



UNIVERSIDADE FEDERAL DO AMAZONAS - UFAM  
INSTITUTO DE COMPUTAÇÃO- ICOMP  
PROGRAMA PÓS-GRADUAÇÃO EM INFORMÁTICA - PPGI

# Do Fenômeno à Reação: Descritores da Mudança de Conceito como Evidências para Estratégias de Reação

Albert França Josué Costa

Manaus  
Dezembro 2025

Albert França Josuá Costa

# Do Fenômeno à Reação: Descritores da Mudança de Conceito como Evidências para Estratégias de Reação

Tese apresentada ao Programa de Pós-Graduação  
em Informática do Instituto de Computação da  
Universidade Federal do Amazonas como requi-  
sito parcial para a obtenção do título de Doutor em  
Informática.

Orientadora

Eulanda Miranda dos Santos, Dra.

Coorientador

Rafael Giusti, Dr.

Universidade Federal do Amazonas - UFAM

Instituto de Computação- IComp

Manaus

Dezembro 2025

Ficha Catalográfica

Elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

---

C837f      Costa, Albert França Josuá  
Do Fenômeno à Reação: Descritores da Mudança de Conceito como  
Evidências para Estratégias de Reação / Albert França Josuá Costa. - 2025.  
147 f. : il., color. ; 31 cm.

Orientador(a): Eulanda Miranda dos Santos.  
Coorientador(a): Rafael Giusti.  
Tese (doutorado) - Universidade Federal do Amazonas, Programa de  
Pós-Graduação em Informática, Manaus, 2025.

1. Mudança de Conceito. 2. Descritores. 3. Estratégia de Reação. 4.  
Classificação Taxonômica. I. Santos, Eulanda Miranda dos. II. Giusti,  
Rafael. III. Universidade Federal do Amazonas. Programa de Pós-  
Graduação em Informática. IV. Título

---



Ministério da Educação  
Universidade Federal do Amazonas  
Coordenação do Programa de Pós-Graduação em Informática

### FOLHA DE APROVAÇÃO

## "DO FENÔMENO À REAÇÃO: DESCRITORES DA MUDANÇA DE CONCEITO COMO EVIDÊNCIAS PARA ESTRATÉGIAS DE REAÇÃO"

**ALBERT FRANÇA JOSUÁ COSTA**

Tese de Doutorado defendida e aprovada pela banca examinadora constituída pelos professores:

Profa. Dra. Eulanda Miranda dos Santos - **Presidente**

Prof. Dr. José Francisco de Magalhães Netto - **Membro Interno**

Prof. Dr. Marco Antônio Pinheiro de Cristo - **Membro Externo**

Prof. Dr. Alceu de Souza Britto Junior - **Membro Externo**

Profa. Dra. Elloá Barreto Guedes da Costa - **Membro Externo**

Manaus, 02 de dezembro de 2025.



Documento assinado eletronicamente por **José Francisco de Magalhaes Netto, Professor do Magistério Superior**, em 09/12/2025, às 14:29, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Eulanda Miranda dos Santos, Professor do Magistério Superior**, em 10/12/2025, às 11:24, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Marco Antônio Pinheiro de Cristo, Professor do Magistério Superior**, em 10/12/2025, às 15:20, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

---



Documento assinado eletronicamente por **Elloá Barreto Guedes da Costa, Usuário Externo**, em 13/12/2025, às 21:19, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

---



A autenticidade deste documento pode ser conferida no site [https://sei.ufam.edu.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufam.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **2862835** e o código CRC **76830A0C**.

---

Avenida General Rodrigo Octávio, 6200 - Bairro Coroadó I Campus Universitário Senador Arthur Virgílio Filho, Setor Norte - Telefone: (92) 3305-1181 / Ramal 1193  
CEP 69080-900, Manaus/AM, [coordenadorppgi@icomp.ufam.edu.br](mailto:coordenadorppgi@icomp.ufam.edu.br)

---

Referência: Processo nº 23105.046910/2025-22

SEI nº 2862835

---

## **Agradecimentos**

---

Aos trabalhadores brasileiros, que financiaram esta tese.

À Professora Doutora Eulanda Miranda dos Santos pela orientação contínua ao longo de uma década.

Ao Professor Doutor Rafael Giusti pela orientação prestada.

À Elda Ely Gomes de Souza, pelo que não cabe nesta tese.

# Do Fenômeno à Reação: Descritores da Mudança de Conceito como Evidências para Estratégias de Reação

Autor: Albert França Josuá Costa

Orientadora: Eulanda Miranda dos Santos, Dra.

Coorientador: Rafael Giusti, Dr.

## Resumo

A rápida evolução dos dados em fluxo impõe desafios significativos aos modelos de aprendizado de máquina, frequentemente expostos ao fenômeno da mudança de conceito, caracterizado por alterações na distribuição conjunta das instâncias ao longo do tempo. Embora abordagens de reação à mudança de conceito sejam amplamente utilizadas, como aprendizado incremental e o retreino completo do classificador, há uma lacuna na literatura quanto ao uso de descritores da mudança de conceito como evidências para orientar a escolha da estratégia de reação mais adequada. Esta tese tem como objetivo demonstrar que descritores da mudança de conceito, incluindo severidade, zona de influência, velocidade, frequência, recorrência e previsibilidade, podem subsidiar a definição de uma estratégia de reação capaz de preservar o desempenho de classificadores em cenários dinâmicos. Para tanto, foram conduzidas três etapas de investigação: (i) revisão sistemática e proposta de uma nova taxonomia de métodos não supervisionados de detecção de mudança de conceito; (ii) análise experimental do impacto dos descritores sobre o desempenho do classificador, utilizando 32 cenários distintos e 5 bases de dados sintéticas; e (iii) avaliação abrangente de cinco estratégias de reação à mudança de conceito (estática, retreino completo, aprendizado incremental, esquecimento e oráculo) sob diferentes configurações de descritores, complementada por testes estatísticos, análise do tamanho de efeito e um estudo de ablação. Os resultados

obtidos evidenciam que a adequação da estratégia de reação está diretamente relacionada às características da mudança. Identificou-se que descritores como severidade e velocidade exercem forte influência no desempenho dos classificadores, enquanto recorrência e frequência se mostram determinantes em cenários de reaparecimento de padrões. Ademais, verificou-se que a estratégia de reação mais amplamente adotada na literatura, o retreino completo, não é plenamente eficaz, sendo que, em determinadas situações, abordagens como o aprendizado incremental com esquecimento ou mesmo a não reação apresentaram desempenho superior. As principais contribuições desta pesquisa incluem: (i) a proposição de uma taxonomia sistemática para métodos não supervisionados de detecção de mudança de conceito; (ii) a caracterização empírica das inter-relações entre descritores e seu impacto no desempenho do classificador; (iii) definição, com base em evidências, da estratégia de reação à mudança; e (iv) a disponibilização de uma estrutura experimental robusta e replicável. Os resultados reforçam a importância de considerar a natureza da mudança de conceito no desenvolvimento de sistemas de aprendizado adaptativos, oferecendo subsídios tanto para a pesquisa acadêmica quanto para aplicações práticas em domínios sensíveis, como saúde, segurança e finanças.

*Palavras-chave:* Mudança de Conceito, Descritores, Estratégia de Reação, Classificação Taxonômica.



# Do Fenômeno à Reação: Descritores da Mudança de Conceito como Evidências para Estratégias de Reação

Autor: Albert França Josuá Costa

Orientadora: Eulanda Miranda dos Santos, Dra.

Coorientador: Rafael Giusti, Dr.

## Abstract

The rapid evolution of data streams poses significant challenges to machine learning models, which are frequently exposed to the phenomenon of concept drift, characterized by changes in the joint distribution of instances over time. Although adaptive approaches such as incremental learning and full retraining are widely employed, the literature still presents a gap regarding the use of concept drift descriptors as evidence to guide the choice of the most suitable reaction strategy. This thesis aims to demonstrate that concept drift descriptors, including severity, influence zone, speed, frequency, recurrence, and predictability, can support the definition of a reaction strategy capable of preserving classifier performance in dynamic scenarios. To this end, three stages of investigation were conducted: (i) a systematic review and the proposal of a new taxonomy for unsupervised concept drift detection methods; (ii) an experimental analysis of the impact of descriptors on classifier performance across 32 distinct scenarios and 5 synthetic datasets; and (iii) a comprehensive evaluation of five reaction strategies to concept drift (static, full retraining, incremental learning, forgetting, and oracle) under different descriptor configurations, complemented by statistical tests, effect size analysis, and an ablation study. The results show that the adequacy of the reaction strategy is directly related to the characteristics of the drift. It was found that descriptors such as severity and speed strongly influence classifier performance, while recurrence and frequency

are determinant in scenarios of pattern reappearance. Furthermore, the widely adopted strategy in the literature, full retraining, proved not to be fully effective, as in certain situations approaches such as incremental learning with forgetting, or even no reaction, yielded superior performance. The main contributions of this research include: (i) the proposition of a systematic taxonomy for unsupervised concept drift detection methods; (ii) the empirical characterization of the interrelationships among descriptors and their impact on classifier performance; (iii) the evidence-based definition of reaction strategies to concept drift; and (iv) the presentation of a robust and replicable experimental framework. The findings highlight the importance of considering the nature of concept drift when developing adaptive learning systems, offering support both for academic research and for practical applications in sensitive domains such as healthcare, security, and finance.

*Keywords:* Concept Drift, Descriptors, Reaction Strategies, Taxonomic Classification.

Figura 1 – Exemplo de mudança de conceito: a fronteira de decisão é deslocada entre os conceitos, o que pode ocasionar perda de desempenho do classificador. A região sombreada indica a diferença entre os conceitos.	26
Figura 2 – Componentes da mudança de conceito: a) Fontes Probabilísticas e b) Descritores.	27
Figura 3 – Fontes probabilísticas da mudança de conceito.	27
Figura 4 – Exemplo de mudança de conceito virtual ocasionada por mudança em $p(x)$ .	30
Figura 5 – Exemplo de mudança de conceito virtual ocasionada por mudança em parte dos atributos.	31
Figura 6 – Exemplo de mudança de conceito virtual ocasionada por surgimento de novo atributo.	32
Figura 7 – Exemplo de mudança de conceito real ocasionada por variação no balanceamento entre as classes.	33
Figura 8 – Exemplo de mudança de conceito real ocasionada por evolução do conceito.	33
Figura 9 – Exemplo de mudança de conceito real ocasionada pelo desaparecimento de uma classe.	34
Figura 10 – Exemplo de mudança de conceito real do tipo I ocasionada por mudança em $p_{t_0}(y x) \neq p_{t_n}(y x)$ e $p_{t_0}(x) \neq p_{t_n}(x)$ .	35
Figura 11 – Exemplo de mudança de conceito real do tipo II ocasionada por mudança em $p_{t_0}(y x) \neq p_{t_n}(y x)$ e $p_{t_0}(x) = p_{t_n}(x)$ .	36
Figura 12 – Exemplo de mudança de conceito real inconsistente - $\exists x(\arg \max p_{t_0}(y x) = c_1 \cap \arg \max p_{t_n}(y x) = c_2)$ .	36

Figura 13 – Exemplo de mudança de conceito real plena - $\forall x(\arg \max p_{t_0}(y x) = c_1 \cap \arg \max p_{t_n}(y x) = c_2)$ . . . . .	37
Figura 14 – Exemplo de mudança de conceito real cruzada - $\exists x(\arg \max p_{t_0}(y x) = c_1 \cap \arg \max p_{t_n}(y x) = c_2) \cap \exists z(\arg \max p_{t_0}(y z) = c_1 \cap \arg \max p_{t_n}(y z) = c_2)$ . . . . .	38
Figura 15 – Mudança de conceito na perspectiva do seus descritores. . . . .	40
Figura 16 – Descritor de severidade. O lado esquerdo representa uma mudança de conceito de baixa severidade: as áreas em amarelo em (a) e (b) indicam, respectivamente, as regiões que devem ser esquecidas e aprendidas após a mudança. O lado direito representa uma mudança de conceito de alta severidade: as áreas em amarelo em (c) e (d) indicam, respectivamente, as regiões que devem ser esquecidas e aprendidas após a mudança. . . . .	42
Figura 17 – Descritor de zona de influência. (a) representa o espaço de entrada do conceito inicial. Nos exemplos (b) e (c), as regiões destacadas em amarelo apresentam probabilidade conjunta de classificar as instâncias diferentes em relação à (a), tornando-as regiões de conflito entre os conceitos. . . . .	44
Figura 18 – Exemplo de mudança de conceito repentina, caracterizada pela substituição do conceito antigo pelo novo em um curto intervalo de tempo, com $V_{mc} = t_{mf} - t_{mi} = 1$ . . . . .	45
Figura 19 – Exemplo de mudança de conceito gradual contínua caracterizada pela substituição suave do conceito antigo pelo conceito novo sem coexistência de conceitos. . . . .	47
Figura 20 – Exemplo de mudança de conceito gradual probabilística caracterizada pela substituição suave do conceito antigo pelo conceito novo com coexistência de conceitos. . . . .	47
Figura 21 – Descritor de velocidade. . . . .	48

Figura 22 – Descritor de Frequência. Dois exemplos de frequência não periódica da mudança de conceito. A primeira mudança entre os conceitos amarelo e verde tem frequência igual a $\frac{1}{4}$ . A segunda mudança tem a frequência definida em $\frac{1}{5}$ (conceito verde e rosa). Na figura $i < j < k < l$ . . . . .	49
Figura 23 – Descritor de Recorrência - exemplo de mudança de conceito recorrente cíclica. As frequências $f_1$ e $f_2$ representam os conceitos 1 e 2, respectivamente. . . . .	50
Figura 24 – Descritor de Recorrência - exemplo de mudança de conceito recorrente acíclica. Os conceitos previamente observados reaparecem em intervalos irregulares. . . . .	51
Figura 25 – Visão geral do experimento. . . . .	67
Figura 26 – Valores das medianas dos erro prequencial nas bases de dados. . . .	73
Figura 27 – Variação da mediana do erro prequencial nos cenários investigados variando o valor de cada descritor considerando os detectores supervisionados. . . . .	76
Figura 28 – Variação da mediana do erro prequencial em todos os cenários investigados variando o valor de cada descritor considerando os detectores semi-supervisionados, não supervisionado e o contexto sem detecção. . . . .	77
Figura 29 – Visão geral do experimento. . . . .	81
Figura 30 – Visão geral das estratégias de reação investigadas neste trabalho. C indica o classificador. . . . .	83
Figura 31 – Valores positivos indicam que a estratégia EP alcançou os menores erros. X indica que não há diferença significativa entre EP e EE, no nível de significância de 0,05. . . . .	85
Figura 32 – Tamanho do efeito dos testes estatísticos. A linha vermelha pontilhada indica o ponto a partir do qual o efeito é considerado forte ( $\eta^2 \geq 0,14$ ). . . . .	86
Figura 33 – O mapa de calor ilustra o desempenho de cada estratégia de reação nos cenários em cada base de dados. . . . .	86

Figura 34 – O mapa de calor ilustra os cenários em que cada estratégia alcançou o melhor resultado. . . . .	90
Figura 35 – Valores positivos indicam que EQ alcançou a menor taxa de erro. X indica que não há diferença significativa entre EQ e EP, no nível de significância 0,05. . . . .	91
Figura 36 – O crescente número de publicações entre os anos de 2009 e 2019 em formato de revisão da literatura sobre detecção de mudança de conceito. . . . .	114
Figura 37 – Taxonomia proposta de métodos não supervisionados de detecção de mudança de conceito. . . . .	117
Figura 38 – Estrutura geral de métodos não supervisionados de detecção de mudança de conceito em lote. . . . .	119
Figura 39 – Estratégias de seleção de instâncias empregadas nos métodos de detecção por lote parcial analisados. . . . .	119
Figura 40 – Estrutura geral de métodos não supervisionados de detecção de mudança de conceito <i>online</i> . . . . .	131
Figura 41 – Cenário DDM. . . . .	144
Figura 42 – Cenário EDDM. . . . .	145
Figura 43 – Cenário DDAL. . . . .	145
Figura 44 – Cenário STUDD. . . . .	146
Figura 45 – Cenário MD3-EGM. . . . .	146
Figura 46 – Cenário DSDD. . . . .	147

---

## Lista de Tabelas

---

Tabela 1 – Resumo das fontes probabilísticas de mudança de conceito. . . . .	39
Tabela 2 – Comparação dos trabalhos relacionados em termos do descritor avaliado. . . . .	60
Tabela 3 – Comparação dos trabalhos relacionados em termos de descritor avaliado. . . . .	65
Tabela 4 – Valores dos descritores em cada cenário. . . . .	71
Tabela 5 – Comparação dos descritores em cada contexto por meio de testes estatísticos. . . . .	75
Tabela 6 – Valores dos descritores para os melhores e os piores resultados. . . .	78
Tabela 7 – Interpretação do tamanho do efeito. . . . .	82
Tabela 8 – Comparando as estratégias de reação em cada cenário em todas as bases de dados utilizando teste estatístico. . . . .	88
Tabela 9 – Comparando as estratégias de reação em cada cenário em todas as bases de dados utilizando teste estatístico. . . . .	89
Tabela 10 – Categorização dos métodos de detecção de mudança de conceito baseados em lote. . . . .	120
Tabela 11 – Categorização dos métodos de detecção de mudança de conceito online.	130
Tabela 12 – Compilação de algumas características de detectores de mudança não supervisionados encontrados na literatura, destacando o tipo de mudança tratada; se há dependência da aplicação; se é realizada autoanotação; e a dependência de dados rotulados. . . . .	139
Tabela 13 – Valores dos parâmetros nas bases de dados <i>Line</i> , <i>Hiperplane</i> , <i>Circle</i> , <i>SineH</i> and <i>Sine</i> por cenário. . . . .	143

<b>1</b>	<b>INTRODUÇÃO</b>	<b>18</b>
<b>1.1</b>	<b>Hipótese de Pesquisa</b>	<b>20</b>
<b>1.2</b>	<b>Problema de Pesquisa</b>	<b>21</b>
<b>1.3</b>	<b>Objetivos</b>	<b>21</b>
<b>1.4</b>	<b>Contribuições da Tese</b>	<b>21</b>
<b>1.5</b>	<b>Publicações</b>	<b>22</b>
<b>1.6</b>	<b>Organização da Tese</b>	<b>23</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>24</b>
<b>2.1</b>	<b>Conceito e Mudança de Conceito</b>	<b>24</b>
<b>2.2</b>	<b>Fontes da Mudança de Conceito</b>	<b>27</b>
2.2.1	Probabilidade <i>a priori</i>	29
2.2.1.1	Probabilidade <i>a priori</i> de $x$	30
2.2.1.2	Probabilidade <i>a priori</i> de $y$	31
2.2.2	Probabilidade condicional de $x$	33
2.2.3	Probabilidade <i>a posteriori</i>	34
2.2.4	Síntese das Fontes de Mudança de Conceito	38
<b>2.3</b>	<b>Descritores da Mudança de Conceito</b>	<b>39</b>
2.3.1	Dimensão Espacial	41
2.3.1.1	Severidade	41
2.3.1.2	Zona de Influência	43
2.3.2	Dimensão Temporal	45
2.3.2.1	Velocidade	45
2.3.2.2	Frequência	49
2.3.2.3	Recorrência	50
2.3.3	Sem Dimensão - Previsibilidade	51
<b>2.4</b>	<b>Considerações Finais</b>	<b>52</b>



<b>3</b>	<b>TRABALHOS RELACIONADOS</b>	<b>53</b>
<b>3.1</b>	<b>Uma Visão Geral dos Métodos de Detecção de Mudança de Conceito</b>	<b>53</b>
<b>3.2</b>	<b>Análise dos Descritores de Mudança de Conceito e seus Impactos.</b>	<b>55</b>
3.2.1	Detectores de Mudança de Conceito	56
3.2.2	Características da Mudanças de Conceitos e Seus Impactos no Desempenho do Classificador	57
3.2.3	Análise crítica-comparativa dos trabalhos relacionados	59
<b>3.3</b>	<b>Mudança de Conceito e Estratégias de Reação: Uma Análise Abrangente com o Uso de Descritores</b>	<b>60</b>
3.3.1	Estudos sobre Detecção de Mudança de Conceito	61
3.3.2	Estudos sobre a Natureza da Mudança de Conceito	62
3.3.3	Uma breve análise comparativa dos trabalhos relacionados	64
<b>3.4</b>	<b>Considerações Finais</b>	<b>65</b>
<b>4</b>	<b>ANÁLISE DOS DESCRITORES DE MUDANÇA DE CONCEITO E SEUS IMPACTOS</b>	<b>66</b>
<b>4.1</b>	<b>Bases de Dados</b>	<b>68</b>
<b>4.2</b>	<b>Cenários</b>	<b>70</b>
<b>4.3</b>	<b>Detectores de Mudança de Conceito</b>	<b>72</b>
<b>4.4</b>	<b>Primeira Série de Experimentos</b>	<b>72</b>
<b>4.5</b>	<b>Segunda Série de Experimentos</b>	<b>74</b>
<b>4.6</b>	<b>Considerações Finais</b>	<b>78</b>
<b>5</b>	<b>MUDANÇA DE CONCEITO E ESTRATÉGIAS DE REAÇÃO: UMA ANÁLISE ABRANGENTE COM O USO DE DESCRITORES</b>	<b>80</b>
<b>5.1</b>	<b>Estratégias de Reação</b>	<b>83</b>
<b>5.2</b>	<b>Bases de Dados e Cenários</b>	<b>84</b>
<b>5.3</b>	<b>Primeiro Estágio do Experimento</b>	<b>84</b>

<b>5.4</b>	<b>Segundo Estágio do Experimento . . . . .</b>	<b>87</b>
<b>5.5</b>	<b>Teste de Ablação . . . . .</b>	<b>90</b>
<b>5.6</b>	<b>Considerações Finais . . . . .</b>	<b>92</b>
<b>6</b>	<b>CONCLUSÕES . . . . .</b>	<b>94</b>
<b>6.1</b>	<b>Limitações e Trabalhos Futuros . . . . .</b>	<b>96</b>
	<b>Referências . . . . .</b>	<b>99</b>
<b>APÊNDICE A</b>	<b>UMA VISÃO GERAL DOS MÉTODOS NÃO SUPERVISIONADOS DE DETECÇÃO DE MUDANÇA DE CONCEITO . . . . .</b>	<b>111</b>
<b>APÊNDICE B</b>	<b>VALORES DOS PARÂMETROS DAS BASES DE DADOS. . . . .</b>	<b>143</b>
<b>APÊNDICE C</b>	<b>MÉTRICAS DE DESEMPENHO DOS DETECTORES EM CADA CENÁRIO POR CONTEXTO. . . . .</b>	<b>144</b>

A rápida evolução dos dados em fluxo desafia os modelos de aprendizado de máquina, que frequentemente enfrentam o fenômeno conhecido como mudança de conceito. Introduzido por [Schlimmer e Granger \(1986\)](#), esse termo descreve alterações na distribuição conjunta das instâncias e suas variáveis-alvo ao longo do tempo, as quais podem impactar significativamente o desempenho dos preditores ([BAYRAM; AHMED; KASSLER, 2022](#)). Essas mudanças podem variar desde mudanças locais menores até mudanças globais, exigindo abordagens que adaptem os modelos para manter alta a acurácia preditiva ([GAMA et al., 2014](#)). Embora os preditores possam ser classificadores ou regressores, este trabalho concentra-se na mudança de conceito em tarefas de classificação.

A perda de desempenho pode tornar o classificador inadequado para sua finalidade, exigindo, portanto, abordagens de adaptação para ajustar o modelo à nova distribuição dos dados. Essas abordagens adaptativas visam preservar ou restaurar o desempenho do classificador a níveis aceitáveis, ajustando sua estrutura ou parâmetros em resposta às mudanças. Exemplos dessas abordagens incluem o aprendizado incremental, no qual o classificador é continuamente atualizado com novas instâncias enquanto mantém o conhecimento previamente adquirido ([LEO; KALITA, 2024](#)); e o retreino periódico, em que o classificador é retreinado utilizando os dados mais recentes, descartando os mais antigos. Essa estratégia parte do pressuposto de que instâncias mais recentes são mais relevantes por representarem melhor o conceito atual, enquanto a relevância dos dados antigos diminui ao longo do tempo ([GEMAQUE et al., 2020](#)).

Na literatura, a estratégia de reação mais comumente empregada é o retreino completo do classificador atual, sem levar em consideração outros aspectos, como a disponibilidade de instâncias rotuladas do novo conceito. Métodos clássicos de detecção

de mudança de conceito como *Drift Detection Method* (DDM) (GAMA et al., 2004) e *Early Drift Detection Method* (EDDM) (BAENA-GARCÍA et al., 2006) utilizam essa estratégia. Embora eficaz em determinados cenários, essa abordagem pode ser onerosa e inadequada, especialmente em ambientes com dados rotulados limitados, altos custos computacionais ou na presença de mudanças de conceito com características complexas, como baixa severidade ou baixo impacto local (COSTA; GIUSTI; SANTOS, 2025).

Uma abordagem mais adequada seria identificar as características das mudanças que afetam os dados, a fim de determinar a estratégia de reação ideal considerando a natureza dessas mudanças. Essas características podem ser capturadas por meio de descritores como severidade, zona de influência, velocidade, frequência, recorrência e previsibilidade. No entanto, existe uma lacuna evidente na literatura em relação a estudos que explorem a natureza da mudança de conceito e a conexão entre as características da mudança, definidas por seus descritores, e a adequação de diferentes estratégias de reação.

Por outro lado, compreender as relações entre os descritores da mudança e estratégias de reação é essencial para enfrentar problemas do mundo real em ambientes dinâmicos, uma vez que, nesses ambientes, a adaptabilidade e a tomada de decisão baseada em dados são requisitos fundamentais. Mouchaweh (2016) estabelece que o ambiente dinâmico é caracterizado pela chegada de dados em fluxo contínuo e potencialmente ilimitado, em alta velocidade, e que a distribuição que os gera pode modificar ao longo do tempo.

