

Aplicações de IA no Setor Elétrico Brasileiro

02 de julho de 2025

IA NO SETOR ELÉTRICO

A Volt Robotics vem atuando fortemente nessa área

- **Brazilian Free-Energy Market Mid-and Long-Term Forecasting Using Multi-Source Ensemble Solution**
International Conference on Machine Learning and Applications (ICMLA) 2024, em Miami, EUA
Artigo: [Brazilian Free-Energy Market Mid-and Long-Term Forecasting Using Multi-Source Ensemble Solution | IEEE Conference Publication | IEEE Xplore](#)
Artigo produzido no âmbito de P&D ANEEL realizado em parceria entre a Volt Robotics e Norte Energia.
- **Inteligência Artificial do Setor Elétrico Brasileiro: Uma revisão das principais aplicações publicadas**
Sociedade Brasileira de Planejamento Energético – CBPE 2024
Artigo apresentado e publicado no XIV Congresso Brasileiro de Planejamento Energético (CBPE 2024).
ISBN: 978-65-992540-2-4
Disponibilizado nos anais do evento: [XIV CBPE](#)
- **Customer Segmentation in the Brazilian Free Energy Market by Machine Learning**
7th IEEE Conference on Energy Internet and Energy System Integration (EI2) 2023, em Hangzhou, China
Artigo: [Customer Segmentation in the Brazilian Free Energy Market by Machine Learning | IEEE Conference Publication | IEEE Xplore](#)
Artigo produzido no âmbito de P&D ANEEL realizado em parceria entre a Volt Robotics e Norte Energia.

IA NO SETOR ELÉTRICO

A Volt Robotics vem atuando fortemente nessa área

- **Multi-source data ensemble for energy price trend forecasting**
Elsevier - Engineering Applications of Artificial Intelligence Volume 133, Part B, July 2024, 108125 · 4 de mar de 2024
Periódico de alto impacto e *qualis* A1
Artigo: [Multi-source data ensemble for energy price trend forecasting - ScienceDirect](#)
Publicação oriunda de P&D ANEEL realizado em parceria entre Volt Robotics, CESP e Auren Energia.
- **Inteligência Artificial aplicada à previsão da tendência no preço de mercado para energia**
XXVII Seminário Nacional de Produção e Transmissão de Energia Elétrica - SNPTEE 2023 · 27 de nov de 2023
Artigo disponibilizado nos anais do evento: [XXVII Snp tee 2023](#)
O artigo, fruto de P&D ANEEL realizado em parceria entre Volt Robotics e Furnas Centrais Elétricas, foi premiado em 1º lugar no Grupo de Estudos de Comercialização, Economia e Regulação do Mercado de Energia Elétrica - GCR no XXVII SNPTEE 2023
- **Feature Selection using Complex Networks to Support Price Trend Forecast in Energy Markets**
2023 IEEE International Joint Conference on Neural Networks (IJCNN), em Gold Coast, Australia · 2 de ago de 2023
Conferência de Computação de alto impacto e com *qualis* A1
Artigo: [Feature Selection using Complex Networks to Support Price Trend Forecast in Energy Markets | IEEE Conference Publication | IEEE Xplore](#)
Artigo oriundo de P&D ANEEL realizado em parceria entre Volt Robotics e Auren Energia.

IA NO SETOR ELÉTRICO

A Volt Robotics vêm atuando fortemente nessa área

- **Machine Learning Approach for Trend Prediction to Improve Returns on Brazilian Energy Market**
2022 IEEE Latin America Congress on Computational Intelligence (LA-CCI), em Montevideo, Uruguai.
Conferência de Computação de impacto alto e com *qualis* B2
Artigo: [Machine Learning Approach for Trend Prediction to Improve Returns on Brazilian Energy Market | IEEE Conference Publication | IEEE Xplore](#)
Artigo oriundo de P&D ANEEL realizado em parceria entre Volt Robotics e Auren Energia.
- **Gestão de Dados e Inteligência Artificial aplicados na previsão de tendências de preços no Mercado Livre de Energia**
XIII Congresso Brasileiro de Planejamento Energético - CBPE 2022 · 26 de ago de 2022
Artigo disponibilizado nos anais do evento: [XIII CBPE](#)
Artigo oriundo de P&D ANEEL realizado em parceria entre Volt Robotics e Auren Energia.

Alguns tipos de aplicação de IA...



DADOS



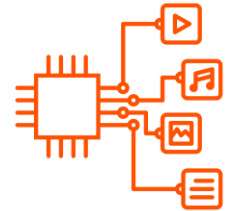
EXPLICAÇÃO



CLUSTERIZAÇÃO



PREVISÃO



GENERATIVA

Alguns tipos de aplicação de IA...



DADOS



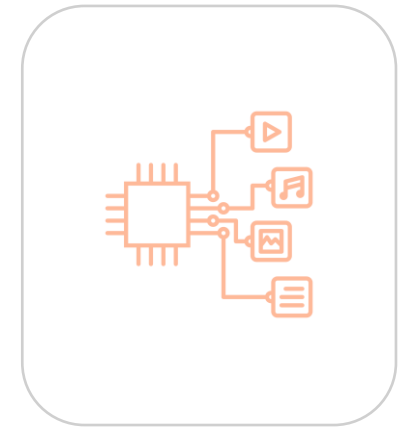
EXPLICAÇÃO



CLUSTERIZAÇÃO



PREVISÃO



GENERATIVA

IBGE

Outros idiomas

Instituto Brasileiro de Geografia e Estatística

Buscar

Acesso à Informação

Dados Abertos

Dados Abertos

Nesta seção, é divulgado o Plano de Dados Abertos (PDA) no âmbito do Instituto Brasileiro de Geografia e Estatística (IBGE).

O Decreto nº 8.777/2016, que instituiu a Política de Dados Abertos do Poder Executo Federal, determinou que os órgãos e as entidades da administração pública federal direta, autárquica e fundacional elaborem e publiquem em sítio eletrônico, até 11 de julho de 2016, o Plano de Dados Abertos.

