

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

Previsão de consumo nos estados do Brasil

Julia Leite da Silva

MONOGRAFIA FINAL

MAC 499 — TRABALHO DE
FORMATURA SUPERVISIONADO

Supervisor: Prof. Dr. Marcelo Finger

São Paulo
2022

*O conteúdo deste trabalho é publicado sob a licença CC BY 4.0
(Creative Commons Attribution 4.0 International License)*

Dedico este trabalho a minha família, meu pai e minha mãe que têm iluminado meu dia e me dado força, alegria e paz. Acima de tudo, dedico a Deus, quem me presentou com capacidade, com uma família maravilhosa e que foi adicionando pessoas no meu caminho para me apoiar e me ajudar em cada passo da jornada.

Agradecimentos

Fight with determination, embrace life and live it with passion. Lose your battles with class and dare to win because the world belongs to those who dare to live. Life is worth too much to be insignificant.

— Charles Chaplin

Queria agradecer, antes de tudo, a Deus por tudo que Ele tem me dado. Uma vez minha mãe me disse que Deus abençoa o trabalho das nossas mãos, que várias vezes onde ela pensou que não daria certo ou não daria conta, ela pedia ajuda e, no final, tudo se endireitava. Ela me disse também que Deus colocava pessoas na vida dela que a ajudavam e orientavam a encontrar um caminho. Fico muito grata de enxergar isso acontecendo na minha vida também.

Agradeço aos meus pais, que têm sido minha força, alegria e lar. Agradeços aos meus amigos e amigas, por colorir meus dias e encher meu rosto de sorrisos.

Quero também agradecer ao Marcelo Finger, meu orientador obrigada para o Felipe, que me surpreendeu com a parceria e a nessa jornada. Muito ajuda. Agradecer também ao Gabriel, pelo apoio para escrever esse trabalho.

Resumo

Julia Leite da Silva. **Previsão de consumo nos estados do Brasil**. Monografia (Bacharelado). Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2022.

[illegible]

Palavras-chave: Palavra-chave1. Palavra-chave2. Palavra-chave3.

Abstract

Julia Leite da Silva. . Capstone Project Report (Bachelor). Institute of Mathematics and Statistics, University of São Paulo, São Paulo, 2022.

[illegible]

Keywords: Keyword1. Keyword2. Keyword3.

Lista de Abreviaturas

MLP	Perceptron Multi-Camadas (<i>Multilayer perceptron</i>)
LSTM	(<i>Long Short Term Memory</i>)
RMSE	Raíz Média de erros Quadrados (<i>Root Mean Squared Error</i>)
MAPE	Erro Percentual Absoluto Médio (<i>Mean Absolute Percentage Error</i>)

Lista de Símbolos

ω	Frequência angular
ψ	Função de análise <i>wavelet</i>
Ψ	Transformada de Fourier de ψ

Lista de figuras

1.1	Relação entre inteligência artificial, aprendizado de máquina e <i>deep learning</i> PATTERSON e GIBSON, s.d.	4
1.2	Exemplo de um modelo simples de regressão linear HYNDMAN e ATHANASOPOULOS, 2021	5
1.3	Ilustração de um neurônio biológico. PATTERSON e GIBSON, s.d. Destaca-se a informação chegando a um dendrito do neurônio por meio de uma sinapse, além de uma sinapse que se inicia no axônio do neurônio e propaga informação a diante.	6
1.5	Ilustração de uma rede neural simples.ACADEMY, 2022	8
1.6	Rectified linear unit (ReLU) PATTERSON e GIBSON, s.d.	9
1.7	Função <i>swish</i> RAMACHANDRAN <i>et al.</i> , 2017	9

Lista de tabelas

2.1	Indicadores de crescimento econômico	11
2.2	Indicadores de inflação e política monetária	11
2.3	Indicadores de inflação e política fiscal	11
2.4	Indicadores sociais	12
2.5	Indicadores de construção civil	12

Lista de programas

Sumário

Introdução	1
1 Fundamentação teórica	3
1.1 Aprendizado de máquina	3
1.2 Problema	4
1.3 Regressão linear	5
1.4 Redes neurais	6
1.4.1 Neurônios artificiais	7
1.4.2 Função de ativação	8
1.4.3 Redes neurais multi-layer perceptrons	9
1.4.4 Redes Neurais Recorrentes	10
1.4.5 Normalização	10
2 Metodologia	11
2.1 Dados	11
2.2 Avaliação de performance	12
2.2.1 Mean absolute error (MAE)	13
2.2.2 Root mean squared error (RMSE)	13
2.2.3 Mean absolute percentage error (MAPE)	13
2.2.4 Delta percentual	13
3 Experimentos	15
3.1 Regressao linear	15
3.2 Redes neurais MLP	15
3.3 Redes recorrentes	15
3.4 regressao linear	15
4 Resultados	17