Por exemplo, na área da saúde, modelos de aprendizado de máquina são utilizados para prever o comportamento de doenças, analisar sinais de dispositivos médicos e interpretar imagens clínicas. Quando empregados em ambientes que sofrem mudanças ao longo do tempo e estão sujeitos à mudança de conceito, os descritores de recorrência e frequência desempenham papel central na definição da estratégia ideal de reação, especialmente em cenários envolvendo doenças sazonais. Nesses casos, retrainar completamente o classificador ao final de cada manifestação da doença pode ser um processo oneroso e desnecessário, pois o mesmo padrão pode reaparecer posteriormente (ROTALINTI et al., 2023; RIOS et al., 2020; KORE et al., 2024). Contudo, em situações

de pandemia, como a COVID-19, ocorre uma mudança abrupta nos padrões de comportamento. Nesses casos, os descritores de velocidade e severidade tornam-se mais importantes na definição da estratégia de reação ideal (SUSNJAK; MADDIGAN, 2023).

Outro exemplo de problema do mundo real propenso à mudança de conceito é a detecção de fraudes (ambientes adversariais). Nesse caso, fraudadores frequentemente alteram seus padrões em curtos intervalos com o objetivo de validar comportamentos fraudulentos. Portanto, qualquer estratégia de reação deve considerar se o tempo necessário para atualizar o classificador é suficiente para torná-lo eficaz antes que ocorra uma nova mudança. Além disso, em muitos cenários reais, a rotulagem de instâncias requer intervenção humana, o que inviabiliza o uso de estratégias como o aprendizado incremental (FARRUGIA et al., 2021; SHAMITHA; ILANGO, 2022), já que esses métodos dependem da disponibilidade contínua de rótulos verdadeiros. Assim, a ausência de diretrizes que associem características específicas da mudança com as estratégias de reação mais adequadas dificulta a adoção efetiva de abordagens adaptativas em cenários do mundo real.

Apesar dos avanços na detecção e caracterização de mudanças de conceito, ainda há uma lacuna significativa na definição de estratégias de reação orientadas por descritores. A literatura apresenta estudos isolados que relacionam alguns descritores, no entanto, carece de uma análise abrangente e sistemática que utilize o conjunto de descritores como evidências para a escolha da estratégia de reação mais adequada.

Dado esse contexto, esta tese assume como hipótese e problema de pesquisa:

## 1.1 Hipótese de Pesquisa

A estimativa e a utilização dos descritores de mudança de conceito, como severidade, zona de influência, velocidade, frequência, recorrência e previsibilidade, permitem selecionar estratégias de reação capazes de preservar ou melhorar, em média, o desempenho do classificador quando comparadas à estratégia de reagir treinando um novo classificador.

## 1.2 Problema de Pesquisa

Como os descritores da mudança de conceito podem ser utilizados como evidências para orientar a seleção da estratégia de reação mais adequada, de modo a preservar o desempenho de classificadores em cenários dinâmicos?

## 1.3 Objetivos

O objetivo principal desta tese é demonstrar que descritores da mudança de conceito podem ser utilizados como evidências para a definição da estratégia de reação à mudança mais adequada.

Os objetivos específicos são:

1. Conduzir uma revisão sistemática da literatura sobre detecção não supervisionada de mudança de conceito.
2. Propor uma nova classificação taxonômica sobre detecção não supervisionada de mudança de conceito;
3. Identificar os descritores que têm maiores impactos no desempenho do classificador;
4. Estabelecer as relações entre os descritores de mudança de conceito e a definição, baseada em evidências, da estratégia de reação mais adequada para preservação do desempenho do classificador em cenários dinâmicos.

## 1.4 Contribuições da Tese

Esta tese fornece importantes contribuições para o campo de mudança de conceito, em especial a classificação taxonômica dos métodos de detecção não supervisionados, a compreensão da natureza da mudança e a definição de estratégias de reação à mudança mais adequadas. As contribuições estão organizadas em três grupos.

### **Classificação Taxonômica**

- Propor uma taxonomia das abordagens do estado-da-arte para detecção não supervisionada de mudança de conceito.

### **Compreensão da Natureza da Mudança de Conceito**

- Classificação dos descritores: A análise dos descritores permitiu classificá-los em duas dimensões (espacial e temporal).
- Identificação de descritores com alto e baixo impacto na mudança de conceito: os resultados desta pesquisa demonstram que as características de uma mudança afetam o desempenho do classificador de maneiras distintas. Assim, identificar quais características têm maior impacto permite priorizá-las na definição de uma estratégia de reação baseada em evidências.

### **Descritores e Estratégia de Reação**

- Análise conjunta de descritores e estratégias de reação: este é o primeiro estudo a analisar simultaneamente quatro descritores e cinco estratégias de reação, possibilitando uma avaliação abrangente do impacto dessas estratégias no desempenho dos classificadores de acordo com as características da mudança de conceito;
- Estrutura experimental: este trabalho propõe uma estrutura experimental robusta, incorporando testes estatísticos e um estudo de ablação, com potencial de replicação em pesquisas futuras; e
- Reação baseada em evidências: este estudo fornece evidências práticas para apoiar a seleção da estratégia de reação mais adequada com base nas características da mudança de conceito.

## **1.5 Publicações**

- GEMAQUE, R. N. et al. An overview of unsupervised drift detection methods. *WIRES Data Mining and Knowledge Discovery*, v. 10, n. 6, p. e1381, 2020. Disponível em: <<https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.1381>>.
- COSTA, A.; GIUSTI, R.; SANTOS, E. M. d. Analysis of descriptors of concept drift and their impacts. *Informatics*, v. 12, n. 1, 2025. ISSN 2227-9709. Disponível em: <<https://www.mdpi.com/2227-9709/12/1/13>>.

- COSTA, A.; GIUSTI, R.; SANTOS, E. M. d. Concept Drift and Reaction Strategies: A Comprehensive Analysis Using Descriptors. Submetido ao periódico Machine Learning. ISSN 0885-6125 (*in peer review*).

## 1.6 Organização da Tese

A partir deste ponto, a tese está organizada da seguinte maneira:

O Capítulo 2 apresenta os conceitos fundamentais para a compreensão do trabalho. O Capítulo 3 reúne e discute a literatura relevante em três grupos: detectores da mudança de conceito; os trabalhos que abordam a natureza da mudança de conceito; e os trabalhos que consideram as características da mudança para a definição de uma estratégia de reação mais adequada. O Capítulo 4 apresenta uma análise sistemática do impacto dos descritores de mudança de conceito em ambientes supervisionados, semi-supervisionado e não supervisionado. O Capítulo 5 analisa a relação entre mudança de conceito e estratégias de reação adotadas. O Capítulo 6 apresenta as conclusões, limitações da tese e sugestões de trabalhos futuros. Por fim, o Apêndice A apresenta uma visão geral dos métodos não supervisionados de detecção de mudança de conceito, propondo uma taxonomia que organiza as principais abordagens existentes.



Neste capítulo, os conceitos fundamentais para o entendimento deste trabalho são apresentados. Inicialmente, o termo conceito e o fenômeno da mudança de conceito são fundamentados. Em seguida, as fontes probabilísticas e os descritores utilizados para caracterizar a mudança são considerados.

## 2.1 Conceito e Mudança de Conceito

Na perspectiva probabilística, o termo conceito é uma tupla composta pela probabilidade *a priori* da variável aleatória  $\mathcal{Y}$  e pela probabilidade condicional de  $\mathcal{X}$  em relação à  $\mathcal{Y}$ , ou seja,  $P(\mathcal{X}|\mathcal{Y})$  (KUNCHEVA, 2004; HOENS; POLIKAR; CHAWLA, 2012; WEBB et al., 2016). Essas variáveis determinam a distribuição da probabilidade conjunta  $P(\mathcal{X}, \mathcal{Y})$ .

No contexto de aprendizado de máquina,  $\mathcal{X}$  e  $\mathcal{Y}$  representam, respectivamente, o conjunto de instâncias descritas por vetores de características que compõem o espaço de entrada e o conjunto de variáveis-alvo. Essas variáveis são definidas nas Equações 2.1 e 2.2:

$$\mathcal{X} = \{x_1, x_2, x_3, \dots, x_n\}, x_n \in \mathbb{R}^d, d \in \mathbb{N}^*, \quad (2.1)$$

$$\mathcal{Y} = \{y_1, y_2, y_3, \dots, y_n\}, y_n \in \mathbb{N}^*, \quad (2.2)$$

de tal forma, um conceito é definido como a distribuição da probabilidade conjunta entre  $\mathcal{X}$  e  $\mathcal{Y}$ , conforme expresso na Equação 2.3:

$$\text{Conceito} = p(\mathcal{X}, \mathcal{Y}). \quad (2.3)$$

O termo mudança de conceito refere-se ao fenômeno de alteração na distribuição dos dados de um problema específico ao longo do tempo (SCHLIMMER; GRANGER, 1986; GEMAQUE et al., 2020; BAYRAM; AHMED; KASSLER, 2022). Diversos autores propuseram definições formais para esse fenômeno (GAMA et al., 2014; ŽLIOBAITĖ; HOLLMÉN, 2015; LOSING; HAMMER; WERSING, 2016; LU et al., 2016; GUO et al., 2022), dentre os quais, a definição de Gama et al. (2014) é amplamente aceita na literatura e está sumarizada na Equação 2.4.

$$\exists x : p_{t_i}(x, y) \neq p_{t_j}(x, y), t_i < t_j, \quad (2.4)$$

em que  $x$  é um vetor de características ( $x \in \mathcal{X}$ ), que descreve uma instância, e  $y$  representa a variável alvo ( $y \in \mathcal{Y}$ ). Uma mudança de conceito ocorre no tempo  $t_j$  se a distribuição conjunta ( $p_{t_j}$ ) difere significativamente da distribuição conjunta ( $p_{t_i}$ ) no tempo  $t_i$ . No contexto de dados em fluxo,  $j$  tende ao infinito (LU et al., 2019).

Webb et al. (2016) propõem uma definição mais abrangente para mudança de conceito ao considerar não somente o paradigma supervisionado, mas também o paradigma de aprendizado não supervisionado (Equação 2.5).

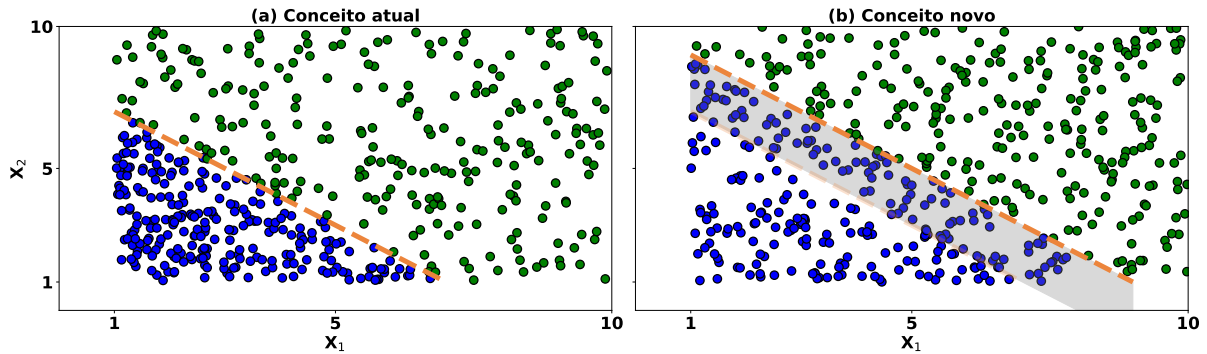
$$\exists x : p_{t_0}(x) \neq p_{t_n}(x), \quad (2.5)$$

a Equação 2.5 difere-se da Equação 2.4 por relaxar a restrição da variável alvo ( $\mathcal{Y}$ ) ser conhecida. Neste trabalho adota-se a definição expressa na Equação 2.4.

Na literatura, há divergências nas terminologias adotadas para identificar a mudança de conceito (BAYRAM; AHMED; KASSLER, 2022). Por exemplo, Quiñonero-Candela et al. (2009) utilizam o termo *dataset shift*, enquanto Widmer e Kubat (1996) empregam o termo *concept shift*. Por sua vez, Moreno-Torres et al. (2012) definem *dataset shift* como uma generalização que abrange os termos *covariate shift*, *prior probability shift* e mudança de conceito.

A mudança de conceito pode impactar o desempenho do classificador ao alterar a fronteira de decisão no espaço de entrada. A Figura 1 apresenta um problema de

Figura 1 – Exemplo de mudança de conceito: a fronteira de decisão é deslocada entre os conceitos, o que pode ocasionar perda de desempenho do classificador. A região sombreada indica a diferença entre os conceitos.



Fonte: Própria.

classificação binária definido pela equação da reta (Equação 2.6).

$$\text{classe} = \begin{cases} \text{azul,} & \text{se } x + y \leq \theta \\ \text{verde,} & \text{se } x + y > \theta, \end{cases} \quad (2.6)$$

onde  $\theta$  assume os valores 8 e 10, respectivamente, nos conceitos atual e novo. A alteração no espaço de entrada ocasionada pela mudança de conceito é representada pela área destacada em cinza na Figura 1.

A mudança de conceito é composta por fontes probabilísticas (IWASHITA; PAPA, 2019; CASTELLANI; SCHMITT; HAMMER, 2021; YAN, 2020; FAHY; YANG; GONGORA, 2022; TAN; LEE; SALEHI, 2022; COSTA; GIUSTI; SANTOS, 2025) e por descritores (MINKU; WHITE; YAO, 2010; MOUCHAWEH, 2016; KHAMASSI et al., 2018; COSTA; GIUSTI; SANTOS, 2025), conforme apresentado na Figura 2. As fontes probabilísticas são divididas nas probabilidades *a priori*, condicional e *a posteriori*. Por sua vez, a literatura estabelece seis descritores: velocidade, frequência, recorrência, severidade, zona de influência e previsibilidade (MOUCHAWEH, 2016; LU et al., 2019; COSTA; GIUSTI; SANTOS, 2025). Esses dois componentes da mudança de conceito são detalhados e exemplificados, respectivamente, nas Seções 2.2 e 2.3.

Figura 2 – Componentes da mudança de conceito: a) Fontes Probabilísticas e b) Descritores.

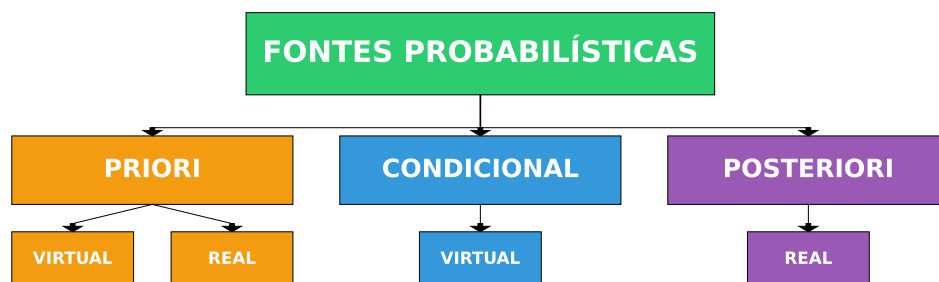


Fonte: Própria.

## 2.2 Fontes da Mudança de Conceito

As fontes probabilísticas da mudança de conceito correspondem às alterações nas componentes da distribuição conjunta formalizadas na Equação 2.4. Nesta seção, essas fontes são identificadas. Em seguida, as mudanças são classificadas conforme sua fonte, e discute-se como cada uma delas impacta o espaço de entrada do problema.

Figura 3 – Fontes probabilísticas da mudança de conceito.



Fonte: Própria.

As fontes probabilísticas da mudança de conceito incluem as alterações nas probabilidades *a priori*, condicional e *a posteriori*. No primeiro caso, o resultado implica em mudança de conceito virtual ou real, dependendo da variável afetada. Já alterações na probabilidade condicional de  $X$  caracterizam mudanças virtuais. Por fim, modificações na probabilidade *a posteriori* geram mudanças de conceito reais (Figura 3).

A probabilidade condicional é definida conforme a Equação 2.7, com a restrição

de  $p(y) > 0$  (BISHOP, 2006; DUDA; HART et al., 2006).

$$p(x|y) = \frac{p(x, y)}{p(y)}. \quad (2.7)$$

Aplicando a propriedade do produto dos meios pelos extremos na Equação 2.7, tem-se,

$$p(x, y) = p(x|y) \cdot p(y), \quad (2.8)$$

e de forma similar, conforme explicitado em Mouchaweh (2016):

$$p(x, y) = p(y|x) \cdot p(x), \quad (2.9)$$

portanto, as probabilidades condicionais da Equação 2.4 podem ser reescritas na forma das Equações 2.8 e 2.9.

Observa-se que os quatro termos que compõem o Teorema de Bayes (2.10) estão presentes nas Equações 2.8 e 2.9 (MOUCHAWEH, 2016; BAYRAM; AHMED; KASSLER, 2022).

$$p(y|x) = \frac{p(x|y) \cdot p(y)}{\sum_{y \in \mathcal{Y}} p(x|y) \cdot p(y)}, \quad (2.10)$$

aplicando a Lei da Probabilidade Total (BISHOP, 2006; DUDA; HART et al., 2006) na Equação 2.10, obtém-se a forma final do Teorema de Bayes (2.11).

$$p(y|x) = \frac{p(x|y) \cdot p(y)}{p(x)}, \quad (2.11)$$

correlacionando às Equações 2.11 e 2.4, identifica-se as quatro fontes probabilísticas da mudança de conceito:

- $p(x)$  a probabilidade *a priori* de  $x$ ,
- $p(y)$  a probabilidade *a priori* de  $y$ ,
- $p(x|y)$  a probabilidade condicional e
- $p(y|x)$  a probabilidade *a posteriori*.

Os autores [Mouchaweh \(2016\)](#), [Bayram, Ahmed e Kassler \(2022\)](#) e [Fahy, Yang e Gongora \(2022\)](#) divergem quanto à possibilidade de uma mudança de conceito ter como fonte uma mudança em  $p(x)$ . [Mouchaweh \(2016\)](#) sustenta que  $p(x)$  é constante para todas as classes, enquanto os demais autores apresentam entendimento oposto. Nesta pesquisa seguimos o entendimento destes últimos, por considerá-lo mais abrangente.

Destaca-se que nas figuras da Seção 2.2, as instâncias que aparecem com transparência representam a posição de cada instância no conceito inicial projetada junto à posição dessas no conceito atual com o objetivo de facilitar a visualização do processo de mudança de conceito. Similarmente, a linha tracejada representa a fronteira de decisão entre as classes.

### 2.2.1 Probabilidade a *priori*

A mudança de conceito pode ter como fonte alterações na probabilidade a *priori* do vetor de características que descreve as instâncias ( $x$ ) ou da variável alvo ( $y$ ). Mudanças em  $p(x)$  correspondem à chamada mudança de conceito do tipo virtual, que se caracteriza por não modificar a fronteira de separação ( $p(y|x)$ ) e, conseqüentemente sem impacto no desempenho do classificador ([MOUCHAWEH, 2016](#); [LU et al., 2019](#); [AGRAHARI; SINGH, 2021](#); [FAHY; YANG; GONGORA, 2022](#); [TAN; LEE; SALEHI, 2022](#)).

Por sua vez, mudanças na probabilidade a *priori* da variável alvo ( $y$ ) podem gerar tanto mudanças de conceito virtual quanto real, a depender de seus efeitos sobre a distribuição condicional ( $p(y|x)$ ). A mudança de conceito real é caracterizada por modificações na fronteira de decisão, o que, por sua vez, impacta o desempenho do classificador ([GAMA et al., 2014](#); [MOUCHAWEH, 2016](#); [AGRAHARI; SINGH, 2021](#); [FAHY; YANG; GONGORA, 2022](#); [TAN; LEE; SALEHI, 2022](#)).

### 2.2.1.1 Probabilidade a *priori* de $x$

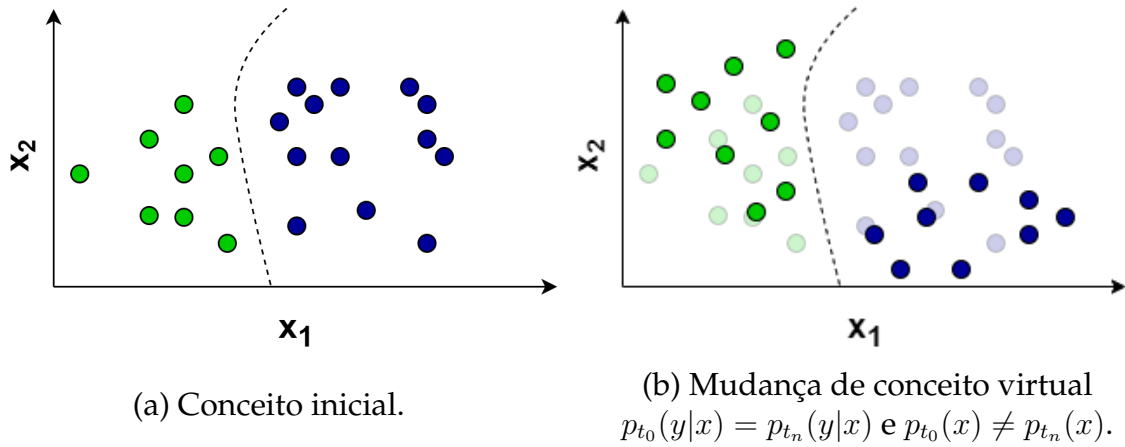
A mudança de conceito que afeta somente a probabilidade a *priori* de  $x$ , sem impactar  $y$ , é formalizada por Bayram, Ahmed e Kassler (2022), conforme as Equações 2.12 e 2.13:

$$p_{t_0}(y|x) = p_{t_n}(y|x), \quad (2.12)$$

$$p_{t_0}(x) \neq p_{t_n}(x). \quad (2.13)$$

A Figura 4 exemplifica uma mudança de conceito virtual. Em (a) apresenta-se o conceito inicial, enquanto em (b) está o conceito surgido após a mudança. Observa-se que nessa situação a mudança não altera a fronteira de separação entre as classes.

Figura 4 – Exemplo de mudança de conceito virtual ocasionada por mudança em  $p(x)$ .



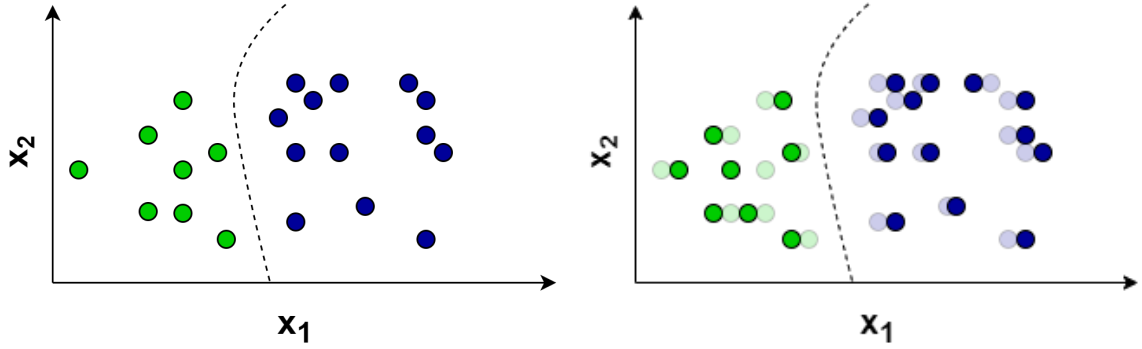
Fonte: Adaptado de Gama et al. (2014), Mouchaweh (2016), Bayram, Ahmed e Kassler (2022), Fahy, Yang e Gongora (2022) e Tan, Lee e Salehi (2022).

Mudanças na probabilidade a *priori* de  $x$  podem gerar dois subtipos: mudança de conceito virtual local e mudança de conceito virtual por evolução de características (BAYRAM; AHMED; KASSLER, 2022). O primeiro é caracterizado por alterações em regiões limitadas do espaço de características e descrito conforme a Equação 2.14 (TSYMBAL et al., 2008; BAYRAM; AHMED; KASSLER, 2022):

$$\begin{aligned} p_{t_0}(y|x) &= p_{t_n}(y|x), \\ p_{t_0}(x_1) &\neq p_{t_n}(x_1), \\ p_{t_0}(x_2) &= p_{t_n}(x_2). \end{aligned} \quad (2.14)$$

A Figura 5 exemplifica a mudança de conceito virtual local, a qual afeta somente o atributo  $x_1$ , enquanto o atributo  $x_2$  permanece inalterado após a mudança de conceito.

Figura 5 – Exemplo de mudança de conceito virtual ocasionada por mudança em parte dos atributos.



(a) Conceito inicial.

(b) Mudança de conceito virtual local

$$p_{t_0}(y|x) = p_{t_n}(y|x),$$

$$p_{t_0}(x_1) \neq p_{t_n}(x_1) \text{ e } p_{t_0}(x_2) = p_{t_n}(x_2).$$

Fonte: Adaptado de [Gama et al. \(2014\)](#), [Mouchaweh \(2016\)](#), [Bayram, Ahmed e Kassler \(2022\)](#), [Fahy, Yang e Gongora \(2022\)](#) e [Tan, Lee e Salehi \(2022\)](#).

A mudança de conceito virtual por evolução de características está associado ao surgimento de novas características com a passagem do tempo, conforme descrito na Equação 2.15 ([MASUD et al., 2010](#)):

$$p_{t_0}(y|x) = p_{t_n}(y|x),$$

$$x_{t_0} = \{x_1, x_2\},$$

$$x_{t_n} = \{x_1, x_2, x_3\},$$

com  $n > 0$ .

