no setor financeiro. No próximo tópico, será

explorada a capacidade desses modelos de detectar padrões de comportamento em dados, destacando como essa habilidade pode transformar a análise preditiva e oferecer insights valiosos para o mercado financeiro.

1.8 O Papel do Aprendizado de Máquina na Análise Econômica e Financeira

O aprendizado de máquina tem revolucionado a forma como dados econômicos e financeiros são analisados, trazendo novas possibilidades para a identificação de padrões e a realização de previsões em cenários complexos e dinâmicos. Com raízes nos métodos estatísticos tradicionais e avanços constantes em aprendizado profundo, os algoritmos modernos oferecem soluções eficazes para lidar com problemas de alta dimensionalidade, não linearidade e dependências temporais. Esse avanço transformou setores como o financeiro, onde variáveis interdependentes, como a curva de juros, requerem abordagens sofisticadas para antecipar comportamentos futuros e gerenciar riscos.

Os primeiros modelos de aprendizado de máquina aplicados à análise econômica baseavam-se em algoritmos como regressões lineares e máquinas de vetores de suporte, que apresentaram resultados satisfatórios para problemas de baixa complexidade. À medida que as demandas por maior precisão e capacidade preditiva cresceram, surgiram métodos mais avançados, como as florestas aleatórias e as redes neurais artificiais. Esses modelos expandiram as possibilidades de análise ao integrar múltiplos indicadores e capturar padrões não lineares, que são frequentemente observados em dados econômicos (Hastie; Tibshirani; Friedman, 2009).

Nos anos 2010, o aprendizado profundo (*deep learning*) consolidou-se como um marco na evolução do campo, introduzindo redes neurais profundas capazes de processar grandes volumes de dados não estruturados e detectar padrões complexos. Modelos como as redes neurais convolucionais (CNNs) e as redes neurais recorrentes (RNNs) destacaram-se na análise de séries temporais, particularmente em cenários financeiros, onde variáveis dependem de dados históricos e de eventos recentes. De acordo com Goodfellow *et al.* (2016), o aprendizado profundo permitiu avanços significativos na previsão de preços de ativos, análise de risco e identificação de anomalias em mercados.

Entretanto, mesmo as redes neurais recorrentes, como as RNNs e LSTMs, enfrentam desafios ao lidar com séries temporais de longa duração. Esses modelos frequentemente sofrem

com o problema do desvanecimento do gradiente, o que limita sua capacidade de capturar dependências de longo prazo em dados sequenciais. Foi nesse contexto que os transformers emergiram como uma solução revolucionária. Introduzido por Vaswani *et al.* (2017), o modelo transformer baseia-se em mecanismos de atenção, que permitem analisar todos os pontos de uma sequência simultaneamente, superando as limitações das arquiteturas anteriores. Isso torna os transformers especialmente adequados para problemas como a previsão de séries temporais financeiras, onde é crucial capturar interações globais e locais entre variáveis.

A arquitetura transformer foi originalmente desenvolvida para processamento de linguagem natural, em tarefas como tradução automática e geração de texto. Contudo, sua flexibilidade e eficiência logo chamaram a atenção de pesquisadores em áreas como a economia e as finanças. Modelos derivados, como o *Temporal Fusion Transformer* (Lim *et al.*, 2021), foram adaptados para prever séries temporais, integrando múltiplas fontes de dados e capturando dependências complexas. Essa capacidade é particularmente valiosa no mercado financeiro, onde variáveis como a curva de juros respondem a uma combinação de fatores, incluindo política monetária, condições macroeconômicas e eventos globais.

No contexto brasileiro, caracterizado por alta volatilidade e condições econômicas desafiadoras, os transformers oferecem uma abordagem promissora para modelar a curva de juros. Sua capacidade de integrar dados históricos, a taxa SELIC e outros indicadores relevantes, ao mesmo tempo em que captura relações de longo e curto prazo, posiciona essa tecnologia como uma ferramenta avançada para a análise preditiva. Além disso, os transformers permitem lidar com grandes volumes de dados e identificar padrões não triviais, criando um diferencial estratégico para instituições financeiras e investidores que operam em mercados de alta incerteza.

Dessa forma, a evolução dos modelos de aprendizado de máquina demonstra como avanços tecnológicos podem transformar a análise econômica. Este trabalho propõe investigar a aplicação de um modelo transformer para o cálculo da curva de juros brasileira, avaliando sua eficácia em prever movimentos futuros e contribuindo para uma melhor compreensão das dinâmicas financeiras do país. No próximo tópico, essa proposta será detalhada, apresentando as motivações e os objetivos específicos do trabalho.

1.9 Proposta de Utilização de um Modelo Transformer para Prever a Curva de Juros Brasileira

Diante da alta variabilidade da curva de juros brasileira e da complexidade de suas interações com variáveis econômicas, surge a necessidade de adotar abordagens tecnológicas inovadoras para sua análise e previsão. Este trabalho propõe o desenvolvimento e aplicação de um modelo transformer para calcular a curva de juros futura, utilizando como base dados históricos da curva, a taxa SELIC e outros indicadores econômicos relevantes. A proposta visa explorar as capacidades avançadas dessa arquitetura de aprendizado de máquina para lidar com séries temporais financeiras em um ambiente marcado por volatilidade e incerteza, como o mercado brasileiro.

Os transformers destacam-se por sua habilidade de capturar padrões complexos e dependências de longo prazo em dados sequenciais, tornando-os uma escolha promissora para a modelagem da curva de juros. Enquanto abordagens tradicionais, como regressões lineares ou modelos econométricos, enfrentam limitações ao lidar com não linearidades e múltiplas variáveis interdependentes, os transformers podem integrar uma ampla gama de dados e identificar relações que escapam às técnicas convencionais. Segundo Vaswani *et al.* (2017), sua arquitetura baseada em mecanismos de atenção oferece uma eficiência superior na análise de grandes volumes de dados, característica essencial para o contexto econômico brasileiro.

A aplicação de um modelo transformer para a previsão da curva de juros apresenta benefícios tanto para a academia quanto para o mercado financeiro. Em termos acadêmicos, este trabalho busca contribuir para o avanço do conhecimento sobre o uso de aprendizado de máquina em economia, especificamente no Brasil. Já do ponto de vista prático, uma ferramenta preditiva robusta pode auxiliar investidores, empresas e gestores de políticas públicas a tomar decisões mais informadas, reduzindo riscos e aproveitando oportunidades de mercado. Além disso, essa abordagem pode oferecer insights valiosos sobre as dinâmicas subjacentes da curva de juros, como a influência da taxa SELIC e de fatores macroeconômicos globais.

Para atingir esse objetivo, este trabalho estrutura-se em etapas que incluem a fundamentação teórica sobre a curva de juros, a política monetária e os transformers; a coleta e o pré-processamento de dados relevantes; o desenvolvimento do modelo preditivo; e a análise dos resultados obtidos. Essas etapas serão detalhadas ao longo deste estudo, com uma ênfase especial

na fundamentação teórica, que será apresentada no próximo capítulo. Essa seção irá abordar os conceitos essenciais para o entendimento do trabalho, como o funcionamento da curva de juros, a política monetária brasileira e a arquitetura transformer, estabelecendo as bases necessárias para o desenvolvimento do modelo proposto.

Assim, o presente estudo não apenas explora o potencial de um modelo transformer para a previsão da curva de juros, mas também reforça a importância de integrar avanços tecnológicos às práticas financeiras e econômicas. A análise dos resultados permitirá avaliar a eficácia desse modelo e suas contribuições para o entendimento e a gestão de variáveis econômicas críticas, como a curva de juros, em um mercado tão dinâmico quanto o brasileiro.

2. Fundamentação Teórica

Esta seção estabelece os conceitos fundamentais que sustentam a análise e o desenvolvimento deste trabalho. A partir de uma revisão bibliográfica abrangente, discute-se a relevância dos títulos públicos e da curva de juros como instrumentos-chave para o funcionamento do mercado financeiro e para a formulação de políticas econômicas. Além disso, exploramos como métodos quantitativos avançados, especialmente aprendizado de máquina, têm sido aplicados em estudos semelhantes, destacando os avanços e limitações encontrados na literatura. Essa base teórica é crucial para justificar a escolha da abordagem metodológica adotada e para estabelecer um marco de referência para os resultados apresentados posteriormente.

2.1 Revisão Bibliográfica sobre Trabalhos Semelhantes

Para fundamentar a proposta deste trabalho e demonstrar a relevância do uso de aprendizado de máquina na previsão de curvas de juros, foi realizada uma revisão bibliográfica abrangente de estudos acadêmicos relacionados. Essa análise tem como objetivo identificar avanços recentes na aplicação de métodos quantitativos no mercado de renda fixa, com foco no uso de modelos de aprendizado profundo para prever variáveis econômicas e financeiras.

Os trabalhos selecionados apresentam aplicações relevantes, como a previsão de curvas de rendimento, a otimização de portfólios de renda fixa e a utilização de dados alternativos para análises mais robustas. Além disso, destacam desafios e limitações das abordagens existentes, como o custo computacional e a dificuldade de interpretar modelos avançados, que ajudam a justificar a escolha de transformers como método principal deste estudo.

A revisão foi organizada de forma a detalhar cada estudo, discutindo seus objetivos, metodologias, resultados e contribuições. Por fim, são estabelecidas conexões entre os trabalhos analisados e a proposta deste TCC, destacando como este trabalho avança o estado da arte ao propor a aplicação de transformers para a previsão da curva de juros brasileira.

2.1.1 Big Data Analytics and Machine Learning in Fixed Income Investment Strategies

O artigo *Big Data Analytics and Machine Learning in Fixed Income Investment Strategies* (Yadav, 2021) explora a integração de técnicas de aprendizado de máquina e big data em estratégias de investimento em renda fixa, com foco em prever curvas de rendimento, gerenciar liquidez e otimizar decisões no mercado financeiro. O trabalho apresenta um panorama abrangente do impacto dessas tecnologias na análise de ativos de renda fixa, destacando como o uso de dados tradicionais e alternativos pode transformar a gestão de investimentos. A seguir, detalha-se a análise desse estudo, aprofundando sua metodologia, aplicações e relevância no contexto acadêmico e prático.

2.1.1.1 Contextualização e Objetivo

O artigo parte da premissa de que os mercados de renda fixa, incluindo títulos soberanos e corporativos, são caracterizados por complexidade, baixa liquidez e alta sensibilidade a fatores macroeconômicos. Esses desafios tornam difícil a previsão precisa de curvas de rendimento e a otimização de carteiras. Nesse contexto, os autores investigam como big data e aprendizado de máquina podem superar essas limitações. O objetivo principal do trabalho é demonstrar que a combinação de dados heterogêneos com algoritmos avançados melhora a eficiência das decisões de investimento, reduzindo riscos e ampliando retornos em mercados de renda fixa.

2.1.1.2 Fontes de Dados Utilizadas

O estudo destaca a importância de incorporar dados variados para melhorar a precisão das análises e previsões no mercado de renda fixa. As fontes de dados incluem:

a) **Dados tradicionais:**

- I. Curvas de rendimento histórico.
- II. Taxas de juros dos títulos.
- III. Indicadores econômicos, como PIB, inflação e desemprego.

b) Dados alternativos:

- I. **Mídias sociais e notícias:** Processamento de linguagem natural (NLP) para extrair sentimentos de publicações sobre mercados e políticas monetárias.
- II. **Imagens de satélite:** Dados indiretos para capturar mudanças econômicas, como atividade industrial.

A integração de dados alternativos representa um avanço significativo, pois permite capturar fatores externos que podem influenciar os mercados financeiros, mas que tradicionalmente não são incluídos nas análises de renda fixa.

2.1.1.3 Metodologias de Machine Learning

O trabalho apresenta várias aplicações de aprendizado de máquina em estratégias de renda fixa. Os métodos incluem tanto abordagens supervisionadas quanto não supervisionadas. Entre as técnicas exploradas, destacam-se:

a) Modelos para prever curvas de rendimento:

- I. Redes neurais recorrentes (RNNs) e LSTMs para analisar séries temporais financeiras.
- II. Algoritmos de regressão para estimar taxas de títulos de longo prazo com base em dados históricos.

b) Gestão de liquidez:

- I. Modelos não supervisionados, como clustering, para identificar padrões em volumes de negociação e prever períodos de baixa liquidez.

c) Análise de sentimento:

- I. NLP aplicado a discursos de bancos centrais e publicações de mercado, permitindo prever impactos em taxas de juros e decisões de política monetária.

Essas metodologias são integradas em sistemas automatizados de análise e decisão, permitindo otimização contínua em carteiras de renda fixa.

2.1.1.4 Resultados e Contribuições

O artigo apresenta resultados que reforçam a eficácia do uso de aprendizado de máquina no mercado de renda fixa:

- a) **Melhoria na previsão de curvas de rendimento:** RNNs e LSTMs apresentaram desempenho superior aos métodos tradicionais, como ARIMA, ao lidar com a complexidade e variabilidade dos dados de títulos.
- b) **Detecção de anomalias na liquidez:** O uso de clustering não supervisionado permitiu identificar padrões de liquidez em períodos críticos, ajudando a mitigar riscos associados à baixa negociabilidade de títulos.
- c) **Impacto da análise de sentimento:** A extração de insights de discursos de bancos centrais demonstrou prever movimentos de mercado com maior precisão, destacando a importância de fatores qualitativos.

Essas contribuições mostram como a combinação de big data e aprendizado de máquina pode transformar o mercado de renda fixa, oferecendo novas perspectivas tanto para gestores de portfólio quanto para reguladores.

2.1.1.5 Limitações e Desafios

Embora os resultados sejam promissores, o estudo reconhece limitações que precisam ser abordadas:

- a) **Dependência de dados alternativos:** O acesso a fontes não convencionais, como imagens de satélite e mídias sociais, pode ser limitado e caro, restringindo sua aplicabilidade em larga escala.
- b) **Complexidade computacional:** Métodos avançados, como RNNs e LSTMs, requerem infraestrutura robusta para treinamento e validação, aumentando os custos de implementação.
- c) **Interpretação de modelos:** Algoritmos como LSTMs são frequentemente criticados por sua natureza de “caixa preta”, dificultando a explicação dos resultados para investidores e reguladores.

Esses desafios indicam que o uso de aprendizado de máquina no mercado de renda fixa ainda está em evolução, exigindo mais estudos para superar essas barreiras.

2.1.2 Machine Learning Applied to Active Fixed-Income Portfolio Management: A Lasso Logit Approach

O artigo *Machine Learning Applied to Active Fixed-Income Portfolio Management: A Lasso Logit Approach* (Hallac; Pareek; Boyd, 2019) explora a aplicação de aprendizado de máquina no gerenciamento ativo de portfólios de renda fixa. O objetivo principal é demonstrar como modelos baseados em regressão Lasso Logit podem otimizar decisões de alocação e duração de títulos, em comparação com abordagens passivas tradicionais. Este estudo combina uma abordagem quantitativa rigorosa com insights práticos, apresentando contribuições significativas para a gestão de ativos financeiros.

2.1.2.1 Contextualização e Objetivo

O artigo parte do princípio de que estratégias ativas de gerenciamento de portfólios podem superar índices de referência (*benchmarks*) em mercados de renda fixa, desde que se utilizem ferramentas avançadas para analisar e prever movimentos de mercado. A regressão Lasso Logit, uma técnica de regularização que elimina variáveis irrelevantes e melhora a interpretabilidade dos modelos, é proposta como uma alternativa eficiente para decisões dinâmicas de alocação de ativos. O estudo concentra-se em títulos do Tesouro dos Estados Unidos, reconhecidos por sua alta liquidez e importância no mercado global.