Plano de Dados Abertos (PDAs)

2024 - 2025

2020 - 2022 (1ª Revisão)

Nos últimos anos, esforços importantes vêm sendo realizados para a organização das bases de dados das instituições ligadas ao governo...

...E com o Setor Elétrico Brasileiro não é diferente...

ACESSIBILIDADE

ALTO CONTRASTE

ENGLISH

Acesso à informação

Política monetária

Estabilidade financeira

BANCO CENTRAL DO BRASIL

Estatísticas

Cédulas e moedas

Publicações e pesquisa

Home

Acesso à Informação

Transparência e prestação de contas

Dados abertos

Dados Abertos

Por meio da disponibilização de dados que estão em sua posse, o Banco Central do Brasil participa da Política de Dados Abertos do Poder Executivo Federal.

A divulgação dos dados gera inúmeros efeitos positivos à sociedade:

fomenta o controle social;

racionaliza gastos; e

estimula a inovação em setores que prestem serviços aos cidadãos.

Em sua essência, o PDA/BC contribui para fortalecer a cultura da transparência, reforça a credibilidade institucional e colabora para o cumprimento da missão institucional de garantir a estabilidade do poder de compra da moeda, zelar por um sistema financeiro sólido, eficiente e competitivo, e fomentar o bem-estar econômico da sociedade.

gov.br

Governo Federal

Órgãos do Governo

Acesso à Informação

Legislação

Acessibilidade

Entrar

Dados Abertos

Conjunto de Dados Ab...

Conjuntos de dados abertos

ONS

Operador Nacional do Sistema Elétrico

DADOS ABERTOS

CONJUNTOS DE DADOS

ORGANIZAÇÕES

GRUPOS

SOBRE

PESQUISAR DADOS

dados abertos

ccee

conjuntos de dados

organizações

grupos

sobre

pesquisar

E-ANEEL

Conjuntos de dados

Organizações

Grupos

Sobre

Pesquisar

Pesquisar dados

Ex.: meio ambiente

epe

20 anos

Empresa de Pesquisa Energética

Dados abertos são dados acessíveis ao público, representados em meio digital, estruturados em formato aberto, processáveis por máquina, referenciados na internet e disponibilizados sob licença aberta que permita sua livre utilização, consumo ou cruzamento, desde que creditada a autoria ou a fonte (conforme Decreto nº 8.777, de 11 de maio de 2016, e Resolução nº 3

Radar

obs

Acesso Restrito

f

YouTube

Instagram

LinkedIn

A EPE

ÁREAS DE ATUAÇÃO

LEILÕES

PUBLICAÇÕES

IMPRENSA

ACESSO À INFORMAÇÃO

TRANSPARÊNCIA E PRESTAÇÃO DE CONTAS

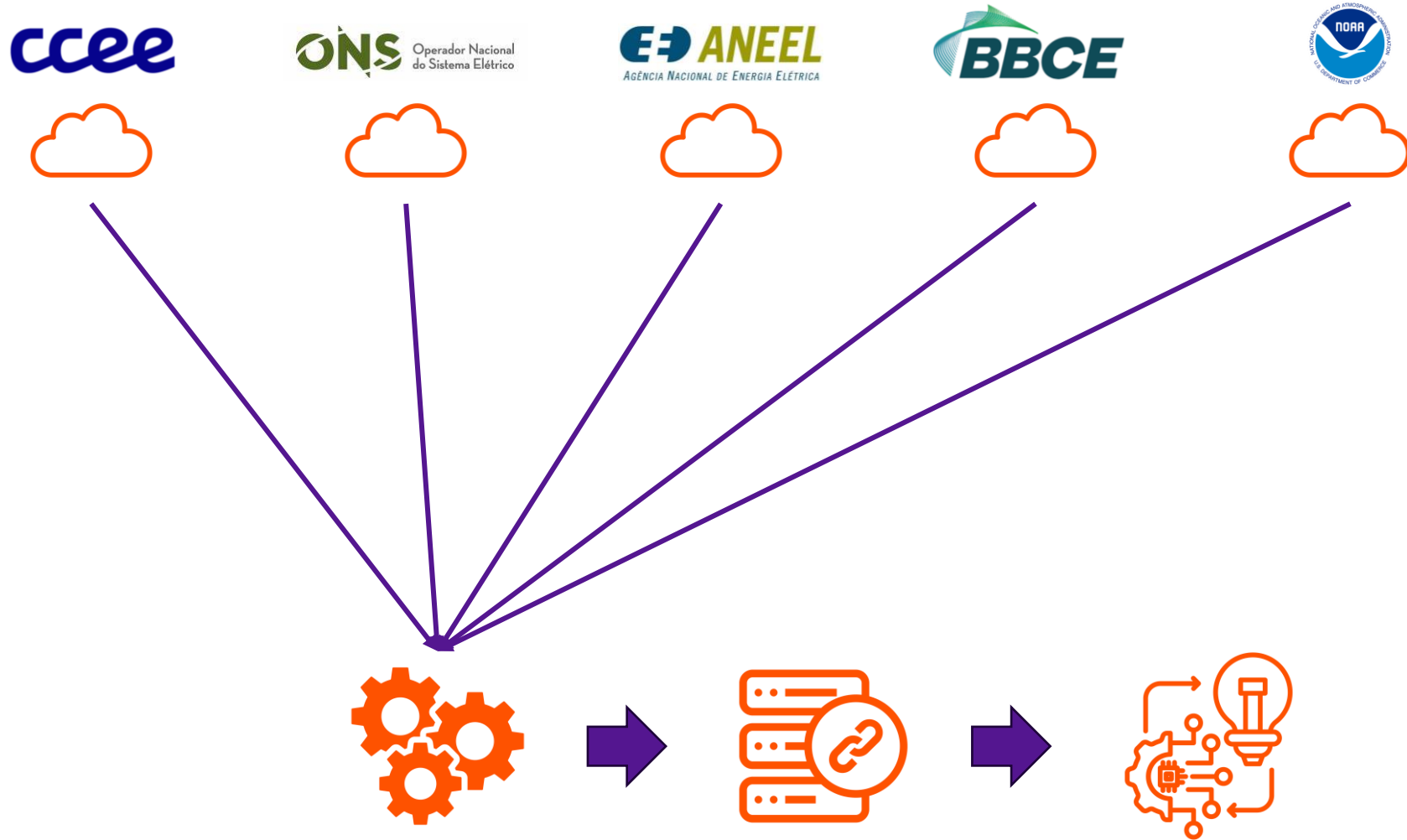
Principal

Publicações

Dados Abertos

Dados Abertos

Conectores...