5 Conclusão	19
Referências	21

Introdução

O que prédios, pontes, hidrelétricas e aeroportos têm em comum? Todos são frutos da indústria da construção civil, um importante componente do investimento brasileiro e, consequentemente, uma das grandes engrenagens responsáveis por movimentar a atividade econômica no Brasil. Em 2021, por exemplo, o Produto Interno Bruto (PIB) desse setor registrou alta de 9,7%, enquanto o PIB do Brasil cresceu 4,6%, assim, o setor da construção figura como importante impulsionador da economia do país. [VASCONCELOS, 2022](#) O cimento, nesse contexto, por ser um ingrediente central da argamassa e do concreto, caracteriza-se como um dos principais insumos da indústria.

Motivação

Contudo, a falta de um modo bem fundamentado para prever a demanda de cimento é uma dor entre as empresas cimenteiras, uma vez que a construção de uma fábrica é custosa e demorada, além disso, aumentar a capacidade de produção de uma fábrica também é um processo custoso.

Dessa forma, um modelo que permitisse prever a demanda a nível de estados do Brasil poderia auxiliar gestores a tomar melhores decisões e a estruturar a estratégia de forma mais embasada, de modo a reduzir os riscos do setor. Além disso, poder-se-ia apoiar órgãos governamentais a direcionar ações para mitigar o impacto ambiental da fabricação desse produto.

Objetivos

Este trabalho, então, propõe-se a aplicar modelos de aprendizado de máquina para determinar qual é mais eficiente para prever a demanda por cimento nos estados do Brasil. Os modelos avaliados são: regressão linear, redes neurais *multi-layer perceptron* (MLP) e redes neurais recorrentes.

Capítulo 1

Fundamentação teórica

1.1 Aprendizado de máquina

Atualmente, inteligência artificial (IA) permeia diversos momentos do cotidiano. É o caso da empresa norte-americana de *streaming* Netflix, que utiliza um conjunto de técnicas de IA para recomendar conteúdo personalizado aos usuários da plataforma de acordo com os interesses particulares de cada um. Dessa forma, proporciona uma experiência única a cada indivíduo que acessa a plataforma com o objetivo de aumentar a satisfação a longo prazo e, conseqüentemente, garantir a retenção dos membros, uma vez que a plataforma é monetizada com assinaturas mensais.

Além disso, não há um modelo ou algoritmo único utilizado para todas as recomendações de conteúdo. Essa tarefa é dividida em subtarefas realizadas por diferentes modelos de acordo com a atividade a ser realizada e os dados disponíveis. Por exemplo, a sub tarefa de decidir qual vídeo será exibido para cada usuário ao logar no perfil da plataforma é executada por um modelo diferente do que o que elenca os vídeos já assistidos que o membro pode continuar a ver. [STECK et al., 2021](#)

Mas a final, o que é inteligência artificial (IA)? O termo "inteligência artificial", *artificial intelligence* em inglês, foi elaborado por John McCarthy e utilizado oficialmente pela primeira vez em 1956 no seminário de Dartmouth, um *workshop* sobre essa área que reuniu os maiores estudiosos do ramo durante dois meses. [RUSSELL e NORVIG, s.d.](#) Esse termo apresenta várias definições, de acordo com o pioneiro Arthur Samuel, pode ser definida como o campo de estudo que dá aos computadores a habilidade de aprender sem serem explicitamente programados. [PATTERSON e GIBSON, s.d.](#)

Aprendizado de máquina, por sua vez, do inglês *machine learning*, são sistemas de IA capazes de adquirir seu próprio conhecimento por meio da extração de padrões dos dados brutos. Configura-se, portanto, como uma sub-área de inteligência artificial. [GOODFELLOW et al., 2016](#). O aprendizado profundo, ou *deep learning*, é uma categoria específica de *machine learning* que compreende modelos de redes neurais com várias camadas de neurônios. [ZHANG et al., 2021](#)