O objetivo do artigo é avaliar se a abordagem Lasso Logit pode identificar variáveis significativas para prever alterações na curva de juros, permitindo ajustes dinâmicos em portfólios de títulos. Adicionalmente, o estudo compara essa abordagem ativa com estratégias passivas, como compra e retenção, para determinar sua eficácia em diferentes cenários de mercado.

2.1.2.2 Metodologia

A metodologia empregada combina dados econômicos e financeiros com técnicas estatísticas avançadas para desenvolver modelos preditivos. Os principais pontos incluem:

a) Base de Dados:

- I. Conjunto de dados de títulos do Tesouro dos EUA, abrangendo diferentes maturidades.
- II. Variáveis macroeconômicas, como inflação, desemprego e PIB, e indicadores financeiros, como volatilidade de mercado e taxas de juros anteriores.

b) Modelo Lasso Logit:

- I. A regressão Lasso Logit é utilizada para selecionar as variáveis mais relevantes e evitar sobreajustes, comum em modelos com muitos predictores.
- II. O modelo avalia a probabilidade de mudanças em parâmetros-chave da curva de juros, como inclinação e convexidade, para ajustar a duração dos portfólios.

c) Comparação de Estratégias:

- I. A abordagem ativa (com base no Lasso Logit) é comparada com estratégias passivas para avaliar o desempenho em termos de retorno ajustado ao risco.

2.1.2.3 Resultados e Contribuições

Os resultados indicam que a abordagem baseada em aprendizado de máquina supera estratégias passivas em vários cenários:

a) Desempenho Superior em Portfólios Ativos:

- I. O modelo Lasso Logit identificou variáveis-chave que impulsionaram retornos superiores em títulos de médio e longo prazo.
- II. Em períodos de alta volatilidade, a abordagem ativa apresentou maior resiliência ao risco, ajustando-se rapidamente às mudanças na curva de juros.

b) Redução de Risco e Melhor Alocação:

- I. A técnica de regularização minimizou o risco de sobreajuste, resultando em portfólios mais estáveis.

- II. Variáveis macroeconômicas específicas, como expectativas de inflação, foram consistentemente classificadas como preditores significativos.

c) Eficácia na Seleção de Títulos:

- I. O modelo permitiu a identificação de oportunidades em títulos subavaliados, aumentando o retorno total da carteira em relação ao benchmark.

O estudo conclui que a aplicação de aprendizado de máquina, em particular o Lasso Logit, é uma ferramenta valiosa para gestores de portfólios de renda fixa que buscam decisões mais informadas e dinâmicas.

2.1.2.4 Limitações e Desafios

Embora os resultados sejam promissores, o artigo também aponta limitações que afetam a implementação prática do modelo:

a) Dependência de Dados de Alta Qualidade:

- I. A eficácia do modelo depende de dados atualizados e confiáveis, especialmente variáveis macroeconômicas que podem ser reportadas com atrasos.

b) Adaptação a Diferentes Mercados:

- I. O modelo foi testado exclusivamente em títulos do Tesouro dos EUA, e sua aplicabilidade a outros mercados, como o brasileiro, requer ajustes.

c) Foco em Risco Sistemático:

- I. O estudo enfatiza variáveis macroeconômicas, mas fatores idiossincráticos, como mudanças políticas ou riscos específicos de emissores, não foram considerados em profundidade.

2.1.3 Machine Learning in Fixed Income Markets: Forecasting and Portfolio Management

O artigo *Machine Learning in Fixed Income Markets: Forecasting and Portfolio Management* (Tao; Smith Chen, 2020) explora a aplicação de técnicas avançadas de aprendizado de máquina em dois aspectos fundamentais dos mercados de renda fixa: a previsão de curvas de rendimento e o gerenciamento de portfólios. O trabalho destaca como modelos baseados em

aprendizado profundo, como redes neurais multicamadas (MLPs) e redes neurais recorrentes (LSTMs), superam abordagens tradicionais em termos de precisão preditiva e eficiência de alocação de ativos.

2.1.3.1 Contextualização e Objetivo

Os mercados de renda fixa, devido à sua sensibilidade a variáveis macroeconômicas e sua complexidade estrutural, apresentam desafios significativos para previsões e estratégias de portfólio. Nesse contexto, o artigo investiga como o aprendizado de máquina pode melhorar a previsão de curvas de rendimento e otimizar a gestão de ativos. Os autores propõem uma abordagem baseada em aprendizado profundo para lidar com séries temporais financeiras, combinando análise preditiva e algoritmos de aprendizado por reforço para a gestão ativa de portfólios.

2.1.3.2 Metodologia

O artigo adota uma abordagem integrada para prever curvas de rendimento e gerenciar portfólios. A metodologia inclui:

a) Previsão da Curva de Juros:

- I. Dados de rendimentos históricos de títulos do Tesouro dos EUA com diferentes maturidades são usados como insumos.
- II. Comparação entre modelos tradicionais (regressão linear) e avançados (MLPs e LSTMs).
- III. As LSTMs foram treinadas para capturar dependências temporais de longo prazo nos dados de rendimento.

b) Gestão de Portfólios com Aprendizado por Reforço:

- I. Desenvolvimento de um sistema baseado em *Deep Deterministic Policy Gradient* (DDPG), um algoritmo de aprendizado por reforço profundo.
- II. O sistema ajusta dinamicamente a alocação de ativos com base em previsões da curva de rendimento e nas condições do mercado.

c) Métricas de Avaliação:

- I. Para a previsão da curva de juros: erro absoluto médio (MAE) e raiz do erro quadrático médio (RMSE).
- II. Para a gestão de portfólios: retorno total ajustado ao risco, medido pelo índice de Sharpe.

2.1.3.3 Resultados

Os resultados confirmam que as abordagens baseadas em aprendizado profundo superam significativamente os métodos tradicionais em vários aspectos:

a) Previsão da Curva de Juros:

- I. **MLPs:** Melhoraram a precisão em comparação com a regressão linear, capturando padrões não lineares entre variáveis.
- II. **LSTMs:** Apresentaram o melhor desempenho, especialmente para previsões de longo prazo, devido à sua capacidade de capturar dependências temporais.

b) Gestão de Portfólios:

- I. O sistema de aprendizado por reforço superou estratégias passivas de compra e retenção, ajustando dinamicamente a alocação de ativos com base em condições de mercado.
- II. Em cenários de alta volatilidade, o DDPG demonstrou resiliência ao risco e conseguiu capturar oportunidades de arbitragem.

c) Impacto Global:

- I. O uso combinado de previsão e aprendizado por reforço resultou em portfólios mais bem ajustados às condições do mercado, maximizando retornos enquanto minimizava riscos.

2.1.3.4 Limitações e Desafios

Embora o estudo apresente resultados robustos, os autores reconhecem limitações:

a) Desafios Computacionais:

- I. A aplicação de LSTMs e aprendizado por reforço requer capacidade computacional significativa, o que pode limitar sua aplicação em instituições menores.

b) Dificuldade de Interpretação:

- I. Modelos de aprendizado profundo, como LSTMs, são frequentemente considerados “caixas pretas”, dificultando a explicação dos resultados para stakeholders.

c) Generalização Limitada:

- I. O estudo foi baseado exclusivamente em dados do mercado norte-americano, e a aplicabilidade em outros mercados, como o brasileiro, pode exigir ajustes significativos.

2.1.4 Machine Learning-Aided Modeling of Fixed Income Instruments

O artigo *Machine Learning-Aided Modeling of Fixed Income Instruments* (Martin; Póczos; Hollifield, 2018) desenvolvido por pesquisadores da Carnegie Mellon University, aborda o uso de aprendizado de máquina para modelar instrumentos de renda fixa, como títulos soberanos e corporativos. O estudo investiga a aplicação de técnicas clássicas e avançadas de aprendizado supervisionado para prever preços de títulos e taxas de rendimento, destacando as vantagens dessas abordagens em um mercado caracterizado por baixa liquidez e transparência limitada.

2.1.4.1 Contextualização e Objetivo

Os mercados de renda fixa desempenham um papel central na economia global, representando mais de US\$ 100 trilhões em valor total, com títulos soberanos considerados os ativos de menor risco em seus respectivos países. No entanto, esses mercados enfrentam desafios, como a falta de transparência em negociações de títulos corporativos e a baixa frequência de negociações. Este cenário aumenta o risco de preços imprecisos e torna as previsões ainda mais desafiadoras.

O artigo tem como objetivo demonstrar como o aprendizado de máquina pode ser utilizado para prever taxas de rendimento de títulos soberanos e corporativos de diferentes maturidades, além de fornecer estimativas mais precisas para preços em cenários de baixa liquidez. A proposta do estudo é superar as limitações das abordagens tradicionais, como precificação por matriz (*matrix pricing*), com modelos mais robustos e adaptáveis.

2.1.4.2 Metodologia

A metodologia do estudo é detalhada e combina técnicas de aprendizado de máquina supervisionado com análises empíricas rigorosas. Os principais elementos incluem:

a) Base de Dados:

- I. Dados históricos de rendimentos de títulos soberanos dos EUA e preços de títulos corporativos, com atributos como taxas de cupom, maturidade, volume negociado e rendimentos passados.
- II. Inclusão de fatores macroeconômicos, como inflação, PIB e volatilidade de mercado.

b) Modelos Utilizados:

- I. Técnicas clássicas: Regressão linear e ARIMA para análise de séries temporais.
- II. Técnicas avançadas: Florestas aleatórias (Random Forests) e redes neurais artificiais (ANNs).
- III. Modelos híbridos: Integração de aprendizado supervisionado com métodos baseados em regressões curvas para melhorar a precisão.

c) Abordagem Experimental:

- I. O modelo é avaliado em dois cenários: previsões de rendimentos de títulos soberanos de curto e longo prazo, e precificação de títulos corporativos com liquidez limitada.
- II. As previsões foram validadas com métricas como erro médio absoluto (MAE) e raiz do erro quadrático médio (RMSE).

2.1.4.3 Resultados

Os resultados destacam a eficácia dos modelos baseados em aprendizado de máquina em comparação com métodos tradicionais:

a) Previsão de Rendimentos:

- I. As redes neurais artificiais apresentaram o melhor desempenho, superando regressões lineares e florestas aleatórias, especialmente para títulos de longo prazo.
- II. O modelo capturou padrões não lineares nos rendimentos, refletindo melhor a influência de fatores macroeconômicos.

b) Precificação de Títulos Corporativos:

- I. A combinação de modelos supervisionados com regressões curvas reduziu significativamente os erros em cenários de baixa liquidez.
- II. O uso de aprendizado de máquina mitigou os problemas associados à precificação por matriz, que depende de spreads arbitrários para estimar preços.

c) Impacto Global:

- I. A aplicação de aprendizado de máquina melhorou a previsibilidade e a precisão em mercados de renda fixa, permitindo uma análise mais robusta de ativos com características variadas.

2.1.4.4 Limitações e Desafios

Apesar dos avanços demonstrados, o artigo reconhece limitações importantes:

a) Qualidade e Disponibilidade de Dados:

- I. A falta de dados detalhados sobre títulos corporativos foi uma barreira significativa, especialmente para prever preços em mercados com pouca transparência.

b) Complexidade Computacional:

- I. Modelos avançados, como redes neurais, exigem capacidade computacional robusta, limitando sua aplicação para pequenas instituições.

c) Generalização dos Modelos:

- I. Embora eficaz nos mercados norte-americanos, os modelos exigem ajustes substanciais para serem aplicados a outros mercados, como o brasileiro, devido a diferenças estruturais.

Os trabalhos revisados destacam a relevância de métodos avançados de aprendizado de máquina para prever curvas de juros e otimizar decisões em mercados de renda fixa. No entanto, o ponto de partida comum em todas essas abordagens é o entendimento profundo da própria curva de juros, sua construção e os fatores que influenciam sua formação. Como a curva de juros reflete as expectativas do mercado sobre taxas de juros futuras e é um elemento central para a gestão de investimentos e políticas econômicas, torna-se essencial explicar seus fundamentos teóricos e como seu cálculo é realizado. A próxima seção abordará, de forma detalhada, o conceito de curva

de juros e os métodos utilizados para sua construção, fornecendo a base necessária para as análises subsequentes.

2.2 Cálculo da Curva de Juros

A curva de juros é uma das ferramentas mais fundamentais e amplamente utilizadas nos mercados financeiros. Representando graficamente a relação entre as taxas de juros e os prazos de vencimento de títulos de dívida, ela reflete as expectativas do mercado sobre a evolução das taxas de juros futuras, os riscos associados a diferentes maturidades e as condições econômicas gerais. Para investidores, formuladores de políticas econômicas e instituições financeiras, a curva de juros serve como um guia essencial para decisões relacionadas à precificação de ativos, estruturação de dívidas e análise de risco.

Dada sua relevância, entender os fundamentos teóricos e os métodos práticos para o cálculo da curva de juros é indispensável para a proposta deste TCC, que busca prever seus movimentos utilizando aprendizado de máquina. Nesta seção, apresentaremos os conceitos básicos, as metodologias de construção e as aplicações práticas da curva de juros, com base no material de Tuckman (2002), além de discutir sua importância no contexto econômico brasileiro.

2.2.1 Conceito e importância da Curva de Juros

A curva de juros, também conhecida como estrutura a termo das taxas de juros, representa graficamente as taxas de rendimento associadas a diferentes prazos de vencimento para instrumentos de dívida de mesma qualidade de crédito. Essa curva é construída a partir dos rendimentos de títulos negociados no mercado, como títulos públicos ou privados, e reflete a relação entre a maturidade de um título e a taxa de juros correspondente.

2.2.1.1 Definição

Formalmente, a curva de juros é um conjunto de taxas spot para diferentes maturidades, organizadas em ordem crescente de prazo. A taxa spot, por sua vez, é a taxa de retorno anualizada obtida a partir de um investimento feito hoje e mantido até o vencimento. Segundo Tuckman (2002), a curva é construída para representar preços livres de risco em mercados eficientes, permitindo uma visão clara das condições de financiamento e investimento ao longo do tempo.

2.2.1.2 Importância

a) Indicador Econômico:

A forma da curva de juros fornece informações sobre as expectativas do mercado para o crescimento econômico e a inflação. Por exemplo:

- I. Uma curva ascendente (normal) sugere crescimento econômico esperado, com taxas de longo prazo maiores que as de curto prazo.
- II. Uma curva invertida pode sinalizar uma recessão iminente, indicando que o mercado espera quedas nas taxas de juros futuras.

b) Precificação de Ativos:

A curva de juros é usada como referência para especificar uma ampla gama de instrumentos financeiros, como swaps de taxas de juros, derivativos e títulos corporativos. Ela permite calcular fatores de desconto que ajustam fluxos de caixa futuros ao seu valor presente.

c) Política Monetária:

A curva é uma ferramenta-chave para bancos centrais, que a utilizam para avaliar a eficácia de suas políticas e transmitir sinais ao mercado. Movimentos na ponta curta da curva, por exemplo, refletem as decisões de política monetária, como alterações na taxa básica de juros.

d) Gestão de Risco e Investimentos:

Investidores e gestores de portfólios utilizam a curva para alinhar suas estratégias de alocação de ativos. A inclinação da curva, por exemplo, pode influenciar decisões sobre a composição de portfólios de títulos de curto e longo prazo.

2.2.2 Principais Tipos de Curvas de Juros

A curva de juros pode assumir diferentes formatos, dependendo das condições econômicas, das expectativas do mercado e das políticas monetárias adotadas. Cada formato fornece informações importantes sobre o estado da economia, as perspectivas de crescimento e a percepção de risco dos investidores. Nesta seção, apresentamos os principais tipos de curvas de juros, suas características e interpretações econômicas.