Lack of Data Freshness Results in Old and Error-Prone Data

Slow and stale data is more than just a nuisance: it spells potential disaster for end users' business decisions. Most data leaders (71%) admit that their end data users make business decisions with old or error-prone data at least some of the time—including 36% who say this happens often or all the time

Data End Users Make Business Decisions With Old Or Error-Prone Data N=300



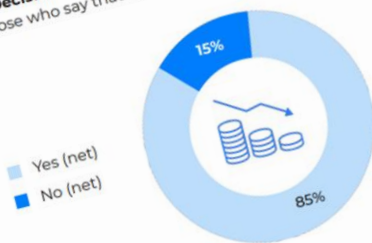
Insights by Company:

- More than 2 in 5 (44%) larger companies (>3,500 employees) say these bad decisions happen often or all the time.
- Younger companies (30 years or less) are twice as likely to say this happens often or all of the time, 51% to 24%.

This Bad Data Is Costing Companies Money

The impact of making business decisions with old or error-prone data is hard to overstate: a massive 85% of data leaders say that it has cost their company money. This applies to companies both large and small.

Business Decisions Made Based On Old Or Error-Prone Data Have Cost Company Money? Among those who say that data end users make business decisions with old or error-prone data n=213



Insights by Company:

This is slightly more common at smaller companies (91%), but more than 4 of 5 data leaders at larger companies (81%) still report that business decisions based on bad data have cost their company money.

A principal matéria prima para qualquer modelo bem sucedido, sobretudo os de **Inteligência Artificial**, são os dados...

...Mas quaisquer dados?

What is GIGO?



Garbage In, Garbage Out

Don't blame the computer if its output is faulty. Output is only as good as input. In other words, the information the computer gives out is only as good as the information that went in. IT stands for Information Technology (computers).

É preciso ter muito cuidado com os dados que são utilizados em nossos modelos...

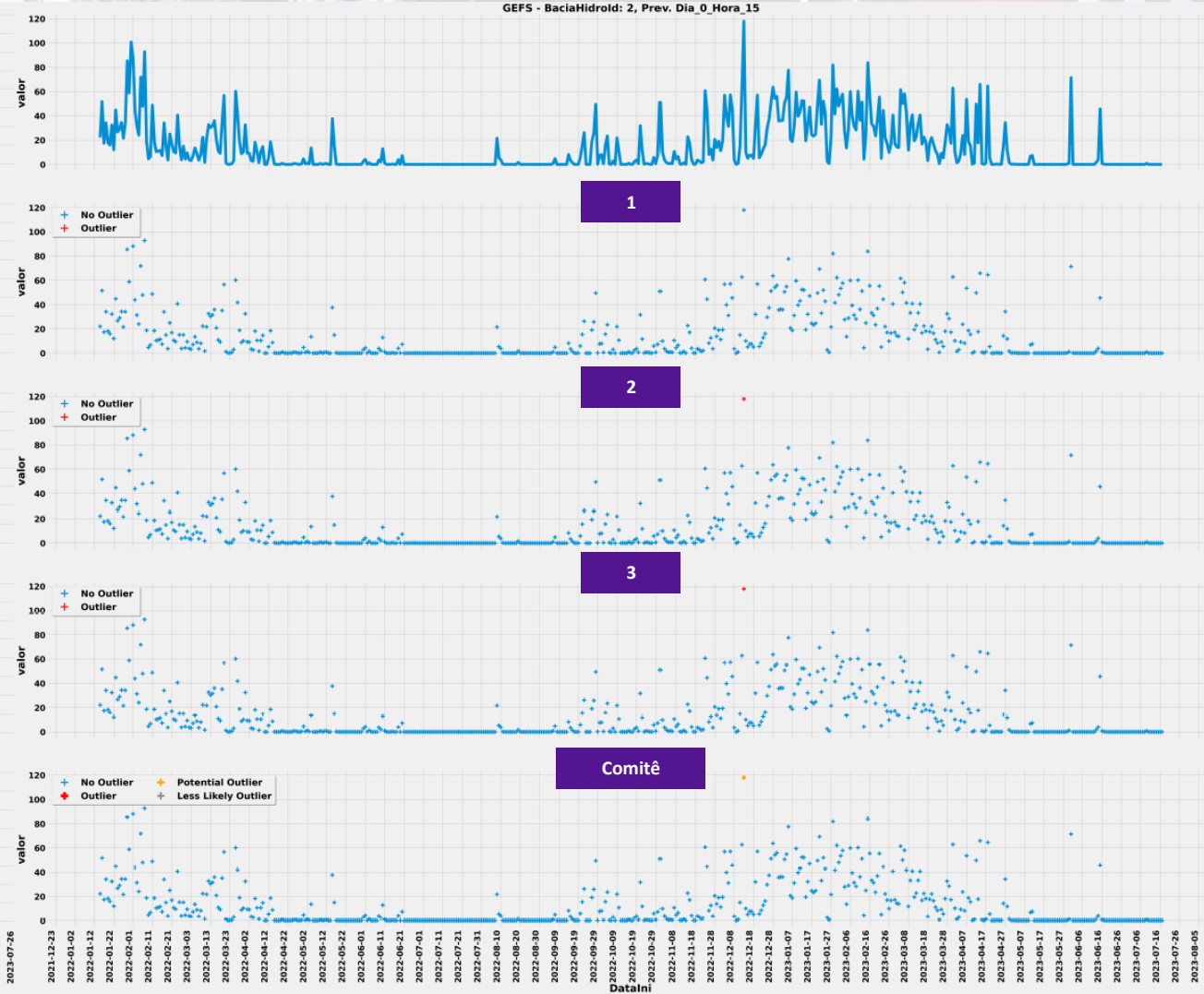
...Dados errados geram padrões de aprendizado errados para os robôs...