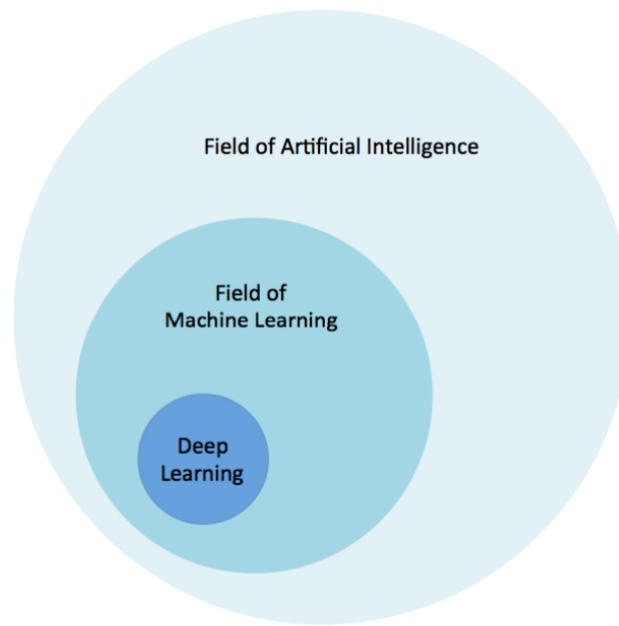


Figura 1.1: Relação entre inteligência artificial, aprendizado de máquina e deep learning *PATTERSON e GIBSON, s.d.*

1.2 Problema

As tarefas de *machine learning* são descritas de acordo com o processamento que o modelo deve realizar a partir de um exemplo de entrada (*input*), em geral, descrito com um vetor $x \in \mathbb{R}^n$, $x = \{x_1, x_2, \dots, x_n\}$. Por exemplo, neste trabalho um exemplo de *input* seriam os dados de março de 2011 do estado de São Paulo, cada entrada x_i , então, corresponde à medição de um dos indicadores econômicos desse estado e mês, logo x_1 pode representar o PIB estadual, x_2 o PIB *per capita* e assim sucessivamente.

Nas tarefas de classificação, o modelo deve prever a qual das k categorias disponíveis um *input* pertence. O algoritmo então cria uma função $f : \mathbb{R}^n \rightarrow 1, \dots, k$, quando $f(x) = y$, o vetor de entrada x foi recebido a categoria y . Um exemplo de tarefa de classificação seria determinar se o consumo de cimento em um estado em um mês específico representa um aumento, queda ou estabilidade em relação ao mês anterior. Neste trabalho, contudo, não se utiliza esse tipo de tarefa. *GOODFELLOW et al., 2016*

Outra categoria de tarefas são as de regressão, aonde o objetivo é, a partir do *input* x , prever um valor numérico. A função criada pelo modelo, então, é $f : \mathbb{R}^n \rightarrow \mathbb{R}$. O problema abordado neste trabalho, então, configura-se como problema de regressão, uma vez que o objetivo é prever o valor do consumo de cimento em um estado e mês específicos a partir dos dados de entrada.

Os problemas de aprendizado de máquina também podem ser divididos entre aprendizado não-supervisionado e supervisionado. No primeiro, o modelo recebe um *dataset* não rotulado e então aprende propriedades da estrutura do *dataset*, pode, então, performar tarefas como a clusterização, que consiste em dividir o conjunto de dados em *clusters*

com exemplos similares. No último, por sua vez, os dados de entrada estão associados a rótulos, resultados conhecidos, chamados de *labels* ou *target* em inglês. Neste trabalho, a utiliza-se aprendizado supervisionado, uma vez que o real consumo de cimento no mês é conhecido. [GOODFELLOW et al., 2016](#)

1.3 Regressão linear

A regressão linear é um modelo de aprendizado de máquina que assume um relacionamento linear entre a variável que será prevista (*target*) e os dados de entrada. Dessa forma, o objetivo é obter uma função linear que receba um vetor $x \in \mathbb{R}^n$, $x = \{x_1, x_2, \dots, x_n\}$ como entrada e devolva a previsão de um escalar $y \in \mathbb{R}$. [GOODFELLOW et al., 2016](#) Seja \hat{y} o valor previsto pelo modelo para \hat{y} , então:

$$\hat{y} = w^T x + b \quad (1.1)$$

onde $w \in \mathbb{R}^n$, $w = \{w_1, w_2, \dots, w_n\}$ é um vetor de parâmetros. Em particular, w corresponde a um conjunto de pesos que determina como cada variável afeta a previsão. Então, se x_i for associado a um peso w_i positivo, aumentar o valor de x_i resulta em um aumento na previsão \hat{y} , se w_i for negativo, por outro lado, um aumento de x_i resulta em diminuição de \hat{y} . Se w_i for igual a 0, por sua vez, a variável x_i não influencia no valor previsto.