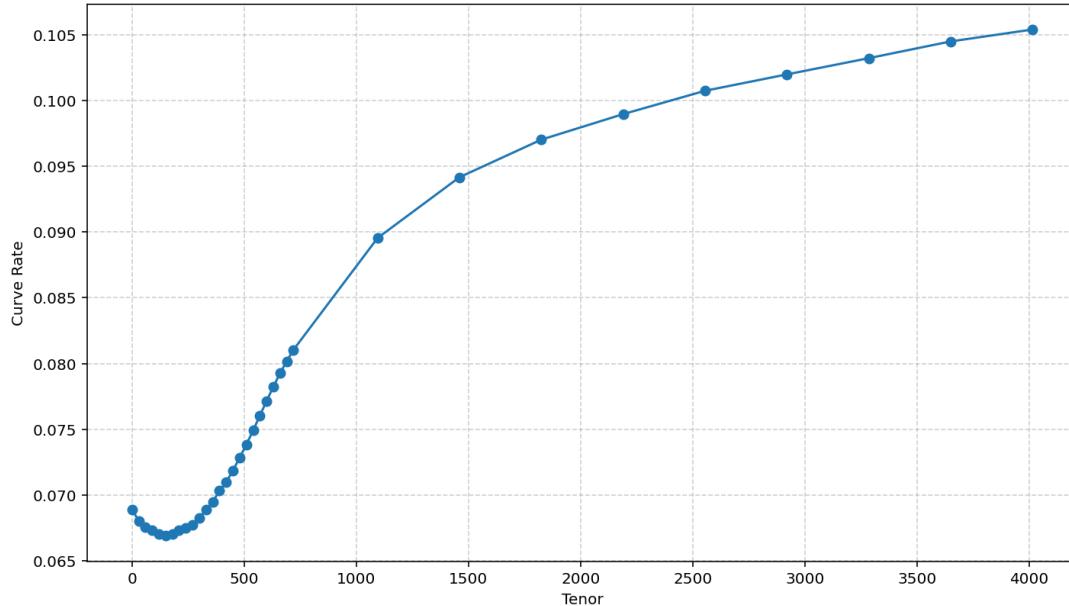
a) Curva Ascendente (Normal)

A curva ascendente, ou normal, é o formato mais comum em economias estáveis, onde os investidores esperam crescimento econômico moderado e taxas de juros futuras mais altas. Nesse caso, títulos de longo prazo oferecem rendimentos maiores do que os de curto prazo, refletindo:

- **Risco de longo prazo:** Investidores exigem um prêmio maior para compensar a incerteza e o risco associados a períodos mais longos.
- **Crescimento econômico saudável:** O mercado espera um aumento na demanda por crédito e na atividade econômica, elevando as taxas futuras, como pode-se ver no Gráfico 1:

Gráfico 1: Curva de juros brasileira ascendente, em 19/01/2018

Curva de Juros Brasileira - 2018-01-19



Fonte: Elaborado pelo autor (2024).

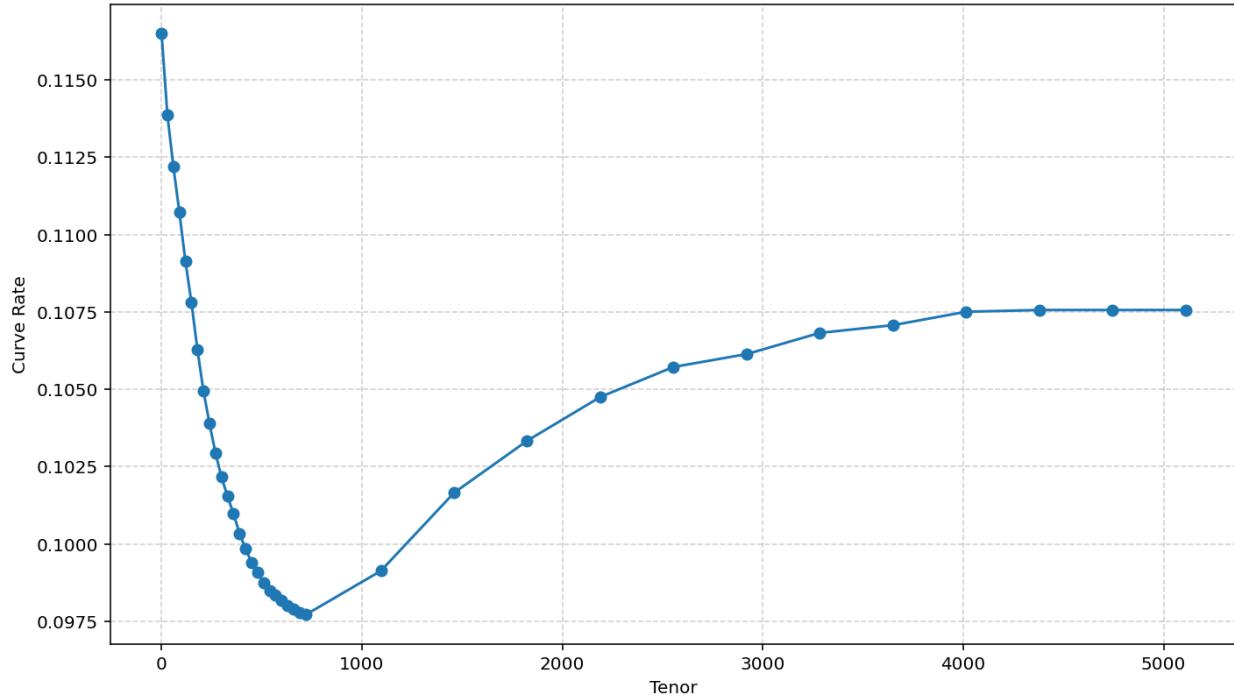
b) Curva Descendente (Invertida)

A curva descendente (ilustrada no Gráfico 2) ocorre quando as taxas de curto prazo são maiores do que as taxas de longo prazo. Esse formato é menos frequente e geralmente sinaliza pessimismo econômico, indicando que o mercado espera uma redução nas taxas futuras.

- **Risco de recessão:** Uma curva invertida frequentemente precede crises econômicas, pois os investidores antecipam cortes nas taxas de juros para estimular a economia.
- **Fuga para a segurança:** Investidores compram títulos de longo prazo como proteção, reduzindo os rendimentos desses ativos.

Gráfico 2: Curva de juros brasileira descendente, em 19/01/2024

Curva de Juros Brasileira - 2024-01-19



Fonte: Elaborado pelo autor (2024).

c) Curva Plana

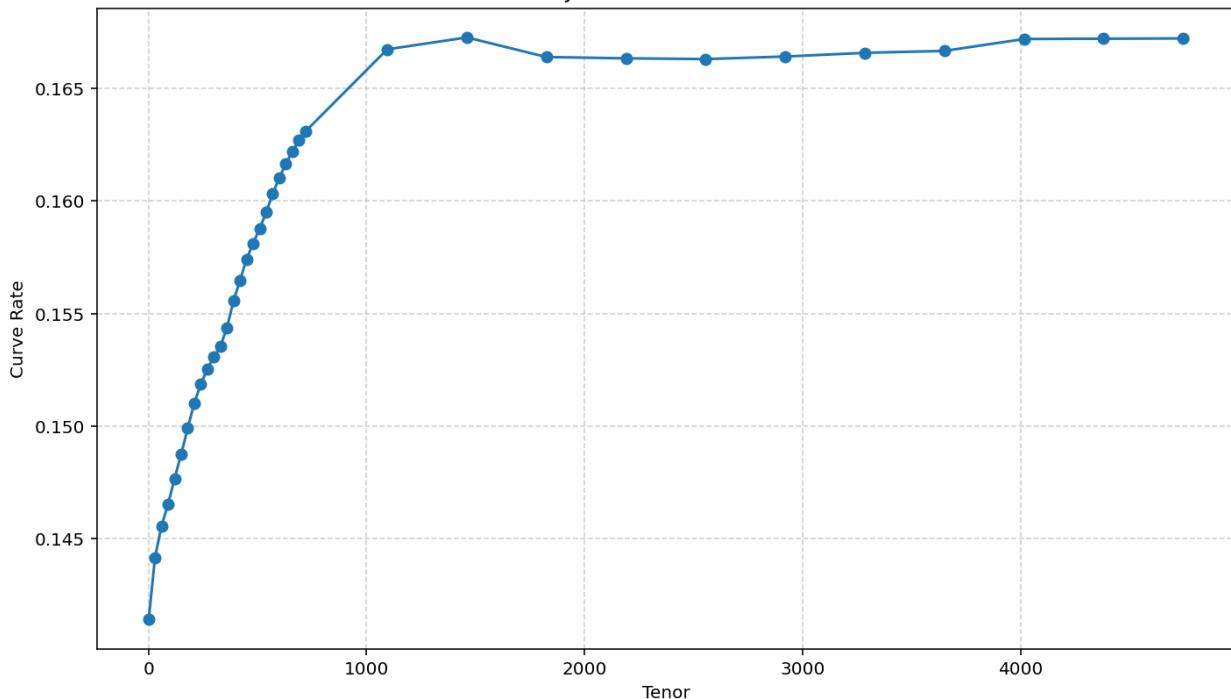
A curva plana (ilustrada no Gráfico 3) ocorre quando as taxas de curto e longo prazo são praticamente iguais. Esse formato reflete uma transição econômica, na qual as expectativas do mercado não indicam direção clara sobre o futuro.

- **Mudança de ciclo:** Pode ocorrer durante a transição entre períodos de crescimento e recessão.

- **Incerteza do mercado:** Reflete equilíbrio temporário entre forças econômicas opostas.

Gráfico 3: Curva de juros brasileira plana, em 19/01/2016

Curva de Juros Brasileira - 2016-01-19



Fonte: Elaborado pelo autor (2024).

2.2.3 Ferramentas Matemáticas e Financeiras para Construção da Curva de Juros

A construção da curva de juros requer uma combinação de ferramentas matemáticas e conceitos financeiros que permitem transformar preços de mercado de instrumentos de dívida em taxas de juros para diferentes prazos de vencimento. Este processo envolve etapas como o cálculo de fatores de desconto, a extração de taxas spot e forward, e o uso de métodos estatísticos para suavização e ajuste da curva. A seguir, exploramos essas ferramentas fundamentais.

2.2.3.1 Fatores de Desconto

Os fatores de desconto são a base para a construção da curva de juros, representando o valor presente de R\$ 1 a ser recebido em uma data futura, com base nas taxas de juros prevalentes no mercado. Formalmente, o fator de desconto para um prazo t é dado pela Fórmula 1:

$$D(t) = \frac{1}{(1+r_t)^t} \quad (1)$$

Onde:

- $D(t)$: Fator de desconto para o prazo t
- r_t : Taxa de juros correspondente ao prazo t

Os fatores de desconto são usados para calcular o preço de títulos a partir de seus fluxos de caixa futuros. A partir desses fatores, derivam-se as taxas spot e forward, que compõem a curva de juros.

2.2.3.2 Taxas Spot

As taxas spot (r_t) representam o rendimento anualizado de um título com vencimento em uma data futura, considerando que o investimento é feito hoje e mantido até o vencimento. São obtidas diretamente a partir dos fatores de desconto, Fórmula 2:

$$(1 + r_t)^t = \frac{1}{D(t)} \leftrightarrow r_t = \left(\frac{1}{D(t)}\right)^{\frac{1}{t}} - 1 \quad (2)$$

As taxas spot são essenciais para a construção da curva de juros, pois fornecem as taxas de retorno livre de risco para diferentes maturidades.

2.2.3.3 Taxas Forward

As taxas forward representam o rendimento esperado para investimentos futuros entre dois períodos específicos. Por exemplo, a taxa forward entre os períodos t_1 e t_2 é calculada pela Fórmula 3:

$$f(t_1, t_2) = \frac{D(t_1)}{D(t_2)} - 1 \quad (3)$$

As taxas forward capturam as expectativas do mercado sobre as taxas de juros futuras, sendo uma ferramenta valiosa para prever mudanças na curva de juros.

2.2.3.4 Métodos de Ajuste e Suavização

A construção da curva de juros frequentemente requer ajustes para garantir que ela seja suave e consistente com os dados de mercado. Os métodos mais comuns incluem:

I. Interpolação Linear:

- a. Utilizado para estimar taxas intermediárias entre pontos conhecidos na curva.
- b. Simples, mas pode gerar descontinuidades em certas situações.

II. Splines Cúbicos:

- a. Modelos mais sofisticados que ajustam a curva de forma suave em toda sua extensão.
- b. Oferecem maior precisão para títulos de maturidades variadas.

III. Modelos Paramétricos (ex.: Svensson e Nelson-Siegel):

- a. Usam equações específicas para descrever a forma da curva, reduzindo a sensibilidade a ruídos nos dados.
- b. Amplamente utilizados por bancos centrais para gerar curvas consistentes.

2.2.3.5 Exemplo Prático

Suponha que o preço de um título zero-cupom com vencimento em 2 anos seja R\$ 940,00. O fator de desconto pode ser calculado preenchendo a Fórmula 1:

$$D(2) = \frac{\text{Preço do Titulo}}{\text{Valor Nominal}} = \frac{940}{1000} = 0,94$$

Com o fator de desconto $D(2) = 0,94$, a taxa spot correspondente será calculada preenchendo a Fórmula 2:

$$r_2 = \left(\frac{1}{D(2)}\right)^{\frac{1}{2}} - 1 = \left(\frac{1}{0,94}\right)^{0,5} - 1 \approx 3,09\% \text{ ao ano}$$

A partir desse cálculo, é possível determinar os demais pontos disponíveis da curva, utilizando os fatores de desconto para diferentes vencimentos. Esses pontos formam uma curva inicial que, posteriormente, será ajustada por meio de métodos de interpolação, como **splines cúbicos** ou modelos paramétricos (por exemplo, Nelson-Siegel). Esse ajuste garante uma curva suave e contínua, fundamental para aplicações práticas como precificação de ativos e análise de risco.

2.3 Modelos Transformer: Arquitetura e Justificativa para o Uso

A introdução dos modelos transformer, apresentada por Vaswani *et al.* (2017) no artigo *Attention is All You Need*, marcou um ponto de virada no aprendizado de máquina. Inicialmente desenvolvidos para processamento de linguagem natural, os transformers demonstraram sua eficácia em tarefas sequenciais complexas, superando limitações de modelos como redes neurais recorrentes (RNNs) e LSTMs. Sua capacidade de capturar padrões complexos e processar grandes volumes de dados em paralelo tornou-os uma escolha natural para a análise de séries temporais, como a curva de juros brasileira.

Nesta seção, exploramos a arquitetura do transformer, seus componentes principais e os motivos que justificam sua escolha para prever a curva de juros.

2.3.1 Estrutura do Modelo Transformer

A arquitetura transformer é composta por módulos que trabalham em conjunto para processar entradas sequenciais, extraíndo padrões relevantes e gerando previsões robustas. O modelo baseia-se no mecanismo de atenção, que permite ao transformer identificar relações importantes entre os elementos da entrada.

I. Interpolação Linear:

Os dados de entrada, como séries históricas de taxas de juros, são convertidos em tensores multidimensionais. Esses tensores representam os dados em formatos que o modelo pode processar eficientemente.

II. Embeddings:

Os *embeddings* são vetores densos que transformam os dados originais em representações matemáticas com dimensões fixas. No contexto deste trabalho, as taxas de juros e outras variáveis econômicas são transformadas em *embeddings* para facilitar os cálculos internos do modelo.

