



TRABALHO DE CONCLUSÃO DE CURSO

**PREDIÇÃO DO PIB TRIMESTRAL BRASILEIRO
VIA EXTREME LEARNING MACHINES**

Otávio Alves Dias

Brasília, Dezembro de 2018

UNIVERSIDADE DE BRASÍLIA

FACULDADE DE TECNOLOGIA

UNIVERSIDADE DE BRASÍLIA
Faculdade de Tecnologia

TRABALHO DE CONCLUSÃO DE CURSO
**PREDIÇÃO DO PIB TRIMESTRAL BRASILEIRO
VIA EXTREME LEARNING MACHINES**

Otávio Alves Dias

*Trabalho de Conclusão de Curso submetido ao Departamento de Engenharia
Elétrica como requisito parcial para obtenção
do grau de Bacharel em Engenharia de Computação*

Banca Examinadora

Prof. Daniel Guerreiro e Silva, Ph.D, FT/UnB

Orientador

Prof. Alexandre Zaghetto, Ph.D, CIC/UnB

Avaliador

Prof. Adolfo Bauchspiess, Ph.D, FT/UnB

Avaliador

FICHA CATALOGRÁFICA

DIAS, OTÁVIO ALVES

PREDIÇÃO DO PIB TRIMESTRAL BRASILEIRO VIA EXTREME LEARNING MACHINES [Distrito Federal] 2018.

xvi, 42 p., 210 x 297 mm (ENE/FT/UnB, Bacharel em, Engenharia Elétrica, 2018).

Trabalho de Conclusão de Curso - Universidade de Brasília, Faculdade de Tecnologia.

Departamento de Engenharia Elétrica

1. Produto Interno Bruto

2. Redes Neurais Artificiais

3. Máquinas de Aprendizado Extremo

4. Nowcasting

I. ENE/FT/UnB

II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

DIAS, O.A. (2018). *PREDIÇÃO DO PIB TRIMESTRAL BRASILEIRO VIA EXTREME LEARNING MACHINES*. Trabalho de Conclusão de Curso, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 42 p.

CESSÃO DE DIREITOS

AUTOR: Otávio Alves Dias

TÍTULO: PREDIÇÃO DO PIB TRIMESTRAL BRASILEIRO VIA EXTREME LEARNING MACHINES.

GRAU: Bacharel em Engenharia de Computação ANO: 2018

É concedida à Universidade de Brasília permissão para reproduzir cópias deste Trabalho de Conclusão de Curso e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. Os autores reservam outros direitos de publicação e nenhuma parte desse Trabalho de Conclusão de Curso pode ser reproduzida sem autorização por escrito dos autores.

Otávio Alves Dias

Depto. de Engenharia Elétrica (ENE) - FT

Universidade de Brasília (UnB)

Campus Darcy Ribeiro

CEP 70919-970 - Brasília - DF - Brasil

Agradecimentos

Este período de graduação foi um marco em minha vida e me ensinou muito além das salas de aula. Portanto, gostaria de agradecer às pessoas que me acompanharam durante esta jornada e que, de maneira direta ou indireta, foram responsáveis pela minha evolução pessoal e profissional.

Agradeço aos meus pais, Maria Regina Alves Dias e Roberto José Dias, pelo amor, companheirismo, apoio e compreensão durante toda a minha vida e principalmente durante este processo de formação. Nada disso teria sido possível sem a presença de vocês.

Aos grandes amigos que acompanharam minha trajetória, dividindo momentos pessoais, acadêmicos e profissionais, de conquistas e de derrotas, de alegrias e de tristezas: Gustavo Henrique Lourenço Ladeira, Lucas Conceição Macedo, Otacílio Antônio de Souza Neto, Anna Carolina Nunes Pinheiro e Túlio Abner de Lima. Obrigado pela lealdade.

À minha namorada Renata Visoná Barbosa, pelo amor, compreensão, paciência e momentos de alegria.

Ao meu orientador, Daniel Guerreiro e Silva, pela paciente, simpática e respeitosa orientação. Obrigado pela confiança em meu potencial e também pelo amparo acadêmico fornecido durante os momentos de dificuldade.

Ao professor Eduardo Peixoto Fernandes da Silva, pela confiança em meu potencial acadêmico, pelas críticas, ensinamentos e orientações de valor imensurável que levarei por toda minha vida.

Aos professores Jacir Luiz Bordim e Alexandre Zaghetto, pela competência, solicitude, conselhos e conversas.

A Deus, pelo amparo nos momentos de dificuldade, dando-me forças para continuar.

Otávio Alves Dias

RESUMO

Decisões políticas e monetárias são feitas baseadas em avaliações do estado econômico e em suas condições futuras. Dado que a maioria dos dados oficiais encontrados são disponibilizados com atraso, conseguir estimar o estado atual da economia antes de suas publicações oficiais é de grande valia para bancos, mercados financeiros e instituições federais.

Com o objetivo de solucionar este problema de atraso da divulgação do PIB, técnicas de aprendizado de máquina têm sido usadas para se estimar o valor de indicadores econômicos a curto prazo, analisando dados históricos em conjunto com dados recentes para a construção de modelos.

Este trabalho estuda a aplicação de aprendizado de máquina à estimativa do Produto Interno Bruto trimestral brasileiro. O processo experimental conduz, a partir de uma base de dados reais, a estimativa deste índice com auxílio de modelos de regressão linear e máquinas de aprendizado extremo, com o intuito de propor uma técnica de implementação simples e com boa capacidade de generalização.

Este trabalho tem como principais contribuições: um estudo comparativo dos modelos preditivos selecionados; e a adoção de um modelo de rede neural artificial como alternativa de solução para problemas de *nowcasting*, apresentando um ganho de precisão de 1.46% em comparação ao modelo de regressão linear.

Palavras-chave: Produto Interno Bruto, Redes Neurais Artificiais, Máquinas de Aprendizado Extremo, *Nowcasting*

ABSTRACT

Monetary policy decisions are made based on evaluations of the economic state and its future conditions. Given that most official data is released with a lag, being able to estimate the current economic state before its official release is of great value to banks, financial markets and federal institutions.

In order to solve this problem, machine learning techniques have been used to estimate the value of economic indicators in short-term periods, analyzing historical data and recent data to build predictive models.

This work studies the application of machine learning to estimate the brazilian quarterly Gross Domestic Product. The experimental process performs, from a real database, the prediction of this index aided by linear regression and extreme learning machine models, leading to a simple technique to implement and with good generalization.

The main contributions of this work are: a comparative study of the selected predictive models; and the adoption of a artificial neural network model as alternative solution to nowcasting problems, with an accuracy gain of 1.46% in comparison to the linear regression model.

Keywords: Gross Domestic Product, Artificial Neural Networks, Extreme Learning Machines, Nowcasting

SUMÁRIO

1	Introdução	1
1.1	Objetivos do trabalho	2
1.2	Organização da monografia	2
2	O Problema de Predição do PIB Trimestral	3
2.1	O Produto Interno Bruto	3
2.1.1	As óticas do produto	4
2.1.2	Definição	6
2.1.3	PIB Real e PIB Nominal	6
2.2	Descrição do Problema de Predição do PIB	7
2.2.1	A metodologia de <i>Nowcasting</i>	7
2.3	Trabalhos Relacionados	8
2.3.1	Avaliação de técnicas de nowcasting	8
2.3.2	Previsões para o crescimento do PIB trimestral brasileiro	9
2.3.3	Métodos de Machine Learning aplicados às projeções de PIB	10
2.4	Síntese	12
3	Regressão pelo mínimo erro quadrático	13
3.1	Aprendizado de Máquina	13
3.1.1	Aprendizado supervisionado	14
3.1.2	Aprendizado por reforço	14
3.1.3	Aprendizado não-supervisionado	14
3.2	Regressão Linear	14
3.2.1	Regressão Univariada	16
3.2.2	Regressão Multivariada	17
3.3	Máquinas de Aprendizado Extremo	19
3.3.1	Redes Neurais Artificiais	19
3.3.2	Definição para Máquinas de Aprendizado Extremo	22
3.4	Síntese	25
4	Estudo de caso e resultados	27
4.1	Conjunto de dados e seleção das variáveis	27
4.1.1	Validação Cruzada	29
4.2	Regressão linear multivariada	30
4.2.1	Resultados	30
4.3	Máquina de Aprendizado Extremo	33
4.3.1	Resultados	33

4.3.2	Análise Comparativa.....	35
4.4	Síntese	37
5	Conclusão e trabalhos futuros	38
5.1	Discussão dos resultados obtidos.....	38
5.2	Trabalhos futuros	39
	REFERÊNCIAS BIBLIOGRÁFICAS	41

LISTA DE FIGURAS

3.1	Exemplo de regressão linear univariada.....	16
3.2	Exemplos de ajuste.....	18
3.3	Modelo de um neurônio.....	20
3.4	Rede <i>feedforward</i> de uma camada	21
3.5	Rede <i>feedforward</i> de múltiplas camadas.....	22
3.6	Rede recorrente	23
4.1	Regressões aplicadas	32
4.2	Predições realizadas com a ELM regularizada	35
4.3	Comparativo entre as metodologias utilizadas	37

LISTA DE TABELAS

4.1	Indicadores selecionados	28
4.2	Valores dos pesos da regressão linear multivariada	31
4.3	Erro absoluto médio para cada conjunto de dados	31
4.4	Relação entre número de neurônios e EAM resultante do processo de <i>K-Fold</i> <i>Cross Validation</i>	34
4.5	Configurações de parâmetros utilizados pela ELM	34
4.6	Comparação dos erros fornecidos pelo método de validação cruzada	36
4.7	Comparação dos resultados de PIB obtidos	36

LISTA DE SÍMBOLOS

Símbolos

x_j	Sinal de entrada da sinapse j em um neurônio
\mathbf{X}	Matriz de variáveis de entrada
b_k	Valor de <i>bias</i> no neurônio k
φ	Função de ativação de um neurônio
y_k	Valor de saída do neurônio k
Y	Conjunto de variáveis de saída
$h_w(x)$	Função hipótese aproximada
$J(w)$	Função custo
\mathbf{w}	Vetor de coeficientes lineares de $h_w(x)$
α	Taxa de aprendizagem
\mathbf{B}	Vetor de pesos das conexões da camada de saída de uma ELM
\mathbf{H}	Saída da camada oculta de uma ELM
$E(Y X)$	Valor esperado de Y dado o valor de X
C	Coefficiente de regularização
N	Quantidade de amostras
m	Quantidade de variáveis
L	Quantidade de neurônios
\mathbf{I}	Matriz identidade
\mathbf{T}	Matriz alvo de treinamento

Sobrescritos

T	Transposição de um vetor ou matriz
-1	Inversão de um vetor ou matriz

Siglas

PIB	Produto Interno Bruto
IA	Inteligência Artificial
ML	<i>Machine Learning</i> (Aprendizado de Máquina)
RNA	Redes Neurais Artificiais
ELM	<i>Extreme Learning Machine</i> (Máquina de Aprendizado Extremo)
MLP	<i>Multi Layer Perceptron</i> (Perceptron de Múltiplas Camadas)
EAM	Erro Absoluto Médio

1 INTRODUÇÃO

Decisões políticas e monetárias são feitas baseadas em avaliações do estado atual da economia e em suas condições futuras [1]. Posto que a maioria dos dados oficiais encontrados são disponibilizados com atraso, instituições financeiras investem parte de seus recursos para prever o contexto político-monetário do país, estimando variáveis do período corrente antes mesmo de suas publicações oficiais [2].

O ato de se "prever" o estado atual da economia é conhecido como *nowcasting*. Uma vez que o contexto político-monetário de um país possui uma importante influência nas tomadas de decisão, instituições mantêm constante vigília sobre índices que ajudem em leituras pontuais e precisas. A busca por métodos capazes de satisfazer essa necessidade tem se intensificado nos últimos anos, direcionando os esforços de bancos e mercados financeiros para o aprimoramento das técnicas empregadas atualmente [3]. A necessidade por informação torna-se ainda mais importante em períodos de crise, pois estes possuem incertezas econômicas associadas a decisões de alto risco [2]. Sendo assim, conhecer a economia a curto prazo é importante pois há um maior grau de previsibilidade do que a longo prazo [1].

O crescimento de um país pode ser observado por um conjunto de fatores, entretanto, o Produto Interno Bruto é um dos índices que melhor elucida o estado econômico de um dado período. Ainda assim, por se tratar de um agregado que resume diversos outros setores, calculá-lo requer reunir uma grande massa de dados de diversas fontes distintas e seu processamento tende a ser oneroso, gerando um atraso entre o período de disponibilização dos dados e o fim do cálculo. Deste modo, há um período de tempo significativo em que os mercados financeiros ficam sem saber como a economia se desenvolve, demorando até um semestre para que informações oficiais sejam divulgadas [2].

Modelos de *nowcasting* visam descobrir mudanças no ciclo econômico conforme elas acontecem, bem como explicitar uma visão transparente de que fatores impactam na variação do crescimento de um país [2]. Existe uma vasta categoria de técnicas que podem ser utilizadas para se estimar uma determinada variável, todavia, a incerteza inerente dos modelos de predição varia de acordo com a técnica utilizada, muitas vezes necessitando do julgamento constante de times especializados [4].

Uma pesquisa conduzida pela *PricewaterhouseCoopers (PwC)* [2] mostrou que a utilização de técnicas de aprendizado de máquina pode diminuir o esforço para o cômputo do Produto Interno Bruto, assim como aumentar ligeiramente a precisão das estimativas. Atualmente, destacam-se as redes neurais artificiais, tendo sido utilizadas em diversos campos do conhecimento por serem capazes de solucionar problemas não triviais [5].

Dentro deste contexto das redes neurais, o trabalho de *Huang et al.* [5, 6] propõe o uso de Máquinas de Aprendizado Extremo como uma alternativa aos métodos tradicionais utilizados,

constatando maior facilidade de utilização e grande capacidade de generalização. Desta forma, um algoritmo simples e robusto que seja capaz de estimar indicadores econômicos pode ser uma alternativa de solução ao problema de *nowcasting* enfrentado pelas instituições financeiras.

1.1 OBJETIVOS DO TRABALHO

Este trabalho tem como objetivo:

1. Estudar e descrever o problema de predição do Produto Interno Bruto trimestral.
2. Estudar e comparar a aplicação de regressões lineares e redes neurais artificiais, mais especificamente as Máquinas de Aprendizado Extremo.
3. Propor uma abordagem alternativa de solução ao problema de predição de PIB utilizando técnicas de aprendizado de máquina.

1.2 ORGANIZAÇÃO DA MONOGRAFIA

O Capítulo 2 introduz o campo de estudo da macroeconomia, dando ênfase a metodologias utilizadas para o cômputo do Produto Interno Bruto de um país, também elucidando o problema de predição deste indicador. O Capítulo 3 apresenta conceitos básicos fundamentais sobre aprendizado de máquina, focando-se em duas técnicas: regressões lineares e redes neurais artificiais.