A constante b , do inglês *bias*, é um parâmetro para mensurar viés, já que o *output* da função é b na ausência de uma entrada. Dessa forma, a equação 1.1 pode ser escrita da seguinte forma:

$$\hat{y} = w_1 x_1 + w_2 x_2 + \dots + w_n x_n + b \quad (1.2)$$

Na ilustração abaixo, um exemplo de um modelo de regressão linear com apenas uma variável:

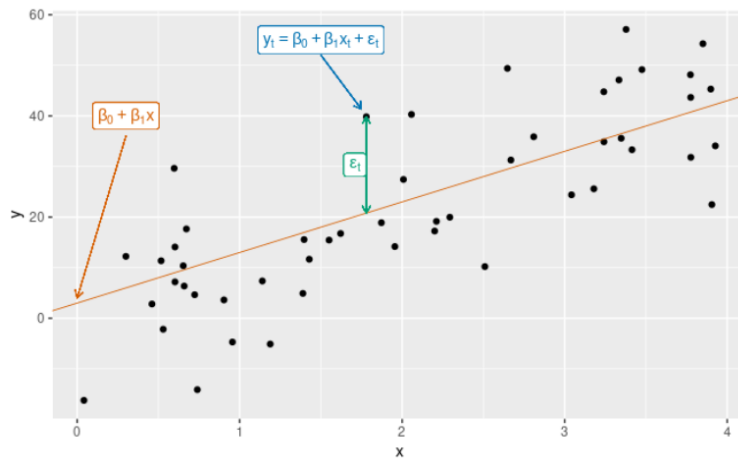


Figura 1.2: Exemplo de um modelo simples de regressão linear [HYNDMAN e ATHANASOPOULOS, 2021](#)

Nessa imagem, as observações estão representadas nos pontos pretos, enquanto a linha em laranja corresponde à previsão realizada pelo modelo. Observa-se que o modelo não prevê com total exatidão os dados observados, há um erro associado a cada previsão, como o destacado em verde na ilustração.

Dessa forma, cada observação y_i possui um erro ε_i associado e pode ser descrita por $y_i = w^T x_i + b + \varepsilon_i$. O vetor de pesos w , então, é escolhido de modo a minimizar os erros em cada previsão.

Por se tratar de um modelo mais simples, é utilizada neste trabalho como base para comparar o desempenho de outros modelos mais robustos.

1.4 Redes neurais

Redes neurais são modelos computacionais inspirados no funcionamento do cérebro animal, onde neurônios trabalham em paralelo sem uma unidade central de controle. Um neurônio biológico é uma célula nervosa que se comunica com outros neurônios e passa impulsos eletro-químicos de uma célula para outra por meio das sinapses. A comunicação entre neurônios, sinapse, ocorre apenas se o impulso for forte o bastante para ativar a liberação de químicos na fenda sináptica.

Um neurônio é composto de vários dendritos, um axônio e corpo celular, como na imagem 1.3.

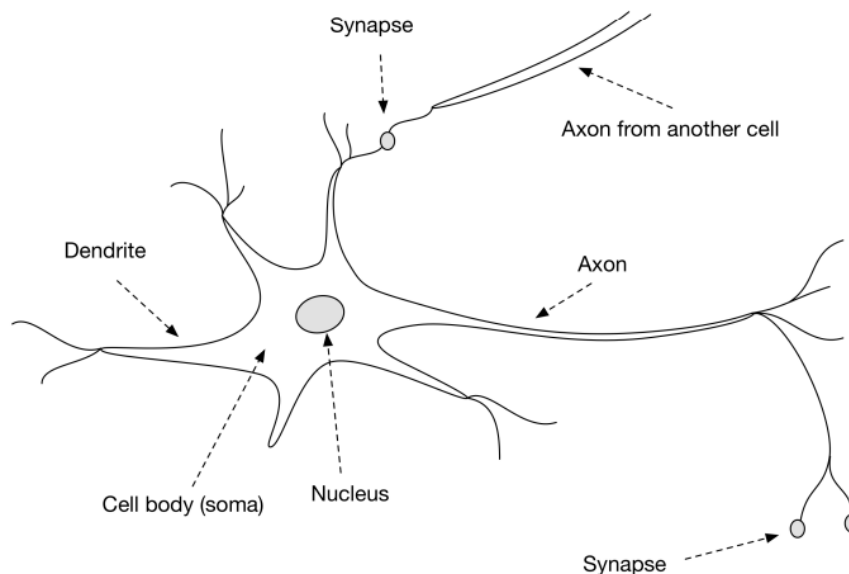


Figura 1.3: Ilustração de um neurônio biológico. *PATTERSON e GIBSON, s.d.* Destaca-se a informação chegando a um dendrito do neurônio por meio de uma sinapse, além de uma sinapse que se inicia no axônio do neurônio e propaga informação a diante.