...Padrões de aprendizado errados levam a conclusões erradas...

...Conclusões erradas levam a decisões erradas...

...Decisões erradas levam a perdas inesperadas!!!

Os dados precisam ser investigados...



Alguns tipos de aplicação de IA...



DADOS



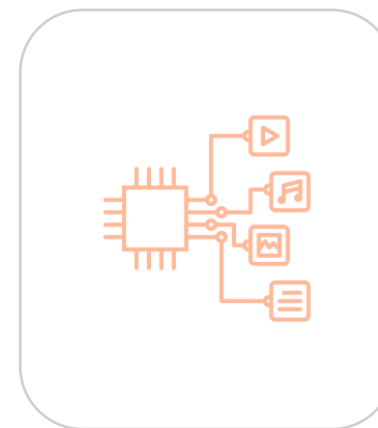
EXPLICAÇÃO



CLUSTERIZAÇÃO



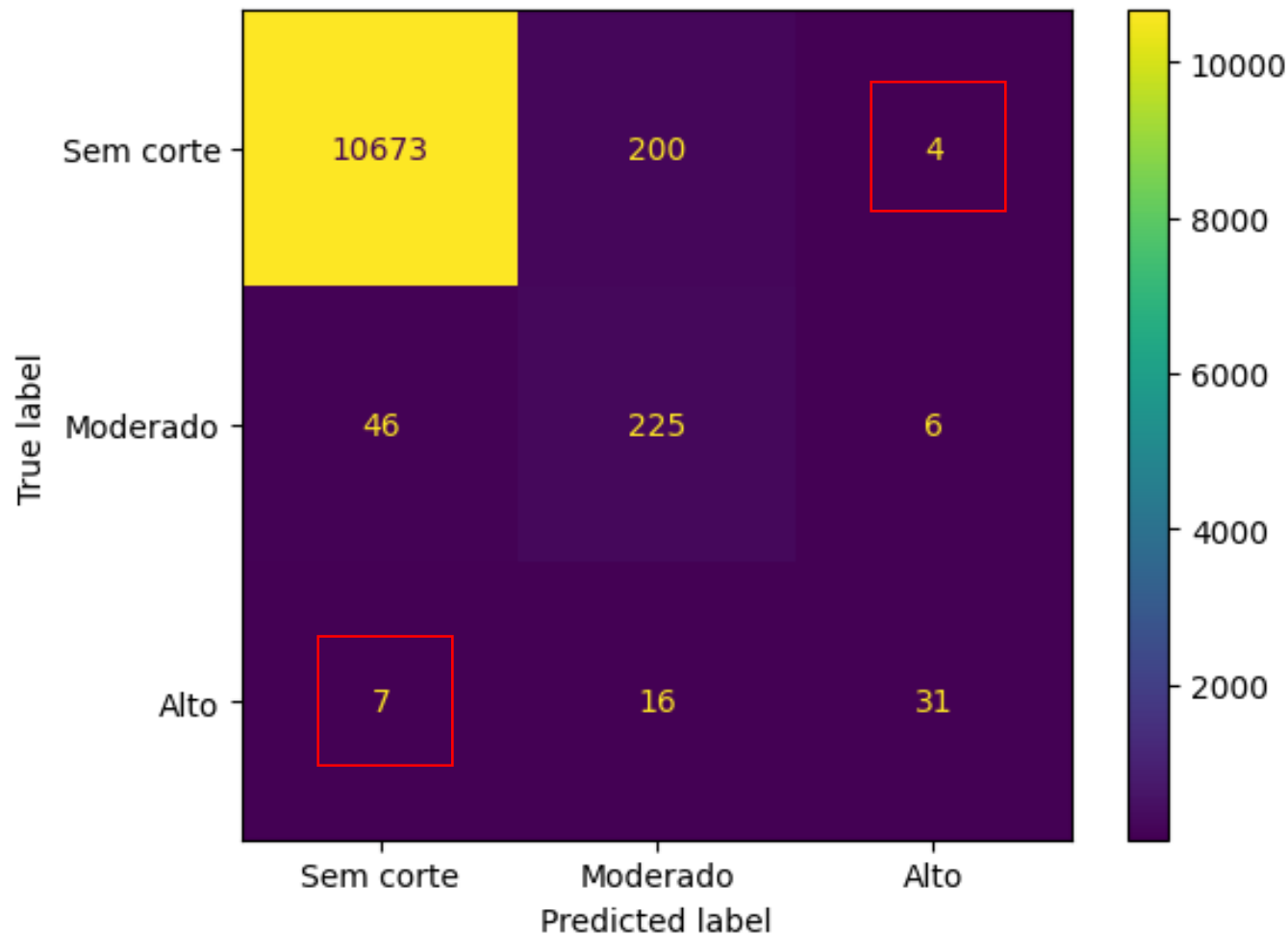
PREVISÃO



GENERATIVA

EXPLAINABLE AI

Explicando curtailment eólico

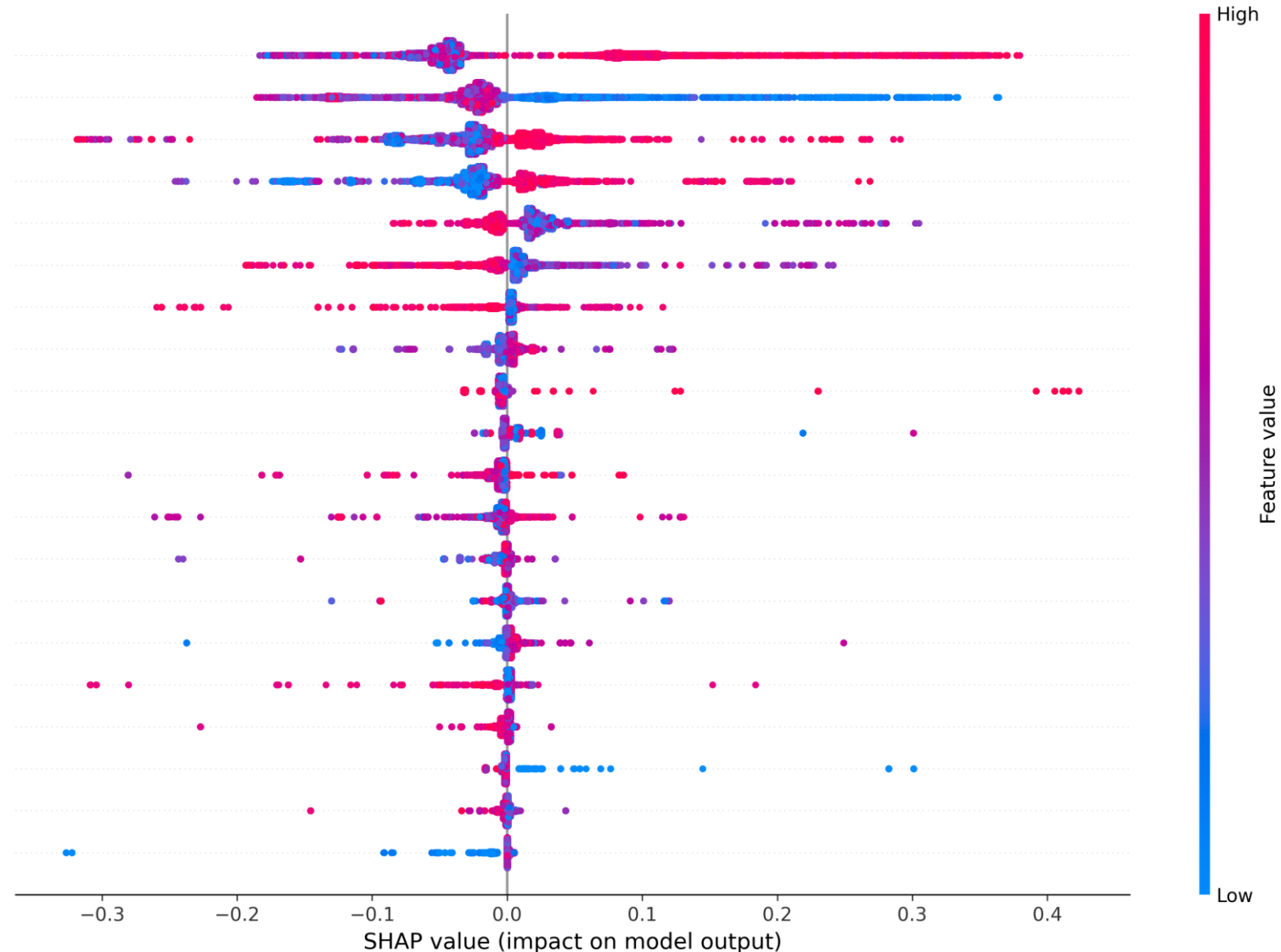


- Acurácias balanceadas:
 - CNF: 80%
 - REL: 74%
 - ENE: 67%
- As acurácias caem à medida que os problemas se tornam mais desbalanceados
- Temos interesse em casos controversos

EXPLAINABLE AI

Explicando curtailment eólico

O que será que é mais importante para prever o Curtailment?



Alguns tipos de aplicação de IA...



DADOS



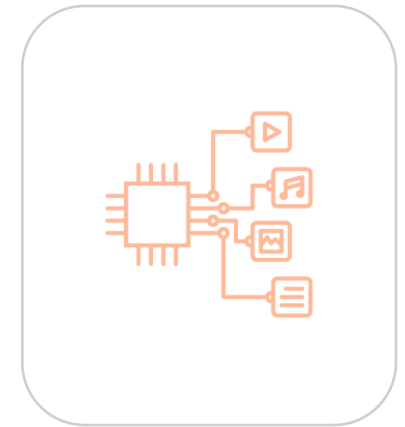
EXPLICAÇÃO



CLUSTERIZAÇÃO



PREVISÃO



GENERATIVA

Clusterização de consumidores

Customer Segmentation in the Brazilian Free Energy Market by Machine Learning
Moisés Santos, Saulo Mastelini, Marcos Basile S. de Paula, Ewerton Guarnier, Donato Silva - Volt Robotics
Maria Santos, Eduardo Campos, Lucas Picarelli - Norte Energia S.A.

Motivation

- The Brazilian Free Energy Market has been experiencing a remarkable transformation led by Deregulation initiatives
- 36% of Brazilian load purchases its energy from Free Market, which tends to grow quickly up to the decade's end