(2.15)

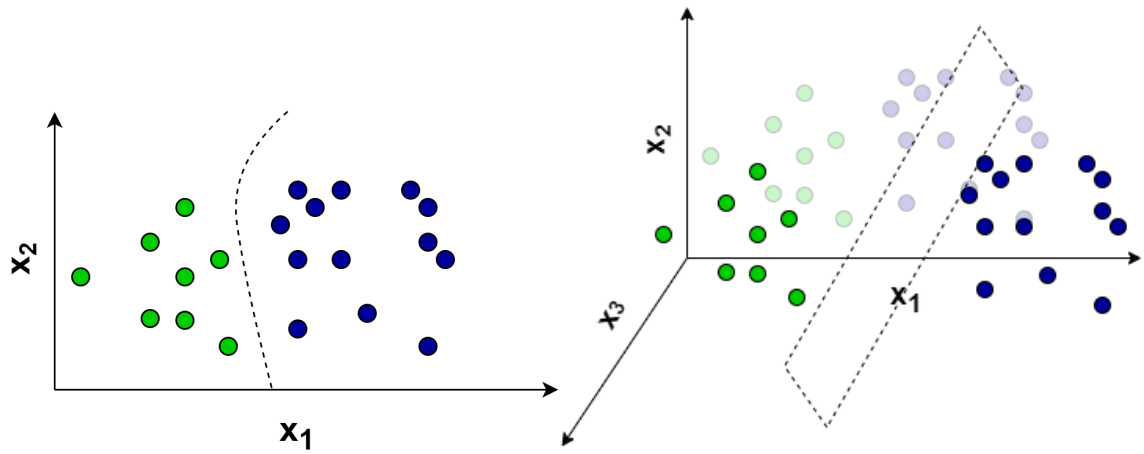
A Figura 6 exemplifica esse subtipo: uma nova características ( $x_3$ ) surge com a passagem do tempo.

#### 2.2.1.2 Probabilidade a priori de $y$

Esse tipo de mudança é caracterizado por variação no balanceamento entre as classes existentes (Figura 7), surgimento de novas classes (Figura 8) ou desaparecimento de



Figura 6 – Exemplo de mudança de conceito virtual ocasionada por surgimento de novo atributo.



(a) Conceito inicial.

(b) Mudança de conceito virtual local

$$p_{t_0}(y|x) = p_{t_n}(y|x), \\ p_{t_0}(x_1) \neq p_{t_n}(x_1) \text{ e } p_{t_0}(x_2) = p_{t_n}(x_2).$$

Fonte: Adaptado de [Gama et al. \(2014\)](#), [Mouchaweh \(2016\)](#), [Bayram, Ahmed e Kassler \(2022\)](#), [Fahy, Yang e Gongora \(2022\)](#) e [Tan, Lee e Salehi \(2022\)](#).

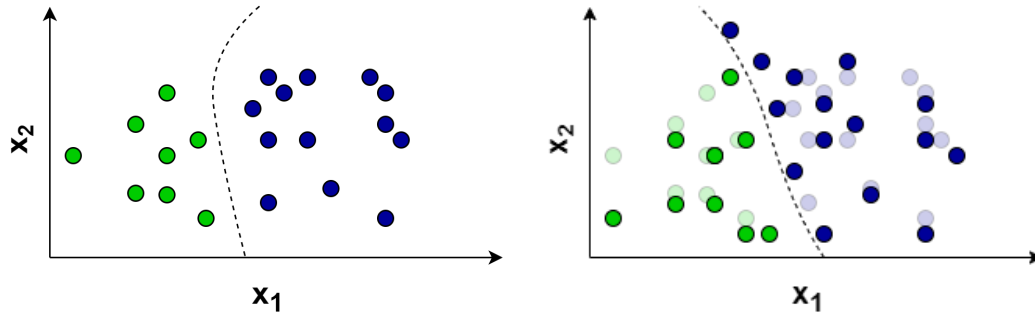
classes existentes (Figura 9). As mudanças nessa fonte (Equação 2.16) ocasionam a degradação no desempenho do classificador ([BAYRAM; AHMED; KASSLER, 2022](#)):

$$p_{t_0}(y) \neq p_{t_n}(y), \quad (2.16)$$

esse tipo de mudança é denominado de mudança de conceito real.

É importante citar que mudanças na probabilidade *a priori* de  $y$  que afetam a probabilidade *a posteriori* ( $p(y|x)$ ) mas não a probabilidade condicional ( $p(x|y)$ ) são chamadas mudança de conceito real do tipo I ([MOUCHAWEH, 2016](#); [BAYRAM; AHMED; KASSLER, 2022](#)). Por sua vez, alterações em  $p(y)$  que não afetam  $p(y|x)$  não caracterizam mudança de conceito e estão relacionadas com o problema de desbalanceamento de classes ([FANGYU et al., 2021](#)). A evolução do conceito e o desaparecimento de classes são, respectivamente, tratados em ([MASUD et al., 2010](#)) e ([ELWELL; POLIKAR, 2011](#)).

Figura 7 – Exemplo de mudança de conceito real ocasionada por variação no balanceamento entre as classes.



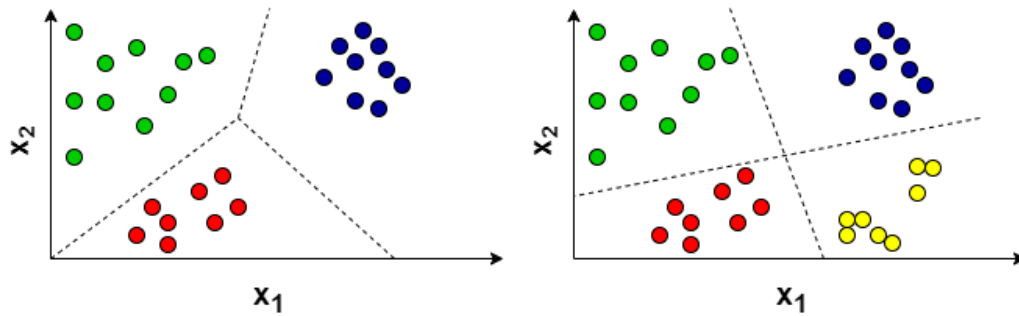
(a) Conceito inicial.

(b) Mudança de conceito real tipo I.

$$p_{t_0}(y) \neq p_{t_n}(y), p_{t_0}(y|x) \neq p_{t_n}(y|x) \\ \text{e } p_{t_0}(x) \neq p_{t_n}(x).$$

Fonte: Adaptado de [Gama et al. \(2014\)](#), [Mouchaweh \(2016\)](#), [Bayram, Ahmed e Kassler \(2022\)](#), [Fahy, Yang e Gongora \(2022\)](#) e [Tan, Lee e Salehi \(2022\)](#).

Figura 8 – Exemplo de mudança de conceito real ocasionada por evolução do conceito.



(a) Conceito inicial.

(b) Mudança de conceito real por evolução do conceito.

$$p_{t_0}(y) \neq p_{t_n}(y).$$

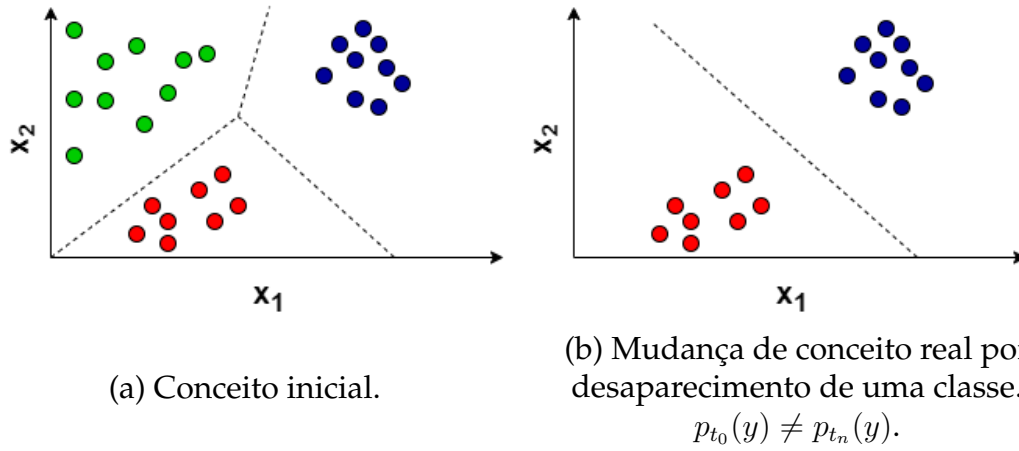
Fonte: Adaptado de [Gama et al. \(2014\)](#), [Mouchaweh \(2016\)](#), [Bayram, Ahmed e Kassler \(2022\)](#), [Fahy, Yang e Gongora \(2022\)](#) e [Tan, Lee e Salehi \(2022\)](#).

### 2.2.2 Probabilidade condicional de x

Esse tipo de mudança ( $p(x|y)$ ) impacta somente a distribuição dentro da mesma classe, sem modificar a fronteira de decisão. Portanto, trata-se de mudança de conceito virtual ([MOUCHAWEH, 2016](#)), descrita conforme a Equação 2.17.

$$\begin{aligned} p_{t_0}(y|x) &= p_{t_n}(y|x), \\ p_{t_0}(x|y) \cdot p(y) &\neq p_{t_n}(x|y) \cdot p(y), \\ p_{t_0}(y) &= p_{t_n}(y) - \text{Restrição}, \end{aligned} \tag{2.17}$$

Figura 9 – Exemplo de mudança de conceito real ocasionada pelo desaparecimento de uma classe.



Fonte: Adaptado de [Gama et al. \(2014\)](#), [Mouchaweh \(2016\)](#), [Bayram, Ahmed e Kassler \(2022\)](#), [Fahy, Yang e Gongora \(2022\)](#) e [Tan, Lee e Salehi \(2022\)](#).

a restrição é necessária porque mudanças em  $p(y)$  impactam a fronteira de separação entre as classes, caracterizando mudanças de conceito real. Além disso, a restrição gera a mesma descrição para mudanças na probabilidade *a priori* de  $x$ . Portanto, as mudanças em  $p(x|y)$  são as mesmas apresentadas na probabilidade *a priori* de  $x$ .

### 2.2.3 Probabilidade *a posteriori*

Essas mudanças afetam a fronteira de separação entre as classes e degradam o desempenho do classificador, o que exige a adoção de estratégias de reação. São classificadas como mudança de conceito real ([GAMA et al., 2014](#); [MOUCHAWEH, 2016](#); [BAYRAM; AHMED; KASSLER, 2022](#); [FAHY; YANG; GONGORA, 2022](#); [TAN; LEE; SALEHI, 2022](#)).

Os seguintes autores sustentam que mudanças em  $p(y|x)$  geram exclusivamente mudanças de conceito real ([GAMA et al., 2014](#); [MOUCHAWEH, 2016](#); [FAHY; YANG; GONGORA, 2022](#); [TAN; LEE; SALEHI, 2022](#)). Por sua vez, [Bayram, Ahmed e Kassler \(2022\)](#) afirmam que mudanças nessa fonte produzem dois tipos principais e três subtipos de mudanças de conceito. Essa classificação é detalhada a seguir.

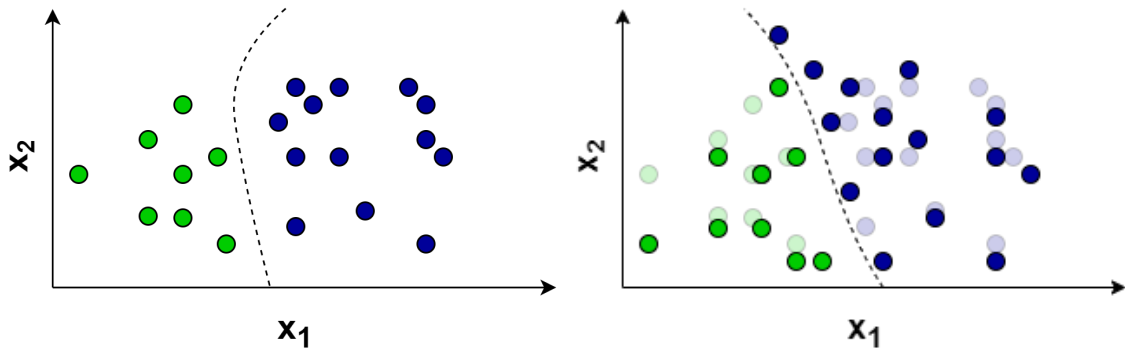
A primeira é denominada de mudança de conceito real do tipo I, descrita con-

forme a Equação 2.18:

$$\begin{aligned} p_{t_0}(y|x) &\neq p_{t_n}(y|x), \\ p_{t_0}(x) &\neq p_{t_n}(x), \end{aligned} \quad (2.18)$$

observa-se que alterações em  $p(y|x)$  estão associadas à alterações em  $p(x)$ . Esse tipo de mudança (Figura 10) afeta a fronteira de decisão (BAYRAM; AHMED; KASSLER, 2022).

Figura 10 – Exemplo de mudança de conceito real do tipo I ocasionada por mudança em  $p_{t_0}(y|x) \neq p_{t_n}(y|x)$  e  $p_{t_0}(x) \neq p_{t_n}(x)$ .



(a) Conceito inicial.

(b) Mudança de conceito real tipo I.  
 $p_{t_0}(y) \neq p_{t_n}(y)$ ,  $p_{t_0}(y|x) \neq p_{t_n}(y|x)$   
e  $p_{t_0}(x) \neq p_{t_n}(x)$ .

Fonte: Adaptado de Gama et al. (2014), Mouchaweh (2016), Bayram, Ahmed e Kassler (2022), Fahy, Yang e Gongora (2022) e Tan, Lee e Salehi (2022).

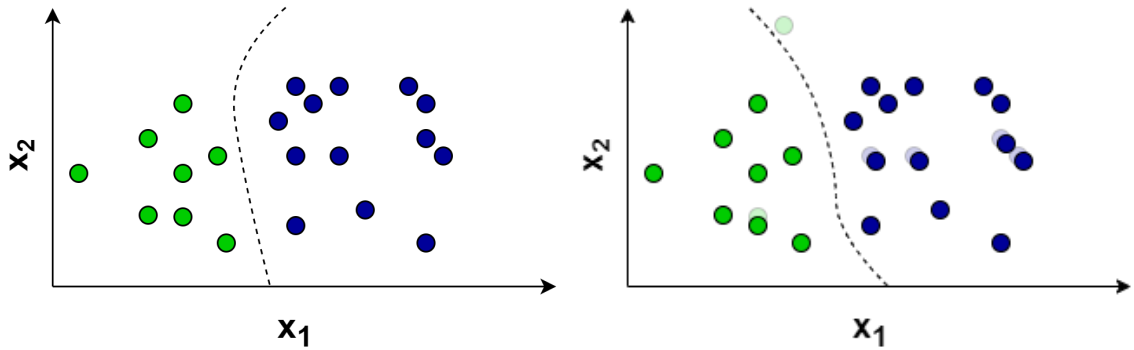
A segunda mudança é denominada mudança de conceito real do tipo II e é descrita conforme a Equação 2.19:

$$\begin{aligned} p_{t_0}(y|x) &\neq p_{t_n}(y|x), \\ p_{t_0}(x) &= p_{t_n}(x), \end{aligned} \quad (2.19)$$

observa-se que alterações em  $p(y|x)$  não estão associadas à alterações em  $p(x)$ . Esse tipo (Figura 11) afeta a fronteira de decisão (BAYRAM; AHMED; KASSLER, 2022).

Por fim, os três subtipos identificados nas mudanças na probabilidade a *posteriori* são denominados: i) mudança de conceito real inconsistente, ii) mudança de conceito real plena e iii) mudança de conceito real cruzada (FORMAN, 2006; MINKU; WHITE; YAO, 2010; WEBB et al., 2016; BAYRAM; AHMED; KASSLER, 2022).

Figura 11 – Exemplo de mudança de conceito real do tipo II ocasionada por mudança em  $p_{t_0}(y|x) \neq p_{t_n}(y|x)$  e  $p_{t_0}(x) = p_{t_n}(x)$ .



(a) Conceito inicial.

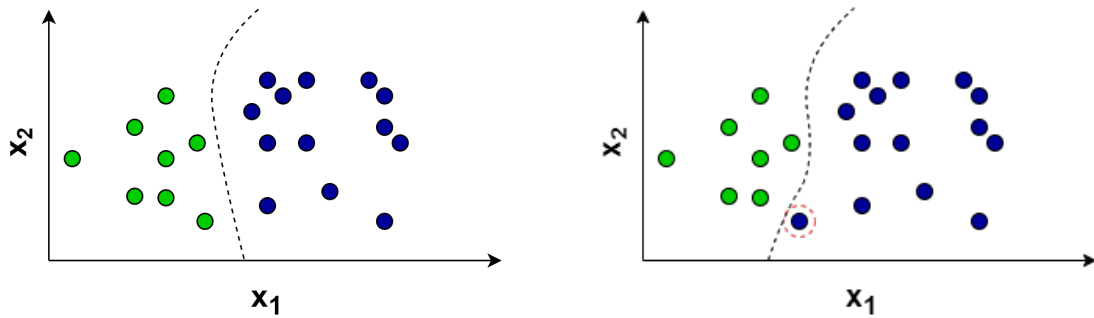
(b) Mudança de conceito real tipo II.  
 $p_{t_0}(y|x) \neq p_{t_n}(y|x)$  e  $p_{t_0}(x) = p_{t_n}(x)$ .

Fonte: Adaptado de [Gama et al. \(2014\)](#), [Mouchaweh \(2016\)](#), [Bayram, Ahmed e Kassler \(2022\)](#), [Fahy, Yang e Gongora \(2022\)](#) e [Tan, Lee e Salehi \(2022\)](#).

A mudança de conceito real inconsistente é caracterizada pela existência de ao menos uma instâncias que modifica sua classe com a passagem do tempo. É descrita conforme a Equação 2.20 ([FORMAN, 2006](#); [BAYRAM; AHMED; KASSLER, 2022](#)):

$$\exists x(\arg \max p_{t_0}(y|x) = c_1 \cap \arg \max p_{t_n}(y|x) = c_2). \quad (2.20)$$

Figura 12 – Exemplo de mudança de conceito real inconsistente -  $\exists x(\arg \max p_{t_0}(y|x) = c_1 \cap \arg \max p_{t_n}(y|x) = c_2)$ .



(a) Conceito inicial.

(b) Mudança de conceito real inconsistente.  
 $\exists x(\arg \max p_{t_0}(y|x) = c_1 \cap$   
 $\arg \max p_{t_n}(y|x) = c_2).$

Fonte: Adaptado de [Gama et al. \(2014\)](#), [Mouchaweh \(2016\)](#), [Bayram, Ahmed e Kassler \(2022\)](#), [Fahy, Yang e Gongora \(2022\)](#) e [Tan, Lee e Salehi \(2022\)](#).

A Figura 12 apresenta um exemplo desse subtipo de mudança. O círculo em

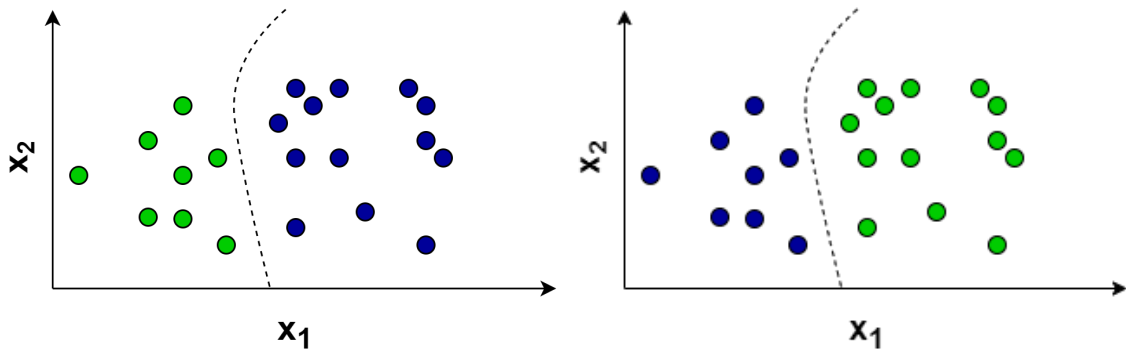
vermelho indica a instância que alterou sua classe e, como consequência, alterou a fronteira de decisão entre as classes com a passagem do tempo.

A mudança de conceito real plena é caracterizada pela alteração de classe de todas as instâncias com a passagem do tempo. É descrita conforme a Equação 2.21 (MINKU; WHITE; YAO, 2010; BAYRAM; AHMED; KASSLER, 2022):

$$\forall x(\arg \max p_{t_0}(y|x) = c_1 \cap \arg \max p_{t_n}(y|x) = c_2). \quad (2.21)$$

A Figura 13 apresenta um exemplo desse subtipo. Todas as instâncias mudaram de classe com a passagem do tempo e, como consequência, a fronteira de decisão entre as classes foi alterada. Esse subtipo também é referenciado na literatura como reversão de conceitos (GAMA et al., 2004; PINAGE; SANTOS, 2015; IWASHITA; PAPA, 2019; GÂLMEANU; ANDONIE, 2021).

Figura 13 – Exemplo de mudança de conceito real plena -  $\forall x(\arg \max p_{t_0}(y|x) = c_1 \cap \arg \max p_{t_n}(y|x) = c_2)$ .



(a) Conceito inicial.

(b) Mudança de conceito real plena.

$$\forall x(\arg \max p_{t_0}(y|x) = c_1 \cap \arg \max p_{t_n}(y|x) = c_2).$$

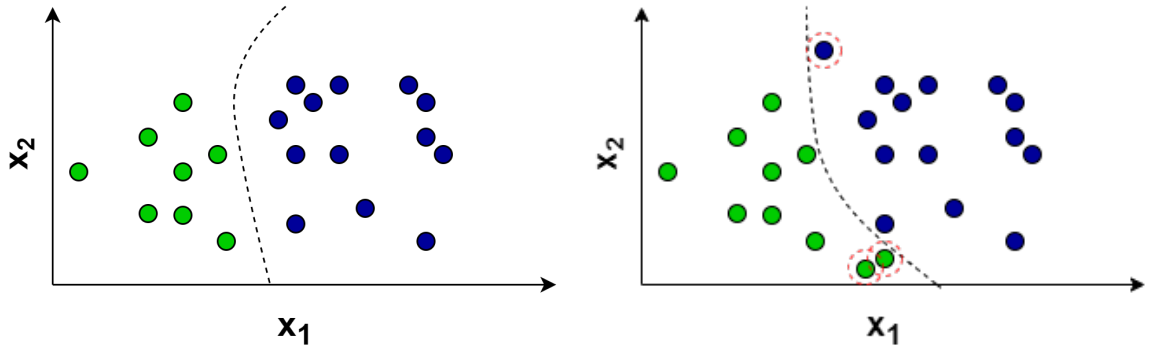
Fonte: Adaptado de Gama et al. (2014), Mouchaweh (2016), Bayram, Ahmed e Kassler (2022), Fahy, Yang e Gongora (2022) e Tan, Lee e Salehi (2022).

A mudança real cruzada implica em modificações de classe de instâncias delimitadas em uma sub-região (BAYRAM; AHMED; KASSLER, 2022). É descrita conforme a Equação 2.22 (WEBB et al., 2016; BAYRAM; AHMED; KASSLER, 2022):

$$\begin{aligned} \exists x(\arg \max p_{t_0}(y|x) = c_1 \cap \arg \max p_{t_n}(y|x) = c_2) \cap \\ \exists z(\arg \max p_{t_0}(y|z) = c_1 \cap \arg \max p_{t_n}(y|z) = c_2). \end{aligned} \quad (2.22)$$

A Figura 14 apresenta um exemplo desse último subtipo. As instâncias circunscritas em um círculo vermelho indicam as sub-regiões do espaço de entrada em que as mudanças ocorrem.

Figura 14 – Exemplo de mudança de conceito real cruzada -  $\exists x(\arg \max p_{t_0}(y|x) = c_1 \cap \arg \max p_{t_n}(y|x) = c_2) \cap \exists z(\arg \max p_{t_0}(y|z) = c_1 \cap \arg \max p_{t_n}(y|z) = c_2)$ .



(a) Conceito inicial.

(b) Mudança de conceito real cruzada.

$$\begin{aligned} &\exists x(\arg \max p_{t_0}(y|x) = c_1 \cap \\ &\arg \max p_{t_n}(y|x) = c_2) \cap \\ &\exists z(\arg \max p_{t_0}(y|z) = c_1 \cap \\ &\arg \max p_{t_n}(y|z) = c_2) \end{aligned}$$

Fonte: Adaptado de [Gama et al. \(2014\)](#), [Mouchaweh \(2016\)](#), [Bayram, Ahmed e Kassler \(2022\)](#), [Fahy, Yang e Gongora \(2022\)](#) e [Tan, Lee e Salehi \(2022\)](#).

#### 2.2.4 Síntese das Fontes de Mudança de Conceito

A Tabela 1 sintetiza a relação entre as fontes de mudança de conceito discutidas nesta seção e os respectivos tipos identificados na literatura. Nota-se que determinados tipos de mudança podem estar associados a mais de uma fonte. Ressalta-se, ainda, que os tipos posicionados acima da linha tracejada correspondem a mudanças virtuais, sem impacto no desempenho do classificador, enquanto aqueles abaixo da linha referem-se a mudanças reais, que comprometem o desempenho do modelo.

Esta seção discutiu o primeiro componente da mudança de conceito. A discussão sobre o segundo componente é apresentada na próxima seção.

Tabela 1 – Resumo das fontes probabilísticas de mudança de conceito.

		Fonte			
		$p(x)$	$p(y)$	$p(x y)$	$p(y x)$
Mudança de conceito	Virtual	✓		✓	
	Virtual local	✓		✓	
	Virtual evolução de características	✓		✓	
	Real tipo I		✓		✓
	Real tipo II				✓
	Real inconsistente				✓
	Real plena				✓
	Real cruzada				✓
	Real evolução de conceitos		✓		
	Real exclusão de conceitos		✓		

Fonte: Própria.

### 2.3 Descritores da Mudança de Conceito

Um descritor é uma propriedade inerente ao fenômeno da mudança de conceito que permite sua caracterização. Na literatura há seis descritores conhecidos, dos quais cinco podem ser agrupados em função da dimensionalidade considerada.