No Capítulo 4 são detalhados os experimentos realizados com as técnicas de aprendizado de máquina, assim como os resultados obtidos. Por fim, no Capítulo 5 são apresentadas as conclusões do trabalho, a avaliação dos resultados e as perspectivas de trabalhos futuros.

2 O PROBLEMA DE PREDIÇÃO DO PIB TRIMESTRAL

Este capítulo apresenta conceitos introdutórios acerca do Produto Interno Bruto (PIB) e os desafios enfrentados para sua predição. São apresentadas definições macroeconômicas fundamentais para a compreensão do contexto em que este trabalho se enquadra. Por fim, descreve-se o processo de predição do indicador em questão, apontando sua aplicação, importância e obstáculos associados. As principais referências para este capítulo são os trabalhos de *Krugman* [7], *Mankiw* [8] e os boletins de 2014 e 2017 elaborados pelo *Banco da Inglaterra* [3, 4].

Para melhor compreensão do contexto deste trabalho, é necessário distinguir os conceitos de macroeconomia e microeconomia. A macroeconomia estuda eventos econômicos e seus comportamentos agregados. Ela analisa, dentro de um período de tempo, o efeito combinado de decisões individuais de diversos setores produtivos sobre o contexto geral de um país. O somatório de ações econômicas individuais podem resultar em impactos maiores que o valor final esperado para o todo. Já a microeconomia refere-se a como indivíduos e firmas tomam decisões e quais são as consequências destas decisões para o mercado.

É importante frisar que a macroeconomia fornece uma visão estratégica, focando-se no gerenciamento de recursos, controle de flutuações, previsão de eventos, prevenção e redução de impactos. Pode-se acrescentar que ela se volta à consolidação de informações derivadas de agregados econômicos, ou seja, mensurações que resumem dados de diversos mercados distintos com o objetivo de escalar o desempenho da economia de um país. Dadas tais descrições, percebe-se que este ramo estuda eventos de longo prazo que servem de insumo na tomada de decisão, possibilitando ações que aumentem a quantidade total de recursos de um ou mais setores e elevem suas respectivas taxas de crescimento.

Portanto, o escopo macroeconômico é o enfoque deste trabalho e cabe ressaltar a importância desta área de pesquisa, uma vez que o estado da economia afeta cada cidadão individualmente. Este ramo de estudo ocupa um papel de relevância no campo das políticas públicas e, consequentemente, é assunto fundamental em debates políticos mundiais, levando em consideração diversos fatores econômicos através da consolidação de dados de diversos setores, com o propósito de serem elaborados modelos que possam descrever e até prever determinados comportamentos. Sendo assim, a mensuração econômica é uma aliada de governos, empresas e também do cidadão: boas políticas dependem de medidas que auxiliem as tomadas de decisão.

2.1 O PRODUTO INTERNO BRUTO

O Produto Interno Bruto (PIB), também conhecido em inglês como *Gross Domestic Product* (GDP), pode ser visto como o termômetro de salubridade econômica de um país, ou seja, este é

um indicador numérico cujas fontes primárias advém de orçamentos e pesquisas governamentais com o intuito de resumir a atividade econômica em um dado período de tempo. No Brasil, essas medidas são calculadas trimestralmente e são divulgadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE) [9].

A contabilidade nacional é a responsável por agregar conjuntos de indicadores de um país. O foco da contabilidade nacional é monitorar rendas e gastos que compõem o fluxo monetário nacional. Há duas maneiras de se interpretar este fluxo: o somatório da renda total ou o somatório dos gastos totais, ambos os pontos de vista são descritos pelo mesmo indicador e partem do princípio que os fluxos devem se igualar. Como questiona Mankiw [8], como poderia o PIB medir tanto os rendimentos quanto os gastos? Para uma economia real, em cada transação há a participação tanto de um comprador quanto de um vendedor, e cada unidade monetária gasta pelo comprador é convertida em renda para o vendedor.

Deve-se ressaltar que há a inclusão de um número variado e vasto de dados a serem computados para composição do PIB. Assim como evidenciado pelo nome, o Produto Interno Bruto é composto de três conceitos importantes, sendo o primeiro chamado de "produto" e voltado ao cálculo de valores monetários acerca da produtividade de um país; o segundo conhecido por "interno" e focado em descrever a área de atuação geográfica do indicador e, por fim, o terceiro denominado "bruto" pois faz referência às depreciações [7, 8].

2.1.1 As óticas do produto

O fluxo monetário circular possui etapas distintas, gerando pontos de vista heterogêneos ao longo do processo e que dependem dos agentes envolvidos. Desta forma, o produto pode ser definido sob três óticas: a ótica do produto, a ótica da renda e a ótica dos dispêndios [7].

A primeira ótica define produto como "o valor total da produção de bens e serviços finais da economia em um determinado período de tempo dentro das fronteiras de um país". O PIB contempla apenas valores de produtos finais para não incorrer no erro de múltipla contagem, dado que o valor de mercado dos insumos já está contido no valor do bem final [7].

O conceito de valor adicionado ilustra a razão de serem somados apenas o valor de bens e serviços finais no cálculo do produto: o valor final constitui-se da adição de todos os valores precedentes, sendo assim, já considera os valores de produção dos insumos anteriores. O valor adicionado pode ser visto como o valor líquido da produção e a soma destes valores em todos os setores produtivos resulta no valor da produção dos bens finais [7].

A segunda ótica define produto como "a soma dos rendimentos pagos em todos os setores produtivos". Os valores adicionados também incluem pagamentos efetuados a quem contribuiu para a transformação dos insumos e agrupam salários, lucros, juros e aluguéis. Os valores adicionados se traduzem nos pagamentos efetuados aos fatores da produção que participaram do processo. Esta ótica postula que o produto não é necessariamente apenas o somatório da produção de um

determinado bem, mas também a renda gerada através dele [7].

A terceira ótica define que o produto também é "o conjunto de dispêndios". Os investimentos privados consistem principalmente de bens de capital comprados pelas empresas. Chama-se de investimento líquido aquele que representa um aumento da capacidade produtiva. O investimento de estoques acumulados é aquele que aumenta os estoques de bens para uso futuro. O investimento bruto é o investimento líquido mais os gastos relativos à depreciação de equipamentos. Os gastos governamentais são aqueles gerados pelo Estado e são de difícil mensuração. Pode-se citar equipamentos militares, trabalhadores do governo como bombeiros e policiais, construções públicas, manutenções de estradas, dentre outros. A exportação líquida (diferença entre exportações e importações) engloba relações comerciais com outras nações [7].

O produto, de acordo com a fórmula abaixo, é resultado da soma de consumos privados (C), investimentos privados (I), gastos governamentais (G) e as exportações líquidas, ou seja, a diferença entre o valor das exportações (E) e importações (M).

$$P = C + I + G + (E - M) \quad (2.1)$$

Mankiw [8] aponta a existência de questões a serem levantadas conforme o cálculo torna-se paulatinamente verossímil: imaginando que um bem seja produzido e vendido por uma loja a um preço fixo, este mesmo valor será adicionado ao cômputo do PIB de um país conforme supracitado. Ainda assim, e se a pessoa que o comprou decidir revendê-lo por um valor diferente do pago inicialmente? Isto não mais representaria um gasto da produção do momento mas sim uma transferência de um bem já existente e de valor previamente estabelecido, logo, o cálculo do PIB não contempla vendas ou trocas de bens usados, apenas gastos de produção de novos bens.

Um segundo questionamento refere-se ao chamado tratamento de inventários: num cenário onde uma empresa aumenta seu inventário de bens como, por exemplo, estocar lotes para venda posterior, o investimento é contabilizado como um gasto e assim ocorre um aumento no PIB juntamente dos custos de produção. Qual a influência disto sobre o PIB? No caso de venda destes mesmos estoques, ocorre uma combinação do custo de compra dos bens com o rendimento gerado pela venda e, desta forma, não há influência sobre o valor do PIB.

O último questionamento levanta o exemplo de serviços de moradia. Determinados bens e serviços não são vendidos amplamente no mercado e, deste modo, não possuem valores de mercado agregados. Para melhor mensuração destes, faz-se necessário o processo de imputação ou estimativa de preços; o aluguel de moradias pode ser visto como a compra de serviços domésticos e produz renda para o locador, renda esta que é computada no valor do PIB. No caso de imóveis próprios, não há aluguel, mas há a estimativa do aluguel que se pagaria. Imputações também são efetuadas sobre serviços governamentais de segurança e saúde ao serem avaliados pelo custo de manutenção necessário. Todavia, vale salientar que o consumo efetuado dentro de uma casa não é computado no valor do PIB, assim como bens e serviços originados de economias informais e ilegais.

2.1.2 Definição

Conforme explanado anteriormente, as definições de produto convergem no sentido de que podem ser consideradas equivalentes, desta forma, uma definição mais ampla de produto é que este pode ser visto como o valor da produção de bens e serviços finais, o gasto em bens e serviços finais produzidos internamente ou a renda dos fatores ganhos das firmas na economia [8].

Um produto é "interno" quando este ignora remessas e recebimentos de rendimentos do exterior, caracterizando apenas a produção no país, ou seja, dentro das fronteiras, desconsiderando bens e serviços produzidos no exterior. Em um contexto real de economia aberta, transações ocorrem entre dois países e estas resultam em custos e rendimentos. O produto interno não considera se os detentores dos fatores de produção são residentes ou não do país, apenas sua produtividade interna [8].

O significado de "bruto" indica que a depreciação está inclusa nos cálculos, ou seja, é a soma do investimento líquido com os dispêndios destinados a repor a produção deteriorada. Uma vez constatado que há uma parte do valor agregado que não constitui rendimento, as empresas separam uma parcela da receita para composição de uma reserva responsável por custear manutenção e reposição de maquinário, ou seja, uma parcela dos bens e serviços é destinada apenas para garantir a capacidade produtiva do capital existente [8].

2.1.3 PIB Real e PIB Nominal

Dado um cenário em que se deseja realizar comparações ao longo do tempo, há problemas referenciais que podem não refletir o cenário real da economia. O PIB pode aumentar ou diminuir de acordo com a flutuação dos preços ou com a flutuação da produtividade e estes não necessitam estar relacionados. A fim de levar em consideração as variações do valor agregado e diminuir as falsas impressões de melhora ou piora econômica, faz-se necessária a diferenciação entre PIB nominal e PIB real.

A soma do valor de todos os bens produzidos em um período considerado é também conhecida como PIB Nominal e avalia a produção a preços correntes. Uma vez que o preço de um bem se eleva, o valor do PIB aumentará proporcionalmente mesmo que não ocorra mudança na quantidade de vendas e neste cenário fica evidente que a economia não está sendo afetada de maneira alguma pela oscilação dos preços. Sendo assim, pode-se perceber que tal medida é inconfiável.

Parte do aumento do valor de um bem pode refletir apenas preços mais altos mas não necessariamente um incremento na produção e, para não incorrerem do erro citado no último parágrafo, faz-se uso do PIB Real, ou seja, uma avaliação da produção de bens e serviços a preços constantes num determinado período de tempo em relação a um ano base de referência e levando-se em conta os valores agregados [7, 8].

2.2 DESCRIÇÃO DO PROBLEMA DE PREDIÇÃO DO PIB

A macroeconomia é uma ciência imperfeita e a habilidade dos economistas de prever o futuro de determinados eventos econômicos equipara-se às tentativas de previsão do tempo, dado que o volume de fatores a serem considerados e a volatilidade destes são altos [8]. Decisões políticas e monetárias são tomadas em tempo real e, em contrapartida, posto que algumas variáveis econômicas são disponibilizadas com um atraso significativo e sofrem revisões posteriores, as informações disponíveis podem ser incompletas, inacuradas e até mesmo inexistentes [1].

2.2.1 A metodologia de *Nowcasting*

A metodologia conhecida como *Nowcasting* utilizada por bancos centrais ao redor do mundo, realiza a análise de dados históricos em conjunto com dados recentes para construção de modelos preditivos de valores de determinados índices do trimestre atual. Dado que métodos estatísticos são aplicados para a realização de tal estimativa e que os modelos devem ser atualizados ao receberem novas informações, questiona-se, em primeiro lugar, se um grande número de dados ampara o processo para que ele se torne mais acurado e, em segundo lugar, quais são os grupos de variáveis que contribuem com maior peso [1].

Diversas metodologias de *nowcasting* aplicadas à estimativa do PIB existem em virtude dos diferentes grupos de variáveis que podem ser utilizadas no processo de modelagem, do período de tempo em que os modelos são aplicados e do país avaliado. O PIB é frequentemente revisado para incorporar novas fontes de dados e sua estimativa pode ser feita com o uso de dados mensais ou trimestrais. É necessário apontar que dados mensais são mais voláteis que os trimestrais e, desta forma, modelos que utilizam o primeiro conjunto tendem a ser menos precisos [3].

Sabendo que cada setor possui seu próprio cálculo de indicadores, a predição do PIB ocorre ao se utilizar o peso de cada um deles. Um dos problemas mais comuns de serem enfrentados durante o cômputo da estimativa é a disponibilidade dos indicadores: a frequência com a qual são encontrados difere, abrangendo dias, semanas, meses e trimestres; além disso, são disponibilizados em intervalos diferentes de tempo. Sendo assim, os modelos tendem a descartar informações que podem ser relevantes e também são obrigados a realizar novos cálculos de predição para alguns dos indicadores, aumentando a margem de erro [3]. Desta forma, ao se utilizar métodos distintos, as incertezas associadas diferem entre si a depender da frequência dos dados e do intervalo de tempo adotado e determinadas estimativas podem ser mais acuradas ao início de um trimestre do que ao fim [4].

Visando alternativas de solução para os desafios citados nos parágrafos anteriores, existem técnicas consolidadas utilizadas na arquitetura de diversos bancos, sendo as mais comuns listadas abaixo [3]:

- *Bridge Equations*: regressões do crescimento trimestral do PIB utilizando indicadores men-

sais;

- *Mixed-Data Sampling* (MIDAS): modelo utilizado para sanar o problema de diferentes frequências entre os dados sem descartar parte da informação de uma ou mais séries;
- *Factor Models*: tendências estatísticas conhecidas como "fatores" são estimadas através de grandes conjuntos de dados.

Conforme já mencionado, as estimativas do PIB são imperfeitas e erros associados aos métodos utilizados já são esperados. Erros demasiadamente altos claramente mostram que a estimativa não resultou em informações satisfatórias, mas não significa que o processo de *nowcasting* é falho. Um evento de larga escala como uma grande depressão, a ocorrência de olimpíadas em um país ou uma crise política intensa são eventos históricos precursores de um incremento das margens de erro e possuem baixas probabilidades de ocorrerem novamente ou são incertos demais para serem antecipados. A decisão de se substituir o modelo utilizado é questionável em face de eventos inesperados [4].

Os modelos de predição geralmente utilizam dados históricos de longos períodos de tempo e conseqüentemente não respondem rapidamente às novas informações. Existem ocasiões onde um indicador diverge do resultado esperado, então o modelo não captura adequadamente as variações dos demais indicadores. De acordo com este cenário, há a verificação dos pesos dos indicadores e mensurações da performance preditiva são efetuadas com o auxílio de um modelo regressivo de *benchmark* que dependa linearmente de valores prévios, uma vez que eles produzem estimativas relativamente acuradas para uma dada janela temporal [4].