III. Positional Encoding:

Como o transformer não processa os dados em sequência, ele utiliza um mecanismo chamado *positional encoding* para adicionar informações sobre a posição relativa dos elementos na entrada. Isso é especialmente relevante para séries temporais, onde a ordem dos dados influencia diretamente as previsões.

IV. Encorder Layers:

Cada camada do encoder realiza duas operações principais:

- I. **Multi-Head Attention:** O mecanismo de atenção identifica os elementos mais relevantes da entrada para a tarefa preditiva. As várias cabeças de atenção permitem que o modelo analise diferentes padrões simultaneamente.
- II. **Feedforward Layer:** Após a atenção, os dados passam por uma camada totalmente conectada, que refina as representações internas.

2.3.2 Componentes Técnicos do Transformer

Os principais componentes do transformer, que definem sua capacidade de aprendizado e eficiência, incluem:

I. Dimensão dos Embeddings (*embed_dim*):

Controla o tamanho dos vetores que representam os dados de entrada. Uma maior dimensão permite ao modelo capturar mais detalhes, mas aumenta o custo computacional.

II. Número de Camadas (*layers*):

Refere-se à profundidade do modelo, ou seja, quantas vezes os dados passam pelo encoder. Camadas adicionais permitem capturar padrões mais complexos.

III. Número de Cabeças de Atenção (*heads*):

Cada cabeça de atenção analisa relações diferentes nos dados. Mais cabeças aumentam a capacidade de aprendizado, mas exigem maior poder computacional.

IV. Feedforward Dimensionality (*ff_dim*):

Determina o tamanho das camadas densas no encoder, influenciando a capacidade do modelo de abstrair informações dos dados.

V. Dropout:

Uma técnica de regularização que desativa aleatoriamente partes do modelo durante o treinamento, reduzindo o risco de sobreajuste.

VI. Taxa de Aprendizado (*learning rate*):

Influencia a velocidade com que o modelo ajusta seus pesos para minimizar o erro. Transformers frequentemente utilizam estratégias adaptativas, como a *learning rate scheduler*, para ajustar dinamicamente essa taxa.

2.3.3 Vantagens do Transformer

Os transformers oferecem vantagens significativas em relação a modelos tradicionais de séries temporais, como RNNs e LSTMs:

I. Processamento Paralelo:

Enquanto as RNNs processam dados sequencialmente, os transformers analisam todos os elementos de entrada simultaneamente, reduzindo o tempo de treinamento e melhorando a escalabilidade.

II. Captura de Dependências Longas:

O mecanismo de atenção permite que o transformer analise relações de curto e longo prazo com a mesma eficiência, superando o problema de gradientes desvanecentes nas RNNs.

III. Flexibilidade em Dados Multivariados:

Os transformers integram facilmente múltiplas variáveis de entrada, como séries temporais e fatores macroeconômicos, permitindo uma modelagem abrangente.

IV. Eficácia Comprovada:

Modelos derivados do transformer, como o Temporal Fusion Transformer (TFT), já demonstraram alta precisão em previsões financeiras e econômicas.

2.3.4 Justificativa para o Uso no Contexto do TCC

A escolha do transformer para este trabalho é fundamentada em suas capacidades técnicas e na complexidade dos dados financeiros brasileiros:

I. Modelagem da Curva de Juros Brasileira:

A curva de juros é influenciada por variáveis macroeconômicas, choques externos e expectativas do mercado. O transformer é ideal para capturar essas interações complexas.

II. Adaptação a Dados Voláteis:

O mercado brasileiro apresenta alta volatilidade e desafios como incertezas fiscais e mudanças abruptas na política monetária. O transformer lida bem com essas características, garantindo previsões robustas.

III. Superação de Limitações de Modelos Tradicionais:

Métodos econômétricos convencionais e modelos de redes neurais recorrentes enfrentam dificuldades em lidar com séries temporais altamente não lineares e multivariadas. O transformer resolve essas limitações, oferecendo maior precisão e interpretabilidade.

Os conceitos apresentados nesta seção serão aplicados diretamente na construção e treinamento do modelo transformer descrito no próximo capítulo (**3.0 Metodologia**). Nessa etapa, serão detalhados os parâmetros específicos do modelo e as etapas de pré-processamento dos dados, além de como o transformer será ajustado para atender às particularidades do mercado financeiro brasileiro.

3. Metodologia

Nesta seção, detalhamos os passos realizados para o desenvolvimento do modelo transformer aplicado à previsão da curva de juros brasileira. O objetivo principal desta etapa foi estruturar os dados e configurar o modelo de forma a capturar as dinâmicas complexas entre os tenores da curva, a taxa SELIC e variáveis auxiliares, como as reuniões do COPOM.

Inicialmente, descrevemos o processo de coleta e pré-processamento dos dados, seguido da configuração do modelo transformer, abordando os parâmetros utilizados e a lógica subjacente à arquitetura. Por fim, explicamos o treinamento do modelo e os critérios adotados para avaliação de seu desempenho.

3.1 Coleta e Pré-processamento dos Dados

A base de dados utilizada neste trabalho foi composta a partir de informações detalhadas sobre os títulos públicos brasileiros, complementadas por variáveis auxiliares, como a taxa SELIC e os ciclos de reuniões do COPOM. A seleção dos títulos e a estruturação dos dados visaram fornecer ao modelo transformer a granularidade necessária para capturar as dinâmicas complexas da curva de juros brasileira.

3.1.1 Fontes de Dados

3.1.1.1 Títulos Públicos Brasileiros

Para a construção da curva de juros, foram utilizados os seguintes títulos públicos, negociados na bolsa de valores brasileira:

a) **LTN (Letras do Tesouro Nacional):**

Títulos pré-fixados de curto a médio prazo, com vencimento típico entre 6 meses e 5 anos. Esses títulos não possuem pagamento de cupons intermediários, oferecendo apenas o valor nominal na data de vencimento. O rendimento é definido pela diferença entre o preço negociado e o *Face Value*.

Prazos médios: De 180 a 1825 dias (aproximadamente 6 meses a 5 anos).

b) NTN-F (Notas do Tesouro Nacional Série F):

Títulos pré-fixados de longo prazo, com vencimentos que variam entre 5 e 10 anos. Diferentemente das LTNs, as NTN-Fs oferecem pagamentos periódicos de cupons, geralmente semestrais, além do valor nominal na data de vencimento.

Prazos médios: De 1825 a 3650 dias (aproximadamente 5 a 10 anos).

c) LFT (Letras Financeiras do Tesouro):

Títulos pós-fixados atrelados à taxa SELIC diária, acumulada até o vencimento. São considerados instrumentos de menor risco no mercado financeiro brasileiro, devido à sua indexação à taxa básica de juros.

Prazos médios: De 360 a 1825 dias (aproximadamente 1 a 5 anos)

Os dados sobre esses títulos foram obtidos na Bloomberg, incluindo:

- I. **Face Value:** Valor nominal do título, correspondente ao montante pago ao investidor na data de vencimento.
- II. **Valor de Fechamento:** Preço pelo qual o título foi negociado no mercado no final de cada dia útil.

Essas informações permitiram calcular as taxas de juros correspondentes para diferentes vencimentos, conforme descrito na seção 2.2.

3.1.1.2 Conversão de Taxas de Juros

As taxas de juros, geralmente expressas em termos anuais, foram convertidas para valores diáridos para proporcionar maior detalhamento e precisão ao modelo. Essa abordagem permitiu capturar as flutuações diárias nos movimentos da curva.

3.1.1.3 Tenores utilizados

A curva de juros foi composta por 36 tenores, cobrindo prazos de 30 a 4745 dias. Para capturar a dinâmica de curto prazo com maior precisão, foi adotada uma densidade maior de pontos nos primeiros 720 dias, com intervalos de 30 dias. Os tenores utilizados foram:

30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 330, 360, 390, 420, 450, 480, 510, 540, 570, 600, 630, 660, 690, 720, 1095, 1460, 1825, 2190, 2555, 2920, 3285, 3650, 4015, 4380 e 4745 dias

Essa estrutura garante que tanto os movimentos de curto prazo quanto as tendências de longo prazo da curva de juros sejam representados de forma adequada.

3.1.1.4 Taxa SELIC e Reuniões do COPOM

- I. **Taxa SELIC:** Extraída da base de dados do Banco Central, representando a taxa básica de juros vigente.
- II. **Reuniões do COPOM:** As datas das reuniões foram utilizadas para criar duas variáveis auxiliares:
 - I. **Coluna Binária:** Indicando se um dia específico era ou não dia de reunião.
 - II. **Coluna Auxiliar:** Representando o número de dias até a próxima reunião, modelando a influência crescente das expectativas de mercado.

3.1.2 Limpeza e Transformação dos Dados

- I. **Tratamento de Valores Nulos:**
 - a. Valores ausentes, resultantes de finais de semana ou feriados, foram descartados para garantir a consistência dos dados.
- II. **Cálculo da Curva de Juros:**
 - a. As taxas de juros diárias foram calculadas com base nos preços dos títulos públicos e organizadas nos 36 tenores definidos.
 - b. A interpolação foi realizada por desconto (*Discount*), utilizando os parâmetros **Day Counter Business 252**.
- III. **Incorporação de Variáveis Auxiliares:**
 - a. A taxa SELIC e as variáveis relacionadas ao COPOM foram integradas à base como insumos adicionais ao modelo, permitindo capturar sua influência na curva de juros.

3.1.3 Estruturação dos Dados para o Modelo

Os dados foram estruturados em dois componentes principais para alimentar o modelo transformer:

I. Curva de Juros:

Representando os 36 tenores calculados para cada dia útil, com valores diários de taxas de juros.

II. Variáveis Auxiliares:

Incluindo a taxa SELIC, a coluna binária de reuniões do COPOM e a contagem regressiva para a próxima reunião, complementando as informações fornecidas ao modelo.

3.2 Configuração do Modelo Transformer

A configuração do modelo transformer foi projetada para prever os valores futuros da curva de juros brasileira com base em seus valores históricos e variáveis auxiliares, como a taxa SELIC e as datas de reuniões do COPOM. O modelo combina diversos componentes, como camadas de atenção, embeddings posicionais e feedforward, que são descritos a seguir.

3.2.1 Estrutura do Modelo

O modelo implementado possui a seguinte arquitetura modular:

I. Camada de Embeddings

- a. Os valores dos tenores da curva de juros e as variáveis auxiliares (taxa SELIC, dias para a próxima reunião do COPOM, etc.) são convertidos em vetores densos por meio de camadas de embeddings.
- b. Essa transformação inicial facilita o aprendizado do modelo, projetando as entradas em um espaço de alta dimensionalidade.

c. Parâmetros:

- i. Dimensão dos embeddings: *128**.
- ii. Separação de embeddings para tenores e variáveis auxiliares.

II. Codificação Posicional

- a. Para que o modelo reconheça a ordem temporal dos dados, foi incorporada uma codificação posicional que adiciona informações sobre a posição de cada elemento na sequência de entrada.
- b. A codificação combina funções seno e cosseno com diferentes frequências para representar a posição de forma única.

III. Camadas de Encoder

- a. O núcleo do modelo consiste em camadas de encoder, compostas por:
 - i. **Atenção Multi-Head:** Identifica relações temporais e interdependências entre os tenores e as variáveis auxiliares em cada sequência.
 - ii. **Rede Feedforward:** Refina as representações internas, permitindo que o modelo aprenda padrões mais complexos.
 - iii. **Normalização e Dropout:** Garantem estabilidade no aprendizado e evitam sobreajuste.
- b. **Parâmetros:**
 - i. Número de camadas: 8*.
 - ii. Número de cabeças de atenção: 8*.
 - iii. Dimensão feedforward: 128*.
 - iv. Taxa de dropout: 0.1.

IV. Camada de Saída

- a. A saída final do modelo é processada por uma camada totalmente conectada, que transforma os vetores processados nas previsões para os 36 tenores da curva de juros.

3.2.2 Penalização por Volatilidade

Uma característica importante do modelo foi a inclusão de uma função de perda personalizada para suavizar a curva de juros gerada. Dados históricos mostraram alta volatilidade entre tenores consecutivos, o que não seria realista em um mercado financeiro. Para corrigir isso:

I. Penalização de Volatilidade:

- a. Diferenças excessivas entre tenores consecutivos foram penalizadas.
- b. Essa suavização foi implementada como um termo adicional na função de perda base (MSE ou Huber Loss).

c. Parâmetros Configuráveis:

- i. Peso da penalização: 0,1.
- ii. Limite de suavidade: 0,02.

3.2.3 Configurações Globais do Modelo

I. Dimensão dos Tenores (Input e Output): 36.

II. Variáveis Auxiliares: 4 (data, SELIC, binário do COPOM, contagem regressiva).

III. Arquitetura:

- o Embeddings para tenores e variáveis auxiliares processados separadamente
- o Codificação posicional adicionada às entradas
- o 8 camadas de encoder com atenção multi-head e feedforward

IV. Parâmetros Fixos:

- o Função de perda inicial: MSE, com posterior substituição por Huber Loss com penalização de suavidade
- o Método de otimização: Adam

3.2.4 Configuração Dinâmica

Para refinar o desempenho do modelo, foram realizados ajustes iterativos em alguns parâmetros durante os testes. Esses ajustes incluíram:

- Dimensão dos embeddings: 128*
- Número de camadas (layers): 8*
- Número de cabeças de atenção: 8*
- Taxa de aprendizado: 1^{-4} *
- Funções de perda: Inicialmente MSE, posteriormente substituída por Huber Loss com suavização*

*Parâmetros configuráveis ajustados ao longo dos experimentos para otimizar a performance do modelo.

3.3 Treinamento do Modelo

O treinamento do modelo transformer foi estruturado em etapas bem definidas, abrangendo a divisão dos dados, o processo de treinamento em si e a validação contínua para otimizar o desempenho. O objetivo foi ajustar os pesos do modelo de forma a capturar os padrões dinâmicos da curva de juros brasileira, garantindo previsões robustas e consistentes.

3.3.1 Divisão dos Dados

Os dados foram divididos em três conjuntos principais para treinamento, validação e teste, seguindo as melhores práticas para séries temporais:

I. Treinamento (70%):

- a. Conjunto utilizado para ajustar os pesos do modelo com base nas entradas e saídas reais.
- b. Inclui a maior parte dos dados para maximizar o aprendizado do modelo.

II. Validação (20%):

- a. Conjunto usado para monitorar o desempenho do modelo durante o treinamento, permitindo ajustes nos hiperparâmetros e na arquitetura.
- b. A validação foi feita utilizando séries temporais subsequentes, sem misturar dados com o conjunto de treinamento para evitar *data leakage*.

III. Teste (10%):

- a. Reservado para avaliar o desempenho final do modelo em dados nunca vistos durante o treinamento ou validação.

3.3.2 Configurações do Treinamento

I. Parâmetros de Treinamento

- a. **Janela Temporal (sequence lenght):** 5 dias
- b. **Lote de Treinamento (batch size):** 4.
- c. **Épocas (epochs):** 100
- d. **Função de perda:**
 - i. Inicialmente, o erro médio quadrático (MSE) foi utilizado para calcular perda
 - ii. Posteriormente, uma penalidade por volatilidade foi adicionada à função de perda, considerando que as diferenças excessivas entre tenores consecutivos não são realistas. A nova função combinou o Huber Loss com uma penalização baseada no limite de suavidade, ajustado para 0,02.
- e. **Otimizador:** Adam, configurado com uma taxa de aprendizado inicial 1^{-4} .