- Dimensão espacial: severidade e zona de influência, e
- Dimensão temporal: velocidade, frequência e recorrência.

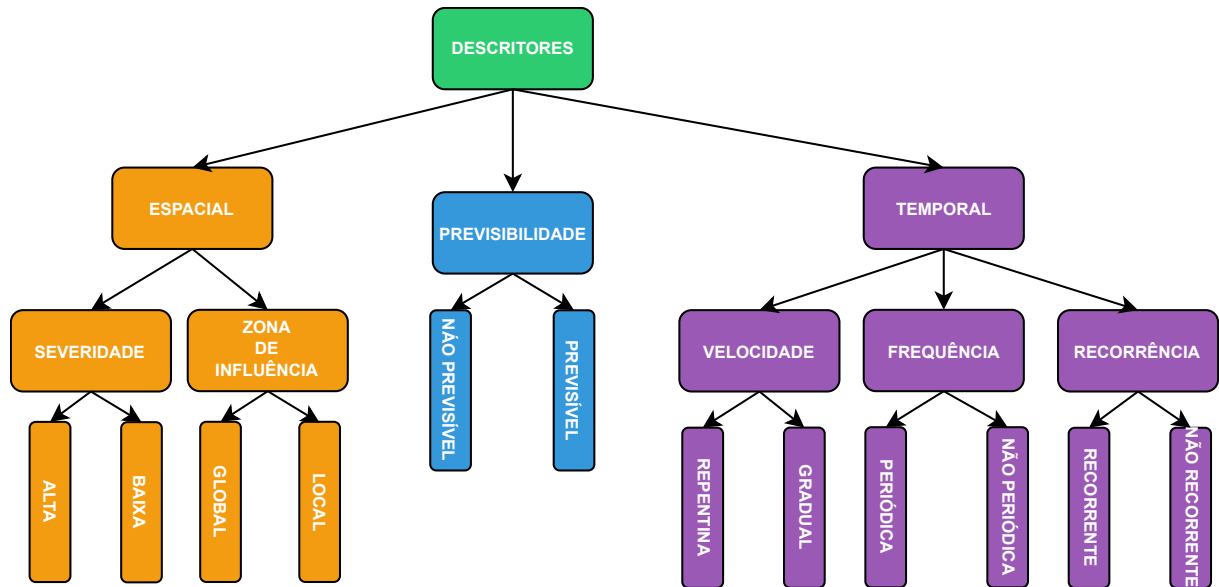
O sexto descritor, previsibilidade, não está relacionado a nenhuma das duas dimensões. A Figura 15 detalha a estrutura dos descritores de mudança de conceito.

O descritor de severidade caracteriza a mudança de conceito como de alta ou baixa severidade, enquanto o de zona de influência a caracteriza como global ou local. No que se refere aos descritores da dimensão temporal, a mudança pode ser: repentina ou gradual (velocidade); periódica ou não periódica (frequência); e recorrente ou não recorrente (recorrência). Por fim, o descritor de previsibilidade indica se a mudança possui propriedades que podem ou não ser previstas.

Os descritores buscam respostas à perguntas específicas a fim de caracterizar as mudanças de conceito, estando essas expressas a seguir. Os descritores de velocidade, severidade e zona de influência permitem compreender a estrutura da mudança gerando respostas, respectivamente, às questões um, dois e três (LU et al., 2019). Quanto



Figura 15 – Mudança de conceito na perspectiva do seus descritores.



Fonte: Adaptado de [Costa, Giusti e Santos \(2025\)](#).

à frequência e à recorrência, esses descritores estabelecem, respectivamente, respostas para as perguntas quatro e cinco ([LU et al., 2019](#); [MOUCHAWEH, 2016](#)). Já o descritor de previsibilidade busca responder a pergunta seis, dado que em ambientes onde as mudanças ocorrem de forma plenamente aleatória, a detecção só é possível após o início dos seus efeitos, como a degradação do desempenho do preditor por exemplo.

1. Qual é a duração da mudança (descritor de velocidade)?
2. Qual a intensidade da mudança (descritor de severidade)?
3. Onde a mudança ocorre (descritor de zona de influência)?
4. Quantas vezes a mudança ocorre em um determinado intervalo de tempo (descritor de frequência)?
5. A mudança tem comportamento cíclico (descritor de recorrência)?
6. Há algum padrão na ocorrência das mudanças (descritor de previsibilidade)?

Portanto, a caracterização da mudança de conceito por meio de seus descritores constitui uma abordagem relevante para compreender sua natureza e orientar a defi-

nição da estratégia de reação mais adequada (MOUCHAWEH, 2016; LU et al., 2019; COSTA; GIUSTI; SANTOS, 2025).

As definições dos descritores apresentadas aqui estão fundamentadas nos trabalhos de Lazarescu, Venkatesh e Bui (2004), Minku, White e Yao (2010) e Žliobaitė (2010), compilados por Mouchaweh (2016); nos trabalhos de Kosina, Gama e Sebastião (2010), Jia, Koh e Dobbie (2017) e Webb et al. (2016), compilados por Goldenberg e Webb (2019); e nas contribuições mais recentes (LU et al., 2019) e (COSTA; GIUSTI; SANTOS, 2025).

### 2.3.1 Dimensão Espacial

As alterações no espaço de entrada decorrentes da mudança de conceito são descritas pelos descritores da dimensão espacial, que fornecem parâmetros relativos à sua magnitude e localização. Em particular, os descritores de severidade e zona de influência permitem mensurar a dissimilaridade entre os conceitos e identificar as regiões do espaço impactadas, respectivamente.

#### 2.3.1.1 Severidade

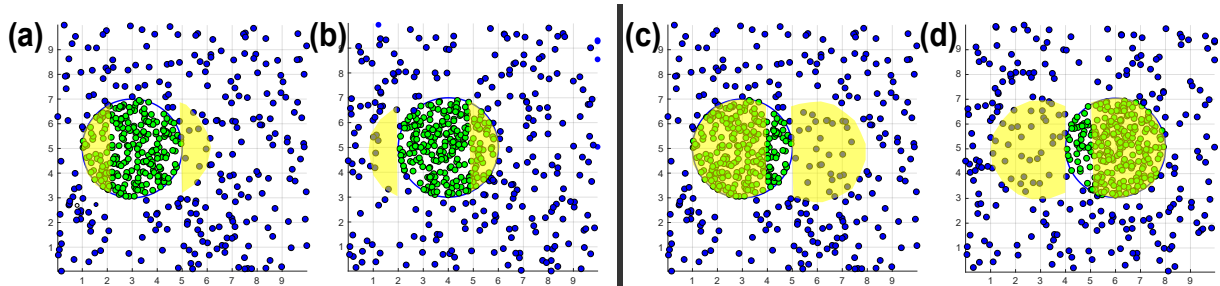
Mensura a dissimilaridade do espaço de entrada entre os conceitos válidos no tempo  $i$  e  $j$ , com  $i < j$  (MOUCHAWEH, 2016; WEBB et al., 2016; LU et al., 2019; COSTA; GIUSTI; SANTOS, 2025). O descritor é formalizado na Equação 2.23.

$$\Delta = \delta(p_{t_i}(\mathcal{X}, \mathcal{Y}), p_{t_j}(\mathcal{X}, \mathcal{Y})), i < j, \quad (2.23)$$

onde  $\delta$  é uma função de dissimilaridade entre duas distribuições de dados, e  $t_i$  e  $t_j$  representam dois pontos distintos no tempo. Geralmente  $\Delta$  é não negativo e, quanto maior o valor de  $\Delta$ , maior é a severidade da mudança (LU et al., 2019).

A mudança de conceito é descrita em termos de severidade em alta (global) ou baixa (parcial). É considerada alta quando o conceito atual é majoritariamente substituído por um novo, enquanto a baixa severidade corresponde a alterações que afetam apenas uma parte da região ocupada pelo conceito atual (MOUCHAWEH, 2016).

Figura 16 – Descritor de severidade. O lado esquerdo representa uma mudança de conceito de baixa severidade: as áreas em amarelo em (a) e (b) indicam, respectivamente, as regiões que devem ser esquecidas e aprendidas após a mudança. O lado direito representa uma mudança de conceito de alta severidade: as áreas em amarelo em (c) e (d) indicam, respectivamente, as regiões que devem ser esquecidas e aprendidas após a mudança.



Fonte: Adaptado de [Costa, Giusti e Santos \(2025\)](#).

A quantificação dessa propriedade pode ser realizada por meio de métricas que avaliam a dissimilaridade entre distribuições de probabilidade, como a divergência de Kullback-Leibler ([KULLBACK; LEIBLER, 1951](#); [DASU et al., 2006](#)). Apesar dessas abordagens, a distinção formal entre os dois níveis de severidade ainda é considerada uma questão em aberto na literatura ([LU et al., 2019](#)).

[Minku, White e Yao \(2010\)](#) classificam a mudança em termos de severidade em severa e cruzada. A severa corresponde à situação em que todas as instâncias modificam seus rótulos devido à mudança de conceito, gerando uma reversão dos conceitos, conforme mostra a Figura 13 (b). Já na mudança de conceito cruzada, uma parte das instâncias do conceito atual mantém seus rótulos, conforme mostra Figura 14.

As classificações propostas por [Minku, White e Yao \(2010\)](#) e [Mouchaweh \(2016\)](#) para as mudanças de conceito de alta severidade assemelham-se, porém são conceitualmente distintas. Para os primeiros autores, deve ocorrer uma reversão completa dos conceitos. Enquanto [Mouchaweh \(2016\)](#) considera as alterações causadas pela mudança de conceito, independentemente de haver mudanças na fronteira de decisão. Este trabalho adota a definição de [Mouchaweh \(2016\)](#), pois essa distingue melhor a severidade da zona de influência.

A Figura 16 apresenta dois exemplos de mudança de conceito com severidades distintas no contexto de um problema de classificação binária, de duas dimensões. No primeiro exemplo, à esquerda na figura, as imagens (a) e (b) representam, respectiva-

mente, o mesmo espaço de entrada antes e depois da mudança de conceito. As regiões destacadas em amarelo indicam as regiões impactadas pela mudança. Observa-se que estas representam uma pequena parte do espaço, caracterizando uma mudança de conceito de baixa severidade. Nesse cenário, reagir à mudança retreinando o classificador significaria esquecer o espaço de entrada e ter que reaprender todas as áreas do espaço de entrada não afetadas pela mudança por causa de uma pequena parte afetada.

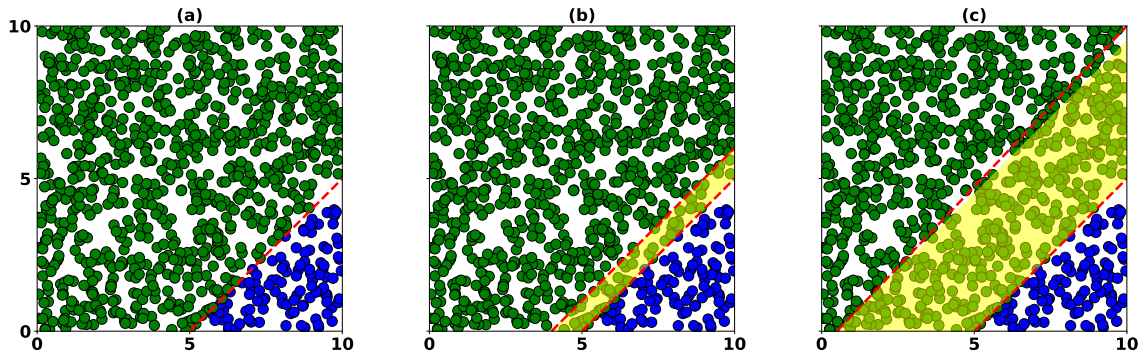
Por sua vez, no exemplo à direita da Figura 16, em que as imagens (c) e (d) também representam o mesmo espaço de entrada antes e depois da mudança de conceito, respectivamente, observa-se que as regiões impactadas pela mudança representam uma grande parte do espaço. Isso caracteriza uma mudança de conceito de alta severidade. Nesse cenário, praticamente toda a classe verde foi afetada, sendo provavelmente o retreino a melhor estratégia para reagir a esse tipo de mudança.

Esse cenário evidencia a importância de se considerar o nível de severidade da mudança de conceito na definição de estratégias de reação. Tomando como exemplo a estratégia mais comumente adotada na literatura, que é baseada no retreino total do preditor, observa-se que, além de apresentar elevado custo computacional, essa abordagem demanda, na maioria dos casos, a obtenção dos rótulos verdadeiros das instâncias. No entanto, o processo de rotulagem pode ser oneroso e, em muitas situações, inviável, a depender da tarefa preditiva em questão. Diante disso, em vez de reagir à mudança independentemente do nível de severidade, é útil estabelecer um limiar de severidade que determine quando a reação se justifica, de modo a executá-la apenas quando estritamente necessário.

#### 2.3.1.2 Zona de Influência

Caracteriza as regiões do espaço de entrada impactadas pela ocorrência da mudança (MOUCHAWEH, 2016). Essas regiões são consideradas regiões de conflito entre o conceito atual e o novo, onde  $p_{t_i}(\mathcal{X}, \mathcal{Y})$  e  $p_{t_j}(\mathcal{X}, \mathcal{Y})$ ,  $i < j$  apresentam diferença significativa (LU et al., 2019). A zona de influência descreve a mudança como local ou global. A primeira ocorre quando uma área restrita do espaço de entrada é impactada, enquanto

Figura 17 – Descritor de zona de influência. (a) representa o espaço de entrada do conceito inicial. Nos exemplos (b) e (c), as regiões destacadas em amarelo apresentam probabilidade conjunta de classificar as instâncias diferentes em relação à (a), tornando-as regiões de conflito entre os conceitos.



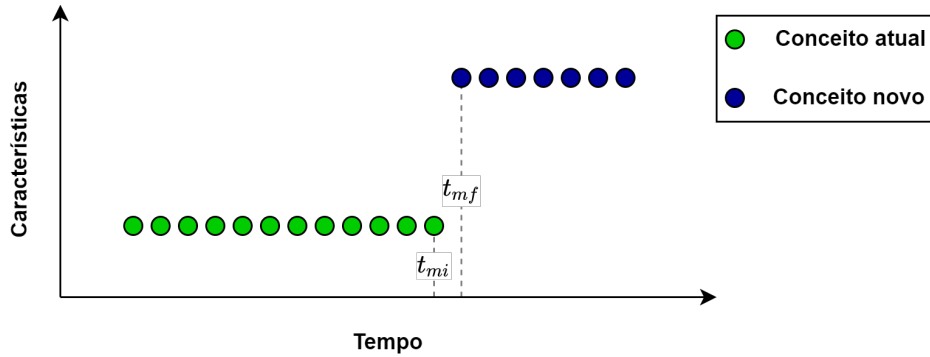
Fonte: Própria.

na segunda uma região significativa do espaço de entrada é afetada, especialmente próxima à fronteira de separação entre as classes (COSTA; GIUSTI; SANTOS, 2025).

A Figura 17 ilustra o descritor zona de influência. A parte (a) representa o conceito inicial, enquanto as partes (b) e (c) destacam, em amarelo, as regiões do espaço de entrada que diferem do conceito original após a mudança. Essas regiões correspondem às áreas de conflito entre os conceitos, nas quais ocorre alteração na probabilidade condicional de classificação.

Conhecer a zona de influência da mudança de conceito é útil para definir uma estratégia de reação que preserve o desempenho do classificador e evite retreinamentos desnecessários. Quando a zona de influência é global (Figura 17(c)), afetando grande parte do espaço de entrada e, geralmente, associada a uma mudança de alta severidade, a estratégia mais adequada tende a ser o retreino completo do modelo. Por outro lado, quando a mudança ocorre de forma local, como ilustrado na Figura 17(b), a estratégia de reação deve ser ajustada conforme outros descritores (GAMA et al., 2014; IWASHITA; PAPA, 2019; BAYRAM; AHMED; KASSLER, 2022; CASTELLANI; SCHMITT; HAMMER, 2021; YAN, 2020; FAHY; YANG; GONGORA, 2022; TAN; LEE; SALEHI, 2022).

Figura 18 – Exemplo de mudança de conceito repentina, caracterizada pela substituição do conceito antigo pelo novo em um curto intervalo de tempo, com  $V_{mc} = t_{mf} - t_{mi} = 1$ .



Fonte: Adaptado de [Minku, White e Yao \(2010\)](#), [Mouchaweh \(2016\)](#) e [Lu et al. \(2019\)](#).

### 2.3.2 Dimensão Temporal

As características temporais da mudança de conceito; como a duração, a frequência e a recorrência; são descritas pelos descritores da dimensão temporal.

#### 2.3.2.1 Velocidade

Esse descritor está relacionado ao tempo de duração do processo de transição entre os conceito atual e novo. A velocidade da mudança é definida como o inverso do tempo necessário para que o conceito atual seja substituído pelo novo ([MINKU; WHITE; YAO, 2010](#)). A velocidade é calculada pela Equação 2.24 ([MOUCHAWEH, 2016](#)):

$$V_{mc} = \frac{1}{t_{mf} - t_{mi}}, \quad (2.24)$$

onde  $t_{mf}$  e  $t_{mi}$  indicam, respectivamente, o ponto de fim e o ponto de início da transição entre os conceitos. Do ponto de vista numérico, a velocidade da mudança tem o valor máximo igual a 1 quando o intervalo entre conceitos é mínimo, e tende a 0 quando esse intervalo tende ao máximo.

O descritor de velocidade classifica as mudanças de conceito em repentina e gradual ([MINKU; WHITE; YAO, 2010](#); [MOUCHAWEH, 2016](#); [LU et al., 2019](#)). A mudança gradual é subdividida em gradual contínua e gradual probabilística ([MOUCHAWEH,](#)

2016; COSTA; GIUSTI; SANTOS, 2025). Considera-se uma mudança repentina quando  $V_{mc} = 1$ , e gradual nos demais casos.

A mudança repentina é caracterizada pela substituição do conceito antigo pelo conceito novo em um curto intervalo de tempo ( $t_{mf} - t_{mi} \approx 1$ ). Em geral, um valor próximo a 1 indica uma mudança repentina (MINKU; WHITE; YAO, 2010; LU et al., 2016). A Figura 18 ilustra a substituição repentina do conceito atual no tempo  $t$  por um novo conceito no tempo  $t+1$ . Nesse exemplo, a mudança ocorre em apenas uma unidade de tempo, e a velocidade de mudança atinge seu valor máximo, com  $V_{mc} = \frac{1}{t_{mf}-t_{mi}} = 1$ .

Por sua vez, os dois subtipos de mudança gradual são caracterizados por um intervalo de transição maior entre os conceitos, o que implica uma menor velocidade de mudança. Esse tipo de mudança tende a ser mais difícil de detectar, uma vez que envolve um período de incerteza entre os conceitos (MOUCHAWEH, 2016).

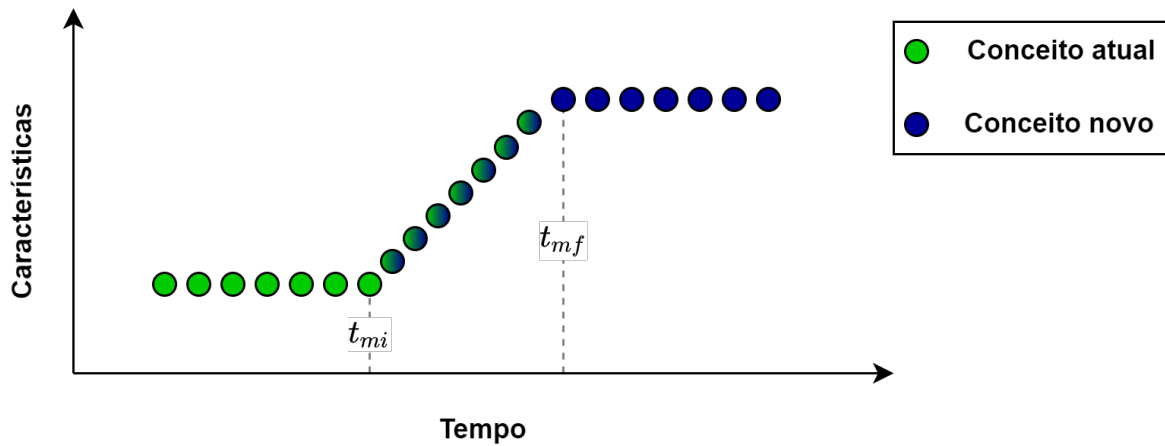
Na mudança gradual contínua, há uma lenta transformação do conceito atual para o novo, havendo um período de transição de tamanho  $n$ , onde instâncias dos dois conceitos aparecerem simultaneamente, gerando  $n$  conceitos intermediários. Não há coexistências entre os conceitos atual e novo, conforme exposto na Figura 19 (BRZEZINSKI, 2010; MOUCHAWEH, 2016; LU et al., 2019). Já na mudança gradual probabilística, a transição entre conceitos ocorre de forma lenta e envolve a coexistência do conceito atual e do novo. Com o tempo, a probabilidade associada ao conceito atual decresce progressivamente até se anular, enquanto a probabilidade do novo conceito aumenta até atingir o valor máximo, conforme exposto na Figura 20.

A Figura 21 apresenta exemplos de mudanças de conceito com diferentes velocidades. De acordo com a definição de Minku, White e Yao (2010), apenas a mudança ilustrada em (b) é considerada repentina. Entretanto, Webb et al. (2016) definem mudança de conceito repentina conforme a Equação 2.25:

$$I_{a+1} - F_a \leq \delta, \quad (2.25)$$

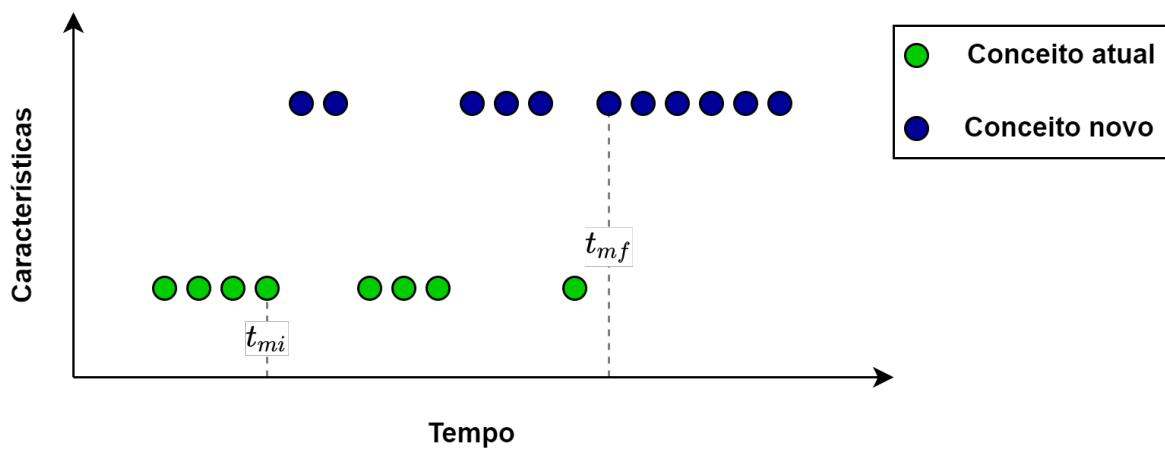
sendo que  $I_{a+1}$  e  $F_a$  representam, respectivamente, funções que retornam o ponto no tempo de início e de fim dos conceitos estáveis  $a$  e  $a + 1$ , e  $\delta \in \mathbb{N}^*$  é um valor definido dependendo do contexto que indica a duração máxima aceita para uma mudança de

Figura 19 – Exemplo de mudança de conceito gradual contínua caracterizada pela substituição suave do conceito antigo pelo conceito novo sem coexistência de conceitos.



Fonte: Adaptado de Minku, White e Yao (2010), Mouchaweh (2016) e Lu et al. (2019).

Figura 20 – Exemplo de mudança de conceito gradual probabilística caracterizada pela substituição suave do conceito antigo pelo conceito novo com coexistência de conceitos.



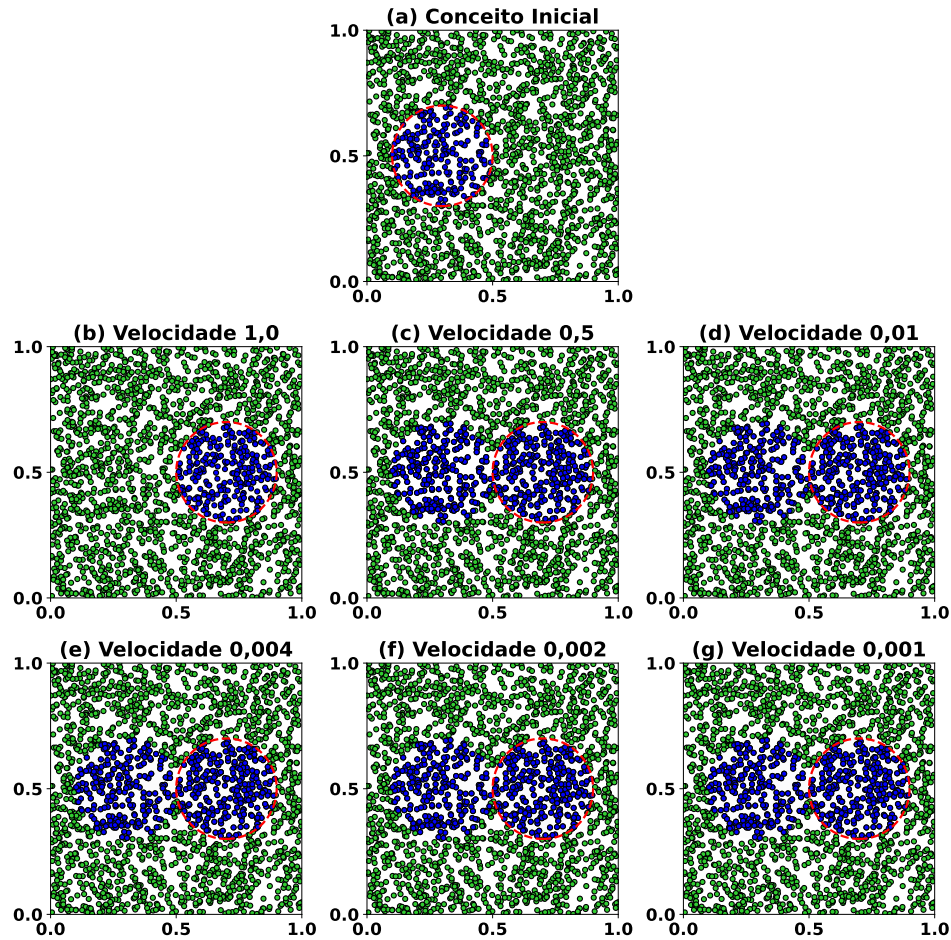
Fonte: Adaptado de Minku, White e Yao (2010), Mouchaweh (2016) e Lu et al. (2019).

conceito repentina.