- Considering this Free load increase, the customer's segmentation based on hourly characteristics and behaviors allows Power Companies to understand their needs and offer more tailored and competitive solutions
- The customer's segmentation simplifies the power companies' process because it is possible to assess few groups of customers rather than thousands of them
- This approach will be used in a more complex Project, which will promote energy portfolio optimization considering both energy trading and consumer optimization
- This R&D project is being developed in partnership with VOLT ROBOTICS and NESA (Norte Energia S.A.) through the Research and Development Program regulated by the National Electric Energy Agency (ANEEL), PD-07427-0722/2022

Experimental Procedures

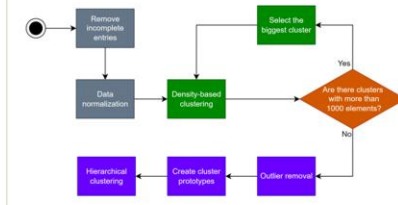
- Dataset: Hourly load consumption from all the Free and Special customers since January 2019, which are available in CCEE (Electric Energy Trading Chamber) - from 2019 to 2022, 12,363 loads have been considered in this analysis
- All consumers are represented by 3 average components: (i) Monthly (47 points), (ii) Hourly - Weekdays (24 points) and (iii) Hourly - Weekend (24 points)
- A step of normalization was used to evaluate all the profiles using the same basis
- Monthly and Hourly distance metrics between consumption profiles were combined to assess the clusters:

$$Md_i = \sum_{m=1}^M |C_{m,i} - \bar{C}_{m,i}|$$
$$Hd_i = \frac{\sum_{h=1}^H |C_{h,i} - \bar{C}_{h,i}| + |C_{wd,i} - \bar{C}_{wd,i}|}{2}$$
$$Td_i = \frac{Md_i + 2Hd_i}{3}$$