2.3 TRABALHOS RELACIONADOS

Dada a importância do tema para a economia de um país e a existência de diversas abordagens para se realizar predições via *nowcasting*, é natural que haja diversas pesquisas voltadas para este tema em específico. Nesta seção são apresentados os trabalhos utilizados como fundamentação teórica para modelos de predição do PIB. Cabe enfatizar que os trabalhos da *PricewaterhouseCoopers (PwC)* [2], de *Vasconcelos* [10] e *Kagohara* [11] foram tomados como inspiração principal no desenvolvimento desta monografia.

2.3.1 Avaliação de técnicas de nowcasting

O trabalho de *Kagohara* [11] propõe a utilização de modelos de *nowcasting* para previsão do PIB real brasileiro. O autor utiliza como fundamentação a premissa de que modelos de previsão de variáveis macroeconômicas apresentam problemas associados aos indicadores utilizados como, por exemplo, falta de dados, diferença temporal entre publicações, falta de acurácia e constantes correções [12]. Desta forma, como alternativa de solução, optou-se pela utilização

de "*bridge equations*" com base em indicadores de maior frequência para aplicação da estimativa proposta [12, 1].

A dissertação de *Kagohara* [11] destaca que a economia brasileira possui instabilidades que dificultam o cálculo de previsão do PIB. O período avaliado foi de janeiro de 1980 a dezembro de 2014 e, de maneira a ilustrar o contexto estudado, são citados exemplos como os da década de 1980, que apresenta 24 trimestres de recessão e 16 de expansão, e 1990 com 20 trimestres de recessão e 20 de expansão. Por questões históricas como a execução do plano Collor e implantação do plano Real, há uma série de tratamentos que devem ser efetuados antes da utilização dos modelos preditivos, pois tais eventos influenciam diretamente nos valores dos indicadores e geram quebras estruturais e temporais que podem refletir num incremento do erro associado. Desta forma, o autor faz uma seleção de 72 variáveis que foram significativas em algum momento da história para explicar a variação do PIB Brasileiro, contendo dados nacionais e internacionais e subdividindo-as em variados grupos que abrangem diferentes setores da economia, a exemplo de produção industrial, mercado financeiro e balança comercial, para elencar os níveis de significância.

O resultado dos modelos aplicados para previsão do PIB brasileiro sugere que técnicas de *nowcasting* são relevantes mesmo considerando uma economia em desenvolvimento. Por fim, partindo do conjunto de dados citado no parágrafo anterior, alcançou-se como resultado um compilado com as variáveis de maior significância preditiva, sendo estas: Ibovespa, IPCA, pessoal ocupado na indústria de São Paulo, IGPDI, exportações, taxa de juros americana e variáveis de produção industrial.

2.3.2 Previsões para o crescimento do PIB trimestral brasileiro

A periodicidade dos índices, conforme já citado, é um dos principais problemas enfrentados ao se realizar estimativas macroeconômicas; desta forma, o trabalho de *Zuanazzi e Ziegelmann* [13] visa a utilização do modelo MIDAS para previsão do PIB brasileiro como alternativa de solução para as diferentes frequências encontradas em variáveis com potencial preditivo. Os autores utilizam séries mensais do índice Bovespa, índice DowJones, taxa de câmbio (dólar PTAX), CDI, produção industrial (geral, extrativa e de transformação), preço do barril do petróleo e valor total das exportações dentro do período do 2º trimestre de 1996 e o 4º trimestre de 2012.

Como primeira etapa, cada regressor foi utilizado de maneira isolada para realização da estimativa do PIB. Os maiores destaques nas previsões foram das variações do Ibovespa, indústria geral, indústria de transformação e exportações. Quando analisados os pesos finais tendo como regressor a variação do Ibovespa, percebeu-se que um crescimento desta variável está associado a um incremento futuro do PIB. Na segunda etapa, foram utilizados regressores múltiplos sob diferentes combinações e variações da produção da indústria geral mostraram associação aos maiores impactos na variação do PIB em conjunto com o Ibovespa.

O trabalho de *Zuanazzi e Ziegelmann* [13] reafirma, ao fazer uso de diferentes técnicas de

nowcasting, o que já fora discutido por *Kagohara* [11]. Percebe-se que a utilização de múltiplos regressores, principalmente àqueles associados à indústria, às exportações e aos índices do mercado financeiro são os mais significativos em relação às variações de crescimento do PIB.

O trabalho de *Mattos* [14] também disserta acerca da utilização de técnicas de *nowcasting* e utiliza o modelo de fatores dinâmicos com dados normalizados de 21 séries temporais de diferentes periodicidades, englobando variáveis como IPCA, preço do barril de petróleo e exportações.

Mattos [14] conclui que a utilização de modelos de *nowcasting* é menos dispendiosa e possui um bom desempenho quando utilizado fora de períodos de recessão econômica. Por fim, evidencia-se que a inclusão de variados indicadores contemplando maiores intervalos de tempo possam diminuir o erro imputado por crises e recessões.

2.3.3 Métodos de Machine Learning aplicados às projeções de PIB

O autor *Vasconcelos* [10] tem como foco a compreensão de metodologias de escolha de modelos e a seleção de variáveis empregando aprendizado de máquina através de um exercício de predição elaborado tendo como contexto o comportamento de economias mundiais. O trabalho é dividido em duas etapas, a primeira volta-se para o estudo da relação de causa e efeito entre investimento e produto; a segunda tem como principal característica antecipar flutuações do produto.

Vasconcelos aborda conceitos fundamentais de aprendizado de máquina, caracterizando os diferentes modelos preditivos mais utilizados e aplicando-os ao escopo econômico. Além disso, comenta que muitas das variáveis apontadas como significativas e utilizadas nos processos preditivos são alvos de crítica por não apresentarem robustez a testes de sensibilidade ou por não possuírem uma relação de causa e efeito evidentes em relação ao crescimento do PIB. Para a primeira análise, concluiu-se que o funcionamento interno da economia de um país é heterogêneo e tal fenômeno pode ser explicado por distorções internas e qualidade das instituições presentes, deste modo, há interferência no desenvolvimento de teorias mais abrangentes e gerais, o que explica o fato de determinadas variáveis apresentarem forte significância para alguns países enquanto que para outros não. Ao exemplo do cenário brasileiro, o país depende fortemente de fluxos de capitais internacionais enquanto os ciclos de países mais ricos são pouco influenciados por investimentos internacionais e, desta maneira, verifica-se que a aversão ao risco internacional é presente no estabelecimento das dinâmicas de um país. Por fim, aponta-se que há mais países em que o investimento causa o crescimento do produto do que o contrário.

A segunda análise realizada propõe uma abordagem baseada em técnicas de aprendizado de máquina para identificar variáveis significativas. O procedimento utiliza indicadores do mercado de trabalho, investimento, da produtividade e da interação do país com o mercado internacional entre o período de treinamento de 2001 a 2015. Com o amparo de modelos de aprendizagem supervisionada, buscou-se por variáveis "*signal*", ou seja, variáveis que possuem informação explicativa acerca da variável dependente do modelo. Sendo assim, a primeira etapa do exercício foi

a criação de uma base de preditores para teste e, em segundo lugar, os resultados da primeira abordagem sofreram transformações para gerar uma base final menor e mais seleta. Como próximo passo, aplica-se um processo de seleção de variáveis empregando algoritmos de aprendizado de máquina, sendo estes MQO, LASSO e regressão em árvore. A última etapa classifica e escolhe modelos por validação cruzada.

Concluiu-se no trabalho supracitado que o aspecto preditivo mais relevante para o PIB são àqueles relacionados às interações com o mercado internacional, especificamente o volume de exportações (bens e serviços).

A análise realizada pela empresa *PricewaterhouseCoopers (PwC)* [2] tem como objetivo realizar a estimativa do PIB do Reino Unido para o segundo trimestre do ano de 2017 como forma de descrever o estado da economia britânica e sua importância. Para a realização de tal proposta, elaborou-se um modelo de *nowcasting* utilizando aprendizado de máquina.

O artigo pontua que informações preliminares acerca da economia são disponibilizadas com um atraso de 4 a 7 meses, sofrendo constantes revisões posteriores. Este cenário é prejudicial para empresas e mercados financeiros que dependem de informações recentes e atualizadas para tomadas de decisão. Com o intuito de caracterizar preliminarmente os dados dos dias mais atuais, é proposto um modelo de estimativa de PIB utilizando uma técnica supervisionada conhecida como regularização "*Elastic Net*", método este caracterizado pelo uso de fatores de penalização como forma de prevenção de *overfitting* e permite que variáveis de alta similaridade sejam utilizadas.

As instituições que tem desenvolvido modelos preditivos utilizando técnicas de *nowcasting* são, em sua grande maioria, bancos centrais. O Banco da Inglaterra destaca-se neste ramo de pesquisa e utiliza como fonte de dados as variáveis de diferentes indústrias como os setores de vendas, serviços e manufatura; em segundo lugar, o Banco Europeu tem produzido uma série de trabalhos relacionados a este tema em específico como, por exemplo, o texto de *Giannone, Modugno e Reichlin* [1] que serviu de fundamentação para todos os trabalhos citados anteriormente. O artigo elaborado pela PwC mostra que a empregabilidade desta nova metodologia é alta, sendo capaz de diminuir pela metade o erro médio quadrático gerado por outras metodologias.

O artigo conclui que indicadores de serviços e produção industrial são os mais fortemente associados às variações do PIB britânico. Por fim, constatam que em um cenário onde a economia inglesa não se encontra com altos índices de expansão, é não só útil como necessário a utilização de um método capaz de descrever o contexto do país através de dados recentes e o modelo empregado na predição realizada apresentou resultados positivos e muito próximos das estimativas de órgãos oficiais do governo.

2.4 SÍNTESE

Neste capítulo conclui-se que o PIB é "o preço de mercado de todos os bens e serviços finais produzidos numa economia em um dado período de tempo" e sua função primordial é mensurar o desempenho econômico de um país. O produto pode ser calculado sob diferentes óticas e os valores agregados incluso por outros setores produtivos devem ser considerados para seu cômputo.

Conforme mencionado, a previsão do PIB de um país possui um atraso associado. Tal atraso existe em decorrência de fatores como a diferença temporal dos dados, falta de informações e sucessivas revisões. Sendo assim, a elaboração de uma metodologia conhecida como *Nowcasting* envolve a estimativa de variáveis para o período corrente de tempo antecipando a publicação de dados oficiais. Conhecer os valores correntes de indicadores macroeconômicos como o crescimento do PIB é uma estatística importante e que descreve o contexto de um país, fornecendo uma poderosa ferramenta ao auxílio da tomada de decisão de governantes e empresas.

Por último, mas não menos importante, os trabalhos utilizados como fundamentação teórica desta monografia apresentaram resultados semelhantes e complementares, auxiliando na escolha das variáveis de maior significância preditiva e reafirmando os principais desafios enfrentados. Ademais, os resultados obtidos foram satisfatórios em sua grande maioria e apresentaram as técnicas de *nowcasting* como solução para predição de dados econômicos recentes.

3 REGRESSÃO PELO MÍNIMO ERRO QUADRÁTICO

Neste capítulo serão explanadas duas técnicas tradicionais de regressão, linear e não-linear, para o uso subsequente na predição do PIB. Os conceitos teóricos abordados orientam-se por aplicações de algoritmos de Aprendizado de Máquina.

3.1 APRENDIZADO DE MÁQUINA

A teoria de Aprendizado de Máquina (em inglês, *Machine Learning*) é multidisciplinar e engloba diversas áreas do conhecimento como Matemática, Estatística, Probabilidade, Economia, Engenharia, Neurociência, entre outras. Em virtude de um contexto tão diversificado e vasto, sua aplicação é ampla e suas definições, por vezes, podem incorporar as de outros campos de estudo [15].

Um algoritmo é uma sequência de instruções que transforma um conjunto de entradas em saídas. O obstáculo se apresenta quando para determinadas tarefas não há uma clara evidência de um algoritmo aplicável. Ainda assim, em alguns casos há um conjunto de dados associados ao problema e desejar-se-ia que um computador fosse capaz de extrair o algoritmo necessário. De tal maneira, acredita-se que há um processo que explica os dados observados e os transforma em informação, ainda que este não seja completamente explícito e baseando-se nesta hipótese, é possível criar uma regra que aproxime e generalize os padrões encontrados [16].

Determinados programas são complexos de serem desenvolvidos e podem tornar-se projetos inviáveis porque não há como encontrar uma solução facilmente. É desejável que os *softwares* possuam a capacidade de aprender, pois os envolvidos não poderiam prever todos os possíveis resultados de um problema. *Machine Learning* (ML) almeja solucionar os desafios descritos e pode ser definido como um método de síntese que busca compreender e automatizar, através de algoritmos, as regras formais computáveis que lidem com incertezas e aprendam a partir de dados e experiências prévias; de forma geral, um agente está aprendendo se possui a capacidade de melhorar seu desempenho em tarefas futuras após observar eventos passados [15].

Uma definição formal para ML foi apresentada por *Mitchell* [17] como:

"A capacidade de um programa de computador aprender com a experiência (E) relacionada a alguma classe de tarefas (T), baseada em uma medida de desempenho (P). Dessa forma, o desempenho em tarefas (T), quando medido por (P), melhora com a experiência (E)"

O contexto de ML possui objetivos definidos e medidas de desempenho associadas que dependem do domínio do problema. Ainda que haja especificidade do caso tratado, determinados

conjuntos podem ser descritos sob diferentes paradigmas de aprendizagem. Os principais paradigmas existentes são classificados como aprendizado supervisionado, aprendizado por reforço e aprendizado não-supervisionado [15].

3.1.1 Aprendizado supervisionado

O aprendizado supervisionado é aquele em que há um "agente" para supervisionar o processo e avaliar se os resultados obtidos estão próximos ou não dos desejados. Há o julgamento constante do modelo, emitindo respostas que ditam como ele deve se comportar a cada iteração de acordo com os dados de entrada fornecidos.

Os algoritmos de treinamento para classificação e regressão são exemplos de aprendizado supervisionado.

3.1.2 Aprendizado por reforço

O aprendizado por reforço interage com os dados e não possui um agente supervisor, ao invés disso, há o papel de um "agente" que julga se o resultado obtido foi o desejado ou não. Desta forma, este tipo de aprendizado é feito através de um índice de desempenho/recompensa que indica ao algoritmo se ele está próximo do comportamento esperado.

Os algoritmos de treinamento para esta forma de aprendizado podem ser vistos em jogos eletrônicos e problemas de navegação robótica.

3.1.3 Aprendizado não-supervisionado

No aprendizado não-supervisionado, o modelo aprende padrões de acordo com os dados de entrada fornecidos sem que haja qualquer forma de julgamento ou valor alvo por parte de um agente. A máquina se auto-ajusta através da identificação de regularidade estatísticas para alcançar o resultado desejado.