II. Estratégia de Validação:

- a. Após cada época de treinamento, o modelo foi avaliado no conjunto de validação
- b. A perda de validação foi monitorada para identificar sinais de *overfitting* ou ajustes necessários nos hiperparâmetros

3.3.3 Procedimentos de Treinamento

I. Treinamento por Época:

O modelo foi treinado iterativamente ao longo de 100 épocas, com o seguinte fluxo:

- a. **Etapa 1:** O modelo foi colocado em modo de treinamento (`model.train()`), e cada lote de dados foi processado.
- b. **Etapa 2:** As previsões para os 36 tenores foram geradas com base na entrada dos 5 dias anteriores.
- c. **Etapa 3:** A perda foi calculada usando a função de perda configurada (com suavização).
- d. **Etapa 4:** O otimizador ajustou os pesos do modelo para minimizar a perda calculada.

II. Validação ao Final de Cada Época:

- a. O modelo foi colocado em modo de avaliação (`model.eval()`), e a perda no conjunto de validação foi calculada.
- b. A perda de validação foi comparada com a de treino para verificar o ajuste do modelo.

III. Monitoramento e Registro de Perdas:

- a. Após cada época, as perdas médias de treino e validação foram registradas para análise.
- b. As perdas foram utilizadas para identificar a época ideal de treinamento, evitando *overfitting*.

3.3.4 Avaliação no Conjunto de Teste

Após o treinamento, o modelo foi avaliado no conjunto de teste para medir sua capacidade de generalização. O procedimento incluiu:

I. Inferência no Último Dia do Conjunto de Teste:

- a. O modelo previu os 36 tenores para o último dia do conjunto de teste.
- b. As previsões foram comparadas aos valores reais por meio de gráficos de análise visual, destacando as diferenças entre as taxas previstas e as observadas.

II. Inferência no Primeiro Dia do Conjunto de Teste:

- a. O modelo também foi avaliado no primeiro dia possível do conjunto de teste para verificar sua precisão em uma posição inicial na sequência temporal.

3.3.5 Visualização dos Resultados

I. Curvas Previsões vs. Reais:

- a. Foram gerados gráficos comparando as previsões do modelo com os valores reais da curva de juros para dias específicos no conjunto de teste.
- b. A análise gráfica destacou a capacidade do modelo de capturar tendências gerais, bem como as limitações em casos de maior volatilidade.

II. Janela de Entrada:

- a. Para facilitar a análise, as curvas dos 5 dias anteriores (janela de entrada) foram visualizadas, mostrando como o modelo usou os dados históricos para gerar previsões.

3.4 Ferramentas e Ambiente de Desenvolvimento

O desenvolvimento do modelo transformer e o processamento dos dados foram realizados utilizando ferramentas amplamente reconhecidas na área de ciência de dados e aprendizado de máquina. A escolha do ambiente e das bibliotecas buscou otimizar o desempenho e facilitar a manipulação do grande volume de dados envolvidos neste trabalho.

3.4.1 Linguagens e Bibliotecas Utilizadas

I. Linguagem de Programação: Python

Python foi a linguagem escolhida para este trabalho devido à sua popularidade na comunidade acadêmica e à ampla disponibilidade de bibliotecas específicas para aprendizado de máquina e manipulação de dados.

II. Bibliotecas de Manipulação e Visualização de Dados:

- a. **Pandas e NumPy:** Utilizadas para a manipulação eficiente dos dados, como o tratamento de valores nulos, pivotagem de tabelas e cálculos matriciais.
- b. **Matplotlib:** Empregada para a visualização de curvas de juros reais e previstas, facilitando a análise gráfica dos resultados.

III. Biblioteca de Aprendizado de Máquina:

- a. **PyTorch:** Escolhida para implementar o modelo transformer devido à sua flexibilidade na construção de arquiteturas personalizadas e na integração com funções de perda customizadas.
- b. A biblioteca foi fundamental para a construção de componentes como camadas de embeddings, atenção multi-head, feedforward e codificação posicional, além do treinamento e inferência do modelo.

IV. Ambiente de Programação:

a. Spyder:

O ambiente **Spyder**, parte do pacote Anaconda, foi utilizado como interface principal para o desenvolvimento do código.

i. Vantagens:

1. **Gerenciamento de Variáveis:** A interface do Spyder permite visualizar as variáveis carregadas em memória, acelerando a programação e o ajuste de scripts em tempo real, o que foi crucial devido ao grande volume de dados manipulados.
2. **Integração:** Compatibilidade com bibliotecas essenciais como Pandas, NumPy e PyTorch.

3.4.2 Ambiente de Execução

O código foi executado em ambiente local, utilizando uma configuração que atendeu às demandas de memória e processamento para o volume de dados e complexidade do modelo:

I. Hardware Utilizado:

- a. Processador: Intel Core i7.
- b. Memória RAM: 32 GB.
- c. Armazenamento: SSD de 512 GB, para leitura e escrita eficiente dos arquivos de dados.
- d. Placa Gráfica (GPU): Não utilizada. Todo o processamento foi realizado em CPU.

II. Sistema Operacional:

- a. Windows 10.

4. Resultados e Discussões

Nesta seção, apresentamos os resultados obtidos ao longo do desenvolvimento do modelo transformer e discutimos o impacto das decisões tomadas nas diferentes etapas do trabalho. Como este projeto foi conduzido de maneira iterativa, a evolução dos outputs do modelo reflete os ajustes progressivos na arquitetura, na escolha dos parâmetros e na preparação dos dados.

Inicialmente, são apresentados os resultados do modelo em sua configuração básica, que utilizava apenas os tenores da curva de juros como entrada e não incluía suavização nas previsões. A análise gráfica evidencia limitações, como a alta volatilidade entre tenores consecutivos, que motivaram ajustes na função de perda e na estrutura do modelo.

Em seguida, detalhamos como as alterações progressivas, incluindo a introdução de uma penalização por volatilidade e a adição de variáveis auxiliares, contribuíram para melhorar os outputs gerados pelo modelo. As previsões passaram a refletir com maior precisão a suavidade e a dinâmica esperada da curva de juros.

Além disso, discutimos o impacto da ampliação das janelas de dados, que permitiu ao modelo capturar tendências de longo prazo, mas também trouxe desafios relacionados ao aumento do tempo de processamento. Por fim, avaliamos os resultados obtidos no conjunto de teste e identificamos as principais limitações enfrentadas, sugerindo caminhos para trabalhos futuros.

A estrutura desta seção é organizada da seguinte forma:

4.1 Progresso Inicial: Configuração Básica do Modelo

No estágio inicial deste trabalho, o modelo transformer foi configurado para processar apenas os tenores da curva de juros como entrada, com uma janela temporal de 30 dias e previsão de 1 dia à frente. Os parâmetros iniciais utilizados foram moderadamente simples, visando validar a capacidade do modelo de capturar padrões gerais, antes de introduzir ajustes mais complexos. As configurações foram as seguintes:

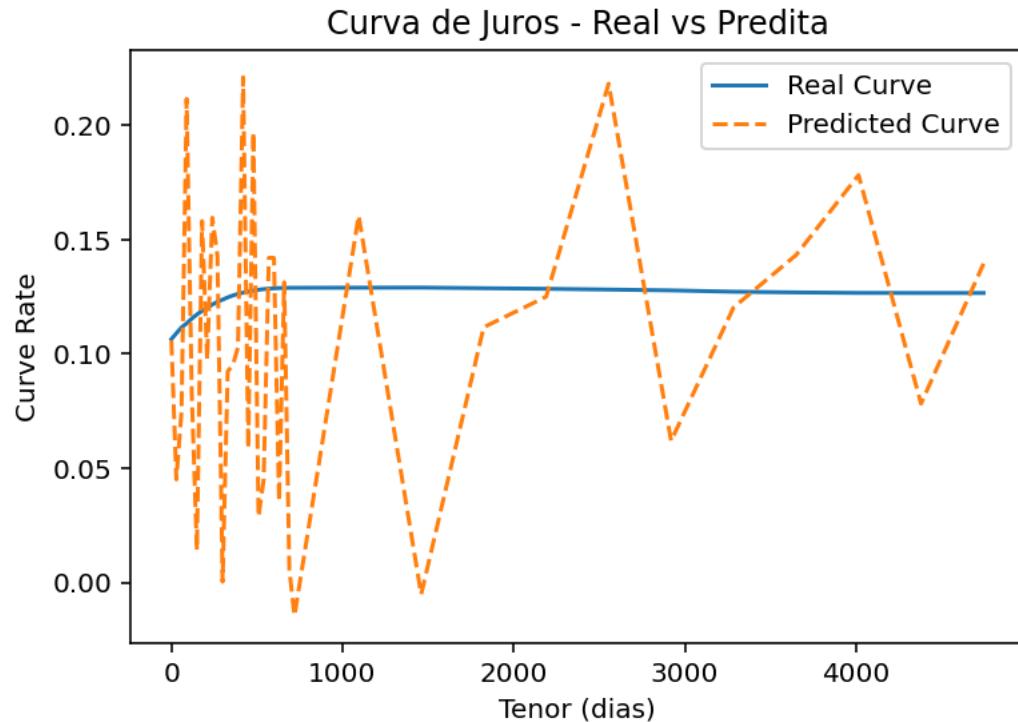
- **Input:** Apenas os tenores das curvas de juros de 02/01/2023 a 01/10/2024.
- **Janela Temporal:** 30 dias.
- **Épocas:** 20.
- **Dimensão dos Embeddings:** 64.

- **Dimensão Feedforward:** 64.
- **Número de Camadas (Layers):** 4.
- **Número de Cabeças de Atenção (Heads):** 4.
- **Output:** Previsão de 1 dia à frente.

Apesar de gerar previsões que refletem parcialmente os padrões dos dados históricos, os resultados iniciais apresentaram limitações significativas. O Gráfico 1 ilustra essas limitações, com destaque para a alta volatilidade nos tenores iniciais.

4.1.1 Gráfico 4: Previsão da Curva de Juros Sem Suavização

Gráfico 4: Previsão inicial da curva de juros sem suavização



Fonte: Elaborado pelo autor (2024).

4.1.2 Análise do Gráfico 4

I. Volatilidade Excessiva nos Tenores Iniciais:

- a. Como mostrado no gráfico, a previsão inicial da curva apresenta flutuações abruptas nos tenores mais curtos (até cerca de 1.000 dias).
- b. Esse comportamento é inconsistente com a realidade, já que a curva de juros no mercado financeiro tende a ser suave e gradual, refletindo as expectativas do mercado sobre as taxas futuras.

II. Tendências de Longo Prazo:

- a. Apesar das instabilidades nos tenores iniciais, o modelo conseguiu capturar parcialmente uma tendência geral nos tenores mais longos. No entanto, a precisão ainda estava aquém do esperado.

III. Limitações Identificadas:

- a. A ausência de suavização nas previsões resultou em inconsistências entre tenores consecutivos.
- b. O uso exclusivo dos tenores como input provavelmente limitou a capacidade do modelo de capturar influências macroeconômicas relevantes, como mudanças na política monetária.

4.1.3 Conclusões e Ajustes Necessários

Com base nos resultados iniciais, foi decidido introduzir ajustes progressivos para melhorar a qualidade das previsões. Os próximos passos incluem:

I. Incorporação de Dados Auxiliares:

- a. Adicionar informações sobre reuniões do COPOM, incluindo:
 - i. Binário indicando dias com reuniões do COPOM.
 - ii. Dias faltantes para a próxima reunião, fornecendo contexto sobre possíveis alterações na taxa SELIC.

II. Mudança na Função de Perda:

- a. Substituir o MSE pela Huber Loss, que é mais robusta a outliers e deve ajudar a suavizar as previsões sem a necessidade imediata de uma penalização explícita por volatilidade.

III. Ajuste do Batch Size:

- a. Aumentar o tamanho do lote de treinamento para 16, permitindo maior estabilidade no ajuste dos pesos do modelo.

IV. Separação de Dados Auxiliares e Tenores:

- a. Configurar entradas separadas no modelo para dados auxiliares e tenores, permitindo que o transformer aprenda as relações entre os diferentes tipos de informação.

Os resultados desses ajustes serão apresentados na subseção **4.2 Ajustes Progressivos na Arquitetura e Função de Perda**, juntamente com gráficos comparativos que destacam as melhorias alcançadas.

4.2 Ajustes Progressivos na Arquitetura e Função de Perda

Após os resultados preliminares apresentados na subseção **4.1**, foram implementados ajustes importantes na arquitetura do modelo e nos dados fornecidos como entrada, com o objetivo de melhorar as previsões e reduzir as flutuações observadas. Entre os ajustes realizados, destacam-se:

I. Inclusão de Variáveis Auxiliares do COPOM:

- a. Foram adicionadas informações relacionadas às reuniões do COPOM, como um binário indicando se o dia era ou não dia de reunião e a contagem regressiva para a próxima reunião. Essas variáveis auxiliares proporcionaram ao modelo um contexto adicional sobre os fatores que influenciam diretamente a curva de juros.

II. Separação de Dados Auxiliares e Tenores:

- a. Os tenores da curva e as variáveis auxiliares foram processados separadamente, permitindo que o modelo aprendesse de forma específica as interações entre os dois tipos de dados.

III. Substituição da Função de Perda:

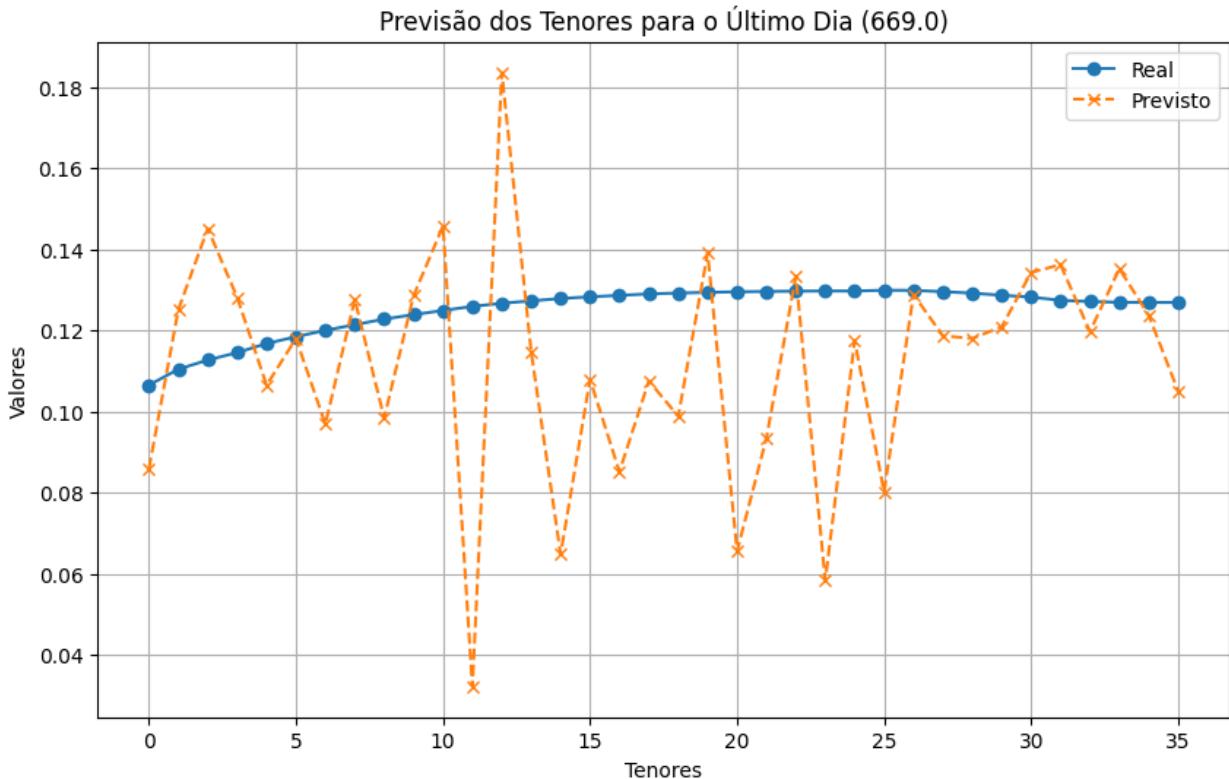
- a. O MSE foi substituído pela Huber Loss, uma função mais robusta a outliers, contribuindo para uma maior estabilidade nas previsões.