- A hybrid Machine Learning Model (HDBSCAN + hierarchical clustering) was used to set clusters
- 3 baselines were considered

Architecture

- Overview of the Hybrid Machine Learning Algorithm:



- The intra-cluster distances (Average and 90th percentile) which were obtained from this hybrid approach were compared to those obtained from 3 baselines:

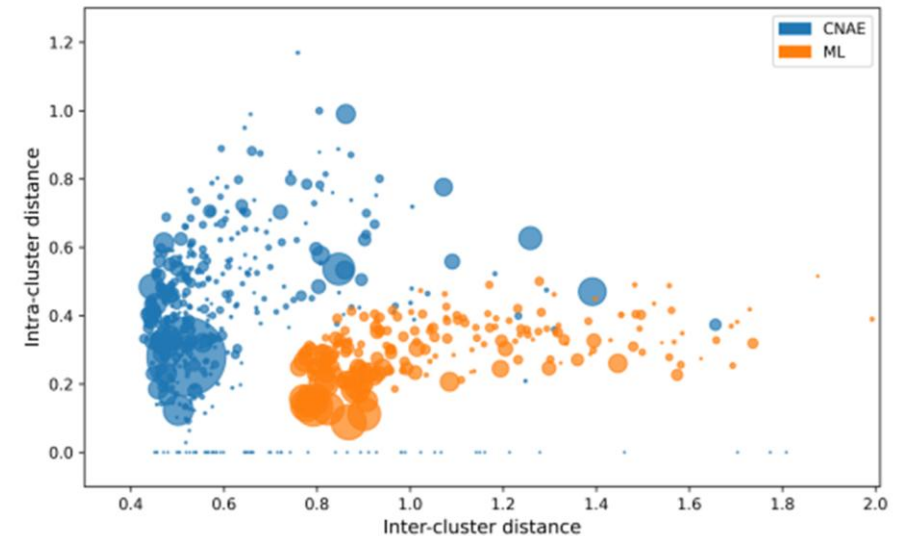
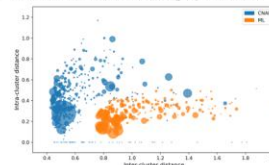
- Flat Load: Consumption constant over time
- Average Load: Average consumption curve calculated considering the whole free load
- CNAE (National Classification of Economic Activities): Official economic categories of activities

Experimental Results

SUMMARIZED RESULTS CONSIDERING THE AVERAGE OF INTRA-CLUSTER DISTANCES AND THE 90-TH PERCENTILE.

Category	Distance	Flat	Average	CNAE	ML
Md	Avg	18.0%	17.1%	14.1%	13.22%
	P90	32.8%	31.0%	25.8%	22.77%
Hd	Avg	23.5%	20.8%	13.5%	7.99%
	P90	60.5%	46.6%	29.5%	13.50%
Td	Avg	21.7%	19.6%	13.7%	9.47%
	P90	48.5%	39.6%	27.1%	15.58%

- The Hybrid Machine Learning approach is notably better than all the baselines, considering both monthly and hourly profiles
- CNAE is a robust baseline, but its dispersion is remarkably worse than Machine Learning solution, mainly in hourly profile - This fact is clearer in the graph below:



Alguns tipos de aplicação de IA...



DADOS



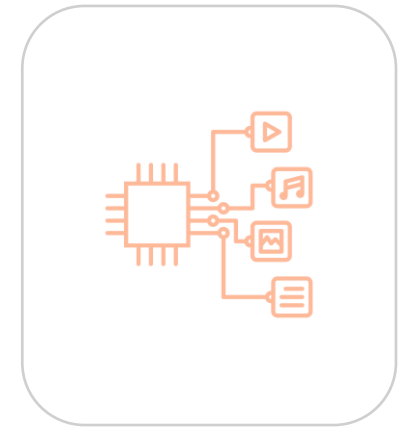
EXPLICAÇÃO



CLUSTERIZAÇÃO



PREVISÃO



GENERATIVA

Por que adotar *Machine Learning* para *Forecasting*?

- Em séries com um elevado número de observações (>200), ML provê resultados tipicamente mais acurados
- A M5 Forecasting Competition é uma competição internacional que avalia modelos de previsão de séries temporais. Na categoria de acurácia, os 50 melhores resultados foram alcançados exclusivamente por modelos de IA

Journal of Intelligent Information Systems (2022) 59:415–433
<https://doi.org/10.1007/s10844-022-00713-9>



A case study comparing machine learning with statistical methods for time series forecasting: size matters

Vitor Cerqueira¹ · Luis Torgo¹ · Carlos Soares^{2,3,4}

Received: 20 January 2022 / Revised: 21 April 2022 / Accepted: 21 April 2022 /
Published online: 16 May 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