Algoritmos de agrupamento com aplicações em sistemas de recomendação ou segmentação de clientes são exemplos para esta forma de aprendizado.

3.2 REGRESSÃO LINEAR

Os modelos de regressão linear foram desenvolvidos por áreas da matemática e estatística com intuits preditivos, são bem estabelecido na literatura e são utilizados como referência na utilização de técnicas mais robustas. De acordo com *Gujarati e Porter* [18], pode-se utilizar a seguinte definição:

"A análise de regressão diz respeito ao estudo da dependência de uma variável, a variável dependente, em relação a uma ou mais variáveis, as variáveis explanatórias, visando estimar e/ou prever o valor médio (da população) da primeira em termos dos valores conhecidos ou fixados (em amostragens repetidas) das segundas."

Na análise de regressão, a dependência estatística é o ponto central de um problema, ou seja, lida-se com variáveis aleatórias, isto é, variáveis probabilísticas que possuem uma incerteza que não pode ser totalmente explicada por modelos determinísticos. A análise de regressão trata a dependência de uma variável em relação a outras e não implica em causalidade. É importante explicar que uma análise de correlação mede o grau de associação linear entre duas variáveis e está relacionada à análise de regressão, porém, diferem conceitualmente uma vez que a regressão almeja estimar o valor de uma variável com base nos valores de outras variáveis [18].

Este trabalho fará uso de séries temporais como fontes de dados de entrada dos modelos. Uma série temporal é um conjunto de valores que uma variável assume em diferentes períodos de tempo, podendo variar entre intervalos diários, semanais, mensais, trimestrais e anuais. A maioria dos estudos pressupõe que as séries são estacionárias no sentido amplo, ou seja, que a média e variância não variam ao longo do tempo [18].

Apesar da abundância de dados, a qualidade é questionável, pois erros de medição são existentes e decorrentes de aproximações e arredondamentos, além disso, os métodos de amostragem para a obtenção de dados variam frequentemente, tornando difícil a comparação amostral. De forma mais específica ao contexto deste trabalho, dados econômicos são apresentados a um nível muito agregado, por exemplo, a sumarização de índices em um único indicador a partir de outros pouco descreve unidades individuais específicas. Como consequência, deve-se ter em mente que a qualidade dos resultados obtidos é um reflexo dos dados de entrada fornecidos e a causa de possíveis insatisfações não necessariamente está associada ao modelo, mas sim à qualidade dos dados.

Uma regressão linear pode ser descrita como uma função $f(x_i)$ que representa um regressor x_i e descreve que o valor esperado de Y , dado X , possui uma relação funcional com x_i . Ou seja, a resposta média de Y varia com X conforme a seguinte relação de esperança:

$$E(Y|X) = f(x_i) \quad (3.1)$$

O termo linear descreve que uma regressão deve ser geometricamente representada por uma reta. De tal maneira, a taxa de variação de Y em relação a X independe de X , ou seja, $\frac{dY}{dX}$ é constante. Uma outra forma de descrever a linearidade pode ser encontrada na linearidade dos próprios parâmetros. Permite-se que sejam aplicadas transformações não lineares nos regressores, preservando a função como uma combinação linear destes.

A análise de dependência de uma variável em relação a um único regressor é conhecida como análise de regressão simples ou univariada. Por outro lado, se a dependência de uma variável for

em relação a mais de um regressor, é conhecida como análise de regressão múltipla ou multivariada.

3.2.1 Regressão Univariada

Uma regressão linear univariada representa uma função linear de primeiro grau para um dado vetor de amostras de entrada $\mathbf{x}^T = (x_1, x_2, \dots, x_N)$, podendo ser descrita pela forma

$$h_w(x_i) = w_1 x_i + w_0 \quad (3.2)$$

onde os coeficientes do vetor \mathbf{w}^T são denominados pesos da hipótese de aproximação $h_w(\mathbf{x})$. O objetivo do processo de regressão é encontrar a hipótese que melhor se aproxima dos dados de modo que os valores de w_1 e w_0 minimizem o erro associado. Tradicionalmente, utiliza-se o critério do *mínimo erro quadrático* (MEQ) como indicador de desempenho:

$$\mathbf{w}^* = \underset{w_0, w_1}{\operatorname{argmin}} \sum_{i=1}^N [y_i - (w_0 + w_1 x_i)]^2 \quad (3.3)$$

onde $\underset{w_0, w_1}{\operatorname{argmin}}$ significa "o argumento mínimo de w_0 e w_1 " e y_i representa o valor desejado da função aproximada [15, 19]. A figura 3.1 ilustra um exemplo de aproximação.

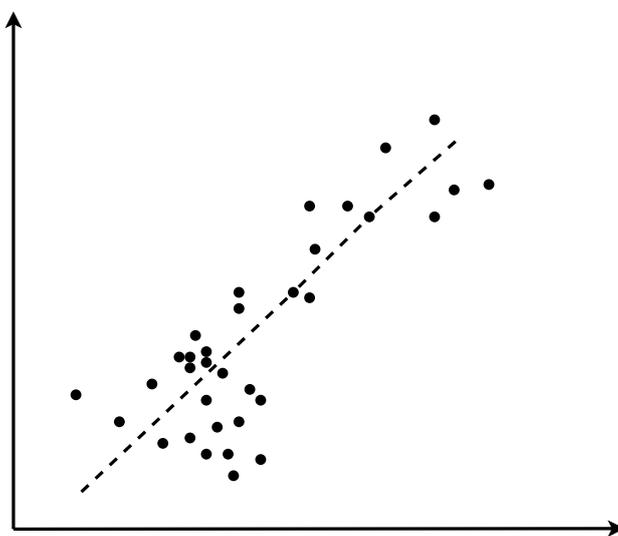


Figura 3.1: Exemplo de modelo linear univariado

A regressão linear necessita de ajuste dos pesos e para a regressão univariada é possível verificar que seu espaço é definido por w_0 e w_1 , desta maneira, sob este contexto, configura-se um problema de otimização que pode utilizar o método de gradientes descendentes a partir de um ponto inicial arbitrário no plano (w_0, w_1) e iterativamente buscará por pontos mais baixos até que haja convergência para um resultado minimizante. O cálculo do gradiente pode ser efetuado ao se tomar as derivadas parciais da função custo $J(w)$ com respeito a w_0 e w_1 .

$$\frac{\partial J(w)}{\partial w_i} = \frac{\partial}{\partial w_i} [y_i - (w_0 + w_1 x_i)]^2 \quad (3.4)$$

Os algoritmos de otimização que fazem uso de gradiente descendente dependem de uma variável conhecida como "taxa de aprendizagem" ou também conhecida como "passo" (α) que definirá a taxa com que a busca pelo valor mínimo deve ocorrer. De maneira geral, define-se o algoritmo pela fórmula:

$$w_i \leftarrow w_i - \alpha \frac{\partial J(w)}{\partial w_i} \quad (3.5)$$

Ainda que o modelo explanado possua bom desempenho, é de se esperar que haja limitações conforme o problema torna-se mais complexo e mais indicadores sejam necessários para auxiliar na descrição do domínio. Sendo assim, o modelo de múltiplas variáveis permite expandir o caso univariado.

3.2.2 Regressão Multivariada

Assumiu-se até aqui que apenas uma variável de entrada afeta uma única saída. No entanto, esta presunção pode ser inviável na prática se existirem outras variáveis que também podem afetar o comportamento do modelo. Dito isso, a regressão linear multivariada expande o espaço dos pesos do modelo apresentado anteriormente e exemplifica um cenário de entrada com x_j elementos de um vetor de m variáveis [15, 19].

É importante frisar que a ausência de colinearidade é necessária e significa que nenhum dos regressores pode ser expresso como uma combinação linear dos demais, ou seja, eles não podem ser linearmente dependentes pois o modelo pode incorrer do erro de utilizar múltiplas variáveis que na verdade se comportam como apenas uma. Sendo assim, não há como relatar a importância independente de um regressor específico e as múltiplas variáveis utilizadas não influenciam no aumento ou redução da perda [18].

A fórmula da hipótese pode ser expressa como:

$$h_{sw}(\mathbf{x}_j) = w_0 + w_1 x_{j,1} + \dots + w_m x_{j,m} = w_0 + \sum_{i=1}^m w_i x_{j,i} \quad (3.6)$$

onde o termo w_0 diferencia-se dos demais por não depender de \mathbf{x} , este peso é conhecido como *interceptor* ou *bias* (b) e para corretude da representação matricial utilizada, inclui-se nos dados de entrada $x_{j,0}$ o vetor unitário [15].

Sendo \mathbf{y} o vetor de saída e \mathbf{X} a matriz dos dados de entrada, é possível encontrar analiticamente os valores de \mathbf{w} para a regressão linear multivariada através da matriz pseudo-inversa

generalizada de *Moore-Penrose*¹, desta forma, a solução que minimiza o erro quadrático pode ser expressa como:

$$\mathbf{w}^* = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (3.7)$$

Um modelo preditivo está sujeito a erros e as fontes mais comuns são *bias* e *variância*. O conceito de *bias* pode ser explanado como a diferença média entre o valor predito e o valor real, ou seja, mensura-se quão distante o resultado da predição executada se encontra do valor correto alvo. Em segundo lugar, conceitua-se *variância* como a variabilidade de uma predição para um dado valor de entrada, em outras palavras, quanto o valor previsto varia de acordo com múltiplas execuções do modelo. Se o valor de *bias* é muito alto, aconselha-se aumentar a complexidade do modelo; por outro lado, se a *variância* é muito alta, deve-se diminuir a complexidade deste [19]. A figura 3.2 ilustra os tipos de ajustes apresentados neste parágrafo.

Em ML, por vezes ocorre o *tradeoff* entre os níveis de complexidade do modelo. Ainda assim, não é possível se ter um modelo de baixa *variância* e alta acurácia. De tal maneira, é comum que haja as situações de subajuste (do inglês, *underfitting*), onde a hipótese é incapaz de capturar os padrões apresentados pelos dados e apresentam altos valores de *bias*, e situações de sobreajuste (do inglês, *overfitting*), onde a aproximação dos valores captura as distorções dos dados e apresentam alta *variância*. Se há o aumento da complexidade, o erro preditivo diminui mas a entrada de novos dados pode gerar resultados ruins, no entanto, se a complexidade é baixa demais, a hipótese não descreve o contexto apresentado. Sendo assim, busca-se pelo modelo que melhor generaliza os dados e apresenta um balanço entre as duas situações discutidas [16].

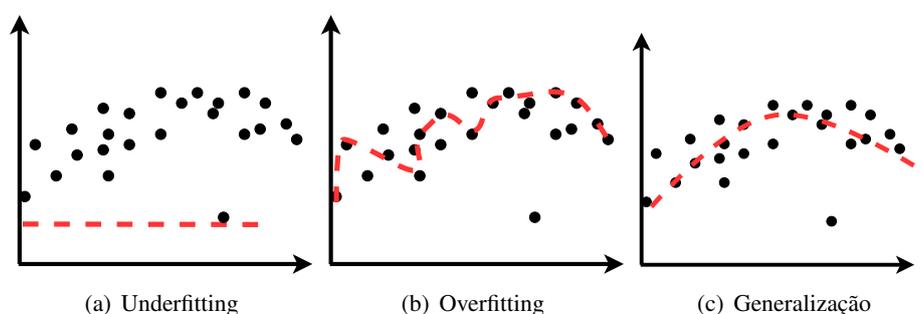


Figura 3.2: Exemplos de ajuste

Conforme explicado por *Gujarati e Porter* [18], ainda que o modelo de regressão linear consiga obter bom desempenho para determinados problemas, a inclusão de mais variáveis não corresponde necessariamente a um aumento de acurácia. O comportamento da aproximação pode ser incompleto devido a fatores como, por exemplo, o desconhecimento da relação entre os regressores e o regressando. Não obstante, conforme discutido na seção 2.2.1 do Capítulo 2, a indisponibilidade de dados e a influência de certos conjuntos de variáveis pode ser tão pequeno

¹A matriz pseudo-inversa de *Moore-Penrose* é definida pela expressão $(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$ e o seu determinante deve ser diferente de zero para que haja uma solução fechada.

que não seja interessante incluí-los no modelo.

Há uma aleatoriedade intrínseca a todos os regressores e, desta forma, de acordo com o princípio da navalha de Occam², o ideal seria formular um modelo simples que explique parte do comportamento de Y utilizando variáveis de maior peso explanatório.

3.3 MÁQUINAS DE APRENDIZADO EXTREMO

Ainda que modelos de regressão linear sejam bem estabelecidos conforme demonstrado pela literatura, para a resolução de problemas mais complexos há uma queda de acurácia conforme o número de parâmetros aumenta em virtude de um incremento proporcional da variância [10].

Conforme citado no início deste capítulo, a neurociência é um dos campos de estudo relacionados a ML. Inspirados pela composição do cérebro humano, pesquisadores do campo da Inteligência Artificial almejam por sua reprodução. O cérebro é capaz de aprender regras conforme exposição ao ambiente e, de maneira geral, uma Rede Neural Artificial (RNA) é uma máquina capaz de modelar a forma como ele executa determinadas tarefas [15].

3.3.1 Redes Neurais Artificiais

De acordo com *Haykin* [20], uma RNA pode ser definida como:

"Um processador massivo, paralelo e distribuído constituído por unidades simples de processamento naturalmente propensas a armazenar conhecimento experiencial e disponibilizá-lo para uso. Assemelha-se ao cérebro em dois fatores:

1. Conhecimento é adquirido pela rede de acordo com o ambiente através de um processo de aprendizado
2. As forças de conexões intraneurais, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido"

Uma RNA armazena unidades de informação em seus neurônios através das intensidades das ligações entre eles. O processo de aprendizado é dado pelo ajuste de tais intensidades para que a rede se adeque ao problema a ser resolvido. As RNAs podem ser lineares ou não-lineares, tolerantes a falhas e podem se adaptar de acordo com os estímulos do ambiente ao alterarem as intensidades das conexões entre seus neurônios conforme necessário.

O neurônio é a unidade primária de processamento de uma RNA. Uma rede de neurônios possui um conjunto de conexões ponderadas denominadas sinapses e cada sinapse possui uma

²O princípio da navalha de Occam postula que explicações simples para um fenômeno são melhores de se trabalhar em detrimento do aumento de complexidade baseado na expectativa de melhores resultados.

ponderação própria. De maneira geral, um sinal de entrada x_j conecta-se a um neurônio k e seu valor é multiplicado pelo peso $w_{k,j}$. Em segundo lugar, através de uma junção somadora há a combinação linear dos sinais de entrada ponderados pelos neurônios. Por fim, uma função de ativação restringe o sinal de saída, limitando sua amplitude normalizada por funções de variados tipos como, por exemplo, as da família sigmoideal.