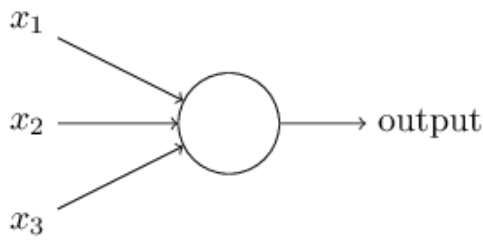
Os dendritos recebem informações de outros neurônios vizinhos, na forma de impulsos elétricos, e conduzi-las até o corpo celular. Ao chegar no corpo celular, a informação é

processada e novos impulsos são gerados. Essa informação é, então, repassada para outro neurônio através do axônio por meio de sinapse. Dessa forma, sinapse é o ponto de contato entre a terminação axônica de um neurônio e o dendrito de outro. [ACADEMY, 2022](#).

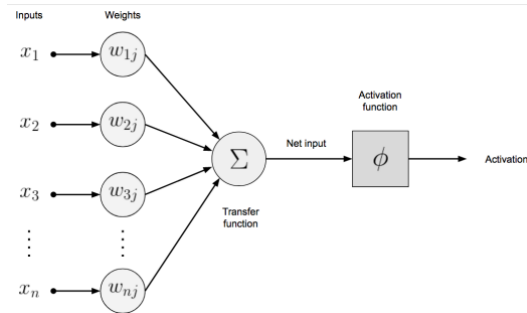
A estrutura e funcionamento dos neurônios biológicos foram base para os cientistas criarem neurônios artificiais, como os *perceptrons*.

1.4.1 Neurônios artificiais

O *perceptron*, foi desenvolvido em 1957 por Frank Rosenblatt, inspirado nos trabalhos de Warren McCulloch e Walter Pitts. Trata-se de um modelo linear de classificação binária que recebe n entradas e produz uma saída binária, como mostrado na ilustração simplificada 1.4a. [ACADEMY, 2022](#) Esse modelo inicial apresentava limitações e foi evoluído com o passar do tempo, as redes neurais atualmente, em geral, utilizam outro modelo de neurônio como ilustrado em 1.4b.



(a) Visão simplificada de um neurônio [ACADEMY, 2022](#)



(b) Ilustração da arquitetura de um neurônio artificial [PATTERSON e GIBSON, s.d.](#)

Como é possível observar na figura 1.4b, cada uma das entradas x_i está associada a um peso w_i . Os pesos w_1, w_2, \dots, w_n expressam a importância das respectivas entradas para o valor de saída, de modo semelhante à regressão linear na seção 1.3. O produto escalar entre os pesos e as respectivas entradas, chamada de *net input* na imagem 1.4b, passa por uma função de ativação ϕ que determina a saída do neurônio. Além disso, um valor de *bias* ou polarização é adicionado ao produto escalar para aumentar a liberdade da função. O *bias* possibilita que um neurônio que possua todas as entradas nulas apresente uma saída não nula, dessa forma, aumenta a capacidade de aproximação da rede. [ACADEMY, 2022](#)

Seja x o vetor das n entradas do neurônio, então $x \in \mathbb{R}^n$, $x = \{x_1, x_2, \dots, x_n\}$, seja também w o vetor com os pesos associados a cada entrada, $w \in \mathbb{R}^n$, $w = \{w_1, w_2, \dots, w_n\}$. Além disso, seja b o valor de *bias* e Φ a função de ativação. Dessa forma, o *output* de um neurônio é dado por:

$$h_{w,b} = \Phi(w \cdot x + b) \quad (1.3)$$

Essa saída é utilizada como uma das entradas dos neurônios na camada seguinte, de modo a formar a estrutura das redes neurais semelhante ao funcionamento do cérebro.

Os neurônios, então, formam a unidade que compõe as redes neurais artificiais, como ilustrado na figura abaixo:

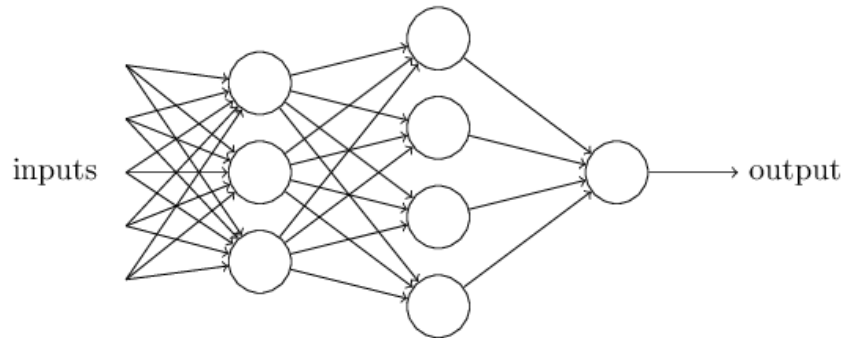


Figura 1.5: Ilustração de uma rede neural simples. *ACADEMY, 2022*

1.4.2 Função de ativação

A função de ativação propaga a saída de um neurônio na rede neural para a camada seguinte. Enquanto os pesos e o *bias* realizam uma transformação linear nos dados de entrada, a função de ativação aplica uma transformação não linear e dessa forma torna possível que a rede neural resolva problemas não lineares e complexos, como reconhecer padrões de escrita. *ACADEMY, 2022*

A função de ativação é um atributo de cada uma das camadas da rede e é escolhida de acordo com a tarefa que será executada, por exemplo, a função sigmóide é recomendada para problemas de classificação. Neste trabalho, foram utilizadas: *rectified linear unit* (ReLU) e *swish*.

Rectified linear unit (ReLU)

A função ReLU, do inglês *rectified linear unit*, é o estado da arte atualmente, uma vez que apresenta bom desempenho em diferentes tarefas. *PATTERSON e GIBSON, s.d.* A função ReLU é dada por:

$$f(x) = \max(0, x) \quad (1.4)$$

Ao utilizar essa função, a derivada, utilizada para atualizar os pesos e *bias* no treinamento da rede, é zero, quando a entrada é nula ou uma constante, como na imagem abaixo. Dessa forma, a ReLU não sofre do problema da dissipação do gradiente como a sigmóide.

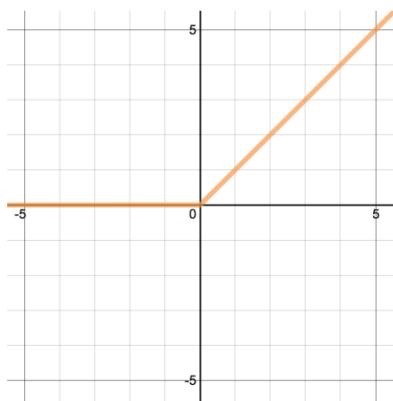


Figura 1.6: Rectified linear unit (ReLU) *PATTERSON e GIBSON, s.d.*

Swish

A função *swish* foi proposta por pesquisadores da Google com a prerrogativa de apresentar melhor desempenho que a ReLU em redes neurais profundas. *RAMACHANDRAN et al., 2017*

A *swish* é uma função não monotônica e suave dada por:

$$f(x) = x \cdot \text{sigmoid}(x) = \frac{x}{1 + e^{-x}} \quad (1.5)$$

O gráfico da função é similar ao da ReLU:

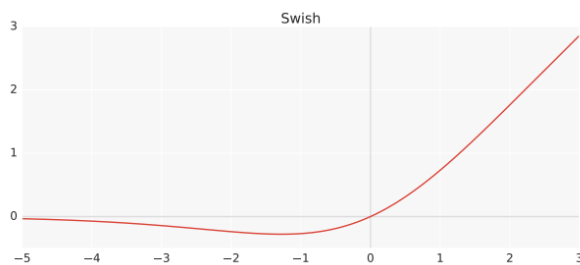


Figura 1.7: Função *swish* *RAMACHANDRAN et al., 2017*

1.4.3 Redes neurais multi-layer perceptrons

Redes neurais são modelos de *machine learning* inspiradas no cérebro humano aonde o aprendizado ocorre ao se agregarem neurônios matemáticos que estabelecem conexões de acordo com o treinamento fornecido. Neste trabalho, aplicaram-se redes *multilayer perceptrons* (MLPs), ou seja, que apresentam múltiplas camadas de neurônios e *feedforward*, onde a saída de uma camada de neurônios é utilizada como entrada para a camada seguinte, sem utilizar retropropagação.

1.4.4 Redes Neurais Recorrentes

Já as redes neurais recorrentes são projetadas para reconhecer padrões nos dados, uma vez que levam tempo e sequência em consideração. Assim, nessas redes, a decisão tomada na etapa anterior influencia a etapa seguinte por conta dos *loops de feedback*, então o presente e o passado recente se combinam para determinar a previsão. Neste trabalho, foram testadas as redes Long Short Term Memory (LSTM), redes Gated Recurrent Unit (GRU) e Bidirecionais.