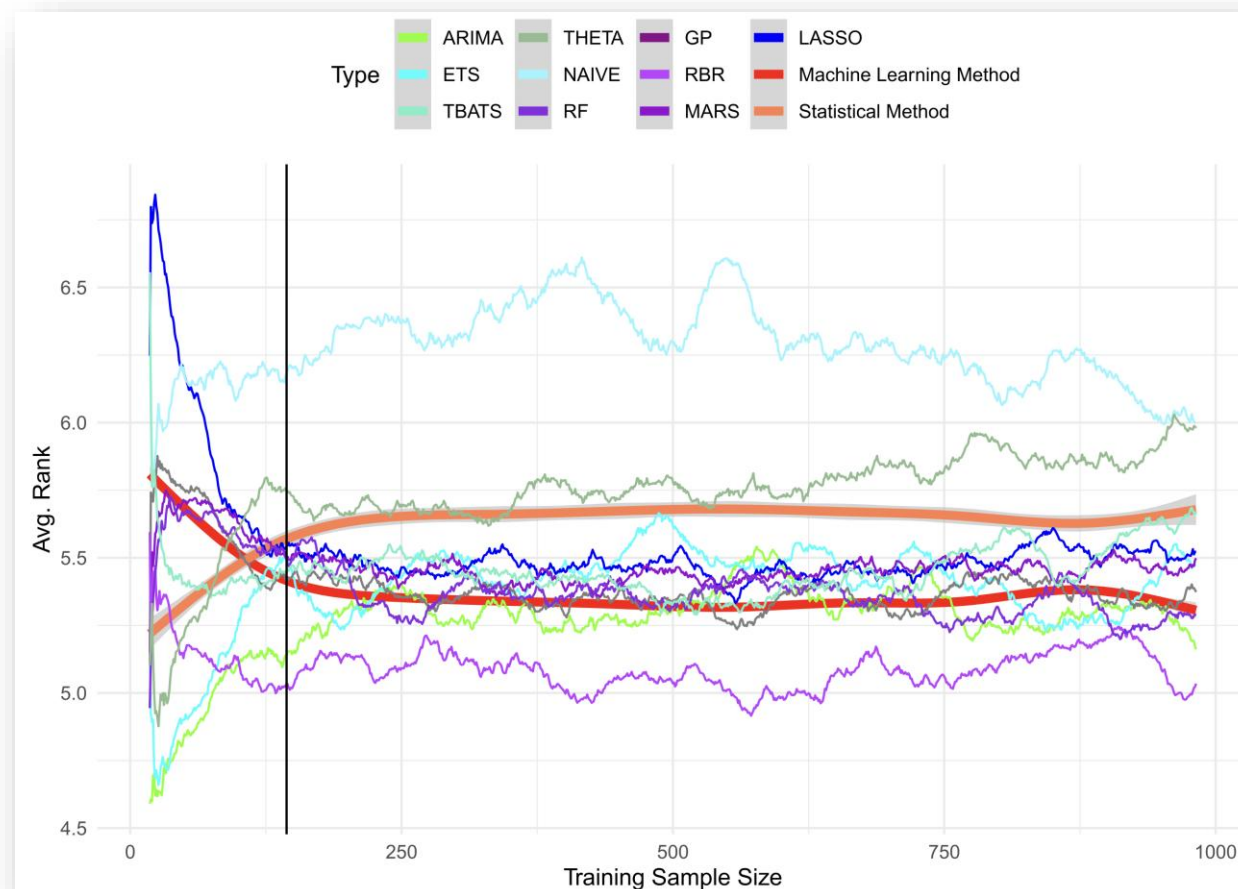
Abstract

Time series forecasting is one of the most active research topics. Machine learning methods have been increasingly adopted to solve these predictive tasks. However, in a recent work, evidence was shown that these approaches systematically present a lower predictive performance relative to simple statistical methods. In this work, we counter these results. We show that these are only valid under an extremely low sample size. Using a learning curve method, our results suggest that machine learning methods improve their relative predictive performance as the sample size grows. The R code to reproduce all of our experiments is available at <https://github.com/vcerqueira/MLforForecasting>.

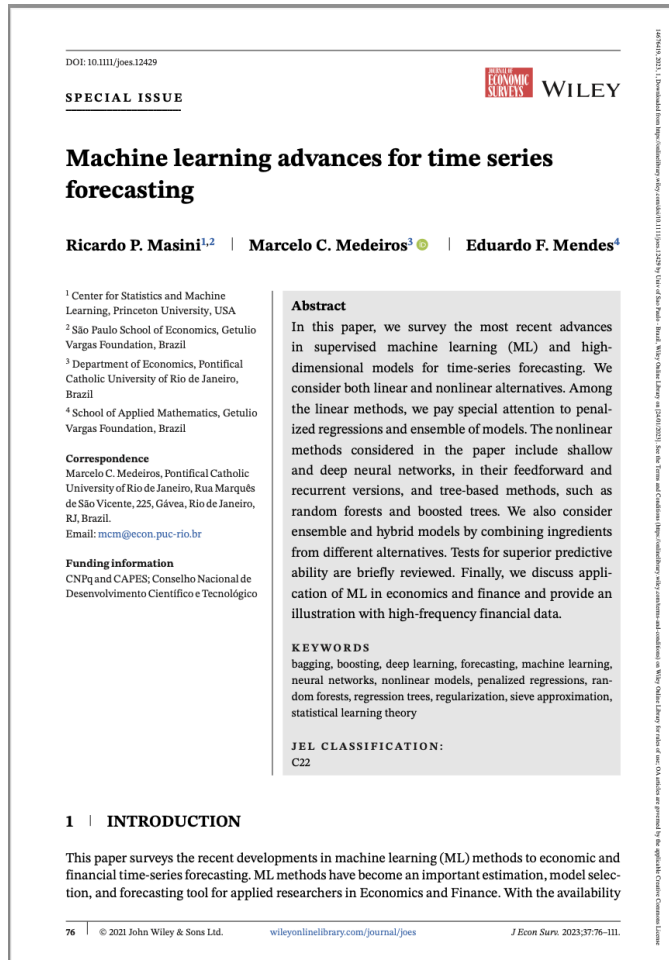
Keywords Time series · Forecasting · Sample size

1 Introduction

Machine learning is a subset of the field of artificial intelligence, which is devoted to developing algorithms that automatically learn from data (Michalski et al., 1983). This area has been at the centre of significant advances in science and technology. This includes problems involving forecasting, such as in the domains of energy (Voyant et al., 2017), healthcare (Lee & Mark, 2010), management (Carbonneau et al., 2008), or climate (Xingjian et al., 2015).



Machine Learning no contexto de previsão de séries temporais



- O uso de *Machine Learning* (ML) para *forecasting* é uma tendência crescente;
- Dentre os grupos de modelos de ML mais utilizados:
 - *Regularized linear regression: Ridge Regression e LASSO*
 - *Artificial Neural Networks (ANN): Multilayer Perceptron e Long Short-Term Memory (LSTM)*
 - *Ensembles: Random Forest e Gradient Boosted Trees*
 - Métodos híbridos
- ANNs são, em geral, mais difíceis de configurar e necessitam de mais dados para treinamento

Disponibilidade de dados e desempenho

Quanto mais dados melhor, mas isso deve ser feito com cautela...