De acordo a definição de Webb et al. (2016), qualquer um dos exemplos da Figura 21 pode ser classificado como uma mudança repentina, a depender do valor atribuído à  $\delta$ . Essa definição permite uma interpretação contextualizada ao problema a ser resolvido, embora haja a necessidade de definir o valor de  $\delta$ . Até o momento, não foram identificadas pesquisas que abordem a definição desse parâmetro. Embora essa definição seja mais flexível, esta pesquisa adotou a definição de Minku, White e Yao



Figura 21 – Descritor de velocidade.

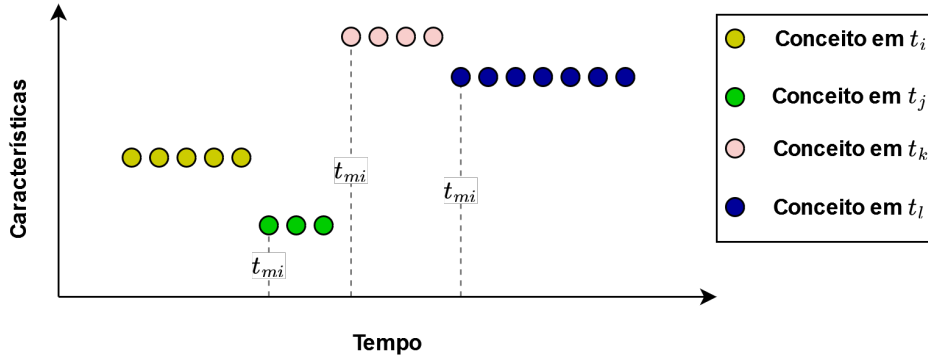


Fonte: Própria.

(2010), por ser amplamente aceita na literatura.

Em relação à estratégia de reação, mudanças de alta velocidade apresentam um desafio maior, pois o período de transição entre conceitos é curto e, como consequência, poucas instâncias estarão disponíveis para adotar a estratégia de retreinar o classificador. Por outro lado, mudanças com período maior de transição entre os conceitos permitem que mais instâncias sejam acumuladas para adotar a estratégia de retreino. Portanto, descrever a mudança em termos de sua velocidade é um elemento importante para definir a melhor estratégia de reação.

Figura 22 – Descritor de Frequência. Dois exemplos de frequência não periódica da mudança de conceito. A primeira mudança entre os conceitos amarelo e verde tem frequência igual a  $\frac{1}{4}$ . A segunda mudança tem a frequência definida em  $\frac{1}{5}$  (conceito verde e rosa). Na figura  $i < j < k < l$ .



Fonte: Própria.

### 2.3.2.2 Frequência

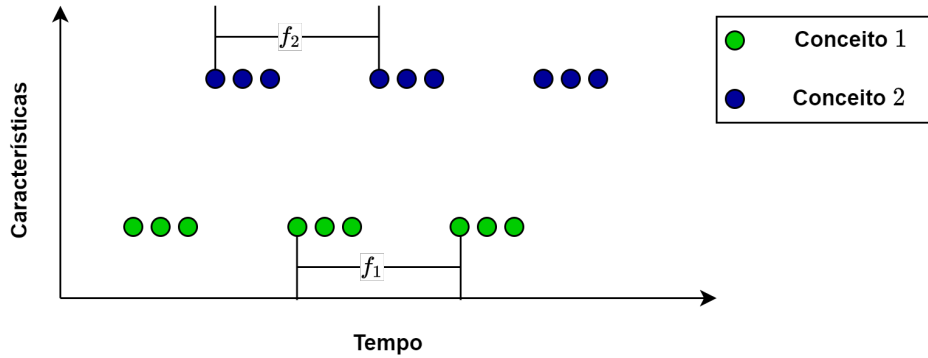
Esse descritor caracteriza o número de ocorrências de mudança em um determinado intervalo de tempo. É definido como o inverso do intervalo entre os pontos de início de dois conceitos consecutivos (KUH; PETSCHÉ; RIVEST, 1990; WIDMER; KUBAT, 1996; WEBB et al., 2016) e formalizado na Equação 2.26 (MOUCHAWEH, 2016):

$$f = \frac{1}{T}, \quad (2.26)$$

onde,  $T$  é o intervalo entre o início de dois conceitos. Assim, quando as mudanças ocorrem em intervalos constantes, são classificadas como periódicas; quando apresentam intervalos variáveis ao longo do tempo, são classificadas como não periódicas (Figura 22) (COSTA; GIUSTI; SANTOS, 2025).

Mudanças com alta frequência praticamente inviabilizam adotar a estratégia de reação baseada em retreino, pois demandam reação constante. Além disso, essas mudanças podem afetar pouco a fronteira de decisão, não necessitando do retreino do classificador. Por sua vez, mudanças de baixa frequência possibilitam que mais instâncias estejam disponíveis, fornecendo mais conhecimento para a definição da estratégia de reação.

Figura 23 – Descritor de Recorrência - exemplo de mudança de conceito recorrente cíclica. As frequências  $f_1$  e  $f_2$  representam os conceitos 1 e 2, respectivamente.



Fonte: Adaptado de Mouchaweh (2016).

### 2.3.2.3 Recorrência

Representa uma mudança que reativa um conceito previamente aprendido, caracterizando a repetição de padrões (NARASIMHAMURTHY; KUNCHEVA, 2007; MINKU; WHITE; YAO, 2010; MOUCHAWEH, 2016; WEBB et al., 2016; RAMÍREZ-GALLEGO et al., 2017; SUÁREZ-CETRULO; QUINTANA; CERVANTES, 2023). É formalizada conforme Equação 2.27, adaptada da definição proposta por Ramírez-Gallego et al. (2017):

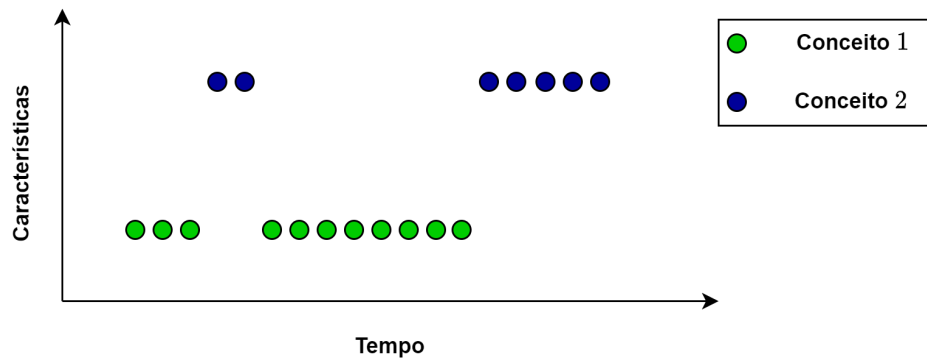
$$\exists x : p_{t_j}(x, y) = p_{t_i}(x, y), t_i < t_j, \quad (2.27)$$

onde  $p_{t_i}$  representa um conceito anteriormente conhecido.

Nesse contexto, mudanças podem ser classificadas em recorrentes ou não recorrentes (MINKU; WHITE; YAO, 2010; MOUCHAWEH, 2016; KHAMASSI et al., 2018). As mudanças recorrentes podem ser cíclicas (Figura 23), quando o ressurgimento de conceitos antigos ocorre em intervalos constantes, ou acíclicas (Figura 24), quando não há um padrão na reativação desses conceitos.

Em relação à reação, ambientes em que as mudanças são periódicas permitem que os padrões de modificações dos conceitos sejam, em tese, conhecidos. Esse conhecimento é útil na definição do tempo ideal para manter o classificador.

Figura 24 – Descritor de Recorrência - exemplo de mudança de conceito recorrente acíclica. Os conceitos previamente observados reaparecem em intervalos irregulares.



Fonte: Adaptado de [Mouchaweh \(2016\)](#).

### 2.3.3 Sem Dimensão - Previsibilidade

A previsibilidade está relacionada ao fato da mudança de conceito possuir algum padrão a ser previsto. Caso contrário, será uma mudança completamente aleatória, não sendo possível antecipar sua ocorrência ([MOUCHAWEH, 2016](#); [KHAMASSI et al., 2018](#); [WEBB et al., 2016](#)). Qualquer aspecto da mudança pode ser previsível, como por exemplo, os pontos de início e de fim da mudança, as regiões afetadas, dentre outros. É importante considerar o descritor de previsibilidade por duas razões. A primeira é que em uma mudança previsível é mais fácil entender a sua origem e seus efeitos. A segunda razão está associada às taxas de atrasos e de falsas detecções mínimas quando a mudança é previsível, sendo essas características desejadas nos detectores de mudança ([KHAMASSI et al., 2018](#)).

A propriedade de previsibilidade é chave na definição da estratégia de reação à mudança. Quando a mudança é plenamente aleatória, onde nenhum dos descritores possuem padrões a serem previstos, a detecção só será possível quando o desempenho do classificador já estiver afetado. Dessa forma, a estratégia de reação deverá lidar com a mitigação da perda de desempenho ([COSTA; GIUSTI; SANTOS, 2025](#)).

## 2.4 Considerações Finais

Neste capítulo, a base teórica que sustenta os conceitos abordados ao longo desta tese foi apresentada. Inicialmente, o termo “conceito” foi definido, bem como o fenômeno da mudança de conceito fundamentado. Por fim, as fontes probabilísticas e os descritores da mudança de conceito foram formalizados e exemplificados.

O próximo capítulo apresenta os trabalhos relacionados, organizados de acordo com as três áreas de contribuições desta tese.

---

**Trabalhos Relacionados**

---

**O**s trabalhos relacionados com as três áreas de contribuições desta tese são apresentados neste capítulo. Inicialmente, os trabalhos sobre detectores de mudança de conceito, em especial os utilizados nos Capítulos 4 e 5. Em seguida, os trabalhos que abordam os descritores da mudança e seus impactos no desempenho do classificador. E por fim, os trabalhos que abordam os descritores da mudança e as estratégias de reação à mudança.

### 3.1 Uma Visão Geral dos Métodos de Detecção de Mudança de Conceito

As abordagens de detecção da mudança de conceito são categorizadas conforme a disponibilidade de instâncias rotuladas. Essa classificação se divide em três tipos: 1) supervisionada, quando os rótulos das instâncias estão plenamente disponíveis; 2) semissupervisionada, quando o acesso aos rótulos verdadeiros é limitado; e 3) não supervisionada, quando não há disponibilidade de rótulos verdadeiros.

Na abordagem supervisionada, a detecção costuma apoiar-se em métricas de desempenho do classificador, como a taxa de erro ou sua variação estatística ao longo do tempo, baseando-se no fato de que essas variações refletem instabilidade no conceito. Na abordagem semissupervisionada, geralmente são utilizados comitês de classificadores e medidas de incerteza ou de pseudoerro para identificar eventuais mudanças de conceito. Por sua vez, na abordagem não supervisionado, a detecção, em geral, baseia-se em mudanças estruturais nas distribuições dos dados, variações na densidade das instâncias ou em discrepâncias no comportamento entre os classificadores, de modo

que possam sinalizar instabilidades no conceito e a ocorrência de mudança (GAMA et al., 2014; WEBB et al., 2016; KRAWCZYK et al., 2017; LU et al., 2019).

No contexto supervisionado, um dos detectores mais populares é o DDM (GAMA et al., 2004). Esse método monitora alterações na taxa de erro do classificador, seguindo o modelo de *Probably Approximately Correct* (PAC), ou Aprendizado Aproximadamente Correto, o qual sugere que um aumento significativo nessa taxa pode sinalizar uma mudança de conceito. Os autores destacam que o DDM apresenta melhor desempenho na detecção de mudanças repentinas. Para aumentar sua eficácia na detecção de mudanças graduais, foi proposto, com base no DDM, o EDDM (BAENA-GARCÍA et al., 2006). Em vez de acompanhar diretamente a taxa de erro, o EDDM monitora o intervalo entre erros consecutivos. Uma redução nesse intervalo pode indicar uma mudança de conceito, enquanto a estabilidade do intervalo sugere um ambiente estacionário.

No ambiente semissupervisionado, temos o método proposto em (PINAGÉ; SANTOS; GAMA, 2020), denominado de *Dynamic Selection Drift Detector* (DSDD). Esse método realiza a seleção dinâmica de classificadores a partir de um comitê inicial e utiliza um detector auxiliar de mudança, guiado por uma taxa de erro aproximada (pseudoerro) como métrica para detectar mudanças de conceito. O DSDD é composto por três módulos: (a) criação do comitê; (b) seleção dinâmica; e (c) detecção de mudança. O primeiro módulo cria um comitê de classificadores com alta diversidade, enquanto o segundo seleciona o classificador mais competente para rotular instâncias desconhecidas. O terceiro módulo monitora a taxa de pseudoerro para detectar mudanças, utilizando detectores como DDM e EDDM.

Por sua vez, o MD3-EGM (SETHI; KANTARDZIC, 2017) utiliza um comitê de classificadores para formar uma margem genérica entre classes, com base na divergência entre os classificadores ao rotular as instâncias. Esse comitê é criado por meio de subespaços aleatórios do espaço original de atributos, para o treinamento de  $n$  classificadores. A detecção de mudança ocorre pelo monitoramento da diversidade entre os classificadores, sinalizando mudança de conceito quando há aumento na discordância entre as atribuições de classe para amostras desconhecidas dentro da janela de detecção.

Por fim, no ambiente não supervisionado, o método *Drift Detection Method Based*

on *Active Learning* (DDAL), proposto em (COSTA; ALBUQUERQUE; SANTOS, 2018), baseia-se na hipótese de que variações na densidade das instâncias mais significativas podem sinalizar uma mudança de conceito. Dessa forma, é estruturado em duas fases: a primeira fase gera um classificador utilizando instâncias dentro de uma janela de referência, enquanto a segunda fase é dividida em três módulos: detecção de mudança, reação e classificação. Nessa segunda fase, o DDAL monitora continuamente a mudança de conceito em cada novo lote de dados não rotulados e, com base nos resultados da detecção, aciona os módulos de reação e classificação conforme necessário.

Outro método de detecção não supervisionada é o *Student-Teacher for Unsupervised Drift Detection* (STUDD), proposto em (CERQUEIRA et al., 2022). O STUDD é estruturado em duas fases. Na fase *offline*, um modelo professor é treinado, e suas previsões são utilizadas para treinar um modelo aluno, que aprende a imitar o professor. Na fase *online*, a discrepância entre as previsões do aluno e do professor serve como métrica para detecção de mudança, utilizando-se um detector como o teste de Page-Hinkley. Quando uma mudança é detectada, o modelo adapta-se por meio do retreinamento com observações recentes, permitindo que o aluno se ajuste ao novo conceito.

## 3.2 Análise dos Descritores de Mudança de Conceito e seus Impactos.

Esta seção apresenta um resumo dos trabalhos que abordam a natureza da mudança de conceito, bem como a compreensão dessa natureza pode fornecer evidências para a definição da estratégia de reação à mudança. Os trabalhos estão divididos em dois grupos, considerando a similaridade com a pesquisa apresentada no Capítulo 4. O primeiro grupo é composto por estudos que abordam a detecção da mudança de conceito, considerando ao menos um descritor. Por sua vez, o segundo grupo engloba os estudos que analisam as características da mudança e seus potenciais impactos no desempenho do classificador.



### 3.2.1 Detectores de Mudança de Conceito

A abordagem proposta por [Guo et al. \(2022\)](#) identifica a velocidade e a recorrência da mudança com base na saída de um detector clássico operando com duas janelas de tempo sobre os dados. A velocidade é calculada conforme definido na Equação 2.24. Enquanto a frequência é obtida a partir da razão histórica entre as acurácias do classificador em cada janela. Trata-se de uma abordagem supervisionada, o que limita sua aplicabilidade em problemas do mundo real.

A abordagem proposta por [Dasu et al. \(2006\)](#) utiliza a divergência de Kullback–Leibler ([KULLBACK; LEIBLER, 1951](#)) para mensurar a dissimilaridade entre duas janelas de dados, identificando as sub-regiões do espaço de entrada com maior divergência e, assim, localizando a região onde ocorre a mudança de conceito. Já em ([HINDER; HAMMER, 2021](#)), o espaço de entrada é decomposto em sub-regiões homogêneas por meio de uma árvore de decisão, cujo critério de divisão é baseado no  $p$ -valor do teste de Kolmogorov–Smirnov ([LOPES, 2011](#)). A árvore resultante é denominada Kolmogorov-Tree ([MBALLO; DIDAY, 2004](#)).

O método *Competence-Based Discrepancy Density Estimator* (CDDE), proposto por [Dong et al. \(2017\)](#), identifica as sub-regiões do espaço de entrada em que ocorre a mudança, por meio do mapeamento das discrepâncias entre duas janelas consecutivas de dados. Em ([DONG et al., 2021](#)), a identificação das regiões segue abordagem semelhante à de [Dong et al. \(2017\)](#), sendo que essas regiões contêm as instâncias mais antigas, que devem ser descartadas. A principal diferença entre os métodos de [Dong et al. \(2017\)](#) e [Dong et al. \(2021\)](#) reside na forma de detecção da mudança.

O método *Disposition-Based Drift Detection Method* (DBDDM), proposto por [Agrahari e Singh \(2022\)](#), adota uma abordagem baseada em duas janelas: uma fixa e outra deslizante. A janela fixa armazena exemplos representativos do conceito atual, enquanto a janela deslizante contém os exemplos mais recentes. A frequência é determinada por um teste aleatório aproximado que avalia se há diferença estatisticamente significativa entre as duas janelas consecutivas. Quando a divergência excede um determinado limiar, um contador é incrementado, representando a frequência da mudança.

O *Interpretable Drift Detector* (IDD), proposto por [Mattos et al. \(2021\)](#), identifica a

causa, as regiões afetadas e a severidade da mudança de conceito. O IDD emprega uma estrutura de árvore de decisão para analisar a acurácia dos nós. Variações na acurácia são utilizadas como base para mensurar a severidade da mudança, e o caminho entre a raiz e o nó indica as regiões afetadas. Por ser um método supervisionado, o IDD apresenta limitações em aplicações práticas nas quais a obtenção de rótulos verdadeiros é restrita.

### 3.2.2 Características da Mudanças de Conceitos e Seus Impactos no Desempenho do Classificador

Um método para descrever a mudança de conceito por meio de fluxos, caracterizando a direção e a magnitude das alterações na distribuição dos dados ao longo do tempo, foi apresentado por [Hinder, Kummert e Hammer \(2020\)](#). A proposta formaliza matematicamente o problema e oferece uma solução inicial baseada em modelos lineares, com o objetivo de possibilitar uma análise explicável das características mais afetadas pela mudança. Por meio de experimentos com dados sintéticos e reais, os autores demonstram a utilidade da abordagem em tarefas como transferência de aprendizado e visualização de mudanças. A pesquisa destaca a relevância de abordagens explicáveis no monitoramento de sistemas técnicos, e indica o uso de redes RBF e técnicas de agrupamento para esse fim.

Em ([AGUIAR; CANO, 2024](#)), uma análise abrangente do impacto da localidade (zona de influência) da mudança de conceito sobre o desempenho de classificadores e detectores é apresentada. Esse estudo também apresentou uma nova proposta de categorização da mudança de conceito com base em sua localização e na quantidade de classes afetadas. Foram propostas quatro categorias: (i) Mudança de Conceito Local de Classe Única: há apenas uma classe afetada pela mudança, impactando no máximo 50% de sua distribuição; (ii) Mudança de Conceito Global de Classe Única: há apenas uma classe afetada pela mudança, impactando ao menos 50% de sua distribuição; (iii) Mudança de Conceito Local de Múltiplas Classes: mais de uma classe é afetada pela mudança, impactando no máximo 50% da distribuição de cada classe; (iv) Mudança

de Conceito Global de Múltiplas Classes: mais de uma classe é afetada pela mudança, impactando ao menos 50% da distribuição de cada classe. Os autores indicam que a localização da mudança influencia o desempenho do classificador, de modo que compreender onde a mudança ocorre é essencial para definir a estratégia de reação mais apropriada, a qual não envolve, necessariamente, o retreinamento.

[Hammer, Vaquet e Hinder \(2024\)](#) apresentam uma análise da localidade no contexto da mudança de conceito, realizada sob a perspectiva da detecção não supervisionada. O estudo fornece uma formalização matemática para identificar as regiões do espaço de dados afetadas pela mudança e destaca a importância de traduzir essas alterações em descrições compreensíveis para operadores humanos. Diversas estratégias são discutidas, como métodos baseados em kdq-tree, técnicas de vizinhança local (LDD-DIS ([LIU et al., 2017](#))) e abordagens centradas em modelos (MB-DL). Experimentos com dados sintéticos demonstram que o método MB-DL alcançou maior acurácia, especialmente em cenários de alta dimensionalidade, desde que técnicas adequadas de pré-processamento fossem aplicadas. Além disso, os autores enfatizam que explicar as mudanças é um passo fundamental para melhorar a interpretação e a resposta em sistemas baseados em aprendizado de máquina. Também apontam que as pesquisas nessa área ainda são incipientes, indicando caminhos promissores para avanços metodológicos e práticos.

A abordagem denominada *Quadruple-Based Approach for Understanding Concept Drift in Data Streams* (QuadCDD), proposta em ([WANG et al., 2024](#)), tem como objetivo não apenas detectar a mudança, mas também fornecer informações relevantes para apoiar a decisão sobre estratégias de reação com base na velocidade e na severidade. O QuadCDD utiliza informações prévias sobre os pontos de início e fim da mudança para definir a velocidade. Também utiliza os rótulos verdadeiros das amostras para mensurar a severidade. Portanto, essa abordagem é totalmente supervisionada, uma vez que pressupõe o conhecimento prévio dos pontos de início e término da mudança. Ressalta-se que, nessa abordagem, a severidade é definida como a diferença na acurácia do classificador. Essa definição diverge daquela apresentada na Seção 2.3.1.1.

### 3.2.3 Análise crítica-comparativa dos trabalhos relacionados

Em relação ao primeiro grupo de trabalhos relacionados, algumas observações podem ser feitas. Primeiramente, os descritores baseados na dimensão espacial são os mais frequentemente abordados, sendo a zona de influência examinada em (DASU et al., 2006), (HINDER; HAMMER, 2021), (DONG et al., 2017) e (DONG et al., 2021), e a severidade em (HINDER; KUMMERT; HAMMER, 2020) e (MATTOS et al., 2021). Isso sugere que há um entendimento mais consolidado da dimensão espacial na literatura que trata do fenômeno da mudança.

Uma segunda observação é que esses estudos, com exceção de (GUO et al., 2022), abordam apenas um descritor, indicando que a combinação de descritores aumenta a complexidade do problema, o que, conseqüentemente, constitui um ponto menos explorado na literatura.

Por fim, nenhum dos trabalhos desse grupo realiza uma análise da natureza da mudança de conceito e de seus potenciais impactos no desempenho do classificador. Diferentemente desses trabalhos, a pesquisa apresentada no Capítulo 4 foca na análise do impacto dos descritores sobre o desempenho do classificador e na identificação de evidências úteis para a definição de uma melhor estratégia de reação. Além disso, esta pesquisa não se restringe aos descritores baseados na dimensão espacial, considerando também a relação entre diferentes descritores, e não apenas de forma isolada.

Por sua vez, os trabalhos englobados no segundo grupo são os mais semelhantes à abordagem apresentada neste trabalho. No entanto, enquanto os trabalhos deste grupo se limitam à análise de dois descritores (WANG et al., 2024) ou de um único descritor (HINDER; KUMMERT; HAMMER, 2020; AGUIAR; CANO, 2024; HAMMER; VAQUET; HINDER, 2024), o presente estudo se destaca por analisar quatro descritores. Quanto ao paradigma de aprendizado, em (HINDER; KUMMERT; HAMMER, 2020), (WANG et al., 2024) e (AGUIAR; CANO, 2024), é analisado apenas o paradigma supervisionado, enquanto (HAMMER; VAQUET; HINDER, 2024) focam exclusivamente no paradigma não supervisionado. Pelo que se tem conhecimento, este é o primeiro trabalho a analisar os paradigmas supervisionado, semi-supervisionado e não supervisionado, distinguindo-se por oferecer uma análise mais abrangente e integrada das

Tabela 2 – Comparação dos trabalhos relacionados em termos do descritor avaliado.

Trabalho	Descritor <sup>1</sup>						Grupo
	VE	SV	ZI	FQ	RR	PV	
Guo et al. (2022)	•				•		1
Dasu et al. (2006)			•				
Hinder e Hammer (2021)			•				
Dong et al. (2017)			•				
Dong et al. (2021)			•				
Agrahari e Singh (2022)				•			
Mattos et al. (2021)		•					2
Hinder, Kummert e Hammer (2020)		•					
Aguiar e Cano (2024)			•				
Hammer, Vaquet e Hinder (2024)			•				
Wang et al. (2024)	•	•					
Este trabalho (Capítulo 4)	•	•		•	•		

<sup>1</sup> VE:Velocidade - SV:Severidade. - ZI:Zona de Influência - FQ:Frequência -RR:Recorrência - PV:Previsibilidade.

características da mudança de conceito.

A Tabela 2 sintetiza a comparação dos trabalhos relacionados.