GRU

LSTM

1.4.5 Normalização

Capítulo 2

Metodologia

2.1 Dados

O objetivo deste trabalho é prever o consumo mensal de cimento nos estados do Brasil. Esse indicador foi obtido do Sindicato Nacional da Indústria do Cimento e apresentava granularidade mensal para cada um dos estados do país.

Indicador	Fonte	Período disponível	Granularidade
PIB a preços constantes	IBGE	1983 até 2019	anual por estado
PIB a preços de mercado	IBGE	1985 até 2019	anual por estado
PIB per capita	IBGE	1985 até 2019	anual por estado
PIB da construção civil	IBGE	1985 até 2019	anual por estado
Desemprego	IBGE	1991 até 2022	anual por estado até 2014 e mensal para o Brasil a partir de 2015

Tabela 2.1: Indicadores de crescimento econômico

Indicador	Fonte	Período disponível	Granularidade
IPCA	IBGE	1981 até 2021	mensal para o Brasil
INCC	FGV	1980 até 2021	mensal para o Brasil
IGP	FGV	1944 até 2021	mensal para o Brasil
SELIC	IBGE	1986 até 2022	mensal para o Brasil

Tabela 2.2: Indicadores de inflação e política monetária

Indicador	Fonte	Período disponível	Granularidade
NFSP	BACEN	1991 até 2022	mensal para o Brasil
estoque de dívida pública	IPEA	1947 até 2019	anual para o Brasil

Tabela 2.3: Indicadores de inflação e política fiscal

Indicador	Fonte	Período disponível	Granularidade
População	IBGE	1991 até 2021	anual por estado
IDH	IBGE	1991 até 2017	irregular

Tabela 2.4: *Indicadores sociais*

Indicador	Fonte	Período disponível	Granularidade
Produção mensal de cimento	SNIC	2003 até 2022	mensal por estado
Valor médio do cimento Portland em reais por quilograma	IPEA	1947 até 2019	anual para o Brasil

Tabela 2.5: *Indicadores de construção civil*

Com o objetivo de direcionar a estratégia de preparação de dados foi realizada uma análise exploratória de cada um dos dados de entrada e da variável resposta. A partir dessa análise, optou-se por utilizar os dados de 2003 até 2019 para o estudo.

Foram adotadas estratégias para garantir dados na granularidade mensal e por estado. Caso os indicadores apresentassem granularidade anual, o valor foi dividido por 12 de modo a obter a média mensal, já caso a granularidade fosse a nível de Brasil, o valor apresentado foi repetido para todos os estados.

A estratégia utilizada para lidar com dados faltantes foi, sempre que possível, repetir o valor anterior que estava disponível nos dados de entrada para preencher a ocorrência. Contudo, alguns indicadores não apresentavam valores mais antigos, então foi usado um valor não presente no intervalo de dados de entrada para marcar como nulo.

Além disso, tomou-se um cuidado para evitar que a previsão fosse realizada com os dados do mês anterior ou do ano anterior no caso dos indicadores anuais. Deslocou-se, portanto, os dados de entrada para frente em um ocorrência de modo a associar os dados de um mês com o consumo no mês seguinte.¹

Por fim, o estado correspondente à medição foi usado como dado de entrada. Como os modelos de inteligência artificial aceitam apenas caracteres numéricos, utilizou-se o método de codificação *one hot* para criar 27 colunas, uma para cada estado, nas quais o valor é 1 quando a linha possui dados daquele estado e é 0 caso contrário.

2.2 Avaliação de performance

Para comparar a eficiência dos modelos mede-se os erros de cada previsão, ou seja, a distância entre o valor previsto pelo algoritmo e o valor do dado real. Neste trabalho, utilizou-se as seguintes métricas estatísticas para mensurar o desempenho: *mean absolute*

¹ os dados correspondentes a, por exemplo, fevereiro de 2004 estão relacionados ao consumo de cimento em março de 2004, com o objetivo de propor um cenário mais pertinente, uma vez que o objetivo do projeto é prever a demanda por cimento no mês seguinte em um estado a partir dos dados do mês atual e, eventualmente, dos anteriores.

error (MAE), *root mean square error* (RMSE) e *mean absolute percentage error* (MAPE). Além disso, foi utilizado o delta percentual (Δ) para avaliar se o modelo tende a subestimar ou superestimar o valor previsto, se é otimista ou pessimista.