IV. Ajuste no Batch Size:

- a. O tamanho do lote de treinamento foi aumentado para 16, o que melhorou a estabilidade no ajuste dos pesos durante o treinamento.

4.2.1 Gráfico 5: Previsão com Dados Auxiliares e Huber Loss

Gráfico 5: Previsão com dados auxiliares e Huber Loss



Fonte: Elaborado pelo autor (2024).

4.2.2 Análise do Gráfico 5:

I. Melhorias Observadas nos Tenores Mais Longos:

- a. Comparado ao gráfico anterior, as previsões para os tenores mais longos (acima de 2.000 dias) apresentam maior estabilidade e seguem mais de perto a tendência dos valores reais.

II. Persistência de Volatilidade nos Tenores Iniciais:

- a. Apesar das melhorias nos tenores mais longos, os tenores iniciais ainda apresentam flutuações significativas, indicando que o modelo tem dificuldade em capturar a suavidade esperada da curva nesses pontos.

III. Impacto da Adição de Variáveis Auxiliares:

- a. A inclusão de informações sobre o COPOM melhorou a capacidade do modelo de capturar mudanças sutis na curva, especialmente em pontos intermediários.

IV. Huber Loss e Estabilidade:

- a. A substituição do MSE pela Huber Loss contribuiu para reduzir o impacto de outliers nas previsões, tornando os outputs mais consistentes nos tenores mais longos.

4.2.3 Conclusão e Ajustes Necessários

Os ajustes realizados nesta etapa trouxeram avanços importantes na qualidade das previsões, especialmente nos tenores de longo prazo. No entanto, algumas limitações ainda persistem, especialmente nos tenores mais curtos. Com base nos resultados, são identificados os seguintes pontos de atenção para os próximos ajustes:

I. Refinamento da Capacidade do Modelo:

- a. A persistência de flutuações nos tenores iniciais sugere a necessidade de aprimorar a arquitetura do modelo, para que ele consiga capturar com maior precisão as dinâmicas de curto prazo da curva.

II. Melhor Exploração de Contexto Econômico:

- a. A introdução de variáveis auxiliares mostrou-se eficaz, mas novas fontes de dados contextuais podem ser exploradas para aumentar ainda mais a robustez das previsões.

III. Ajustes nos Parâmetros de Treinamento:

- a. Parâmetros como a janela temporal, o número de camadas e o número de cabeças de atenção podem ser revisados para equilibrar o aprendizado em diferentes tenores.

Essas limitações serão abordadas nos próximos experimentos, cujos resultados e análises serão apresentados na subseção **4.3 Adição de Variáveis Auxiliares e Dados de Longo Prazo**.

4.3 Adição de Variáveis Auxiliares e Redução da Janela

Após os ajustes realizados na subseção **4.2**, foi testada uma nova configuração do modelo com o objetivo de explorar mais detalhadamente o impacto das variáveis auxiliares e refinar o comportamento da curva nos tenores iniciais. Nesta etapa, os seguintes ajustes foram realizados:

I. Adição da Taxa SELIC como Variável Auxiliar:

- a. A taxa SELIC foi incorporada como um dado auxiliar adicional, fornecendo ao modelo informações relevantes sobre a política monetária que influencia diretamente a curva de juros.

II. Redução da Janela Temporal:

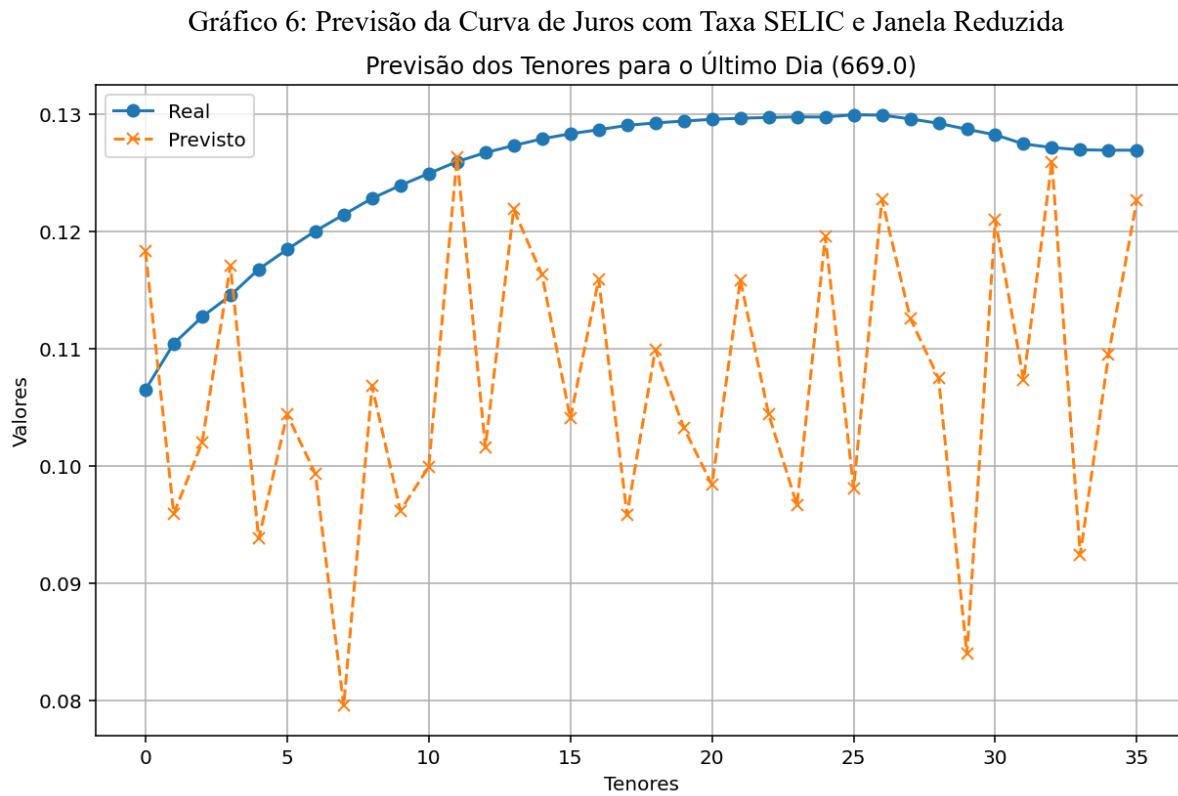
- a. A janela de dados foi reduzida de 30 dias para 7 dias. Essa mudança visou aumentar a sensibilidade do modelo a padrões de curto prazo, que são críticos para capturar as dinâmicas dos tenores iniciais.

III. Configuração de Treinamento:

- a. O batch size foi ajustado para 8, equilibrando estabilidade e eficiência.
- b. A arquitetura foi mantida com 4 layers, 4 heads, e dimensões do embedding e feedforward em 128.

Esses ajustes foram aplicados mantendo as variáveis auxiliares já introduzidas anteriormente (dias com reunião do COPOM e contagem regressiva para a próxima reunião). Os resultados obtidos são apresentados no Gráfico 3.

4.3.1 Gráfico 6: Previsão com Taxa SELIC e Janela Reduzida



Fonte: Elaborado pelo autor (2024).

4.3.2 Análise do Gráfico 6

I. Impacto da Janela Reduzida:

- A redução da janela para 7 dias permitiu ao modelo capturar melhor os padrões de curto prazo, especialmente em relação aos tenores intermediários, como os próximos 90 a 360 dias.
- No entanto, os tenores mais curtos (30 a 90 dias) ainda apresentam flutuações que são inconsistentes com o comportamento real da curva de juros.

II. Contribuição da Taxa SELIC:

- A inclusão da taxa SELIC como variável auxiliar ajudou o modelo a alinhar as previsões de longo prazo com os valores reais.

- b. Isso pode ser observado na redução das discrepâncias para os tenores acima de 2.000 dias, que agora seguem uma tendência mais próxima à realidade.

III. Flutuações nos Tenores Curtos:

- a. Embora as flutuações tenham sido parcialmente reduzidas, a falta de suavidade entre os tenores consecutivos continua a ser um desafio significativo.

4.3.3 Conclusão e Ajustes Necessários

Os ajustes realizados nesta etapa trouxeram avanços importantes, como a melhora na precisão dos tenores intermediários e uma maior estabilidade nas previsões de longo prazo. A inclusão da taxa SELIC também contribuiu para alinhar o modelo às dinâmicas reais da curva de juros. No entanto, as flutuações entre os tenores mais curtos ainda indicam inconsistências, sugerindo que ajustes adicionais são necessários.

Para lidar com essas limitações, propõe-se os seguintes ajustes na próxima etapa:

I. Aumento do Peso Computacional:

- a. Expandir a arquitetura do modelo com mais camadas (*layers*) e cabeças de atenção (*heads*), aumentando sua capacidade de aprendizado.

II. Penalização Explícita por Volatilidade:

- a. Adicionar um termo de penalização por volatilidade na função de perda, incentivando previsões mais suaves entre os tenores consecutivos.

III. Aumento do Número de Épocas:

- a. Estender o treinamento para mais épocas, permitindo ao modelo refinar melhor seus pesos e capturar padrões mais complexos.

Os resultados desses ajustes e a análise comparativa serão apresentados na subseção **4.4 Avaliação com Penalização por Volatilidade**.

4.4 Refinamento da Arquitetura com Penalização por Volatilidade

Nesta etapa, foram implementados novos refinamentos no modelo com o objetivo de reduzir as flutuações nos tenores curtos e alinhar as previsões de forma mais consistente ao comportamento real da curva de juros. Entre os principais ajustes, destaca-se a introdução de uma

penalização explícita por volatilidade, combinada com alterações na arquitetura e no treinamento. As seguintes alterações foram implementadas:

I. Penalização de Volatilidade:

Para reduzir flutuações abruptas entre tenores consecutivos, foi incorporada uma penalização na função de perda. Essa abordagem combina a **Huber Loss** com um termo adicional que incentiva suavidade nas previsões.

a. A lógica por trás dessa penalização é simples:

- i. As diferenças entre os valores previstos para tenores consecutivos são calculadas.
- ii. Caso a diferença exceda um limite aceitável de suavidade (denominado *smoothness limit*), uma penalidade proporcional é aplicada.
- iii. A penalidade é então somada à perda base para formar a função de perda total.

b. A função de perda total pode ser descrita matematicamente pela Fórmula 4:

$$L_{Total} = L_{Base} + \lambda * \frac{1}{n} \sum_{i=1}^{n-1} \max(|\hat{y}_{i+1} - \hat{y}_i| - \varepsilon, 0) \quad (4)$$

i. Onde:

1. L_{Base} é a função de perda base (Huber Loss)
2. λ é o peso da penalização por volatilidade
3. n é o número de tenores
4. \hat{y}_i representa o valor previsto para o i -ésimo tenor
5. ε é o limite de suavidade
6. O termo $\max(|\hat{y}_{i+1} - \hat{y}_i| - \varepsilon, 0)$ aplica penalidade apenas quando a diferença entre valores consecutivos ultrapassa ε

c. Os hiperparâmetros utilizados foram:

- i. Volatility Weight (λ): 0.1.
- ii. Smoothness Limit (ε): 0.02 (2% de variação entre um tenor e o outro)

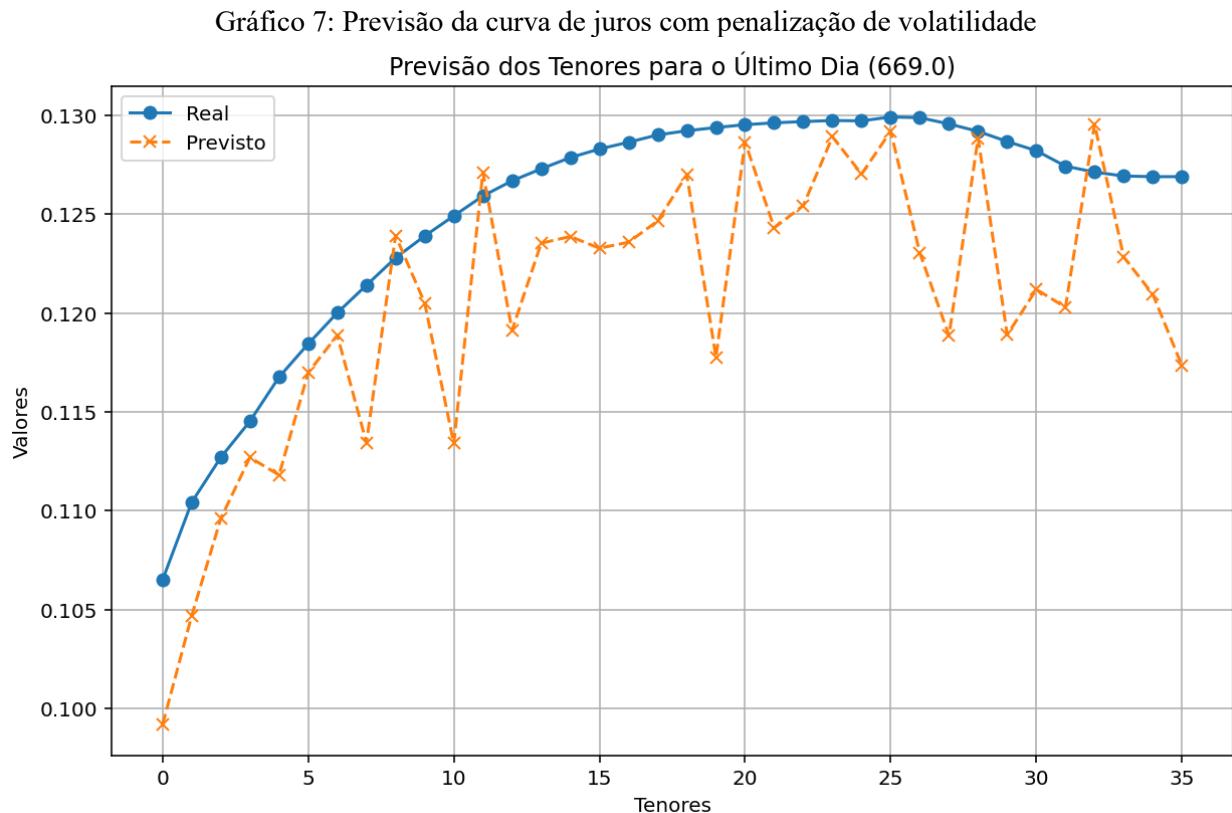
II. Aumento do Peso Computacional:

Para aumentar a capacidade do modelo de capturar padrões complexos, foram utilizadas 8 camadas (*layers*) e 8 cabeças de atenção (*heads*), mantendo as dimensões do embedding e feedforward em 128.

III. Aumento do número de Épocas:

O treinamento foi estendido para 100 épocas, permitindo ao modelo refinar melhor seus pesos e alcançar maior precisão nas previsões.

4.4.1 Gráfico 7: Refinamento com Penalização de Volatilidade



Fonte: Elaborado pelo autor (2024).