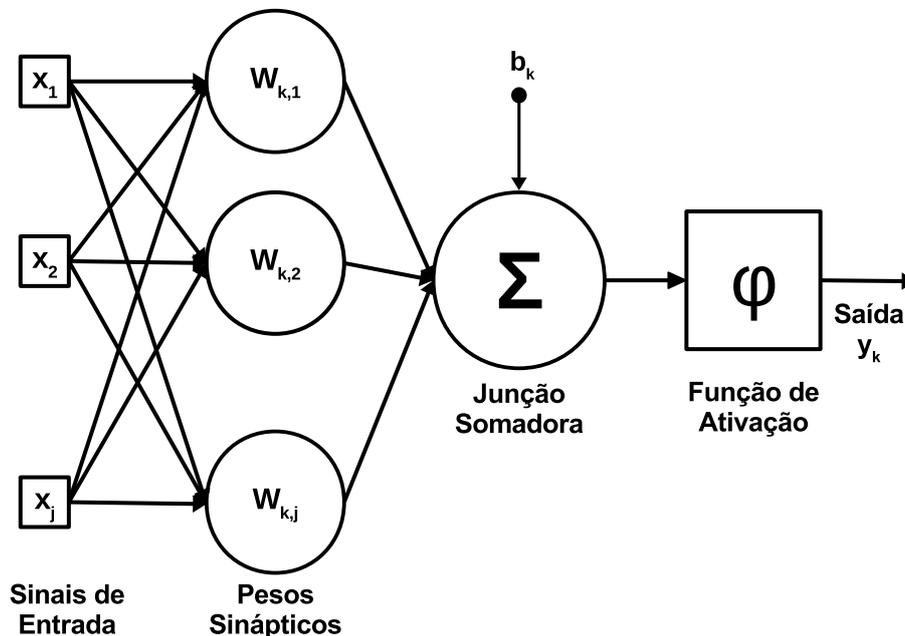


Figura 3.3: Modelo de um neurônio

O valor da junção das entradas ponderadas é definido pela equação 3.8

$$v_k = \sum_{j=0}^m w_{k,j}x_j \quad (3.8)$$

e assim como na regressão linear, o modelo neural também contém um valor de bias (b_k) com o propósito de aumentar ou diminuir o resultado da função de ativação (φ). De tal maneira, a representação matemática de um modelo neural é feita pela equação 3.9 e sua representação gráfica pode ser vista na figura 3.3.

$$y_k = \varphi\left(\sum_{j=1}^m w_{k,j}x_j + b_k\right) \quad (3.9)$$

As RNAs podem constituir diversos tipos de arquiteturas que definem composições distintas para a estrutura de conexões entre seus neurônios, organizando-as por camadas de elementos. As classificações tradicionais englobam a rede *feedforward* de uma camada, a rede *feedforward* de múltiplas camadas e a rede recorrente.

Rede *feedforward* de uma camada

Em uma rede dividida por camadas, a representação mais simples não possui realimentação (*feedforward*) e é composta por uma única camada de sinais de entrada e uma camada de sinais de saída onde ocorre o processamento. Conforme ilustrado pela figura 3.4, a rede é caracterizada por possuir uma única camada responsável por computar os dados.

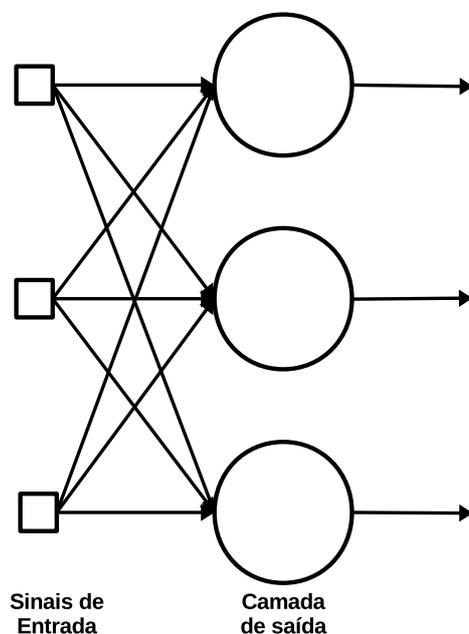


Figura 3.4: Modelo de uma rede *feedforward* de uma camada

Rede *feedforward* de múltiplas camadas

Uma rede *feedforward* de múltiplas camadas não possui realimentação e é composta por uma ou mais camadas intermediárias de processamento. As camadas de processamento também são denominadas "camadas ocultas" por não serem visíveis diretamente pela entrada ou pela saída do modelo. A figura 3.5 ilustra um exemplo para esta categoria de rede.

A inclusão de mais camadas aumenta a complexidade do algoritmo e proporciona resultados estatísticos mais globais. De maneira geral, as saídas dos neurônios de uma camada oculta são utilizadas como entradas para os neurônios da próxima camada e assim por diante. O Perceptron de Múltiplas Camadas (do inglês, Multi Layer Perceptron, MLP) é o exemplo mais conhecido de RNA que utiliza a arquitetura de múltiplas camadas.

A capacidade que a rede tem de estimar ao se deparar com novos dados de entrada que não foram previamente apresentados ao modelo durante o treino é chamada de capacidade de generalização. A configuração dos pesos sinápticos de uma rede que melhor satisfaz a aproximação realizada é aquela que possuir o menor erro associado e, de tal maneira, o treinamento de uma rede MLP é visto como um problema de otimização. O algoritmo de treinamento mais utilizado é

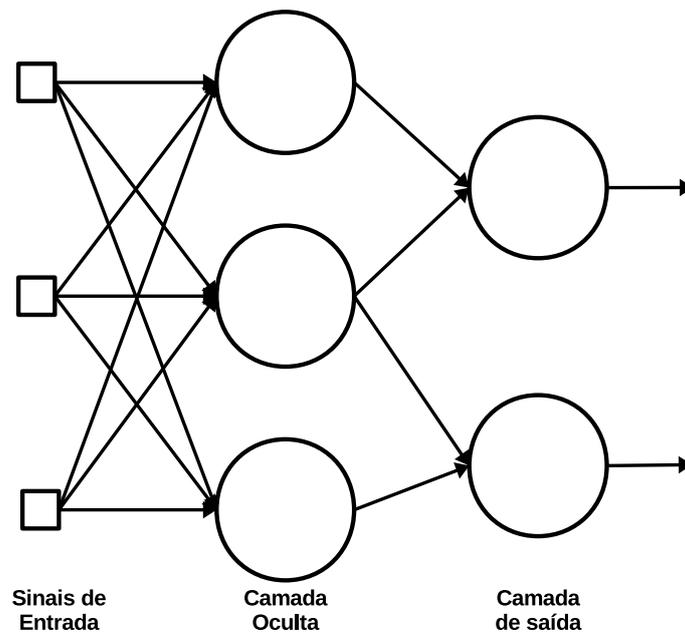


Figura 3.5: Modelo de uma rede *feedforward* de múltiplas camadas

o de gradiente descendente com retropropagação do erro (do inglês, *backpropagation*), sendo este proposto para resolver o problema de cálculo do gradiente da função custo em relação ao espaço dos pesos das camadas ocultas [21].

Quando a rede apresenta indícios de sobreajuste, não consegue atingir um bom desempenho devido ao treinamento excessivo. Uma técnica utilizada para evitar que isso aconteça é conhecida como "validação cruzada", onde há a divisão dos dados disponíveis num conjunto de treinamento e num conjunto de validação, sendo o último usado para medir a capacidade de generalização da rede [21].

Rede recorrente

Uma rede recorrente possui pelo menos um laço de realimentação, podendo assim interligar os valores da saída de uma camada às entradas de neurônios de uma camada anterior. A existência de tais laços interfere na aprendizagem e no desempenho da rede, resultando em um comportamento dinâmico não-linear. A figura 3.6 ilustra um exemplo de rede recorrente.

3.3.2 Definição para Máquinas de Aprendizado Extremo

As redes *feedforward* possuem a capacidade de aproximar funções não-lineares diretamente através dos sinais de entrada e de fornecer modelos para problemas que não poderiam ser resolvidos por modelos tradicionais. Em contrapartida, os algoritmos de aprendizado costumam ser lentos, pois a máxima generalização está associada a modelos otimizados, ou seja, os números de

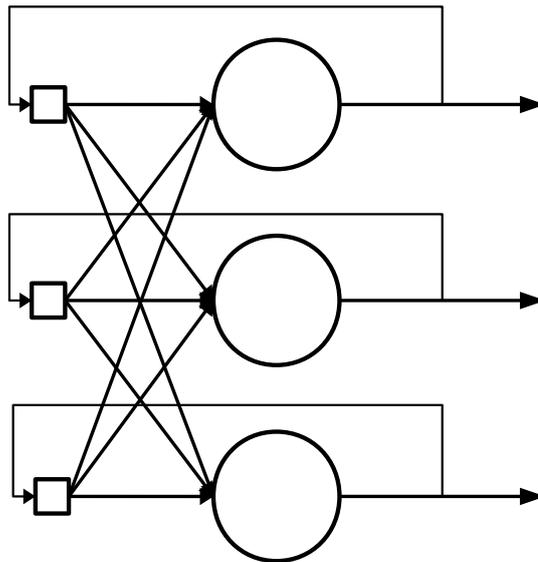


Figura 3.6: Exemplo de um modelo de uma rede recorrente

neurônios e de pesos sinápticos dependem de regularização.

A otimização de redes *feedforward* pelo uso de algoritmos baseados em gradiente descendente com retropropagação apresenta problemas como:

- Para uma baixa taxa de aprendizagem, a convergência é mais lenta, pois o número de iterações é maior. Entretanto, se a taxa for muito alta, o algoritmo tende a divergir;
- A presença de mínimos locais pode gerar soluções subótimas;
- As redes podem ser treinadas demasiadamente a ponto de sofrerem *overfitting*;
- Longo tempo de treinamento.

De acordo com o trabalho de *Huang et al.* [5, 6], uma rede *feedforward* de uma única camada oculta pode ter os pesos desta escolhidos aleatoriamente e ainda assim resultar em um modelo de capacidade de generalização universal. Os autores demonstram que as Máquinas de Aprendizado Extremo (do inglês, *Extreme Learning Machines*) tendem a possuir melhor capacidade de generalização que algoritmos baseados em gradiente descendente com retropropagação e, na maioria dos casos, apresentam menores erros de treinamento.

Uma *Extreme Learning Machine* (ELM) de L neurônios na camada oculta, N amostras, saída única, função de ativação $g(x)$ e matriz de entrada $\mathbf{X}^T = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_N]^T$ pode ser definida pela fórmula

$$t_j = \sum_{k=1}^L \beta_k g(\mathbf{w}_k \cdot \mathbf{x}_j + b_k), j = 1, 2, \dots, N, \quad (3.10)$$

onde $\mathbf{w}_k = [w_{k,1}, w_{k,2}, \dots, w_{k,j}]^T$ é o vetor de pesos sinápticos que conecta o k-ésimo neurônio da camada oculta ao j-ésimo vetor de entrada, β_k é o peso da conexão do k-ésimo neurônio com a camada de saída e b_k é o valor de *bias* associado ao k-ésimo neurônio.

A equação 3.10 pode ser escrita também como

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T}, \quad (3.11)$$

onde \mathbf{H} é a matriz de saída da camada oculta

$$\mathbf{H} = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \dots & g(\mathbf{w}_L \cdot \mathbf{x}_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \dots & g(\mathbf{w}_L \cdot \mathbf{x}_N + b_L) \end{bmatrix}_{N \times L}, \quad (3.12)$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_L \end{bmatrix}_{L \times 1} \quad (3.13)$$

e \mathbf{T} é a matriz alvo de treinamento

$$\mathbf{T} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}_{N \times 1}. \quad (3.14)$$

De tal maneira, diferentemente dos modelos tradicionais de RNAs que requerem ajustes nos pesos de entrada e nos valores de *bias*, estes parâmetros podem ser aleatoriamente distribuídos de forma que a matriz \mathbf{H} permaneça inalterada após o ajuste. Sendo assim, para valores fixos de \mathbf{w} e b , resolve-se o sistema linear $\mathbf{H}\boldsymbol{\beta} = \mathbf{T}$ ao se encontrar a solução de mínimos quadrados $\boldsymbol{\beta} = \mathbf{H}^\dagger \mathbf{T}$ onde \mathbf{H}^\dagger é a matriz pseudoinversa de *Moore-Penrose*, assim como visto na seção 3.2.2 para a resolução da regressão linear multivariada.

O algoritmo para uma ELM clássica proposto por *Huang et al.* [5, 6] consiste nos seguintes passos:

1. Atribuir aleatoriamente valores para os pesos \mathbf{w}_k e *bias* b_k ;
2. Calcular a matriz \mathbf{H} de saída da camada oculta;
3. Calcular a matriz de pesos $\boldsymbol{\beta}$.

Regularização

Diferentemente da formulação inicial de uma ELM, os autores propõem um sistema de regularização que retrabalha os pesos de modo a penalizar valores de grande magnitude. O cálculo de β resume-se à resolução de um problema de regularização de mínimos quadrados, isto é, um mecanismo de controle da complexidade do modelo com o intuito de melhorar a capacidade de generalização. Desta forma, a complexidade poderia se ajustar através da norma dos pesos da camada de saída:

$$\beta^* = \operatorname{argmin}_{\beta} \frac{1}{2} \|\beta\|^2 + \frac{C_i}{2} \sum_{i=1}^L \|t_i^T - \mathbf{H}_i^T \beta^T\|^2 \quad (3.15)$$

O problema exemplificado acima é conhecido como *ridge regression* ou "mínimos quadrados regularizados" onde C_i é um coeficiente de ponderação. Através da equação 3.15, são obtidas relações entre a quantidade de neurônios L e a quantidade de amostras N onde \mathbf{I} é a matriz identidade.

$$\beta^* = \begin{cases} (\frac{\mathbf{I}}{C_i} + \mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{T}, & \text{Se } N \leq L \\ \mathbf{H}^T (\frac{\mathbf{I}}{C_i} + \mathbf{H} \mathbf{H}^T)^{-1} \mathbf{T}, & \text{Se } N > L \end{cases} \quad (3.16)$$

3.4 SÍNTESE

Neste capítulo são apresentados conceitos introdutórios sobre aprendizado de máquina, sua definição e os tipos de aprendizado. Foram abordadas as técnicas de Regressão Linear assim como fundamentos de Redes Neurais Artificiais.

A análise de regressão estuda a relação entre variáveis dependentes e variáveis explanatórias. Tal relação é representada por um modelo matemático linear, podendo este ser univariado ou multivariado a depender da quantidade de regressores.

As RNAs são utilizadas na solução de diversos tipos de problemas de aproximação por serem tolerantes a falhas e adaptativas, tendo assim uma boa capacidade de generalização. Uma vertente recentemente explorada a partir do modelo tradicional de uma RNA é conhecida como Máquina de Aprendizado Extremo.

As Máquinas de Aprendizado Extremo são redes que possuem apenas uma camada oculta e possuem os pesos sinápticos de entrada atribuídos aleatoriamente. Tendem a possuir melhor performance, menor erro de treinamento e menor norma dos pesos sinápticos em relação às redes tradicionais por não fazerem uso de algoritmos iterativos como o gradiente descendente com retropropagação.