### 3.3 Mudança de Conceito e Estratégias de Reação: Uma Análise Abrangente com o Uso de Descritores

Há diversos trabalhos na literatura que abordam unicamente a detecção de mudança de conceito, como os descritos em (IWASHITA; PAPA, 2019),(GEMAQUE et al., 2020) e (HINDER; VAQUET; HAMMER, 2024). Entretanto, há um número reduzido de estudos que exploram a natureza desse fenômeno, bem como as relações entre suas características, os mecanismos de detecção e as estratégias de reação. Nesta seção, alguns dos trabalhos que exploram essa lacuna são detalhados. Os trabalhos foram classificados de acordo com sua abordagem principal e estão organizados em dois grupos: (i) trabalhos que abordam principalmente a detecção da mudança de conceito e (ii) trabalhos que abordam principalmente a natureza da mudança de conceito.

### 3.3.1 Estudos sobre Detecção de Mudança de Conceito

O QuadCDD (WANG et al., 2024), em complemento a descrição realizada na Seção 3.2.2, é uma abordagem para a detecção e compreensão da mudança de conceito. A proposta visa superar limitações de métodos tradicionais que focam apenas no ponto de início da mudança, incorporando quatro dimensões: início, fim, severidade e tipo da mudança ( $D_s, D_e, D_v, D_t$ ). A arquitetura do QuadCDD combina pré-treinamento, ajuste fino adaptativo, detecção e tomada de decisão, utilizando redes neurais com camadas bidirecionais LSTM para gerar quádruplos informativos. A severidade é calculada com base na queda de acurácia entre os conceitos, enquanto que a velocidade é classificada como repentina ou incremental. O modelo demonstrou elevadas taxas de predição de pontos de mudança em experimentos com bases de dados artificiais e reais, alcançando acurácia média de 93,9% na identificação do tipo de mudança e reduzindo atraso na detecção. A proposta oferece um avanço ao integrar diagnóstico detalhado e mecanismos de reação adaptativos, com o aprendizado incremental e o retreino total, mantendo estabilidade e precisão preditiva em ambientes de dados dinâmicos.

Por sua vez, Yang et al. (2025) propõem o *Causal Drift Detection and Rationalization Method* (CDDRM), um método unificado para detecção e racionalização da mudança de conceito, baseado na análise de relações causais entre os atributos e a variável alvo. Diferentemente das abordagens tradicionais que se baseiam em medidas de erro ou associação, o CDDRM detecta mudanças significativas na intensidade e importância das relações causais, utilizando o algoritmo NOTEARS para descoberta causal e o teste qui-quadrado ( $\chi^2$ ) para validação estatística. A mudança é classificada em quatro tipos: mudança de causa (Tipo 1 e 2), mudança na importância (Tipo 3) e mudança na intensidade da relação causal (Tipo 4). O método demonstrou desempenho superior aos métodos clássicos, como DDM, HDDM, ADWIN, PHM e LLDD, com F1 score médio de até 0,91 em experimentos com as bases de dados STAGGER e Agrawal, incluindo cenários de mudança real e virtual. A abordagem também foi aplicada a dados reais de mineração de carvão, destacando mudanças na relação entre pressão hidráulica e energia micro-sísmica como potenciais indicadores de risco operacional. O CDDRM utiliza a estratégia de retreino total adotada pelos detectores comparado.

### 3.3.2 Estudos sobre a Natureza da Mudança de Conceito

Adicionalmente a descrição realizada na Seção 3.2.2, [Hinder, Kummert e Hammer \(2020\)](#) propuseram um método para explicação de mudança de conceito por meio de fluxos vetoriais, isto é, mudanças de direção e magnitude no espaço de dados ao longo do tempo. O enfoque central é estimar a direção média da mudança como uma função vetorial linear que representa a transformação do conjunto de dados de um instante para o seguinte. O método formaliza o problema por meio da minimização da divergência de Kullback-Leibler entre distribuições observadas e previstas, considerando tanto dados determinísticos quanto probabilísticos. A abordagem utiliza diferentes variantes da média vetorial para acomodar distintos cenários de amostragem. Testes empíricos em dados sintéticos e reais mostraram que a técnica permite identificar variáveis mais afetadas pela mudança e oferece desempenho competitivo em tarefas de transferência de aprendizado, além de permitir interpretação direta da dinâmica da mudança.

Em um novo trabalho, [Hinder et al. \(2022\)](#) propuseram uma estrutura formal para a localização da mudança, com o objetivo de identificar pontos de dados específicos e regiões do espaço de entrada afetadas pela mudança. Diferentemente da maioria das abordagens que se concentram na detecção ou quantificação global da mudança, o método proposto reformula a localização da mudança como um problema de classificação supervisionada, permitindo o uso de modelos de aprendizado de máquina clássicos (por exemplo, árvores de decisão, k-NN e florestas aleatória) para aproximar a probabilidade de ocorrência da mudança em um determinado ponto. O núcleo do método é um teste estatístico baseado em métricas de informatividade derivadas da incerteza do classificador, como a divergência de Kullback-Leibler entre distribuições preditas e uniformes. A abordagem superou o kdq-Tree ([DASU et al., 2006](#)) e o LDD-DIS ([LIU et al., 2017](#)) em bases de dados de referência como SEA e Blobs, alcançando identificação mais precisa das regiões afetadas pela mudança. Adicionalmente, o método permite a análise de relevância de atributos, auxiliando na interpretação das origens da mudança.

[Aguiar e Cano \(2024\)](#) também investigam a localidade da mudança de conceito, destacando a intensidade (severidade) e da localização (zona de influência) da mudança no desempenho dos classificadores. Os autores propõem uma nova categorização



baseada no número de classes afetadas e na severidade da mudança (mudança local vs. global; monoclasses vs. multiclasse), e constroem um banco de testes com 2.760 bases de dados sintéticas, geradas a partir dos algoritmos Random RBF e Random Tree. Nove detectores supervisionados são comparados, incluindo ADWIN ([BIFET; GAVALDÀ, 2007](#)), PH ([PAGE, 1954](#)) e FHDDM ([PESARANGHADER; VIKTOR, 2016](#)), sob diferentes níveis de dificuldade, abrangendo mudanças repentinas, graduais e incrementais. Os resultados revelam que detectores como PH e ADWIN apresentam desempenho mais equilibrado entre precisão e tempo de detecção, enquanto outros, como EDDM e STEPD ([NISHIDA; YAMAUCHI, 2007](#)), obtêm alta sensibilidade com custo de muitos falsos positivos. A análise evidencia que mudanças incrementais e locais são mais desafiadoras de detectar, especialmente em cenários multiclasse, e propõe diretrizes para futuras pesquisas, incluindo a necessidade de detectores menos sensíveis à distribuição de erro e mais eficazes em cenários realistas.

[Hammer, Vaquet e Hinder \(2024\)](#) apresentam uma revisão abrangente sobre métodos para localização e explicação da mudança de conceito, com ênfase em abordagens não supervisionadas. Os autores formalizam rigorosamente o conceito de mudança em processos de distribuição ao longo do tempo e introduzem a noção de *drift locus*, a menor região no espaço de dados cuja remoção tornaria o processo estacionário. O artigo analisa criticamente quatro estratégias principais de localização: kdq-tree ([DASU et al., 2006](#)), LDD-DIS ([LIU et al., 2017](#)), segmentação por deriva e métodos baseados em modelos supervisionados. Para cada abordagem, são discutidos os quatro estágios operacionais (aquisição de dados, construção do descritor, cálculo de dissimilaridade e normalização) e suas implicações práticas. A análise experimental revela que métodos baseados em modelos, especialmente aqueles que utilizam florestas aleatórias, superam as abordagens baseadas em partições fixas (exemplo, kdq-tree), especialmente em cenários com mudanças de alta severidade. O estudo também discute as limitações das técnicas tradicionais de detecção baseadas no desempenho do modelo e defende o uso de técnicas explicativas como complemento essencial para operadores humanos, especialmente em sistemas críticos.

Por fim, [Komorniczak e Ksieniewicz \(2024\)](#) investigaram a capacidade de meta-



atributos na identificação de conceitos em dados não estacionários, com foco na detecção implícita da mudança de conceito. Os autores avaliaram mais de 160 meta-atributos provenientes de nove categorias (complexidade, estatística, teoria da informação, entre outras), em bases de dados sintéticos, semi-sintéticos e reais, com mudanças de diferentes velocidades (repentina, gradual e incremental). Cinco experimentos foram conduzidos para verificar a efetividade desses meta-atributos na identificação de conceitos, incluindo análises de similaridade, classificação, seleção de atributos e análise de variância. Os resultados apontam que os meta-atributos da categoria estatística são altamente eficazes em dados sintéticos e semi-sintéticos, enquanto nos dados reais se destacam meta-atributos de teoria da informação, agrupamento e complexidade. Um conjunto reduzido de 17 meta-atributos foi proposto como promissor para reconhecimento de conceito, sendo validados em múltiplos classificadores. Os autores concluem que a origem dos dados influencia diretamente o valor preditivo dos meta-atributos e recomendam a seleção personalizada conforme o tipo de dado.

### 3.3.3 Uma breve análise comparativa dos trabalhos relacionados

A Tabela 3 apresenta uma comparação entre os trabalhos relacionados quanto aos descritores de mudança abordados, estratégias de reação adotadas e suas respectivas abordagens. Observa-se que poucos estudos consideram descritores variados.

O trabalho de [Wang et al. \(2024\)](#), por exemplo, incorpora os descritores de velocidade e severidade, com estratégias de reação do tipo incremental e de retreino total. Outros estudos, como ([HINDER; KUMMERT; HAMMER, 2020](#)), ([HINDER et al., 2022](#)) e ([HAMMER; VAQUET; HINDER, 2024](#)), focam em apenas um descritor específico, sem propor estratégias de reação. Por sua vez, [Aguiar e Cano \(2024\)](#) consideram a zona de influência, mas com abordagem reativa limitada. Destacam-se ainda os trabalhos que propõem abordagens alternativas: baseado em causalidade ([YANG et al., 2025](#)) e que adota meta-atributos como substitutos aos descritores tradicionais ([KOMORNICZAK; KSIENIEWICZ, 2024](#)).

A pesquisa descrita no Capítulo 5 diferencia-se por integrar quatro descritores

(VE, SV, FQ, RR) e analisar cinco estratégias de reação (estática, padrão, incremental, esquecimento e oráculo), contribuindo para uma análise mais abrangente da mudança e reforçando a importância de alinhar a detecção à natureza do fenômeno observado.

Tabela 3 – Comparação dos trabalhos relacionados em termos de descritor avaliado.

Trabalho	Descritor <sup>1</sup>						ER <sup>2</sup>	Abordagem
	VE	SV	ZI	FQ	RR	PV		
Wang et al. (2024)	•	•					Incremental, Padrão	Detecção
Yang et al. (2025)			Causalidade <sup>3</sup>				×	
Hinder, Kummert e Hammer (2020)		•					×	
Hinder et al. (2022)			•				×	Análise
Aguiar e Cano (2024)			•				Incremental, Padrão	
Hammer, Vaquet e Hinder (2024)			•				×	
Komorniczak e Ksieniewicz (2024)			Meta-atributos <sup>4</sup>				×	
Este trabalho (Capítulo 5)	•	•		•	•		Estática, Padrão, Incremental, Esquecimento, Oráculo	Análise

<sup>1</sup> VE:Velocidade - SV:Severidade. - ZI:Zona de Influência - FQ:Frequência -RR:Recorrência - PV:Previsibilidade.

<sup>2</sup> ER: Estratégia de Reação.

<sup>3</sup> Em vez de utilizar descritores tradicionais como velocidade, severidade, zona de influência ou frequência da mudança de conceito, o estudo propõe uma abordagem baseada em causalidade para detectar e racionalizar a mudança de conceito.

<sup>4</sup> Em vez de utilizar descritores tradicionais como velocidade, severidade, zona de influência ou frequência da mudança de conceito, o estudo propõe uma abordagem baseada em meta-atributos para caracterizar a mudança de conceito.

### 3.4 Considerações Finais

Neste capítulo, os trabalhos relacionados às contribuições desta tese foram brevemente apresentados, analisados e comparados com o presente trabalho. O próximo capítulo apresenta uma análise dos descritores de mudança de conceito e seus impactos.

---

### Análise dos Descritores de Mudança de Conceito e seus Impactos

---

Neste capítulo nós analisamos a influência de descritores na mudança de conceito por meio da variação dos valores de cada descritor. Para alcançar esse objetivo foram conduzidos experimentos com cinco bases de dados sintéticas em nove contextos distintos: seis diferentes detectores de mudança amplamente conhecidos na literatura - dois supervisionados, DDM ([GAMA et al., 2004](#)) e EDDM ([BAENA-GARCÍA et al., 2006](#)); dois semi-supervisionados, DSDD ([PINAGE; SANTOS, 2015](#)) e MD3-EGM ([SETHI; KANTARDZIC, 2017](#)); e dois não-supervisionados, DDAL ([COSTA; ALBUQUERQUE; SANTOS, 2018](#)) e STUDD ([CERQUEIRA et al., 2022](#)).

Para tentar avaliar os descritores em um ambiente de detecção perfeita, versões oráculos dos detectores supervisionados foram incorporadas ao experimento (DDM-O e EDDM-O). Um oráculo representa um detector ideal, que conhece previamente todos os pontos de ocorrência de mudança, não havendo perdas, falsas detecções ou atrasos nas detecções. O nono contexto é o sem detecção (SD), que foi analisado para prover uma visão do impacto dos descritores no cenário em que não ocorre detecção de mudança.

Foram investigadas 32 variações de valores dos descritores, chamadas aqui de cenários. Considerando as cinco bases de dados estudadas, um total de 160 cenários distintos foram testados nos nove contextos descritos acima. Em todos os cenários foi utilizada uma Árvore de Decisão como classificador, sendo ao todo analisadas 1.440 combinações distintas. É importante mencionar que todos os detectores foram aplicados usando os valores padrões definidos por seus autores.

Os experimentos foram divididos em duas séries (Figura 25). A primeira série é constituída por quatro etapas e tem como objetivo mensurar o desempenho do classificador (C) em termos de erro frequencial, e o desempenho do detector (D) em relação aos percentuais de falsas detecções, perdas e atraso das detecções. Nesse cenário, uma

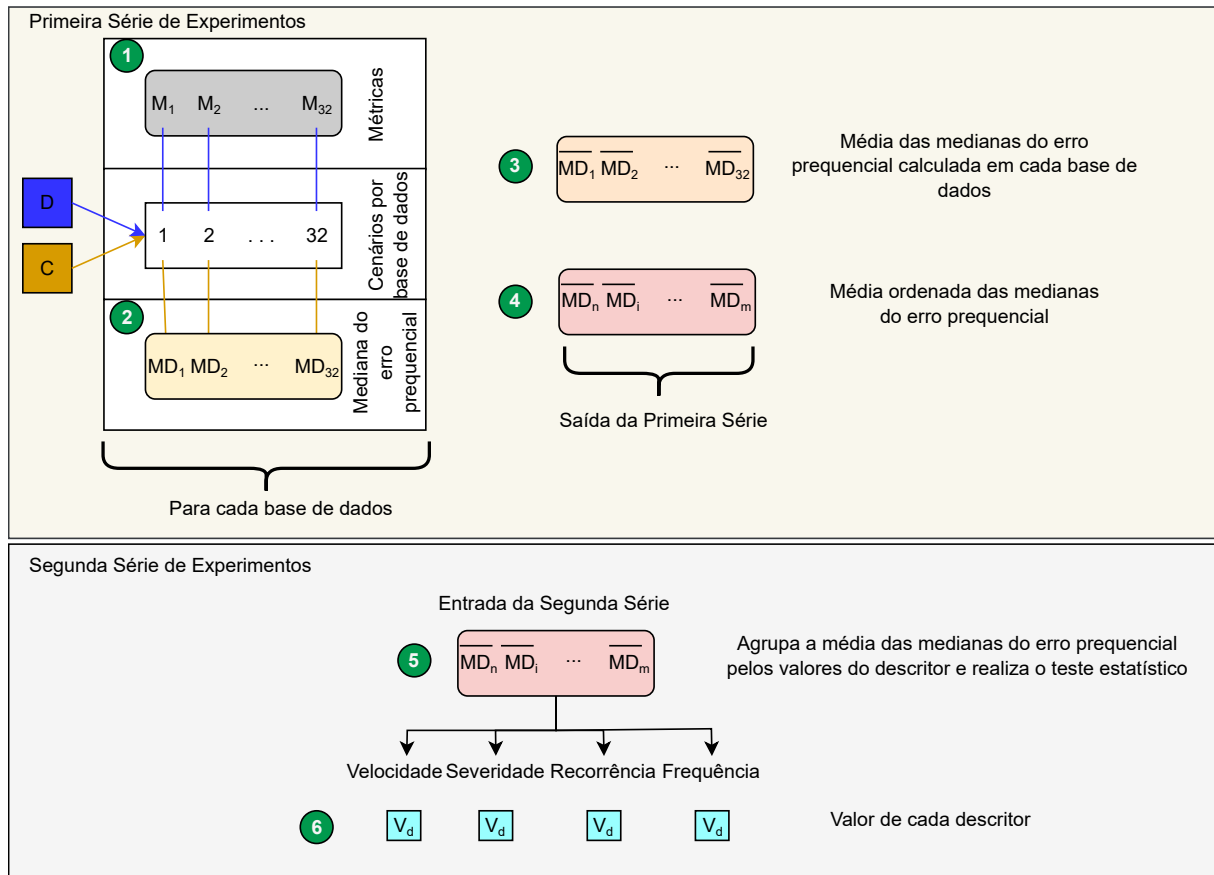


Figura 25 – Visão geral do experimento.

falsa detecção é definida como sendo uma detecção que ocorre antes de uma mudança conhecida; uma perda ocorre quando uma mudança conhecida não é detectada; e o atraso é definido como sendo uma detecção que ocorre após uma mudança conhecida mas que não se caracteriza como uma falsa detecção.

Os resultados obtidos na primeira série são utilizados como entrada para a segunda série de experimentos. Esta é constituída por duas etapas e tem como objetivo identificar as combinações de valores dos descritores de mudança de conceito que geram os melhores e os piores resultados considerando o ranqueamento dos cenários feito com base na mediana do erro prequencial. É importante destacar que ambas as séries são repetidas para cada um dos nove contextos.

Na primeira etapa, as métricas ( $M_i$ ) de desempenho do detector são calculadas para cada um dos 32 cenários de cada uma das cinco bases de dados utilizadas. Do mesmo modo, para cada cenário é computada a mediana do erro prequencial ( $MD_i$ )

do classificador na segunda etapa. Na terceira etapa, a média das ( $MD_i$ ) é calculada considerando as medianas do erro frequencial do mesmo cenário  $i$  para cada uma das bases de dados. As médias formam uma lista do desempenho mediano do classificador em cada cenário nas diferentes bases. Finalmente, a lista gerada na etapa anterior é ordenada de forma crescente na última etapa. Essa lista indica em quais cenários o classificador e o detector obtiveram os melhores e os piores desempenhos. Observa-se que, por se tratar de erro frequencial, quanto menor o valor, melhor é o desempenho.

A segunda série dos experimentos tem como foco fazer análise estatística dos resultados. Para tanto, inicia com a quinta etapa, que divide os cenários ranqueados na etapa anterior de acordo com os valores dos descritores para cada um dos quatro descritores analisados. Para os descritores de velocidade, severidade e recorrência, os subgrupos formados são comparados por meio do teste estatístico de Mann-Whitney (MANN; WHITNEY, 1947) e os subgrupos formados no descritor de frequência são comparados por meio do teste estatístico de Friedman (FRIEDMAN, 1937), todos na etapa seis. Em ambos os testes, o nível de significância foi definido em 0,05. No final, os resultados gerados indicam qual valor por descritor leva a um melhor ou pior desempenho ( $V_d$ ), considerando os cenários analisados (Seção 4.2).

Os resultados obtidos nas duas séries de experimentos são discutidos nas Seções 4.4 e 4.5. Inicialmente, apresenta-se as bases de dados sintéticas utilizadas, em sequência os cenários analisados.

## 4.1 Bases de Dados

No contexto de mudança de conceito, as bases de dados reais apresentam limitações significativas, principalmente devido à dificuldade em determinar com precisão a localização das mudanças ou mesmo em comprovar sua existência. Além disso, há um número restrito de bases de dados reais conhecidas que efetivamente contêm mudanças (LU et al., 2019; SOUZA et al., 2020; KOMORNICZAK; KSIENIEWICZ, 2024).

Algumas abordagens sugerem a injeção de mudanças em bases de dados reais. Entretanto, essas estratégias geralmente se limitam à alteração do descritor de veloci-

dade (MASTELINI et al., 2018; STEVANOSKI et al., 2024), o que é insuficiente para os objetivos deste estudo. Este trabalho foca nos descritores de velocidade, severidade, frequência e recorrência, cuja análise demanda maior controle sobre as características das mudanças. Bases de dados sintéticas, por sua vez, permitem modificar diretamente os valores das instâncias para simular alterações nos descritores, oferecendo uma oportunidade de compreender melhor o impacto de cada um.

Por essa razão, optou-se pela utilização exclusiva de bases de dados sintéticas nos experimentos, uma vez que a avaliação do impacto dos descritores seria limitada ou mesmo inviável em muitos casos com bases de dados reais. Ressalta-se que essa escolha não constitui uma limitação do estudo, mas sim uma decisão metodológica que visa assegurar a execução adequada da pesquisa e a consistência dos resultados.

Portanto, no presente trabalho foram geradas 32 variações para cada uma das cinco bases de dados sintéticas fornecidas em (MINKU; WHITE; YAO, 2010). Cada base representa um problema de classificação binária com 10.000 instâncias, balanceado, sem ruído e com dois atributos, com exceção da base *Hiperplane* que contém três atributos.

As bases *Line* e *Hiperplane* são definidas pela Equação 4.1:

$$\sum_{i=1}^d a_i x_i, \quad (4.1)$$

sendo  $d = 1$  e  $d = 2$  para as bases *Line* e *Hiperplano*, respectivamente. Nos experimentos, o valor de  $a_i$  foi variado. Já as bases *SineH* e *SineV* são definidas pela Equação 4.2:

$$y = a \sin(bx) + c, \quad (4.2)$$

sendo que o valor de  $b$  foi variado em *SineH*, afetando o eixo horizontal, enquanto em *SineV* foi variado o valor de  $c$ , afetando o eixo vertical. Por fim, a base *Circle* é definida pela Equação 4.3:

$$(x - a)^2 + (y - b)^2 = r^2, \quad (4.3)$$

com o valor de  $r$  sendo variado em nossos experimentos.

Adicionalmente, para cada base de dados, parâmetros específicos foram ajustados para simular diferentes tipos de mudança de conceito. Por exemplo, na base *Line*,

mudanças de alta severidade foram criadas variando o valor de  $a_0$  em 0,70 para cada conceito, enquanto para mudanças de baixa severidade, os valores foram variados em 0,15. Similarmente, na base *Circle*, mudanças de alta severidade foram criadas deslocando o valor do raio do círculo ( $r$ ) por 0,30 para cada novo conceito, enquanto para mudanças de baixa severidade o valor foi deslocado em 0,10.

Em todas as bases de dados, mudanças recorrentes foram simuladas repetindo os conceitos a partir da metade das mudanças conhecidas. As frequências da mudança de conceito foram simuladas ajustando o intervalo entre as ocorrências das mudanças. Por exemplo, nos cenários de alta frequência, uma mudança ocorre a cada 250 instâncias, enquanto que em cenários de baixa frequência, a mudança ocorre a cada 3,000 instâncias. Por último, mudanças repentinas foram simuladas substituindo o conceito atual pelo novo conceito após uma única instância, enquanto que mudanças graduais foram simuladas por uma substituição progressiva do conceito atual a cada 250 instâncias. A Tabela 13 (Apêndice B) fornece um resumo detalhado do ajuste de cada parâmetro. As variações dos valores em cada base estão indicadas nas colunas Início, Fim e Passo.

## 4.2 Cenários

Os cenários avaliados no experimento estão descritos na Tabela 4 e foram gerados por meio da variação dos valores dos descritores de mudança de conceito. Quatro descritores foram analisados: velocidade, severidade, recorrência e frequência. O descritor Zona de Influência não foi incluído na análise devido à limitação em simular com precisão a localização da mudança dentro do fluxo de dados.

Na Tabela 4, a primeira coluna apresenta o identificador do cenário. A segunda coluna exibe os descritores de mudança, seguindo a ordem: Velocidade – Severidade – Recorrência – Frequência. Os valores de cada descritor são representados por suas iniciais e estão detalhados no rodapé da tabela. A coluna Qtd. Mudanças indica o número de mudanças presentes, e as duas últimas colunas representam, respectivamente, a velocidade e a frequência da mudança. Cenários em que a frequência apresenta apenas um valor indicam que a mudança é periódica. Mudanças não periódicas são indicadas

Tabela 4 – Valores dos descritores em cada cenário.