2.2.1 Mean absolute error (MAE)

O MAE, sigla do inglês para *mean absolute error* ou média do erro absoluto mede o erro absoluto de cada previsão e é dado por: [HEWAMALAGE et al., 2022](#)

$$MAE = \frac{\sum_{i=1}^n |\hat{y}_i - y_i|}{n} \quad (2.1)$$

2.2.2 Root mean squared error (RMSE)

A RMSE, sigla para *root mean squared error* é semelhante à MAE, contudo eleva os erros ao quadrado antes de somá-los e tira a raiz logo depois. A RMSE é, por tanto, mais sensível a *outliers*. [HEWAMALAGE et al., 2022](#)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (2.2)$$

2.2.3 Mean absolute percentage error (MAPE)

Foi utilizada também a MAPE, *Mean absolute percentage error*, para mensurar a escala do erro em relação ao tamanho das medições.

$$MAPE = \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (2.3)$$

2.2.4 Delta percentual

O delta percentual, Δ , é utilizado para mensurar se o modelo apresenta tendência de subestimar ou superestimar a variável, se é otimista ou pessimista.

$$\Delta = \frac{\hat{y}_i - y_i}{y_i} \quad (2.4)$$

Capítulo 3

Experimentos

Para cada um dos modelos selecionados foram realizados experimentos alterando parâmetros e os dados visando melhoria de performance. Os experimentos realizados variam de acordo como o modelo, como pode ser visto a seguir:

3.1 Regressao linear

A regressão linear, por se tratar de um modelo mais simples, possui menos parâmetros agregados. Testou-se contudo as seguintes funções de normalização e transformação nos dados de entrada

3.2 Redes neurais MLP

3.3 Redes recorrentes

3.4 regressao linear

Capítulo 4

Resultados

blafgdf .

Capítulo 5

Conclusão

ahahahahahh

Referências

- [ACADEMY 2022] Data Science ACADEMY. *Deep Learning Book*. disponível em: <https://www.deeplearningbook.com.br>. Acesso em: 17 Dezembro. 2022. 2022 (citado nas pgs. 7, 8).
- [GOODFELLOW *et al.* 2016] Ian GOODFELLOW, Yoshua BENGIO e Aaron COURVILLE. *Deep Learning*. <http://www.deeplearningbook.org>. MIT Press, 2016 (citado nas pgs. 3–5).
- [HEWAMALAGE *et al.* 2022] Hansika HEWAMALAGE, Klaus ACKERMANN e Christoph BERGMEIR. *Forecast Evaluation for Data Scientists: Common Pitfalls and Best Practices*. 2022. DOI: [10.48550/ARXIV.2203.10716](https://doi.org/10.48550/ARXIV.2203.10716). URL: [5Curl%7Bhttps://arxiv.org/abs/2203.10716%7D](https://arxiv.org/abs/2203.10716) (citado na pg. 13).
- [HYNDMAN e ATHANASOPOULOS 2021] Rob J HYNDMAN e George ATHANASOPOULOS. *Forecasting: principles and practice*. 3ª ed. OTexts.com/fpp3. Accessed on 17 December 2022. OTexts, 2021 (citado na pg. 5).
- [PATTERSON e GIBSON s.d.] Josh PATTERSON e Adam GIBSON. *Deep Learning: A practitioner's approach* (citado nas pgs. 3, 4, 6–9).
- [RAMACHANDRAN *et al.* 2017] Prajit RAMACHANDRAN, Barret ZOPH e Quoc V. LE. *Searching for Activation Functions*. 2017. DOI: [10.48550/ARXIV.1710.05941](https://doi.org/10.48550/ARXIV.1710.05941). URL: <https://arxiv.org/abs/1710.05941> (citado na pg. 9).
- [RUSSELL e NORVIG s.d.] Stuart J. RUSSELL e Peter NORVIG. *Artificial Intelligence: A Modern Approach, third edition* (citado na pg. 3).
- [STECK *et al.* 2021] Harald STECK *et al.* “Deep learning for recommender systems: a netflix case study”. Em: *AI Magazine* 42.3 (nov. de 2021), pgs. 7–18. DOI: [10.1609/aimag.v42i3.18140](https://doi.org/10.1609/aimag.v42i3.18140). URL: <https://ojs.aaai.org/index.php/aimagazine/article/view/18140> (citado na pg. 3).
- [VASCONCELOS 2022] Ieda VASCONCELOS. *Informativo Econômico PIB*. 2022. URL: <https://cbic.org.br/wp-content/uploads/2022/03/informativo-economico-pib-04-marco-2022.pdf> (citado na pg. 1).

- [ZHANG *et al.* 2021] Aston ZHANG, Zachary C. LIPTON, Mu LI e Alexander J. SMOLA. “Dive into deep learning”. Em: *arXiv preprint arXiv:2106.11342* (2021) (citado na pg. 3).