Com base nesta fundamentação teórica, o próximo capítulo apresentará as metodologias utilizadas na realização deste trabalho, demonstrando o processo experimental para predição do PIB brasileiro através de *nowcasting*.

4 ESTUDO DE CASO E RESULTADOS

Conforme elucidado no Capítulo 2, a estimativa de indicadores macroeconômicos é um dos principais tópicos de estudo dos bancos centrais, que utilizam diversas técnicas de *nowcasting* para descrever o estado da economia. As RNAs possuem um amplo leque de aplicações como, por exemplo, problemas de regressão para aproximação de funções, desta forma, podem ser exploradas como alternativa aos métodos já utilizados atualmente. Sendo assim, este trabalho propõe a aplicação de Máquinas de Aprendizado Extremo para previsão do PIB trimestral brasileiro. Neste capítulo serão apresentados os resultados experimentais obtidos através dos algoritmos implementados, assim como os ajustes dos parâmetros dos modelos utilizados.

Os algoritmos estudados por este trabalho foram implementados na linguagem de programação Python [22], versão 3.6.5, amparado pelas bibliotecas Numpy [23], versão 1.14.5; Pandas [24], versão 0.23.3; e Matplotlib [25], versão 2.2.2. O desenvolvimento foi feito utilizando o editor Visual Studio Code [26], versão 1.26.0. Os algoritmos foram executados em um computador com processador Intel Core i5 4440 de frequência 3.10GHz, com 8GB de memória RAM e sistema operacional Windows 7 Professional 64-bit SP1.

4.1 CONJUNTO DE DADOS E SELEÇÃO DAS VARIÁVEIS

De acordo com a seção 2.2 do Capítulo 2, um país possui detalhes intrínsecos à sua economia interna que influenciam os modelos preditivos. No caso do Brasil, este possui disponibilidade mais restrita de dados, dependência de fluxos de capitais internacionais e períodos variados de recessão entre 1980 e 1999 ocasionados por uma série de mudanças políticas ao longo dos anos como tentativa de reestabelecimento da economia nacional. Este trabalho escolheu especificamente o período pós plano Real para não ter de lidar com flutuações ocasionadas pela mudança de moeda e também por se tratar de um período de aparente estabilização econômica.

Tendo como ano base 1995, os dados foram selecionados e coletados através do Sistema Gerenciador de Séries Temporais (SGS) do Banco Central [27], da base de dados aberta fornecida pelo IPEA Data [28], e do sistema de séries estatísticas e históricas do IBGE [9]. Os dados são de domínio público e podem ser acessados por qualquer cidadão brasileiro.

Os índices foram escolhidos por serem significativos em termos de economia nacional. A amostra deste trabalho é composta por 12 índices econômicos com dados observados entre o período do quarto trimestre de 1995 até o quarto trimestre de 2017, totalizando 89 registros obtidos. O período de tempo foi delimitado segundo a disponibilidade dos dados das variáveis utilizadas, cuja escolha foi feita de acordo com dois fatores. O primeiro fator diz respeito à quantidade de citações daquela variável entre os autores utilizados como base teórica para este

trabalho; o segundo fator está relacionado aos valores dos pesos que melhor minimizam o erro absoluto médio (EAM) de aproximações realizadas pela regressão linear multivariada.

A tabela 4.1 contém dados advindos do Sistema de Contas Nacionais Trimestrais e apresenta os índices de volume com base em 1995 para o PIB a preços de mercado, compreendendo informações econômico-financeiras da renda bruta de cada setor [9]. Ademais, contém também o índice BOVESPA, sendo este um indicador de desempenho médio dos ativos mais negociados e representativos do mercado de ações brasileiro [29].

Tabela 4.1: Indicadores selecionados

Grupo	Indicador	Periodicidade	Fonte	Significância
Produção industrial	Indústria geral	Trimestral	IBGE	Forte
Produção industrial	Indústria de transformação	Trimestral	IBGE	Forte
Produção industrial	Indústrias extrativas	Trimestral	IBGE	Forte
Produção industrial	Indústria de construção	Trimestral	IBGE	Média
Serviços	Serviços (Total)	Trimestral	IBGE	Média
Mercado internacional	Exportação (Total)	Trimestral	IBGE	Forte
Mercado internacional	Importação (Total)	Trimestral	IBGE	Média
Consumo	Consumo do governo	Trimestral	IBGE	Fraca
Consumo	Consumo das famílias	Trimestral	IBGE	Fraca
Mercado Financeiro	Bovespa	Mensal	BM e FBOVESPA	Forte
Produção agropecuária	Agropecuária (Total)	Trimestral	IBGE	Fraca
Impostos	Arrecadação de impostos	Trimestral	IBGE	Fraca

Os indicadores foram divididos entre três categorias de significância de acordo com a quantidade de referências:

- Forte: citado por três ou mais autores;
- Média: citado por um ou dois autores;
- Fraca: utilizado no processo experimental de um ou mais autores mesmo sem apresentar resultados de relação direta com a variação do PIB.

Os indicadores da categoria de produção industrial foram amplamente citados como significantes por *Kagohara* [11], *Zuanazzi e Ziegelmann* [13], *PwC* [2] e *Mattos* [14]; índices da categoria de mercado financeiro foram fortemente referenciados por *Zuanazzi e Ziegelmann* [13], *Kagohara* [11] e *Vasconcelos* [10], destacando-se principalmente o índice Bovespa; o setor de serviços mostrou-se presente durante todos os processos experimentais apesar de ser considerado de baixo impacto pela *PwC* [2] e *Vasconcelos* [10]; variáveis de relação com o mercado internacional apresentaram relações evidentes com as variações do PIB, destacando-se o volume de exportações conforme dito por *Kagohara* [11], *Zuanazzi e Ziegelmann* [13], *PwC* [2] e *Vasconcelos* [10], não obstante, com menor impacto, as importações foram apenas citadas por *Vasconcelos* [10]; os consumos do governo e também das famílias são considerados no cálculo do PIB conforme

visto na seção 2.1.1 do Capítulo 2 e fizeram-se presentes durante os processos de seleção de variáveis dos autores supracitados mas não apresentaram relação direta com o PIB; demais variáveis que englobem grupos de impostos ou produção agropecuária não foram amplamente citadas mas, ainda assim, foram escolhidas com o intuito de se avaliar a influência destes regressores no comportamento do algoritmo implementado.

O índice BOVESPA necessitou de duas modificações para que pudesse ser utilizado. O primeiro ajuste foi feito sobre a periodicidade, foi necessário calcular uma média aritmética simples sobre o fechamento mensal a cada três meses como forma de conversão da periodicidade. O segundo ajuste foi feito sobre a mudança no cálculo do indicador ao longo dos anos, em relação a 1994, o valor atual está defasado em 10 vezes; em relação a 1997, o valor atual também está defasado em 10 vezes; deste modo, os valores entre 1995 e 1997 foram multiplicados por 100.

Como passo final no tratamento dos dados, aplicou-se uma normalização onde cada amostra da variável fora dividida pelo seu desvio padrão com o objetivo de se padronizar a escala das variáveis.

Os autores utilizados como base para este trabalho também citaram outros indicadores que poderiam ser incluídos, todavia, durante o processo de prospecção foram encontradas divergências quanto às datas de início e fim, em razão da variabilidade de metodologias aplicadas; diferentes periodicidades; por fim, indicadores derivados de outros indicadores, que indicavam correlação.

4.1.1 Validação Cruzada

Ainda que uma hipótese possua uma baixa proporção de erros, não significa que a generalização é adequada e há a possibilidade de sobreajuste, sendo assim, faz-se necessário testá-la frente a dados ainda não apresentados para que seja selecionada a melhor configuração de parâmetros do modelo aplicado [15].

O método de validação cruzada (do inglês, *cross validation*) é utilizado quando se deseja selecionar um modelo de melhor generalização do modelo. O objetivo é selecionar a hipótese que melhor minimiza o erro de generalização resultante quando apresentada a um conjunto de dados de teste [20].

Os passos utilizados pela técnica de validação cruzada são [21]:

1. Dividir os dados em conjunto de treinamento e conjunto de validação;
2. Treinar o modelo utilizando os dados do conjunto de treinamento;
3. Utilizar o modelo treinado para realizar estimativas com os dados pertencentes ao conjunto de validação;
4. Calcular o erro das estimativas feitas para o conjunto de validação, estimando a capacidade de generalização do modelo.

Uma variante desta técnica é chamada de *k-fold cross-validation*. Aplicada principalmente quando a quantidade de dados disponíveis é pequena, o intuito é de utilizá-los tanto como conjunto de treinamento como de validação [15]. Os passos abordados são:

1. Embaralhar os dados aleatoriamente;
2. Dividir os dados em K subconjuntos iguais;
3. Para cada grupo único de 1 até K, selecioná-lo como conjunto de validação;
4. Selecionar o restante dos K-1 grupos como conjunto de treinamento;
5. Treinar o modelo com o conjunto de treinamento;
6. Aplicar o modelo ao conjunto de validação;
7. Calcular o erro das estimativas feitas para o conjunto de validação;
8. Calcular a média aritmética dos K erros obtidos.

O embaralhamento dos dados é necessário também para incluir períodos de diferentes contextos econômicos (recessão, expansão). Apesar de o conjunto selecionado possuir um período bem delimitado, as oscilações dos regressores podem indicar subperíodos de recessão que não seriam vistos pelo modelo durante o processo de treinamento.

4.2 REGRESSÃO LINEAR MULTIVARIADA

O processo experimental foi dividido em duas etapas e executado utilizando uma base de dados econômicos e financeiros compilados de acordo com a tabela 4.1.

A primeira etapa experimental consiste em, a partir de uma base conhecida, definir o conjunto dos dados que resulta no menor erro absoluto médio para a aproximação. Uma regressão linear multivariada foi aplicada sobre cinco conjuntos de dados contendo diferentes combinações de indicadores, objetivando avaliar o comportamento dos pesos de acordo com o método avaliativo de referência.

4.2.1 Resultados

A partir da tabela 4.2, são resumidos os impactos dos regressores selecionado de acordo com os diferentes conjuntos de dados utilizados no processo preditivo. É evidente que os índices Bovespa, da indústria geral e de arrecadação de impostos contém os maiores valores absolutos em relação aos demais.

Tabela 4.2: Valores dos pesos da regressão linear multivariada

Indicador	Valor do peso				
	Conjunto 1	Conjunto 2	Conjunto 3	Conjunto 4	Conjunto 5
Indústria geral	1.49	-0.02	3.92	3.08	0.26
Indústria de transformação	-1.53	-0.11	-0.21	-1.47	-
Indústrias extrativas	-0.14	-0.18	0.03	-	-
Indústria de construção	0.29	0.59	-	-	-
Serviços (Totais)	0.79	0.82	-5.31	-0.22	-
Exportações (Totais)	-1.01	-0.84	-0.14	-1.21	-0.97
Importações (Totais)	-2.17	-2.06	-1.06	-2.71	-1.7
Consumo do governo	0.58	-	-0.23	-	-
Consumo das famílias	-0.15	-1.35	0.60	-	-
BOVESPA	2.04	2.21	2.94	2.22	2.34
Agropecuária (Total)	1.15	-	-	0.89	-
Arrecadação de impostos	2.37	2.71	-	2.22	1.80

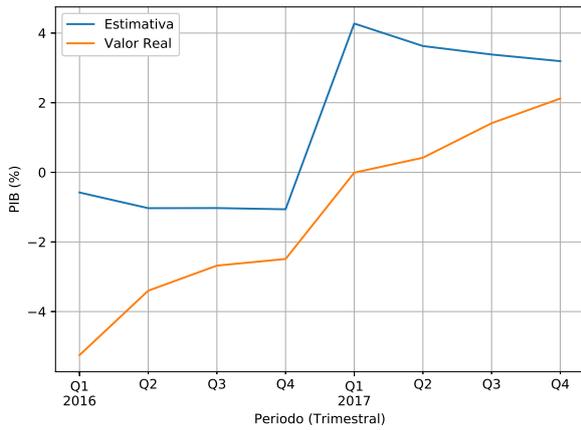
Utilizando como referência o método de validação cruzada com o valor de $K = 5$, a segunda etapa do experimento implica em utilizar o conjunto dos regressores que apresentaram melhor desempenho, fornecendo-os como dados de entrada do modelo preditivo. A regressão linear fez uso de 90% dos registros da base para treinamento, correspondendo ao período de 1994 a 2015, restando 10% dos dados para teste, correspondendo aos trimestres compreendidos entre os anos de 2016 e 2017. Uma vez que o processo de *nowcasting* visa prever os eventos mais recentes da economia, um pequeno conjunto de testes é suficiente para o experimento. As estimativas aplicadas são exibidas pelo conjunto de gráficos da figura 4.1. Pode-se notar pela tabela 4.3 que o conjunto 4 apresentou o melhor desempenho geral ao evidenciar o menor EAM.

A tabela 4.3 apresenta duas medidas de erro absoluto, sendo a primeira referente ao método de *K-Fold Cross Validation* e a segunda relativa ao EAM da estimativa realizada. Apesar de o gráfico exibido pela figura 4.1(e) aparentar melhor aproximação visual, as variáveis deste conjunto 5 demonstram-se piores segundo os valores de erro apresentados pela validação cruzada aplicada. Sendo assim, o conjunto de dados preferido para este trabalho foi o de número 4 por possuir melhor desempenho no experimento de validação cruzada. A predição realizada utilizando este conjunto específico pode ser vista na figura 4.1(d).