4.4.2 Análise do Gráfico 7:

I. Melhorias na Suavidade dos Tenores Curtos:

- a. A penalização por volatilidade foi altamente eficaz na redução de oscilações nos primeiros tenores (30 a 180 dias). As previsões agora refletem o comportamento contínuo esperado da curva de juros.

II. Estabilidade nos Tenores Intermediários e Longos:

- a. As previsões para tenores acima de 360 dias continuam estáveis e alinhadas aos valores reais, indicando que o aumento no número de camadas e cabeças de atenção foi eficaz.

III. Impacto do Aumento do Número de Épocas:

- a. O treinamento prolongado permitiu que o modelo refinasse ainda mais suas previsões, especialmente nos pontos intermediários e de inflexão da curva.

4.4.3 Conclusões e Preparação para a Análise Final

Os resultados desta etapa demonstram que o modelo está cada vez mais próximo de reproduzir a dinâmica real da curva de juros. A introdução da penalização por volatilidade se mostrou fundamental para corrigir o comportamento inconsistente observado nos tenores curtos, enquanto o aumento da capacidade computacional e do número de épocas aprimorou a precisão geral do modelo.

Até o momento, todas as etapas (4.1 a 4.4) foram conduzidas utilizando janelas de dados menores para permitir maior dinamismo nos experimentos, otimizando o tempo de processamento e possibilitando ajustes rápidos. Essa estratégia foi crucial para refinar progressivamente o modelo antes de expandir para análises mais robustas.

Na próxima etapa, apresentada na subseção **4.5 Análise Final com Janela Expandida**, será utilizada uma janela de dados significativamente maior. Esse experimento permitirá ao modelo capturar tendências de longo prazo e validar a robustez dos ajustes realizados, embora com maior tempo de processamento devido à quantidade ampliada de dados.

4.5 Análise Final com Janela Expandida

Nesta etapa final, o modelo foi treinado utilizando uma janela de dados significativamente maior, abrangendo toda a série histórica disponível a partir de 2016. Essa abordagem buscou capturar tendências de longo prazo e validar a robustez dos ajustes realizados nas etapas anteriores. Contudo, os resultados indicaram desafios adicionais e limitações importantes que precisam ser considerados. Impactos do Aumento na Janela de Dados

I. Maior Complexidade dos Dados:

- a. A ampliação da janela trouxe uma maior diversidade de contextos econômicos e eventos históricos, como:
 - i. **Mudanças políticas significativas no Brasil:** A troca de governos e alterações na condução da política monetária impactaram diretamente o mercado financeiro.
 - ii. **Pandemia de COVID-19 (2019-2020):** Este evento global causou flutuações sem precedentes nas taxas de juros, na curva de juros e nos parâmetros macroeconômicos subjacentes.
- b. Essas dinâmicas adicionaram ruído aos dados e dificultaram a identificação de padrões consistentes pelo modelo, especialmente com um conjunto limitado de variáveis auxiliares.

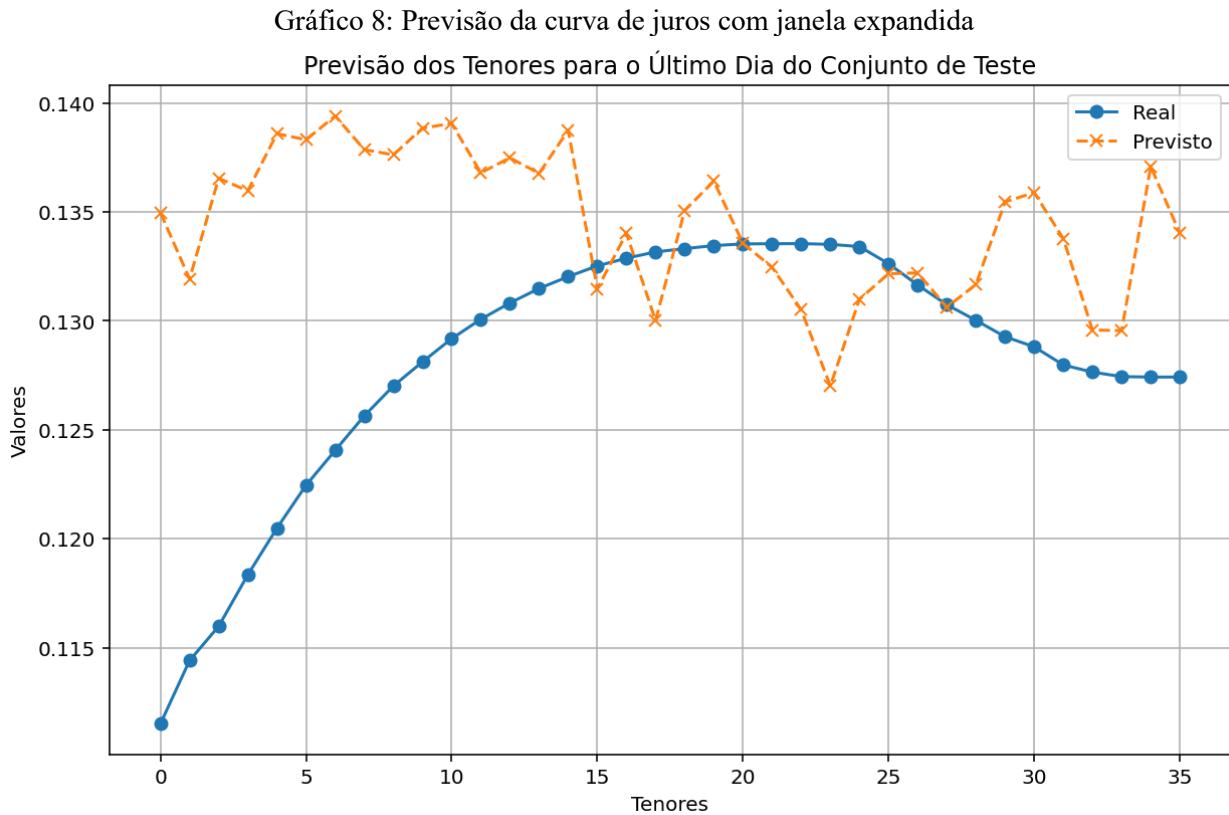
II. Impacto nas Previsões:

- a. O gráfico resultante indica que, embora o modelo mantenha alguma estabilidade nos tenores longos, as flutuações aumentaram significativamente nos tenores curtos e intermediários. Isso sugere que o modelo foi incapaz de generalizar para períodos com alta volatilidade, como aqueles afetados pela pandemia ou por mudanças abruptas na política monetária.

III. Efeitos do Período Expandido:

- a. O aumento no período de teste ampliou as flutuações naturais da curva, mas também destacou que as variáveis utilizadas podem ser insuficientes para explicar a complexidade do mercado financeiro brasileiro nesse intervalo de tempo.

4.5.1 Gráfico 8: Previsão com Janela Expandida



Fonte: Elaborado pelo autor (2024).

4.5.2 Análise do Gráfico 8

I. Instabilidade nos Tenores Curtos e Intermediários:

- a. O modelo apresenta flutuações notáveis nos tenores curtos (30 a 720 dias) e intermediários (até 3.000 dias), indicando que a maior diversidade de eventos e contextos econômicos dificultou a capacidade do modelo de realizar previsões consistentes.

II. Estabilidade Relativa nos Tenores Longos:

- a. Apesar das dificuldades nos períodos mais curtos, o modelo conseguiu manter alguma estabilidade nos tenores mais longos (acima de 3.000 dias), sugerindo que as tendências de longo prazo são menos afetadas por eventos pontuais ou alta volatilidade.

III. Limitações da Abordagem:

- a. A inclusão de poucas variáveis auxiliares limitou a capacidade do modelo de contextualizar e ajustar suas previsões para períodos marcados por choques externos, como a pandemia de COVID-19 ou mudanças bruscas na política monetária brasileira.
- b. A maior janela de dados também ampliou o impacto do ruído nos dados históricos, dificultando a captura de padrões consistentes.

4.5.3 Conclusões Preliminares

Os resultados obtidos nesta etapa final indicam que o modelo transformer, embora tenha demonstrado boa capacidade em cenários com dados mais controlados e períodos menores, enfrenta limitações significativas ao lidar com uma quantidade maior de dados históricos e períodos marcados por eventos extraordinários. Algumas observações importantes incluem:

I. Maior Sensibilidade à Volatilidade:

- a. O aumento da janela de dados e a maior diversidade de eventos econômicos ampliaram a complexidade do problema, expondo a dificuldade do modelo em lidar com alta volatilidade e mudanças bruscas no mercado.

II. Insuficiência das Variáveis Auxiliares:

- a. As variáveis auxiliares utilizadas (taxa SELIC e informações do COPOM) não foram suficientes para contextualizar adequadamente o modelo em um cenário tão complexo. A ausência de dados como indicadores de mercado, inflação ou variáveis globais contribuiu para as previsões inconsistentes.

III. Lições e Próximos Passos:

- a. A análise destaca a importância de incluir mais variáveis macroeconômicas e financeiras no modelo para capturar de forma mais abrangente as dinâmicas do mercado.
- b. Além disso, métodos mais robustos de tratamento de ruído ou segmentação dos dados por períodos econômicos específicos poderiam ser explorados para melhorar os resultados.

Essas reflexões serão discutidas mais detalhadamente no capítulo 5, onde serão apresentadas as principais conclusões deste trabalho e sugestões para futuros estudos.

5. Conclusões e Sugestões para Trabalhos Futuros

O presente capítulo busca sintetizar os principais resultados e reflexões deste trabalho, abordando tanto os avanços alcançados quanto as limitações enfrentadas. Além disso, serão discutidas sugestões para futuros estudos, visando aprimorar a abordagem de modelagem preditiva da curva de juros brasileira por meio de aprendizado de máquina. Por fim, uma reflexão final será apresentada, destacando a relevância do tema no contexto atual e os desafios de aplicar técnicas modernas, como os modelos transformers, a sistemas financeiros complexos.

5.1 Conclusões Gerais

Ao longo deste trabalho, foi possível explorar a aplicação de um modelo transformer para prever a curva de juros brasileira, partindo de variáveis macroeconômicas como a taxa SELIC, reuniões do COPOM e a própria dinâmica da curva ao longo do tempo. Os experimentos realizados e os resultados obtidos permitem destacar as seguintes conclusões:

I. Eficácia do Modelo Transformer em Cenários Controlados:

O modelo demonstrou boa capacidade preditiva em cenários com janelas de dados menores e períodos econômicos menos voláteis. Nessas condições, foi possível observar uma convergência razoável entre os valores reais e as previsões do modelo, especialmente nos tenores mais longos da curva, onde as flutuações tendem a ser menos intensas.

II. Insuficiência de Variáveis Auxiliares em Contextos Complexos:

Ao expandir a janela temporal para incluir períodos marcados por alta volatilidade econômica e política, o modelo enfrentou dificuldades significativas. Entre 2016 e 2024, o Brasil vivenciou uma sequência de eventos que aumentaram substancialmente a incerteza no mercado financeiro, com destaque para a pandemia de COVID-19 e mudanças políticas relevantes.

a. Impactos da Pandemia de COVID-19 (2020-2021):

O COVID-19 gerou um choque global que afetou profundamente as economias de todo o mundo, incluindo o Brasil. A pandemia resultou em uma crise sanitária e econômica sem precedentes, com lockdowns prolongados, colapso de cadeias de suprimentos e uma desaceleração econômica acentuada. No mercado financeiro, isso levou a uma

volatilidade extrema, revisões frequentes na política monetária e fiscal, além de intervenções emergenciais para conter os efeitos adversos. O impacto na curva de juros foi evidente, com oscilações significativas e imprevisíveis nas expectativas de curto e longo prazo, dificultando a modelagem.

b. Mudanças Políticas no Brasil (2016-2024):

Durante o período analisado, o Brasil passou por um cenário de instabilidade política que contribuiu para o aumento da incerteza no mercado. A sequência de eventos — desde o impeachment de Dilma Rousseff em 2016, passando pelo governo Bolsonaro (2019-2022), marcado por crises institucionais e a condução controversa durante a pandemia, até o retorno de Luiz Inácio Lula da Silva à presidência em 2023 — gerou flutuações nos indicadores financeiros e tirou a confiança de muitos investidores. Essa instabilidade política, combinada com dúvidas sobre a condução econômica e fiscal, impactou diretamente o comportamento da curva de juros.

Esses fatores revelam que um conjunto limitado de variáveis auxiliares, como taxa SELIC e reuniões do COPOM, não é suficiente para capturar a complexidade de um mercado tão dinâmico e sujeito a choques externos e internos.

III. O Mercado Financeiro como um Sistema Caótico:

O mercado financeiro brasileiro, como outros ao redor do mundo, pode ser considerado um sistema caótico e dinâmico. Ele é altamente influenciado por uma miríade de fatores interdependentes, como política econômica, eventos globais, mudanças na percepção de risco e crises inesperadas. Essa característica torna o comportamento do mercado difícil de modelar com um número reduzido de variáveis. O período analisado reforça essa conclusão, pois eventos como a pandemia e as crises políticas ampliaram a imprevisibilidade, dificultando a identificação de padrões consistentes e exacerbando as limitações do modelo transformer, que precisa de contexto amplo para fazer previsões mais robustas.

IV. Impactos da Volatilidade em Janelas Maiores:

A ampliação da janela temporal destacou os desafios associados à utilização de períodos mais longos. Embora os tenores mais longos tenham demonstrado relativa estabilidade, os tenores curtos e intermediários apresentaram flutuações significativas, expondo as limitações do modelo em generalizar para contextos mais diversos e voláteis. Além disso, a inclusão de

dados de períodos marcados por choques externos e mudanças estruturais no mercado gerou maior ruído, dificultando a convergência das previsões para os valores reais.

Essas conclusões reforçam a importância de avançar em duas frentes principais: ampliar o conjunto de variáveis auxiliares utilizadas nos modelos para englobar fatores econômicos e financeiros mais amplos e adaptar as técnicas de modelagem para lidar com a natureza não linear e caótica do mercado financeiro brasileiro.

5.2 Aprendizados e Reflexões do Estudo

Este estudo permitiu não apenas explorar a aplicação de um modelo transformer para prever a curva de juros brasileira, mas também identificar desafios e limitações intrínsecos ao processo de modelagem em um mercado financeiro tão dinâmico e volátil. As principais lições obtidas ao longo do desenvolvimento deste trabalho são apresentadas a seguir:

I. A Importância de Iterações Progressivas:

Durante o desenvolvimento, a abordagem inicial com janelas menores de dados e ajustes incrementais da arquitetura do modelo foi fundamental para compreender os impactos de cada alteração nas previsões. Essa estratégia possibilitou identificar rapidamente limitações, como a necessidade de penalização por volatilidade e a inclusão de variáveis auxiliares.

II. Limitações de Modelos com Dados Reduzidos:

Uma das principais reflexões deste trabalho foi perceber que a eficácia de um modelo de aprendizado de máquina está diretamente ligada à qualidade e à abrangência das variáveis fornecidas. Embora o modelo transformer tenha demonstrado bom desempenho em cenários controlados, a utilização de um conjunto limitado de variáveis auxiliares comprometeu sua capacidade preditiva em períodos mais complexos, marcados por choques econômicos e políticos.

III. A Influência de Eventos Externos no Mercado:

A pandemia de COVID-19, as instabilidades políticas no Brasil e outros eventos externos destacaram como o mercado financeiro é influenciado por fatores exógenos e imprevisíveis. Esse aprendizado evidenciou a necessidade de incluir dados que reflitam mudanças estruturais e choques externos para melhorar a robustez das previsões.