Cenário	Descrição <sup>1</sup>	Qtd. Mudanças	Velocidade	Frequência
1	R-A-R-PA	39	1	250
2	R-A-R-PM	9	1	1000
3	R-A-R-PB	3	1	3000
4	R-A-R-NP	9	1	Alt <sup>2</sup>
5	R-A-NR-PA	39	1	250
6	R-A-NR-PM	9	1	1000
7	R-A-NR-PB	3	1	3000
8	R-A-NR-NP	9	1	Alt <sup>2</sup>
9	R-B-R-PA	39	1	250
10	R-B-R-PM	9	1	1000
11	R-B-R-PB	3	1	3000
12	R-B-R-NP	9	1	Alt <sup>2</sup>
13	R-B-NR-PA	39	1	250
14	R-B-NR-PM	9	1	1000
15	R-B-NR-PB	3	1	3000
16	R-B-NR-NP	9	1	Alt <sup>2</sup>
17	G-A-R-PA	39	250	250
18	G-A-R-PM	9	250	1000
19	G-A-R-PB	3	250	3000
20	G-A-R-NP	9	250	Alt <sup>2</sup>
21	G-A-NR-PA	39	250	250
22	G-A-NR-PM	9	250	1000
23	G-A-NR-PB	3	250	3000
24	G-A-NR-NP	9	250	Alt <sup>2</sup>
25	G-B-R-PA	39	250	250
26	G-B-R-PM	9	250	1000
27	G-B-R-PB	3	250	3000
28	G-B-R-NP	9	250	Alt <sup>2</sup>
29	G-B-NR-PA	39	250	250
30	G-B-NR-PM	9	250	1000
31	G-B-NR-PB	3	250	3000
32	G-B-NR-NP	9	250	Alt <sup>2</sup>

<sup>1</sup> Descrição: Velocidade-Severidade-Recorrência-Frequência. Velocidade (R) Repentina (G) Gradual. Severidade (A) Alta (B) Baixa. Recorrência (R) Recorrente (NR) Não Recorrente. Frequência (PA) Periódica Alta (PM) Periódica Média (PB) Periódica Baixa (NP) Não-Periódica.

<sup>2</sup> Alt = [1000 4000 4250 4500 4750 5000 8000 9000 9250].

por "Alt", e seus valores estão detalhados em nota de rodapé. Como exemplo, no cenário 1, a mudança de conceito é definida como (R) Repentina, (A) Alta, (R) Recorrente e (PA) Periódica Alta, respectivamente para os descritores de velocidade, severidade, recorrência e frequência. Nesse cenário, ocorrem 39 mudanças, todas com velocidade igual a 1, ocorrendo a cada 250 instâncias.



### 4.3 Detectores de Mudança de Conceito

A análise apresentada neste capítulo abrange os ambientes supervisionado, semisupervisionado e não supervisionado, considerando dois detectores para cada um dos ambientes, o que representa um diferencial em relação ao trabalho de [Aguiar e Cano \(2024\)](#), que se limita ao ambiente supervisionado. Os detectores utilizados no presente experimento foram descritos na Seção 3.1.

### 4.4 Primeira Série de Experimentos

Os resultados da primeira série de experimentos estão resumidos nas figuras do Apêndice C. Destaca-se que os cenários DDM-O, EDDM-O e SD não estão representados nas figuras porque a detecção é perfeita (oráculo) nos dois primeiros; e não há detecção no último.

Ao considerar os detectores supervisionados, o DDM e o EDDM apresentaram as menores taxas de atraso na detecção. No entanto, o DDM alcançou taxas de perda de detecção geralmente superiores a 60%, além de uma alta porcentagem de detecções falsas, com algumas exceções. Em relação ao EDDM, a porcentagem de perda de detecção não segue um padrão claro e depende do cenário, enquanto a taxa de detecção falsa geralmente permanece abaixo de 50%.

Nos cenários não supervisionados, surpreendentemente, o DDAL e o STUDD não detectaram mudança de conceito. Consequentemente, as porcentagens de detecção falsa e atrasada foram zero, enquanto a perda de detecção foi máxima. O desempenho dos detectores não supervisionados foi, portanto, semelhante ao cenário sem detecção. Por fim, os dois métodos semisupervisionados apresentaram resultados divergentes entre si. Por um lado, o MD3-EGM e o DSDD tiveram desempenhos opostos em relação à perda de detecção. Por outro, o MD3-EGM apresentou uma menor taxa de detecção com atraso, enquanto o DSDD obteve uma menor taxa de falsa detecção.

A comparação direta entre os seis detectores indica uma relação de ordem entre eles. DDM e EDDM apresentaram o melhor desempenho, seguidos por DSDD e MD3-EGM. Os piores foram os métodos não supervisionados, pois não detectaram nenhuma mudança de conceito. Esse desempenho insatisfatório pode ser atribuído aos valores dos hiperparâmetros dos modelos. O sucesso de métodos não supervisionados geralmente

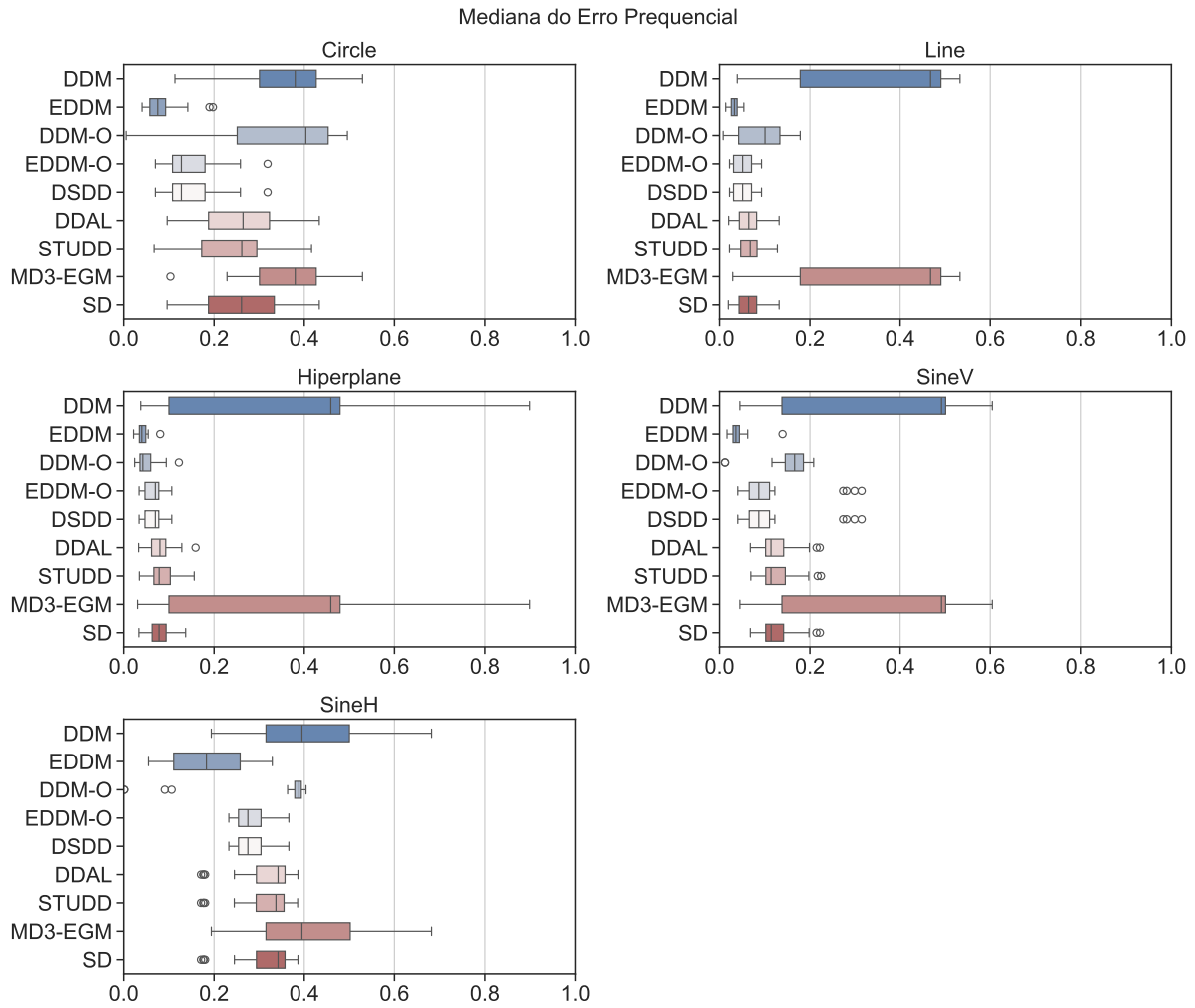


Figura 26 – Valores das medianas dos erro prequencial nas bases de dados.

depende do ajuste fino de diversos hiperparâmetros. No entanto, nos experimentos, foram utilizados os valores padrão dos métodos, sem ajuste-fino.

Ao analisar as medianas do erro prequencial do classificador em cada contexto, apresentadas na Figura 26, é possível destacar o seguinte. Os contextos DDM e MD3-EGM apresentaram o pior desempenho, conforme indicado por suas maiores medianas de erro prequencial em todos os conjuntos de dados. Além disso, o DDM-O também apresentou uma alta taxa de erro no conjunto de dados *Circle*. Resultados opostos são observados para o EDDM, que alcançou a menor mediana de erro em todos os cenários. Esse resultado é consistente com as taxas de falsa detecção e de atraso obtidas pelo EDDM (Figura 42).

Também é importante comparar os detectores supervisionados com suas versões

oráculo em termos de erro frequencial. O DDM-O superou o DDM, exceto no conjunto de dados *Circle*. Por sua vez, o EDDM apresentou desempenho melhor ou igual ao do EDDM-O. Considerando o contexto sem detecção, seu desempenho foi semelhante aos demais contextos. Isso sugere que a melhor forma de reagir à mudança de conceito nem sempre é retrainar os modelos. Essa observação é reforçada quando se consideram os métodos não supervisionados, cujos desempenhos são semelhantes aos observados em cenários sem detecção de mudança.

## 4.5 Segunda Série de Experimentos

Os resultados desta série de experimentos são apresentados na Tabela 5 e nas Figuras 27 e 28. O objetivo desta etapa é comparar os valores dos descritores a fim de identificar aqueles com maior influência sobre a mudança de conceito, considerando todos os cenários e conjuntos de dados analisados. A Figura 27 ilustra essa comparação entre os detectores supervisionados e suas versões oráculo, enquanto a Figura 28 apresenta os resultados obtidos com os demais detectores.

A Tabela 5 apresenta a mediana do erro frequencial obtida mediante a variação dos valores dos descritores. Os valores em **negrito** indicam o valor do descritor que gerou a menor mediana de erro frequencial, ou seja, o melhor desempenho. Os valores sublinhados indicam diferenças estatísticas significativas entre as medianas de erro frequencial associadas a diferentes valores do descritor, conforme identificado por meio de teste estatístico apropriado. Quanto menor a mediana do erro frequencial associada a um valor de descritor, mais facilmente a mudança de conceito é tratada no cenário gerado por esse valor. A partir Tabela 5, podem ser feitas algumas observações:

Tabela 5 – Comparação dos descritores em cada contexto por meio de testes estatísticos<sup>1</sup>.

Contexto	Velocidade			Severidade			Recorrência			Frequência				
	R	G	$\rho$ -Valor	A	B	$\rho$ -Valor	R	NR	$\rho$ -Valor	PA	PM	PB	NP	$\rho$ -Valor
DDM	16,9	<b>16,6</b>	0,98	18,3	<u>14,5</u>	< 0,05	17,5	<b>16,0</b>	0,34	<b>14,2</b>	17,1	17,8	17,2	0,89
EDDM	<b>16,2</b>	17,0	0,78	19,8	<b>15,1</b>	0,19	20,4	<u>12,7</u>	< 0,05	22,1	20,2	<u>9,4</u>	14,3	< 0,05
DDM-O	19,3	<b>15,4</b>	0,13	19,2	<u>14,9</u>	< 0,05	<u>14,3</u>	19,0	< 0,05	18,7	18,4	<u>9,2</u>	16,5	< 0,05
EDDM-O	17,6	<b>15,8</b>	0,11	19,5	<u>13,6</u>	< 0,05	18,5	<u>14,5</u>	< 0,05	17,0	<b>15,4</b>	16,7	15,7	0,61
DDAL	15,9	<b>14,5</b>	0,58	19,5	<u>12,2</u>	< 0,05	16,5	<b>15,1</b>	0,28	23,9	16,8	<b>10,2</b>	13,3	< 0,05
STUDD	16,7	<b>14,9</b>	0,68	20,2	<u>12,1</u>	< 0,05	17,5	<b>13,8</b>	0,20	23,6	18,0	<u>10,0</u>	11,8	< 0,05
MD3-EGM	16,8	<b>16,9</b>	0,89	18,5	<u>14,8</u>	< 0,05	17,5	<b>15,9</b>	0,32	<b>14,4</b>	17,5	17,9	17,2	0,78
DSDD	17,6	<b>15,8</b>	0,11	19,5	<u>13,6</u>	< 0,05	18,5	<u>14,5</u>	< 0,05	17,0	<b>15,4</b>	16,7	15,7	0,61
SD	15,9	<b>14,7</b>	0,49	19,8	<u>12,0</u>	< 0,05	16,6	<b>14,8</b>	0,23	23,2	16,9	<u>10,0</u>	13,0	< 0,05

<sup>1</sup> Valores em negrito indicam a menor mediana do erro frequencial entre os cenários da primeira série de experimentos. Valores sublinhados indicam diferença estatística significativa segundo os testes estatísticos aplicados. Os testes foram conduzidos com nível de significância de 0,05.

- Para o descritor de velocidade, o valor gradual apresentou o melhor desempenho em todos os contextos, exceto no EDDM. Essa exceção pode estar associada ao fato de que o EDDM foi proposto especificamente para lidar com mudanças repentinas. No entanto, não houve nenhum contexto em que se observou diferença estatística significativa entre os valores repentino e gradual, o que sugere que a velocidade exerce pouco impacto sobre a mudança de conceito.
- O descritor de severidade apresentou comportamento oposto, uma vez que mudanças de baixa severidade resultaram em melhor desempenho em todos os contextos. O teste estatístico indicou diferenças significativas em todos os contextos, com exceção do EDDM. Esse resultado pode ser interpretado como uma evidência robusta de que a severidade exerce forte impacto sobre a mudança de conceito.
- Em relação à recorrência, os cenários não recorrentes apresentaram melhor desempenho, com exceção do DDM-O. O teste estatístico indicou diferenças significativas em quatro dos nove contextos avaliados, sugerindo que a recorrência possui impacto limitado sobre a mudança de conceito, sendo relevante apenas em situações específicas.

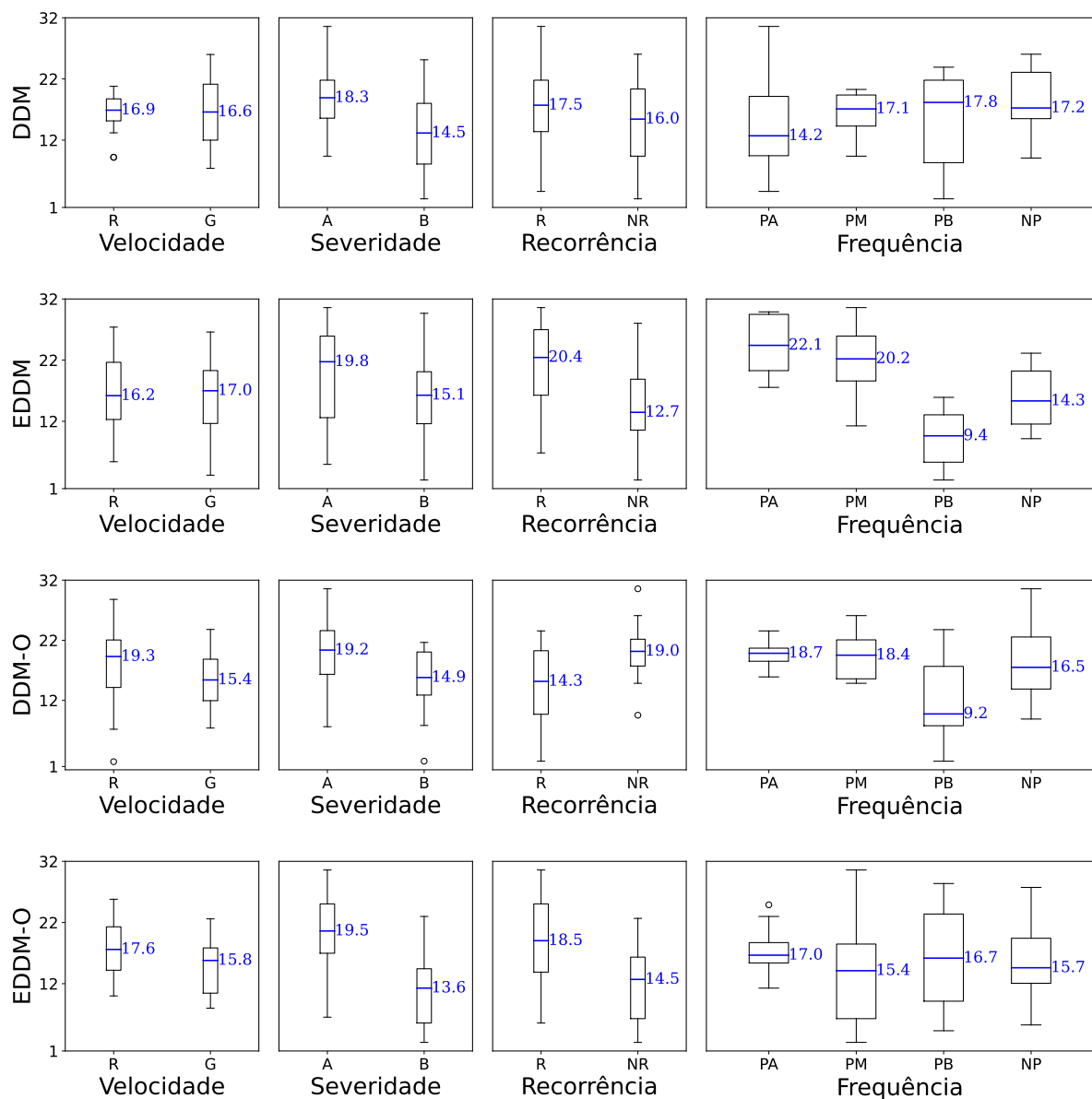


Figura 27 – Variação da mediana do erro prequencial nos cenários investigados variando o valor de cada descritor considerando os detectores supervisionados.

- Uma situação semelhante foi observada em relação ao descritor frequência. Nesse caso, os melhores resultados foram geralmente associados a cenários com baixa frequência. O teste estatístico indicou diferenças significativas em cinco dos nove contextos avaliados. Todas as diferenças significativas foram obtidas em cenários de baixa frequência, o que sugere que mudanças com baixa frequência tendem a impactar positivamente o desempenho da detecção. Contudo, influências ocasionais de outros descritores podem atenuar o impacto do descritor frequência.

Por fim, com base nos resultados apresentados na Tabela 5 e nas Figuras 27 e

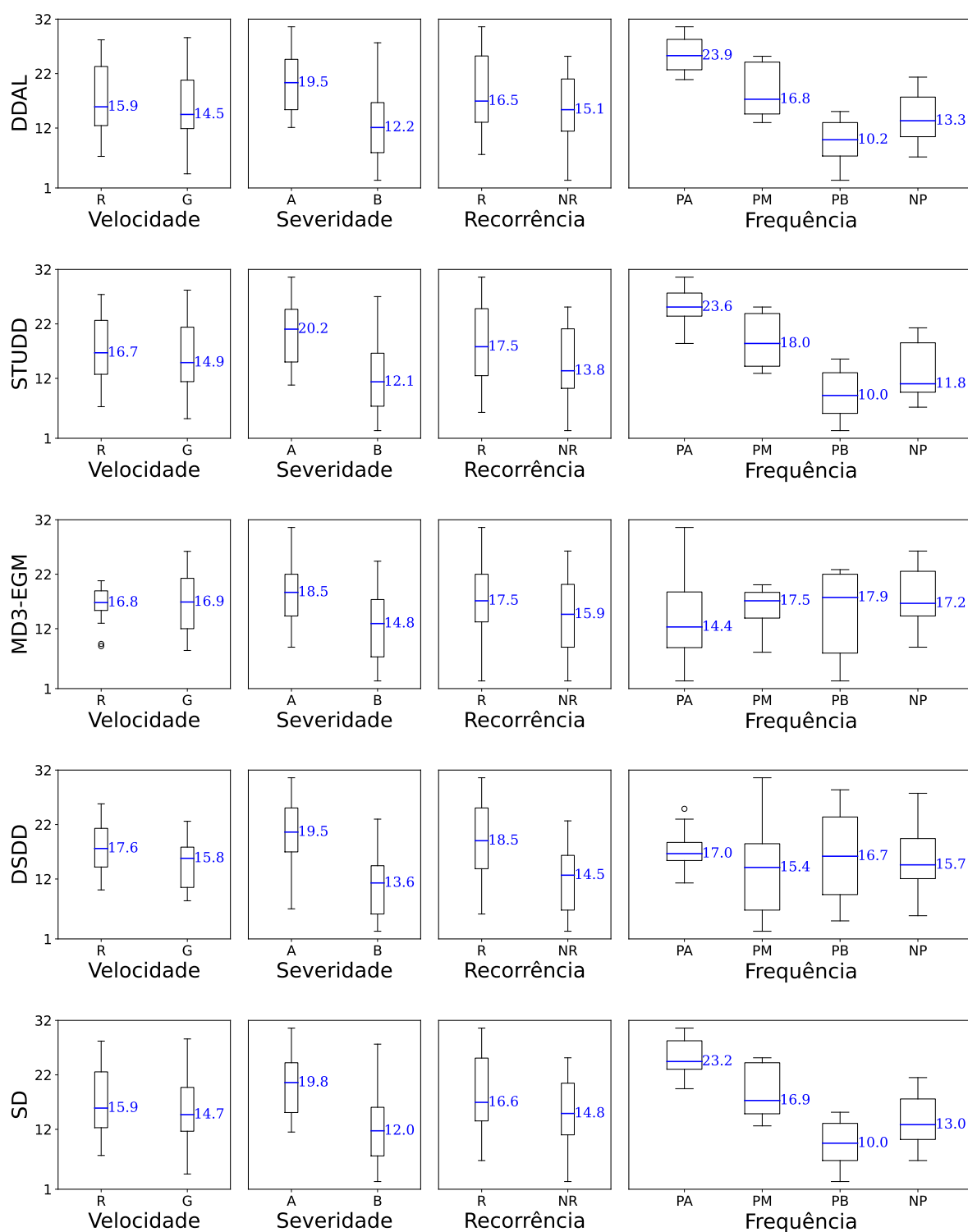


Figura 28 – Variação da mediana do erro prequencial em todos os cenários investigados variando o valor de cada descritor considerando os detectores semi-supervisionados, não supervisionado e o contexto sem detecção.

28, a Tabela 6 resume as combinações de valores dos descritores que resultaram nos melhores desempenhos, bem como aquelas que conduziram aos piores desempenhos no

tratamento da mudança de conceito. Por exemplo, entre os nove contextos analisados, os melhores resultados foram obtidos em oito deles quando o valor do descritor velocidade foi gradual.

Tabela 6 – Valores dos descritores para os melhores e os piores resultados.

Resultados <sup>1</sup>	Velocidade		Severidade		Recorrência		Frequência			
	R	G	A	B	R	NR	PA	PM	PB	NP
Melhor	1/9	<b>8/9</b>	0/9	<b>9/9</b>	1/9	<b>8/9</b>	2/9	2/9	<b>5/9</b>	0/9
Pior	<b>8/9</b>	1/9	<b>9/9</b>	0/9	<b>8/9</b>	1/9	<b>7/9</b>	0/9	2/9	0/9

<sup>1</sup> Valores em negrito indicam a maior frequência de ocorrência por valor de cada descritor. A fração representa a frequência com que o descritor assumiu o respectivo valor nos cenários de melhor e pior desempenho, considerando os diferentes contextos analisados.

## 4.6 Considerações Finais

Neste capítulo, foi analisado o impacto de quatro descritores da mudança de conceito sobre o desempenho de classificadores base e detectores de mudança, com o objetivo de identificar os descritores mais relevantes para o processo de reação à mudança. Os experimentos foram conduzidos em cinco conjuntos de dados distintos, considerando 32 variações nos valores dos descritores, oito detectores diferentes e um contexto sem detecção, totalizando 1.440 combinações distintas avaliadas. Os resultados indicaram que mudanças de conceito caracterizadas por baixa velocidade, baixa severidade, não recorrência e baixa periodicidade geraram menores impactos no desempenho dos classificadores. Por outro lado, mudanças com alta velocidade (repentina), alta severidade, recorrência e alta periodicidade resultaram em maior impacto sobre o desempenho dos classificadores.

A primeira conclusão deste estudo é que a reação à mudança de conceito não pode ser uniforme para todos os tipos de mudança. Ao contrário, ela deve considerar as características específicas de cada mudança e, em alguns casos, pode-se argumentar que a melhor resposta é, simplesmente, não reagir. Ao detalhar os resultados apresentados na Tabela 6, observou-se que os descritores de velocidade, severidade e recorrência apresentaram valores bem definidos nos cenários com melhor e pior desempenho,

enquanto a frequência demonstrou um comportamento muito menos definido. Isso sugere que os três primeiros descritores exercem maior influência sobre a mudança de conceito. No entanto, ao combinar esses resultados com os da Tabela 5, observa-se que apenas os resultados relacionados à severidade apresentam, predominantemente, diferenças estatisticamente significativas, seguidos pelos descritores de frequência e recorrência.

A segunda conclusão é que os descritores de severidade, recorrência e frequência exercem, respectivamente e em ordem decrescente, os impactos mais relevantes sobre a mudança de conceito. Destaca-se que o impacto isolado da severidade se mostra mais expressivo do que o dos demais descritores. Por sua vez, a velocidade não apresentou indícios de exercer influência significativa sobre o fenômeno. Diante disso, qualquer estratégia de reação deve considerar, prioritariamente, os descritores com maior impacto. Por fim, a terceira conclusão é que há uma necessidade de incorporar mecanismos de descrição da mudança de conceito ao processo de tratamento desse fenômeno. Esse mecanismo pode ser executado de forma concorrente à detecção, sendo que as propriedades extraídas a partir da descrição devem ser utilizadas para orientar a definição da estratégia de reação mais eficaz a ser adotada.

Com base nessas conclusões, o próximo capítulo é focado em identificar a estratégia de reação à mudança de conceito mais adequada, considerando as diferentes características das mudanças.