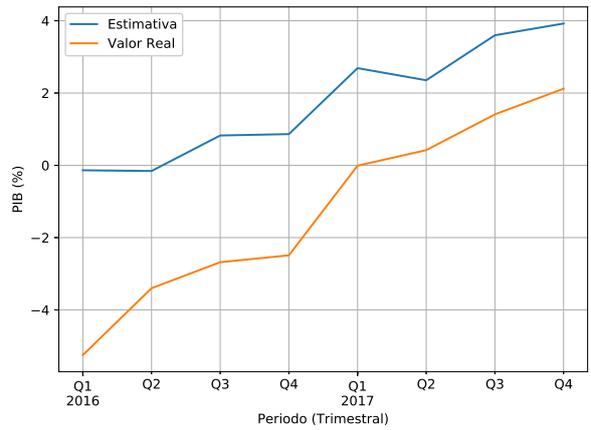
Tabela 4.3: Erro absoluto médio para cada conjunto de dados

Erro Absoluto Médio	Conjunto 1	Conjunto 2	Conjunto 3	Conjunto 4	Conjunto 5
<i>K-Fold Cross Validation</i>	1.49 ± 0.43	1.46 ± 0.18	1.73 ± 0.19	1.37 ± 0.15	1.53 ± 0.24
Estimativa	2.58	2.97	2.03	2.42	2.38

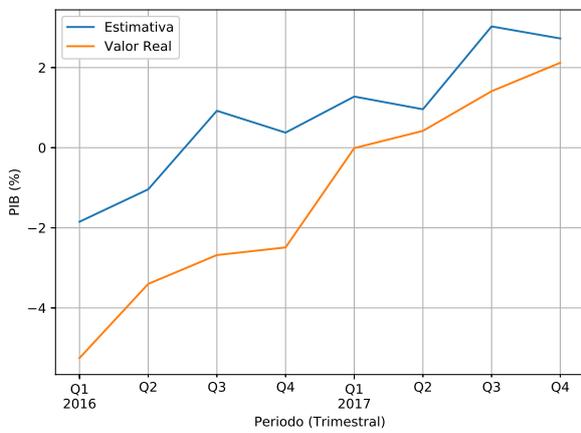
Assim como explanado em capítulos anteriores, o modelo de regressão linear é naturalmente utilizado como referência de comparação para modelos mais complexos. Este trabalho propõe a



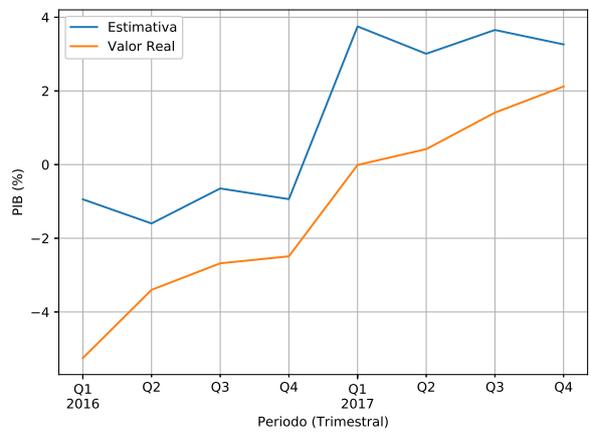
(a) Conjunto 1



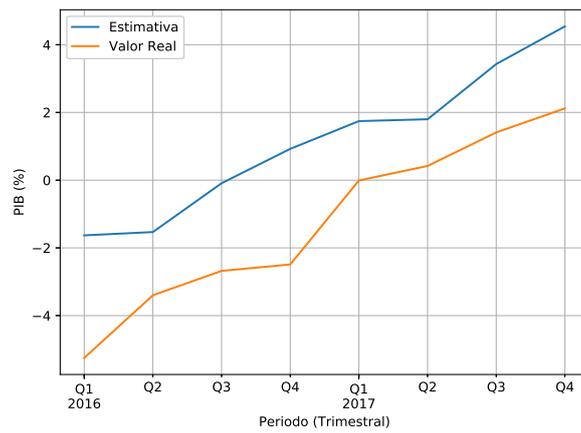
(b) Conjunto 2



(c) Conjunto 3



(d) Conjunto 4



(e) Conjunto 5

Figura 4.1: Regressões aplicadas

implementação de uma técnica de *nowcasting* que faça uso de RNAs, mais especificamente de Máquinas de Aprendizado Extremo, de tal maneira, os resultados das estimativas efetuadas com o modelo de regressão linear serão utilizados para nortear os experimentos subsequentes.

4.3 MÁQUINA DE APRENDIZADO EXTREMO

Como alternativa de regressão, este trabalho implementa uma ELM regularizada. A literatura de referência explica que tais RNAs alcançam rapidez e performance tanto em modelos preditivos como de classificação. As teorias de ELMs afirmam que os pesos sinápticos de entrada dos neurônios podem ser aleatoriamente atribuídos e os pesos de saída podem ser encontrados através da resolução de um sistema linear simples, sendo assim, a implementação deste algoritmo pode ser efetuada com baixo custo computacional. Por fim, a regressão linear multivariada foi utilizada para investigar a relação entre os regressores e as variáveis explanatórias, auxiliando este experimento na seleção das variáveis de entrada.

Para a elaboração da ELM, foi necessário ajustar a distribuição dos pesos sinápticos de entrada, a função de ativação (φ) e a quantidade de neurônios da camada oculta (L). A configuração dos pesos w foi feita aleatoriamente de acordo com uma distribuição normal de valor médio igual a 0 e desvio padrão igual a 1. Em segundo lugar, escolheu-se a função logística como função de ativação. Por fim, a quantidade de neurônios foi selecionada empiricamente, variando o valor do parâmetro de acordo com a variação do EAM.

$$\varphi(\cdot) = S(x) = \frac{1}{1 + e^{-x}} \quad (4.1)$$

O experimento foi dividido em duas etapas. A primeira consistindo na aplicação de uma ELM clássica ao conjunto de dados selecionados com o objetivo de avaliar a relação entre a quantidade de neurônios e o erro absoluto médio. A segunda etapa efetua a regularização do modelo, o foco desta fase é configurar os parâmetros L e C que melhor generalizem o problema.

4.3.1 Resultados

Ao ser aplicado o método de validação cruzada com $K = 5$ em relação à variação da quantidade de neurônios do modelo, foram obtidos os erros absolutos médios de cada treinamento realizado conforme pode ser verificado pela tabela 4.4.

É evidente a existência de uma relação entre diminuição do EAM ao se ampliar a quantidade de neurônios do modelo. Ainda assim, para que a acurácia aumente, é necessário também que a quantidade de neurônios aumente, resultando em maior tempo de processamento. Visando incrementar o desempenho do modelo da ELM, a regularização foi inserida no processo preditivo e novos ajustes foram efetuados de acordo com o valor do coeficiente de regularização.

Tabela 4.4: Relação entre número de neurônios e EAM resultante do processo de *K-Fold Cross Validation*

Quantidade de Neurônios	Erro Absoluto Médio
50	4833.89 ± 7637.72
100	2499.05 ± 2142.97
200	687.73 ± 446.28
300	29.91 ± 39.62
400	15.61 ± 14.20
500	7.92 ± 2.65
600	6.15 ± 3.30
700	3.25 ± 0.76
800	2.40 ± 0.76
900	3.19 ± 0.45
1000	2.98 ± 0.75
2000	2.29 ± 0.14
3000	1.75 ± 0.35
5000	1.95 ± 0.27

Ainda que esteja evidente a relação entre L e EAM, os valores de C podem influenciar no processo preditivo e, de tal maneira, faz-se necessário a verificação da melhor combinação de parâmetros C e L para que se reduza EAM.

De acordo com a tabela 4.5, nota-se que após a utilização da regularização, os valores tornaram-se mais estáveis, assim como também houve uma diminuição do desvio padrão. Ainda que os pesos de entrada sejam modificados de maneira aleatória, ocasionando uma maior variabilidade dos valores tabelados, é possível perceber que ao se utilizar L = 100 há maiores chances de se diminuir o erro associado. Não obstante, para C = 10 também fica evidente uma maior tendência de melhor generalização. Sendo assim, a ELM implementada escolheu o conjunto supracitado para a realização das predições deste trabalho.

Tabela 4.5: Configurações de parâmetros utilizados pela ELM

L \ C	C							
	0.001	0.01	0.1	1	10	100	1000	
50	1.63 ± 0.12	1.47 ± 0.18	1.35 ± 0.14	1.44 ± 0.18	1.56 ± 0.21	1.35 ± 0.20	1.43 ± 0.15	
100	1.48 ± 0.22	1.53 ± 0.27	1.37 ± 0.18	1.38 ± 0.19	1.35 ± 0.07	1.50 ± 0.17	1.65 ± 0.17	
200	1.45 ± 0.22	1.40 ± 0.21	1.54 ± 0.24	1.43 ± 0.36	1.50 ± 0.31	1.40 ± 0.20	1.47 ± 0.24	
300	1.55 ± 0.15	1.60 ± 0.10	1.44 ± 0.15	1.44 ± 0.20	1.33 ± 0.30	1.45 ± 0.25	1.53 ± 0.33	
400	1.48 ± 0.40	1.60 ± 0.35	1.40 ± 0.13	1.50 ± 0.36	1.48 ± 0.16	1.68 ± 0.19	1.47 ± 0.27	
500	1.48 ± 0.27	1.61 ± 0.14	1.47 ± 0.19	1.46 ± 0.16	1.41 ± 0.51	1.56 ± 0.15	1.40 ± 0.26	
600	1.53 ± 0.30	1.52 ± 0.24	1.42 ± 0.26	1.53 ± 0.25	1.58 ± 0.19	1.56 ± 0.37	1.47 ± 0.23	
700	1.79 ± 0.29	1.45 ± 0.23	1.44 ± 0.08	1.45 ± 0.23	1.60 ± 0.20	1.65 ± 0.25	1.54 ± 0.19	
800	1.58 ± 0.31	1.40 ± 0.14	1.51 ± 0.28	1.61 ± 0.20	1.45 ± 0.29	1.51 ± 0.31	1.43 ± 0.39	
900	1.57 ± 0.16	1.39 ± 0.22	1.70 ± 0.42	1.49 ± 0.16	1.34 ± 0.18	1.46 ± 0.33	1.53 ± 0.19	
1000	1.42 ± 0.17	1.52 ± 0.22	1.40 ± 0.16	1.53 ± 0.14	1.41 ± 0.32	1.61 ± 0.28	1.54 ± 0.23	

Uma vez selecionado o conjunto de parâmetros, a ELM foi treinada com dados até o ano de 2015 e testada com dados dos trimestres dos anos de 2016 e 2017. É importante notar que os pesos sinápticos da camada de entrada são gerados aleatoriamente e, deste modo, foram executadas quatro predições utilizando o mesmo conjunto de dados escolhido sob as mesmas configurações para se avaliar o comportamento da estimativa frente a variabilidade dos pesos.

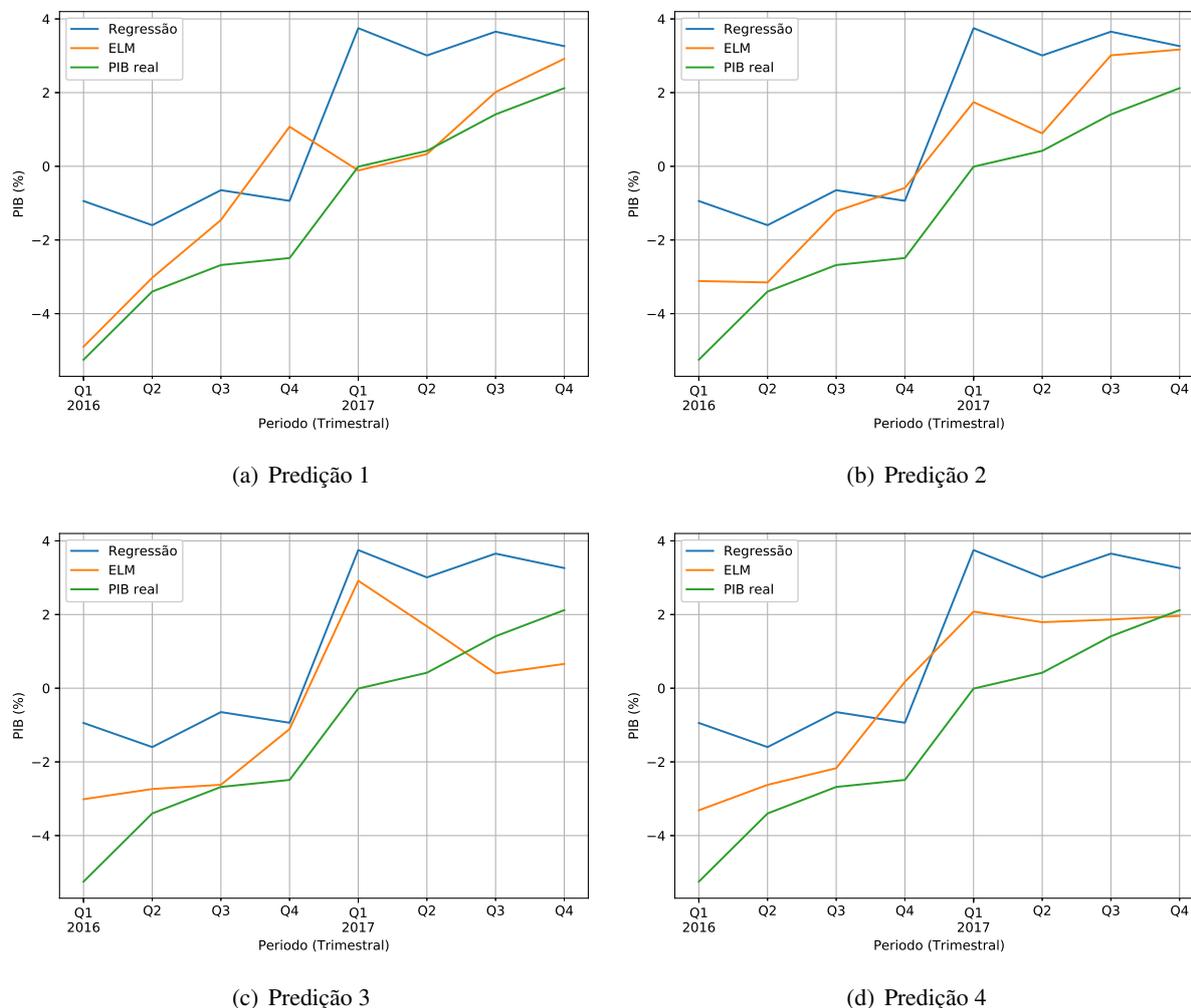


Figura 4.2: Predições realizadas com a ELM regularizada

Nota-se que a ELM foi capaz de apresentar melhor ajuste se comparada com a regressão linear multivariada. O gráfico 4.2(a) em especial apresentou um bom resultado de aproximação, acompanhando de maneira clara e evidente a tendência do crescimento do PIB.

4.3.2 Análise Comparativa

A investigação efetuada por este trabalho avalia se o comportamento que variações econômicas e financeiras apresentam por meio de determinados indicadores macroeconômicos pode servir como premissa para a estimativa do PIB brasileiro. Tendo como referência a regressão linear multivariada, deseja-se julgar se a empregabilidade de redes neurais artificiais para processos

de *nowcasting* pode gerar melhores resultados ao diminuir o erro absoluto médio.

De acordo com os valores de EAM encontrados na tabela 4.6, retornados pelo método de validação cruzada executado nas seções 4.2.1 e 4.3.1, é possível comparar o desempenho de ambos os modelos preditivos empregados. Nota-se que o desempenho da ELM foi, em média, superior ao da regressão linear.

Tabela 4.6: Comparação dos erros fornecidos pelo método de validação cruzada

Modelo	EAM
ELM	1.35 ± 0.07
Regressão	1.37 ± 0.15

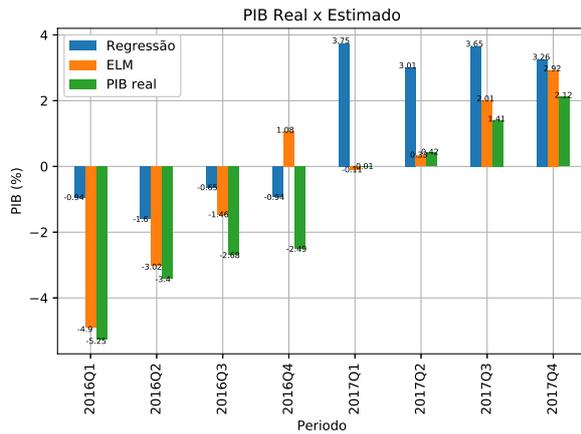
Ainda que os resultados apresentem uma diferença pequena, a tabela 4.7 é capaz de evidenciar as diferenças numéricas de cada modelo para cada valor real do PIB. A primeira predição efetuada pela ELM consegue descrever melhor as tendências de crescimento do PIB em vista da proximidade do cenário real.