IV. A Necessidade de Contexto Econômico Mais Amplo:

Durante os experimentos, ficou claro que as variáveis atualmente utilizadas, como taxa SELIC e reuniões do COPOM, são insuficientes para capturar a complexidade do mercado financeiro. Indicadores adicionais, como inflação, PIB, câmbio e dados globais, poderiam fornecer ao modelo informações cruciais para entender melhor os padrões e relações subjacentes.

V. Impactos de Janelas Temporais Maiores:

O uso de janelas maiores revelou desafios significativos, como o aumento do ruído nos dados e a dificuldade do modelo em lidar com períodos econômicos de alta volatilidade. Apesar disso, também reforçou a necessidade de segmentar dados ou ajustar o modelo para considerar as diferenças entre períodos históricos distintos.

VI. Interação entre Machine Learning e Economia:

Um aprendizado valioso deste estudo foi perceber que, embora técnicas avançadas de machine learning, como transformers, tenham um enorme potencial, seu sucesso depende de uma integração cuidadosa com princípios econômicos e financeiros. Isso sugere a importância de desenvolver abordagens híbridas que combinem métodos estatísticos tradicionais com aprendizado de máquina.

Esses aprendizados servem como base para os ajustes futuros, mas também abrem caminhos para estudos mais aprofundados, que combinem técnicas avançadas de modelagem com uma maior riqueza de variáveis contextuais e econômicas.

5.3 Sugestões para Trabalhos Futuros

O presente estudo demonstrou a viabilidade e as limitações de usar modelos transformers para prever a curva de juros brasileira. Com base nos resultados obtidos e nas reflexões apresentadas, propõem-se as seguintes sugestões para futuros trabalhos que desejem explorar abordagens semelhantes:

5.3.1 Ampliação das Variáveis Auxiliares

Uma das principais limitações observadas foi a insuficiência de variáveis auxiliares para explicar adequadamente as dinâmicas complexas da curva de juros. Para enriquecer o modelo e fornecer um contexto mais amplo, recomenda-se a inclusão de variáveis adicionais, tais como:

I. Indicadores Macroeconômicos Internos:

- a. Inflação (IPCA).
- b. PIB (Produto Interno Bruto) trimestral e crescimento econômico.
- c. Taxas de câmbio, com ênfase na cotação do dólar e do euro.
- d. Taxa de desemprego e índices de confiança do consumidor.

II. Indicadores do Mercado Financeiro:

- a. IBOVESPA e outros índices da bolsa de valores brasileira.
- b. Volume de aportes estrangeiros no mercado financeiro brasileiro.
- c. Curva de juros futuros e comportamento dos derivativos locais.

III. Indicadores Globais:

- a. Yield da Treasury norte-americana.
- b. Dados do mercado de commodities, como petróleo e minério de ferro, que influenciam diretamente a economia brasileira.
- c. Taxas de juros de bancos centrais globais, como o Federal Reserve (EUA) e o Banco Central Europeu.

Essas variáveis podem oferecer uma visão mais rica das condições macroeconômicas que afetam a curva de juros e melhorar significativamente a capacidade preditiva do modelo.

5.3.2 Segmentação Temporal

O uso de períodos amplos para modelagem mostrou-se desafiador devido ao aumento do ruído e às mudanças estruturais no mercado ao longo do tempo. Uma abordagem alternativa seria segmentar os dados históricos em intervalos específicos, como:

- I. Períodos de crise, como a pandemia de COVID-19 (2019-2020) ou a crise política de 2016.

- II. Períodos de estabilidade econômica, onde as variáveis tendem a ter um comportamento mais previsível.

A análise segmentada pode ajudar o modelo a ajustar seus parâmetros de forma mais precisa para diferentes contextos, melhorando a capacidade de generalização.

5.3.3 Integração de Modelos Híbridos

Embora o transformer tenha demonstrado grande potencial, a complexidade do mercado financeiro pode exigir abordagens que combinem diferentes métodos. Sugere-se a exploração de:

- I. **Modelos Estatísticos Tradicionais:** Utilizar econometria para analisar relações históricas entre variáveis macroeconômicas e a curva de juros, como nos modelos de regressão multivariada ou VAR (Modelos Autorregressivos Vetoriais).
- II. **Redes Neurais Reais:** Integrar redes neurais recorrentes (RNNs) ou LSTMs para capturar dependências temporais de longo prazo.
- III. **Abordagens de Ensemble:** Combinar previsões de transformers, redes neurais e modelos estatísticos para criar um modelo híbrido que aproveite as vantagens de cada método.

5.3.4 Expansão do Horizonte de Dados

O uso de dados de apenas um país, como o Brasil, limita a capacidade do modelo de capturar padrões globais que afetam a curva de juros. Futuros estudos podem considerar:

- I. **Inclusão de Dados Internacionais:** Incorporar dados de outros mercados financeiros e bancos centrais para avaliar o impacto de eventos globais na curva de juros brasileira.
- II. **Análises Comparativas:** Comparar o comportamento da curva de juros brasileira com curvas de outros países emergentes ou economias avançadas.

5.3.5 Abordagem Contextualizada de Crises

Futuros trabalhos podem explorar técnicas que reconheçam explicitamente o impacto de crises econômicas e eventos inesperados, como:

- I. Utilizar métodos que identifiquem outliers e adaptem o modelo para lidar com períodos de crise.
- II. Treinar o modelo separadamente para períodos "normais" e períodos de crise, avaliando a eficácia em cada cenário.

Essas sugestões visam não apenas aprimorar o desempenho preditivo, mas também contribuir para uma compreensão mais profunda das dinâmicas econômicas e financeiras que afetam a curva de juros. Trabalhos futuros que incorporem essas recomendações poderão superar as limitações observadas neste estudo e avançar significativamente no uso de aprendizado de máquina para mercados financeiros.

5.4 Limitações do Estudo

Embora este trabalho tenha demonstrado o potencial de modelos transformers para prever a curva de juros brasileira, algumas limitações impactaram os resultados e merecem ser destacadas para contextualizar os desafios enfrentados. As principais limitações observadas foram:

I. Escassez de Variáveis Auxiliares:

- a. O modelo foi alimentado apenas com a taxa SELIC, reuniões do COPOM e informações internas da curva de juros. Apesar de sua relevância, essas variáveis são insuficientes para capturar plenamente a complexidade do mercado financeiro brasileiro, especialmente em períodos marcados por crises ou eventos globais.

II. Impactos de Eventos Extraordinários:

- a. O período analisado (2016-2024) foi marcado por eventos extraordinários, como a pandemia de COVID-19 e mudanças políticas significativas. Esses eventos aumentaram a volatilidade dos dados e introduziram padrões atípicos que o modelo teve dificuldade de compreender com o conjunto limitado de informações fornecidas.

III. Ruído em Janelas Temporais Longas:

- a. A utilização de janelas de dados mais longas, especialmente na etapa final, ampliou o impacto do ruído nos dados históricos. Esse ruído dificultou a generalização do modelo e destacou a necessidade de segmentação dos dados ou de ajustes mais específicos para períodos de alta volatilidade.

IV. Limitações Computacionais:

- a. O aumento do peso computacional, necessário para treinar o modelo em janelas maiores e com configurações mais robustas (como mais camadas e cabeças de atenção), tornou o treinamento mais demorado e menos eficiente. Isso limitou a possibilidade de testar outras configurações ou realizar múltiplas iterações em janelas temporais amplas.

V. Foco Exclusivo no Brasil:

- a. Embora o objetivo principal fosse modelar a curva de juros brasileira, o mercado financeiro global influencia diretamente a dinâmica local. A ausência de dados externos, como o comportamento do mercado norte-americano (ex.: Treasury de 10 anos) ou fluxos de capitais estrangeiros, reduziu a capacidade do modelo de contextualizar as flutuações observadas na curva de juros.

VI. Restrições do Transformer em Mercados Caóticos:

- a. Apesar de sua capacidade avançada de aprendizado, o modelo transformer demonstrou limitações ao lidar com mercados caóticos e interdependentes, como o mercado financeiro. A alta sensibilidade a eventos externos e a necessidade de variáveis macroeconômicas amplas revelaram que modelos complementares ou híbridos podem ser necessários para resultados mais robustos.

5.5 Reflexão Final

Este trabalho explorou a aplicação de um modelo transformer para prever a curva de juros brasileira, oferecendo insights relevantes sobre os avanços e os desafios de usar aprendizado de máquina em mercados financeiros complexos. Ao longo das etapas, foi possível identificar não apenas o potencial dessa abordagem, mas também as limitações inerentes à modelagem de sistemas tão dinâmicos e interdependentes.

Os resultados obtidos demonstraram que, em cenários controlados e com janelas de dados mais curtas, o modelo apresentou bom desempenho, refletindo a capacidade dos transformers de capturar padrões em séries temporais. Contudo, à medida que a complexidade dos dados aumentou, os desafios de lidar com a volatilidade do mercado, eventos extraordinários e um conjunto limitado de variáveis se tornaram evidentes. Isso reforça a necessidade de integrar técnicas mais robustas, variáveis auxiliares diversificadas e métodos híbridos para lidar com as nuances de sistemas caóticos, como o mercado financeiro.

Além disso, o trabalho destaca como o contexto econômico e político brasileiro entre 2016 e 2024 adicionou camadas de incerteza ao problema. A pandemia de COVID-19, mudanças de governos e crises institucionais criaram um ambiente de extrema volatilidade, dificultando a tarefa de prever a dinâmica da curva de juros. Este contexto ilustra a importância de considerar não apenas variáveis técnicas, mas também fatores exógenos que impactam diretamente o comportamento do mercado.

Por fim, este estudo reforça a relevância de combinar avanços tecnológicos com uma compreensão profunda das dinâmicas econômicas e financeiras. O uso de modelos transformers para prever curvas de juros é um campo promissor, mas que ainda requer refinamentos, experimentações e integrações com métodos tradicionais para alcançar seu pleno potencial. As reflexões e sugestões apresentadas neste trabalho visam contribuir para futuros estudos, incentivando a exploração de abordagens mais completas e integradas para enfrentar os desafios do mercado financeiro.

6. Referências

- BARBOSA, F. de H. A longa batalha contra a inflação: uma análise da história recente da economia brasileira. **Estudos Econômicos**, v. 40, n. 2, p. 241-266, 2010.
- BISHOP, C. M. **Pattern recognition and machine learning**. Springer: [s. n.], 2006.
- BOGDANSKI, J.; TOMBINI, A.; WERLANG, S. R. **Implementing inflation targeting in Brazil**. [S. l.]: Banco Central do Brasil, 2000. (Working Paper Series, n. 1).
- CARVALHO, C.; CASTRO, J. M. P. de. Autonomia do Banco Central: fundamentos e implicações. **Estudos Econômicos**, v. 51, n. 1, 2021.
- CARVALHO, C. E. Política monetária e inflação no Brasil: uma análise dos regimes desde o Plano Real. **Estudos Avançados**, v. 35, n. 101, p. 29-48, 2021.
- CAMPBELL, J. Y. Some lessons from the yield curve. **Journal of Economic Perspectives**, v. 9, n. 3, p. 129-152, 1995.
- DUFFEE, G. R. Information in (and not in) the term structure. **Review of Financial Studies**, v. 31, n. 3, p. 553-598, 2018.
- FERGUSON, N. **A ascensão do dinheiro: a história financeira do mundo**. São Paulo: Campus, 2008.
- FRANCO, G. H. B. **O Plano Real e outros ensaios**. Rio de Janeiro: Campus, 2000.
- FRIEDMAN, M. The role of monetary policy. **The American Economic Review**, v. 58, n. 1, p. 1-17, 1968.
- GIAMBIAGI, F.; ALÉM, A. C. **Finanças públicas: teoria e prática no Brasil**. 2. ed. Rio de Janeiro: Elsevier, 2001.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. [S. l.]: MIT Press, 2016.
- GU, S.; KELLY, B. T.; XIU, D. Empirical asset pricing via machine learning. **Review of Financial Studies**, v. 33, n. 5, p. 2223-2273, 2020.
- HALLAC, D.; PAREEK, A.; BOYD, S. Machine learning applied to active fixed-income portfolio management: a Lasso Logit approach. **Journal of Financial Data Science**, v. 1, n. 3, p. 64-81, Fall 2019.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning: data mining, inference, and prediction**. 2. ed. [S. l.]: Springer, 2009.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436-444, 2015.

LIM, B. *et al.* Temporal fusion transformers for interpretable multi-horizon time series forecasting. **International Journal of Forecasting**, v. 37, n. 4, p. 1748-1764, 2021.

LOBO, M. de O. **Mercados de capitais no Brasil imperial: bancos, títulos e finanças públicas**. São Paulo: Alameda, 2006.

LUIS, M. de; RODRÍGUEZ, E.; TORRES, D. **Machine learning applied to active fixed-income portfolio management**: a lasso logit approach. Madrid: Banco de España, 2023. (Working paper 2324).

MARTIN, D.; PÓCZOS, B.; HOLLIFIELD, B. Machine learning-aided modeling of fixed income instruments. In: NIPS 2018 WORKSHOP ON CHALLENGES AND OPPORTUNITIES FOR AI IN FINANCIAL SERVICES: THE IMPACT OF FAIRNESS, EXPLAINABILITY, ACCURACY, AND PRIVACY, 2018. **Proceedings** [...]. Montréal: [s. n.]: 2018. p. 1-7.

MISHKIN, F. S. **The economics of money, banking, and financial markets**. 12. ed. Boston: Pearson, 2019.

NUNES, M. C. M. **Machine learning in fixed income markets**: forecasting and portfolio management. 2022. Tese (Doutorado em Engenharia e Ciências Físicas) – University of Southampton, Southampton Business School, 2022.

OREIRO, J. L. **Macroeconomia do desenvolvimento**: uma perspectiva keynesiana. Rio de Janeiro: LTC, 2017.

PASTORE, A. C. **Inflação e crises**: o papel da política econômica. São Paulo: Nobel, 1994.

REINHART, C. M.; ROGOFF, K. S. **This time is different**: eight centuries of financial folly. Princeton University Press, 2009.

SILVA, A. M.; FURTADO, M. A.; SILVA, M. de A. O sistema SELIC e sua importância no mercado financeiro brasileiro. **Revista Brasileira de Economia**, v. 65, n. 3, 2011.

TAO, D.; SMITH, J. P.; CHEN, Y. Machine learning in fixed income markets: forecasting and portfolio management. **Quantitative Finance and Economics**, v. 3, n. 2, p. 110-138, 2020.

TUCKMAN, B. **Fixed income securities**: tools for today's markets. Hoboken: Wiley, 2002.

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

VILLELA, A.; SUZIGAN, W. **Política do governo e crescimento da economia brasileira: 1889-1945**. Rio de Janeiro: IPEA/INPES, 1973.

YADAV, P. S. Big data analytics and machine learning: transforming fixed income investment strategies. **North American Journal of Engineering and Research**, v. 1, n. 2, p. 1-16, Apr.-June 2021.

DE LUIS, Mercedes; RODRÍGUEZ, Emilio; TORRES, Diego. **Machine Learning Applied to Active Fixed-Income Portfolio Management: A Lasso Logit Approach**. *Documentos de Trabalho – Banco de Espanha*, n. 2324, set. 2023.

NUNES, Manuel Clemente Mendonça. **Machine Learning in Fixed Income Markets: Forecasting and Portfolio Management**. Tese (Doutorado em Engenharia e Ciências Físicas) – University of Southampton, Southampton Business School, 2022.

MARTIN, Daniel; PÓCZOS, Barnabás; HOLLIFIELD, Burton. **Machine learning-aided modeling of fixed income instruments**. *Proceedings of the Neural Information Processing Systems Workshop on Challenges and Opportunities for AI in Financial Services*, Montréal, Canada, 2018.