Colunas = variáveis explicativas/dimensões

Quanto mais, melhor...mas CUIDADO!!!

Linhas =
amostras/histórico

*Quanto mais,
melhor!*



Data	VARIÁVEL A SER PREVISTA (Y)	X1	X2	X3	X4	X5	...	Xn
01/09/2023	149	1	9	20	20	9	...	12
31/08/2023	16	2	7	16	4	18	...	5
30/08/2023	3	5	19	17	2	2	...	7
29/08/2023	2	9	14	14	17	1	...	7
28/08/2023	16	11	2	7	4	3	...	11
27/08/2023	1	13	13	14	13	19	...	8
26/08/2023	9	19	16	11	17	3	...	6
25/08/2023	5	12	17	15	6	2	...	4
24/08/2023	2	14	10	6	18	18	...	12
23/08/2023	6	7	1	10	17	12	...	16
22/08/2023	5	20	1	12	13	8	...	8
21/08/2023	15	9	3	2	18	20	...	15
20/08/2023	9	12	13	18	9	18	...	10
19/08/2023	15	8	4	9	17	20	...	18
18/08/2023	6	4	8	6	12	13	...	9
17/08/2023	8	19	19	5	2	14	...	17
16/08/2023	7	4	5	2	19	12	...	7
15/08/2023	17	10	19	7	5	3	...	13
14/08/2023	20	14	13	2	7	14	...	5
13/08/2023	2	13	8	13	4	3	...	14
12/08/2023	15	5	18	7	6	19	...	10
11/08/2023	10	15	3	3	6	11	...	5
...	...	...	...	...	...	...	...	...
DD/MM/AAAA	10	16	14	9	13	11	...	11

Redução de Dimensionalidade: Extração ou Seleção de Atributos?

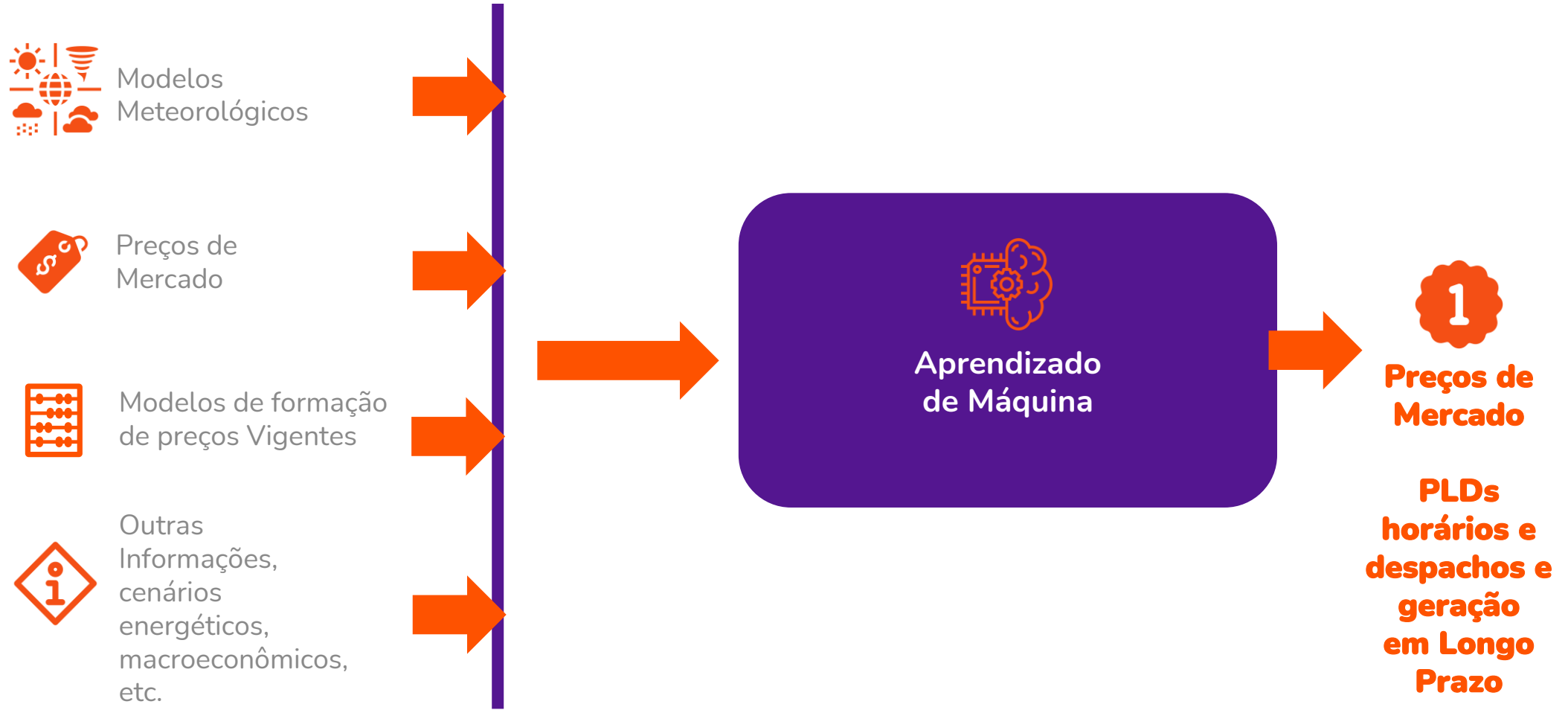
Seleção de Atributos

- Seleção de um subconjunto de variáveis explicativas
- Critérios que buscam reduzir redundâncias e levar o problema à menor dimensão possível
- Tentativa de se excluir de antemão variáveis que não exercem influência no resultado

Extração de Atributos

- Criação de uma nova representação de conjuntos de dados, convertendo para menor dimensão
- Possibilidade de criação de algoritmos que sintetizem variáveis de entrada a partir do conjunto original, perdendo, porém, explicabilidade original

ML para previsão de preços e geração



Muito obrigado!!!

Marcos.basile@voltrobotics.com.br