Tabela 4.7: Comparação dos resultados de PIB obtidos

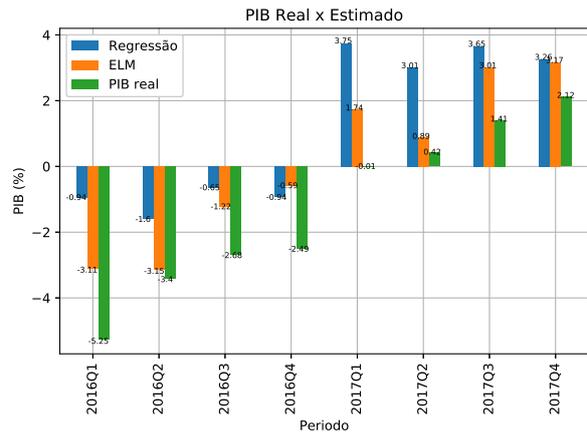
Período	PIB Real (%)	Rede Neural (ELM)				Regressão
		Predição 1	Predição 2	Predição 3	Predição 4	
2016Q1	-5.25	-4.90	-3.11	-3.01	-3.31	-0.94
2016Q2	-3.4	-3.02	-3.15	-2.74	-2.62	-1.60
2016Q3	-2.68	-1.46	-1.22	-2.62	-2.17	-0.65
2016Q4	-2.49	1.08	-0.59	-1.11	0.17	-0.94
2017Q1	-0.01	-0.11	1.74	2.92	2.08	3.75
2017Q2	0.42	0.33	0.89	1.68	1.79	3.01
2017Q3	1.41	2.01	3.01	0.40	1.86	3.65
2017Q4	2.12	2.92	3.17	0.66	1.96	3.26
EAM		0.88 ± 0.15	1.32 ± 0.23	1.37 ± 0.26	1.24 ± 0.18	2.42 ± 0.32

A figura 4.3 explica graficamente a diferença entre os valores de PIB previstos em cada período que os modelos foram aplicados.

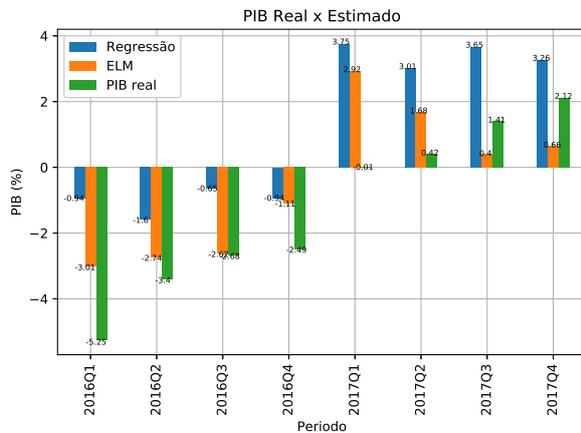
Como observado acima, de maneira geral, o desempenho de generalização da ELM foi 1.46% superior a da regressão linear. A ELM conseguiu obter menores valores de erro ainda que dentro de pequenas variações se comparados com os erros da regressão. A regularização demonstrou um aumento significativo de desempenho quando aplicada na rede neural, diminuindo a necessidade de um número cada vez maior de neurônios na camada oculta. De tal maneira, os resultados obtidos demonstram que a aplicabilidade da ELM para problemas econômicos de regressão é factível e gera resultados mais acurados que a modelagem linear.



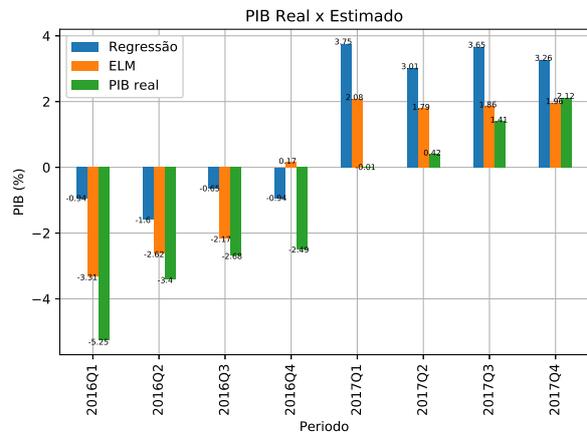
(a) Predição 1



(b) Predição 2



(c) Predição 3



(d) Predição 4

Figura 4.3: Comparativo entre as metodologias utilizadas

4.4 SÍNTESE

Neste capítulo são apresentados os resultados dos experimentos elaborados. De forma geral, são explicados os processos de seleção de variáveis, validação da hipótese e ajuste dos parâmetros da ELM.

Os resultados das duas etapas implementadas indicam forte impacto do índice Bovespa na economia brasileira. Os indicadores utilizados foram escolhidos com o auxílio da regressão linear multivariada conforme os valores dos pesos de maior impacto preditivo.

Os métodos aplicados apresentaram resultados similares mas, ainda assim, a ELM regularizada foi capaz de superar a regressão linear.

O capítulo seguinte realiza a conclusão deste trabalho, além de discutir o processo experimental utilizado.

5 CONCLUSÃO E TRABALHOS FUTUROS

Este trabalho de graduação consistiu na descrição do problema de *nowcasting* do PIB brasileiro e na análise de redes neurais artificiais, especificamente as Máquinas de Aprendizado Extremo, como alternativa de solução. A metodologia utilizada para a avaliação das variáveis que melhor caracterizam as variações de crescimento do PIB trimestral brasileiro fundamentou-se em trabalhos correlatos, bem como o emprego de regressões lineares multivariadas. A proposta de máquinas de aprendizado extremo foi o modelo de aprendizado de máquina escolhido para análise por ser um algoritmo simples e com grande capacidade de generalização.

Com o auxílio de trabalhos sobre o tema, pôde-se analisar o problema de *nowcasting*, compreendendo sua motivação e os desafios associados. Com o objetivo de estudar e comparar a aplicação de regressões lineares e máquinas de aprendizado extremo, a parte experimental deste trabalho executa experimentos com uma base de dados abertos e reais, envolvendo os algoritmos supracitados. Não obstante, os experimentos realizados basearam-se no método de *K-Fold Cross Validation* para avaliação dos resultados obtidos.

Os dados foram coletados com o auxílio de portais oficiais de instituições nacionais e mantiveram-se dentro do intervalo de 1995 a 2017 em virtude da disponibilidade de informações e padronização das metodologias utilizadas pelos órgãos governamentais envolvidos. Além disso, o custo computacional de se aplicar tanto o modelo de regressão linear quanto o de rede neural foi baixo, visto que a quantidade de dados foi pequena, assim como a quantidade de variáveis selecionadas.

A discussão dos resultados obtidos e suas respectivas conclusões são apresentadas a seguir, de acordo com os objetivos propostos na Seção 1.1.

5.1 DISCUSSÃO DOS RESULTADOS OBTIDOS

Com base nos resultados obtidos, conclui-se que o desempenho da ELM foi ligeiramente superior ao da regressão linear multivariada, logo, o objetivo de se propor a utilização de RNAs para solução de problemas de *nowcasting* mostrou-se factível. Nota-se também que o EAM do método de validação aplicado para a ELM foi menor que o da regressão linear.

Os resultados experimentais permitem afirmar os seguintes pontos:

- Os indicadores de produção agropecuária e de arrecadação de impostos totais acabaram surpreendendo. Apesar de terem sido utilizados em alguns dos trabalhos referenciados, não foram considerados como amplamente significativos e, ainda assim, tiveram importância no processo preditivo deste trabalho conforme pôde ser visualizado pelos pesos resultantes da regressão linear.

- Pode-se considerar que variáveis de produção industrial e relacionadas ao mercado financeiro são regressores de alta importância e devem ser exploradas. Sendo assim, comprovam-se as afirmações feitas nos trabalhos de *Zuanazzi e Ziegelmann* [13], *Kagohara* [11] e *Vasconcelos* [10].
- A regularização da ELM aumentou o desempenho do modelo, melhorando a qualidade das estimativas e diminuindo o EAM.
- Ao se utilizar bases de dados mais robustas, o modelo de máquinas de aprendizado extremo é recomendado para fornecer resultados mais precisos que o modelo de regressão linear.

Por fim, o estudo de metodologias de *nowcasting* que utilizem técnicas de aprendizado de máquina provou-se eficaz, sendo capaz de realizar previsões de índices econômicos de maneira precisa, mostrando-se uma área promissora a ser explorada.

5.2 TRABALHOS FUTUROS

Este trabalho propõe as seguintes perspectivas a serem abordadas futuramente:

- Como forma de aumentar a capacidade de generalização do regressor proposto, aconselha-se o emprego de uma quantidade maior de dados que incorporem maiores períodos de tempo e, desta maneira, forneçam ao modelo um histórico econômico mais vasto, provendo informações sobre períodos de recessão ou expansão da economia.
- Com o propósito de melhorar a acurácia do modelo, é interessante o emprego de um número maior de variáveis para aumentar a robustez do método. Assim como citado nos trabalhos relacionados, o uso de indicadores que explorem as variações do dólar, índices do mercado financeiro, índices do mercado internacional e impostos sobre o consumo podem diminuir o erro associado.
- A fim de diminuir o erro associado, recomenda-se explorar o uso de outras técnicas de regularização que sejam capazes de penalizar os índices mais ruidosos.
- Uma vez que os pesos sinápticos de entrada da ELM são atribuídos aleatoriamente, recomenda-se o estudo de distribuições que melhor otimizem o modelo selecionado.
- Com o intuito de se produzir estimativas de melhor qualidade, será necessário realizar o estudo comparativo do esforço computacional necessário para processamento dos dados, sendo que a quantidade de neurônios de uma RNA pode aumentar consideravelmente o tempo de treinamento necessário.
- Dado que a quantidade de fontes governamentais é vasta, torna-se interessante a utilização de *Big Data* para melhoria da generalização dos modelos utilizados, bem como seleção

de variáveis mais significativas. De tal maneira, seria interessante a implementação de metodologias de computação distribuída capazes de paralelizar o treinamento de modelos com grandes quantidades de dados.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] GIANNONE, D.; REICHLIN, L.; SMALL, D. H. Nowcasting gdp and inflation: The real-time informational content of macroeconomic data releases. *Working Paper Series 633*, 2006.
- [2] HINDS, S.; RIMMINGTON, L.; DANCE, H.; GILLHAM, J.; SENTANCE, A.; HAWKSWORTH, J. A machine learning approach to estimating current gdp growth. *UK Economic Outlook*, p. 35–47, 2017.
- [3] BELL, V.; CO, L. W.; STONE, S.; WALLIS, G. Nowcasting uk gdp growth. *Quarterly Bulletin*, p. 58–68, 2014.
- [4] ANESTI, N.; HAYES, S.; MOREIRA, A.; TASKER, J. Peering into the present: the bank’s approach to gdp nowcasting. *Quarterly Bulletin*, p. 122–133, 2017.
- [5] HUANG, G.-B.; ZHU, Q.-Y.; SIEW, C.-K. Extreme learning machine: Theory and applications. *Neurocomputing*, v. 70, p. 489–501, 2006.
- [6] HUANG, G.; HUANG, G.-B.; SONG, S.; YOU, K. Trends in extreme learning machines: A review. *Neural Networks*, v. 61, p. 32–48, 2015.
- [7] KRUGMAN, P.; WELLS, R. *Introdução a Economia*. : Elsevier Academic, 2007.
- [8] MANKIW, N. G. *MACROECONOMICS*. : Worth Publishers, 2009.
- [9] IBGE. <<https://www.ibge.gov.br/estatisticas-novoportal/todos-os-produtos-estatisticas.html>>. Accessed: 2018-10-11.
- [10] VASCONCELOS, B. F. B. de. *Poder preditivo de métodos de Machine Learning com processos de seleção de variáveis: uma aplicação às projeções de produto de países*. Tese (Doutorado) — UNIVERSIDADE DE BRASÍLIA, Brasília, Brasil, 2017.
- [11] KAGOHARA, D. M. *AVALIANDO TÉCNICAS DE NOWCASTING: UMA APLICAÇÃO DO PIB BRASILEIRO*. Dissertação (Mestrado) — FUNDAÇÃO GETÚLIO VARGAS, São Paulo, Brasil, 2015.
- [12] CASTLE, J. L.; HENDRY, D. F.; KITOV, O. I. Forecasting and nowcasting macroeconomic variables: A methodological overview. *DISCUSSION PAPER SERIES*, v. 674, 2013.
- [13] ZUANAZZI, P. T.; ZIEGELMANN, F. A. Previsões para o crescimento do pib trimestral brasileiro com séries financeiras e econômicas mensais: Uma aplicação de midas. *Economia Aplicada*, v. 18, p. 295–318, 2014.
- [14] MATTOS, P. M. *NOWCASTING BRAZILIAN GDP*. Dissertação (Mestrado) — FUNDAÇÃO GETÚLIO VARGAS, São Paulo, Brasil, 2017.
- [15] RUSSEL, S.; NORVIG, P. *Artificial Intelligence: A Modern Approach*. : Pearson, 2009.
- [16] ALPAYDIN, E. *INTRODUCTION TO MACHINE LEARNING*. : The MIT Press, 2014.
- [17] MITCHELL, T. *MACHINE LEARNING*. : McGraw-Hill, 1997.
- [18] GUJARATI, D.; PORTER, D. *Econometria Básica*. : Mc Graw Hill, 2011.
- [19] HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. : Springer, 2009.

- [20] HAYKIN, S. *Neural Networks and Learning Machines*. : Pearson, 2008.
- [21] SILVA, D. G. E. *Uso de Aprendizado de Máquina para estimar esforço de execução de testes funcionais*. Dissertação (Mestrado) — UNIVERSIDADE ESTADUAL DE CAMPINAS, Campinas, Brasil, 2009.
- [22] PYTHON. <<https://www.python.org/>>. Accessed: 2018-09-27.
- [23] NUMPY. <<http://www.numpy.org/>>. Accessed: 2018-09-27.
- [24] PANDAS. <<https://pandas.pydata.org/>>. Accessed: 2018-09-27.
- [25] MATPLOTLIB. <<https://matplotlib.org/>>. Accessed: 2018-11-02.
- [26] VISUAL Studio Code. <<https://code.visualstudio.com/>>. Accessed: 2018-11-02.
- [27] BACEN SERIES. <<https://www3.bcb.gov.br/sgspub/>>. Accessed: 2018-10-11.
- [28] IPEA Data. <<http://www.ipeadata.gov.br/>>. Accessed: 2018-10-10.
- [29] IBOVESPA. <http://www.b3.com.br/pt_br/>. Accessed: 2018-10-11.