---

## Mudança de Conceito e Estratégias de Reação: Uma Análise Abrangente com o Uso de Descritores

---

O objetivo deste capítulo é identificar a estratégia de reação à mudança de conceito mais adequada, considerando as diferentes características das mudanças. Os experimentos realizados foram estruturados em dois estágios, seguidos de um teste de ablação. A Figura 29 resume esses dois estágios e o teste de ablação. Os experimentos foram conduzidos utilizando os mesmos cinco conjuntos de dados sintéticos investigados no capítulo anterior, cada um com 32 variações (também denominadas cenários), e cinco estratégias distintas de reação (Figura 30). O teste de ablação foi realizado para avaliar a influência do detector nos resultados obtidos. No total, 1.000 foram testes realizados.

O EDDM ([BAENA-GARCÍA et al., 2006](#)) foi selecionado como detector para a primeira e a segunda etapas dos experimentos em razão de seu desempenho superior demonstrado no experimento apresentado no Capítulo 4. Para o teste de ablação, o DDM ([GAMA et al., 2004](#)) foi utilizado por ser amplamente empregado na literatura e apresentar semelhanças operacionais com o EDDM. Em todas as etapas, a árvore de decisão de Hoeffding ([BIFET; GAVALDÀ, 2009](#)) foi adotada como classificador, uma vez que sua capacidade de aprendizado incremental é exigida por algumas das estratégias testadas. Os detectores e o classificador foram aplicados com seus valores de parâmetros padrão, conforme definidos por seus autores originais.

O objetivo da primeira etapa foi comparar a Estratégia Estática (EE), que não utiliza um detector, com a Estratégia Padrão (EP), que substitui o classificador por um novo treinado com as instâncias mais recentes. Essa comparação visou identificar cenários em que as diferenças decorrentes da mudança de conceito não justificariam a adoção de estratégias de reação que alterassem o conhecimento do classificador. O

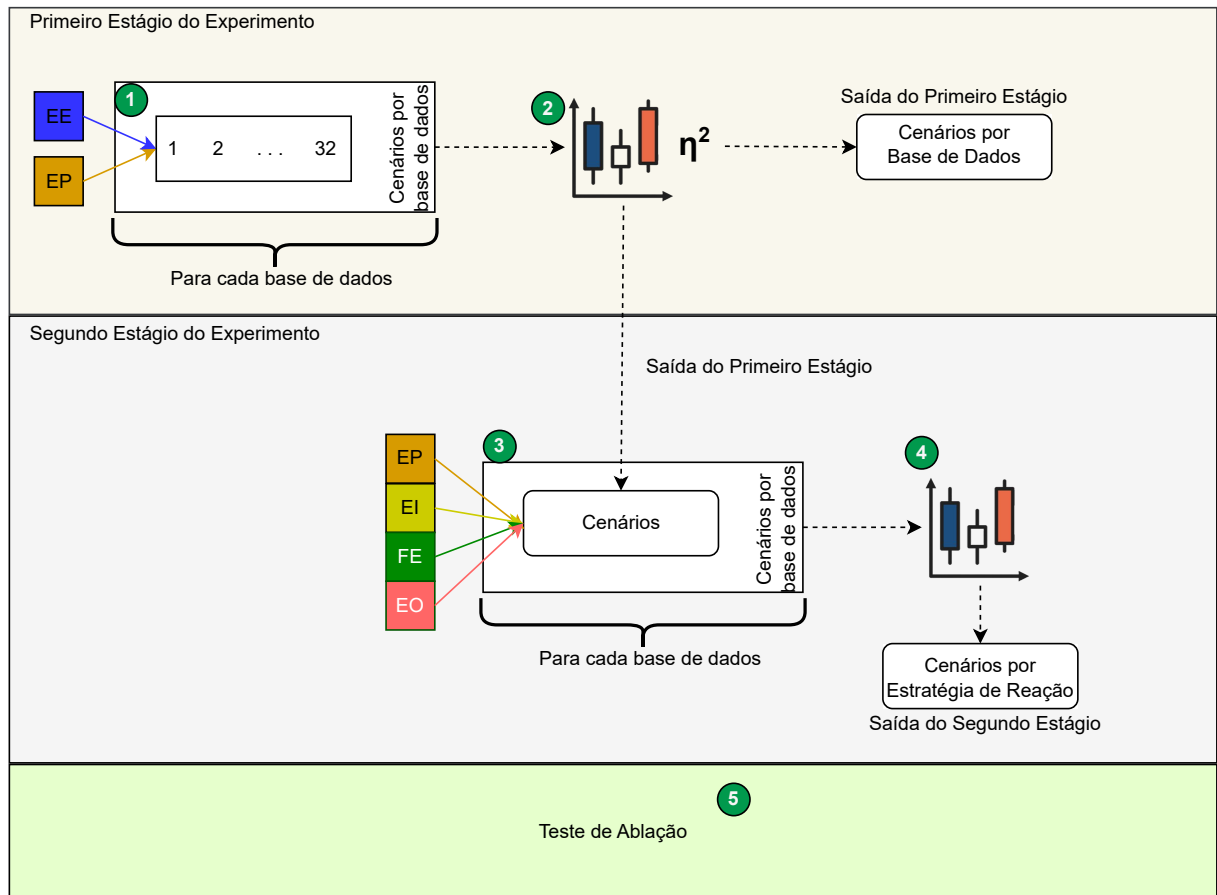


Figura 29 – Visão geral do experimento.

principal critério para essa avaliação foi o custo de reaprender regiões do espaço de entrada além da zona de influência da mudança. Esse aspecto é especialmente crítico quando não há garantia de obtenção de instâncias suficientes para possibilitar o reaprendizado, o que pode, potencialmente, resultar em uma degradação do desempenho do novo classificador em relação ao anterior.

O erro frequencial (DAVID; VOVK, 1999) foi utilizado como métrica de desempenho. Os resultados foram comparados por meio do teste estatístico de Mann-Whitney (MANN; WHITNEY, 1947), com nível de significância de 0,05. É importante destacar que a estratégia EP incorre em custo de retreinamento, enquanto a EE não. Portanto, a análise isolada do resultado do teste estatístico não é suficiente para determinar a estratégia de reação mais apropriada. Por esse motivo, a magnitude da diferença de desempenho foi mensurada utilizando o tamanho de efeito  $\eta^2$  (RICHARDSON, 2011; WILCOX, 2022). O tamanho de efeito é uma estatística que reflete a magnitude quan-

Tabela 7 – Interpretação do tamanho do efeito.

Intervalo de $\eta^2$	Efeito
$< 0,01$	Insignificante
$0,01 \leq \eta^2 < 0,06$	Pequeno
$0,06 \leq \eta^2 < 0,14$	Médio
$\geq 0,14$	Forte

titativa de um fenômeno. Em síntese, ele complementa o  $\rho$ -valor, que indica se um efeito existe, abordando o significado prático e mostrando quão grande o efeito precisa ser para ter relevância no mundo real (HEDGES; VEVEA, 2001; PARDO; ANTIVILO-BRUNA; MIGUEZ, 2020).

A interpretação de  $\eta^2$  proposta por Cohen (1988) foi adotada neste trabalho, conforme apresentado na Tabela 7. Ao final do primeiro estágio, a estratégia EP foi considerada a mais apropriada apenas nos cenários em que o teste estatístico indicou diferença significativa com tamanho de efeito forte ( $\eta^2 \geq 0,14$ ). Esses cenários foram, então, testados no segundo estágio do experimento. Nos demais cenários, a estratégia ES foi considerada a mais adequada. Nesses casos, nenhuma reação foi realizada.

O objetivo do segundo estágio dos experimentos foi definir a melhor estratégia de reação com base nos descritores de mudança de conceito. Quatro estratégias de reação foram investigadas: Estratégia Padrão (EP), Estratégia Incremental (EI), Estratégia com Esquecimento (EQ) e Estratégia Ótima (EO), detalhadas na Seção 5.1.

Seus desempenhos, em termos de erro prequencial, foram comparadas utilizando o teste estatístico de Kruskal-Wallis (KRUSKAL; WALLIS, 1952) e o teste *post-hoc* de Mann-Whitney (MANN; WHITNEY, 1947), ambos com nível de significância em 0,05. Por fim, o teste de ablação teve como objetivo analisar a influência do EDDM nos resultados. Para isso, o EDDM foi substituído pelo DDM, e os cenários do segundo estágio experimental foram reavaliados, considerando especificamente as estratégias que apresentaram melhor desempenho nesse estágio. Os resultados obtidos no primeiro e no segundo estágios dos experimentos e no teste de ablação são discutidos, respectivamente, nas Seções 5.3, 5.4 e 5.5.

## 5.1 Estratégias de Reação

Conforme mencionado anteriormente, cinco estratégias de reação são investigadas neste trabalho. A Figura 30 ilustra essas estratégias. A primeira, denominada EE, opera sob a suposição de um ambiente estável, ou seja, não utiliza um detector. O objetivo de testar essa estratégia é identificar cenários nos quais a diferença entre os conceitos não seja significativa o suficiente para justificar os custos de atualização do classificador. A segunda estratégia é a EP, considerada padrão na literatura. Ela consiste em substituir o classificador por um novo, treinado com as instâncias mais recentes, denominadas aqui de janela de alerta. Esse método é adotado por detectores clássicos como o DDM (GAMA et al., 2004), EDDM (BAENA-GARCÍA et al., 2006) e HDDM (FRÍAS-BLANCO et al., 2015).

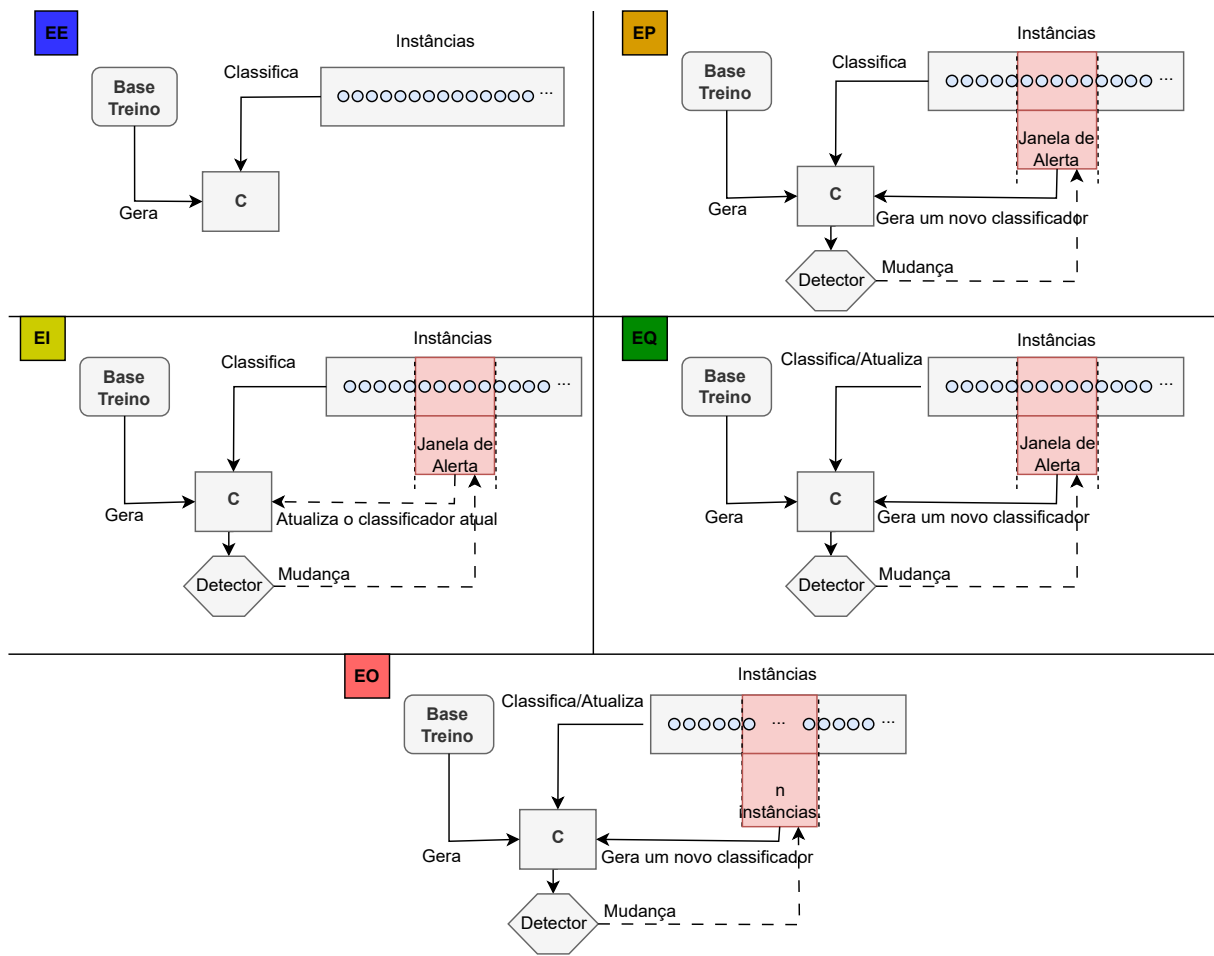


Figura 30 – Visão geral das estratégias de reação investigadas neste trabalho. C indica o classificador.

A terceira é a Estratégia Incremental (EI), que consiste na atualização incremental do classificador com instâncias da janela de alerta, sempre que uma mudança for detectada. Portanto, os conceitos previamente aprendidos não são esquecidos. O objetivo de testar a EI é analisar o impacto da ausência de esquecimento completo na estratégia de reação. Por outro lado, na Estratégia de Esquecimento (EQ), o classificador é atualizado incrementalmente com todas as instâncias, ou seja, à medida que as instâncias chegam de forma contínua e sequencial, elas são classificadas e utilizadas para atualizar o classificador atual. Quando uma mudança é detectada, o classificador atual é substituído por um novo, treinado com as instâncias da janela de alerta, assim como é feito na EP. O objetivo da EQ é analisar o impacto do esquecimento parcial do classificador na estratégia de reação. A estratégia EQ é empregada pelo método de detecção de mudança baseado em dissimilaridade (PINAGE; SANTOS, 2015).

Por fim, a Estratégia Oráculo (EO) considera a existência de um detector ideal, capaz de identificar com precisão todas as mudanças. Ao detectar uma mudança, um novo classificador é treinado com as primeiras  $n$  instâncias do novo conceito, sendo  $n$  um hiperparâmetro da estratégia de reação. O objetivo de testar essa estratégia é minimizar as influências do conceito anterior, bem como reduzir falsas detecções, perdas e atrasos na detecção.

## 5.2 Bases de Dados e Cenários

As bases de dados e os cenários testados nesta pesquisa estão descritos nas Seções 4.1 e 4.2; respectivamente.

## 5.3 Primeiro Estágio do Experimento

Os resultados deste estágio são apresentados nas Figuras 31, 32 e 33. A Figura 31 indica que, como esperado, a estratégia EP apresentou desempenho superior à estratégia EE na maioria dos casos. As exceções são o cenário 6 da base *Circle*; os cenários 8, 20, 27 e 28 da *Line*; o cenário 26 da *Hiperplane*; e os cenários 11, 15, 18, 24 e 27 da base

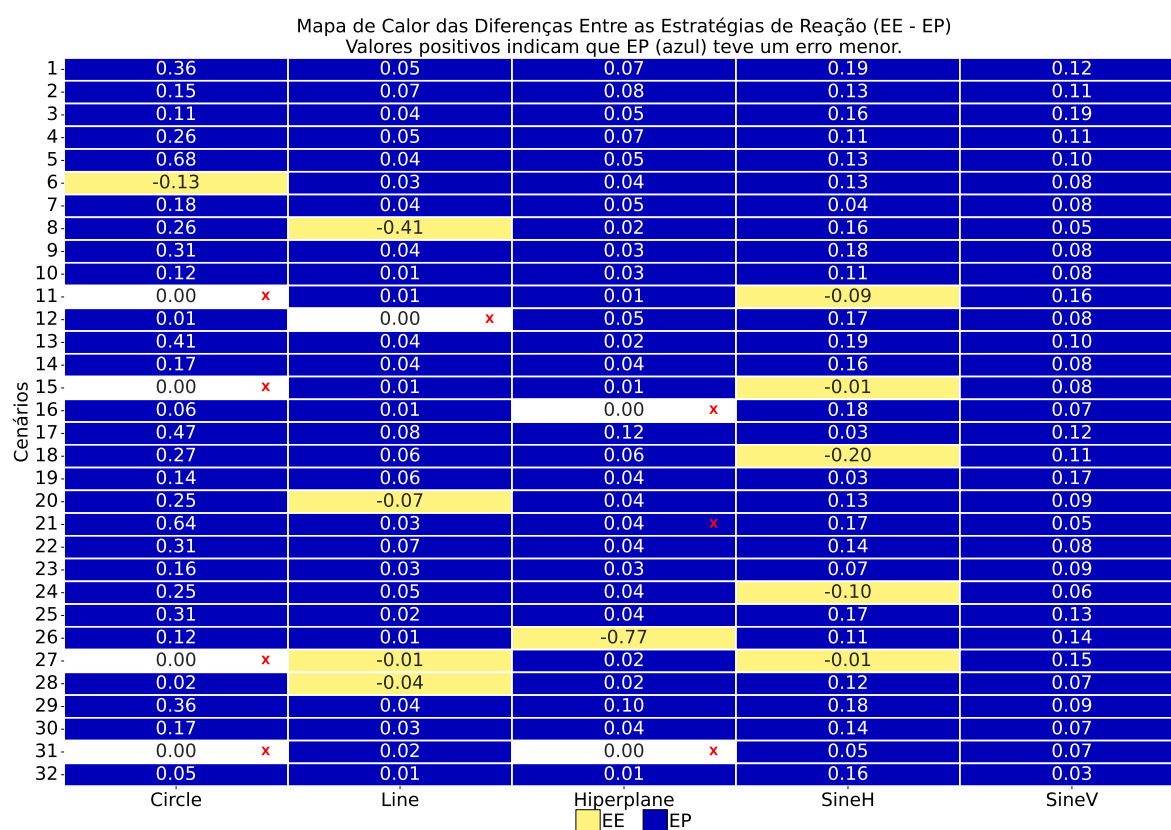


Figura 31 – Valores positivos indicam que a estratégia EP alcançou os menores erros. X indica que não há diferença significativa entre EP e EE, no nível de significância de 0,05.

*SineH*. Diferentemente da EE, a estratégia EP envolve um custo de retreinamento, como previamente mencionado. Portanto, uma análise mais precisa e justa dos resultados deve considerar o tamanho do efeito da diferença estatística.

O tamanho do efeito calculado para cada cenário em cada base de dados está resumido na Figura 32. A linha vertical tracejada vermelha indica o ponto a partir do qual a diferença passa a ter um efeito forte, e os círculos em azul ou vermelho indicam o valor do efeito calculado. Aqueles localizados à esquerda da linha representam situações em que não houve diferença significativa entre as estratégias ou em que a diferença apresenta efeito desprezível ( $\eta^2 < 0,01$ ), efeito pequeno ( $0,01 \leq \eta^2 < 0,06$ ) ou efeito médio ( $0,06 \leq \eta^2 < 0,14$ ). Nesses casos, assume-se EE como a estratégia mais apropriada. Já os círculos à direita da linha indicam situações em que a diferença significativa entre as estratégias apresenta um grande efeito, sendo a EP considerada a estratégia de reação mais adequada.

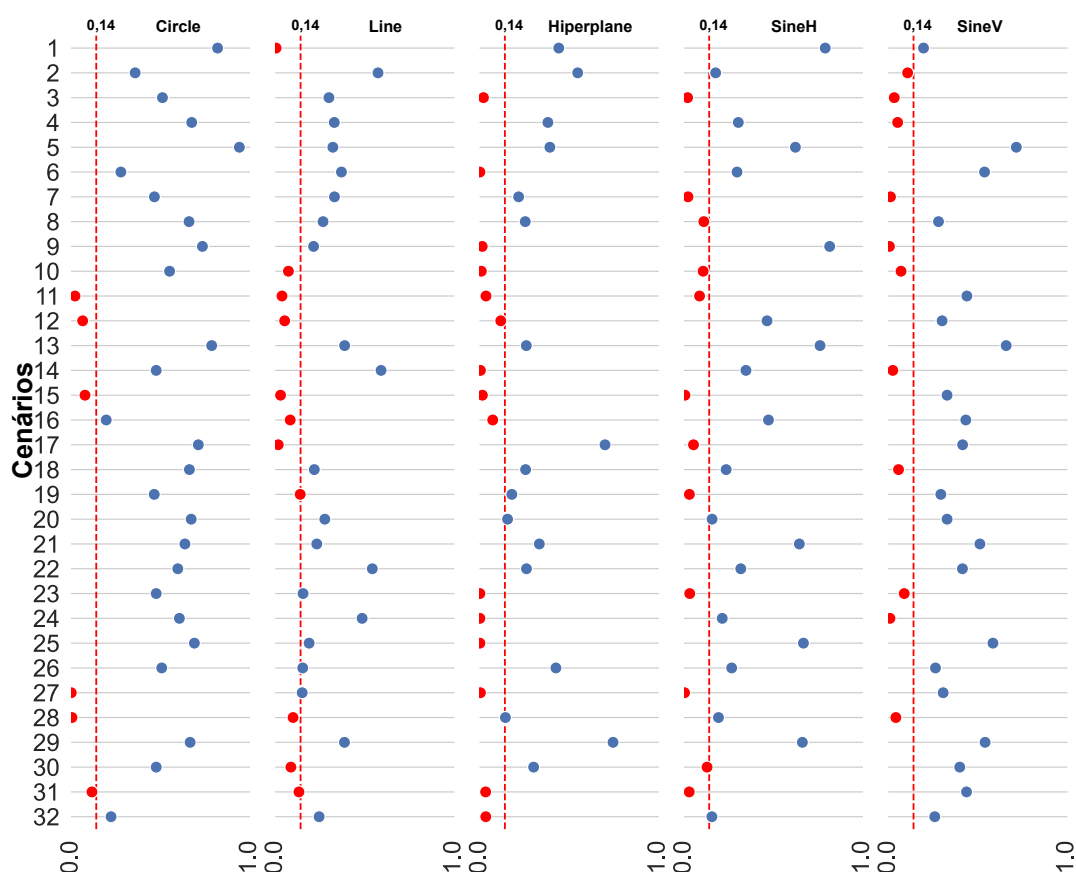


Figura 32 – Tamanho do efeito dos testes estatísticos. A linha vermelha pontilhada indica o ponto a partir do qual o efeito é considerado forte ( $\eta^2 \geq 0,14$ ).

O resultado do primeiro estágio dos experimentos está resumido na Figura 33. Em um determinado cenário, uma estratégia é considerada superior se obtiver os melhores resultados na maioria simples das bases de dados avaliadas. Como exemplo, a estratégia EP, representada pelo quadrado azul, apresentou melhor desempenho em quatro das cinco bases de dados utilizadas no cenário 1 (*Circle*, *Hiperplane*, *SineH* e *SineV*), enquanto a estratégia EE, representada pelo quadrado amarelo, obteve melhor desempenho apenas na base de dados *Line*.

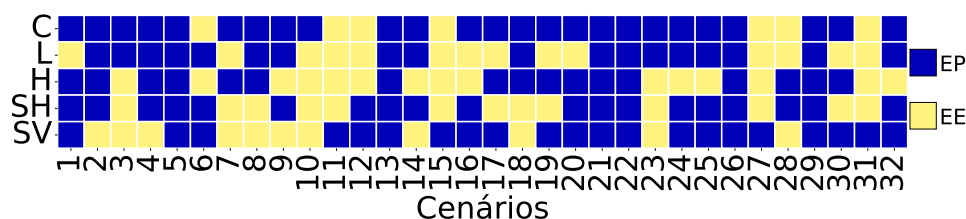


Figura 33 – O mapa de calor ilustra o desempenho de cada estratégia de reação nos cenários em cada base de dados.

Considerando cada cenário por base de dados, a estratégia EE superou a estratégia EP nos cenários 3, 7, 10, 11, 12, 15, 23, 27, 28 e 31. Ao analisar esses cenários em termos de seus descritores, observou-se que todos apresentam frequência periódica baixa. Além disso, em sete dos oito cenários com frequência periódica baixa, a estratégia EE obteve melhor desempenho. Esses resultados indicam que, para mudanças de conceito caracterizadas por frequência periódica baixa, não retrainar o modelo tende a ser a estratégia de reação mais apropriada. Essa constatação representa uma importante conclusão deste trabalho. Nos demais cenários, há evidências estatísticas de que a estratégia de reação mais adequada é o retreinamento do modelo. No entanto, nesses casos, é necessário definir a forma ideal de retreinamento. Para esse fim, esses cenários são utilizados como entrada para o segundo estágio dos experimentos.

## 5.4 Segundo Estágio do Experimento

Os resultados obtidos no segundo estágio dos experimentos estão resumidos nas Tabelas 8 e 9, e na Figura 34. Esse estágio tem como objetivo identificar a estratégia de reação à mudança de conceito mais apropriada para os cenários em que os descritores da mudança indicam a necessidade de retreinamento do classificador. Os desempenhos foram comparados por meio do teste estatístico de Kruskal-Wallis (KRUSKAL; WALLIS, 1952) e do teste de Mann-Whitney como procedimento *post-hoc* (MANN; WHITNEY, 1947), ambos com nível de significância em 0,05. As estratégias analisadas neste estágio são EP, EI, EQ e EO. O erro frequencial foi utilizado como métrica de desempenho.

Os valores das medianas do erro frequencial obtidos nos cenários e nas bases de dados investigados para as quatro estratégias de reação analisadas são apresentados nas Tabelas 8 e 9. Os valores em negrito indicam a estratégia com a menor mediana do erro frequencial, enquanto os valores sublinhados indicam diferença significativa entre os desempenhos das estratégias, de acordo com o teste estatístico aplicado. O teste de Kruskal-Wallis indicou diferenças significativas em todas as comparações ( $p$ -valor = 0,00). Por sua vez, o teste *post-hoc* de Mann-Whitney indicou, em sua maioria, diferenças significativas nas comparações pareadas entre as estratégias, com exceção de



Tabela 8 – Comparando as estratégias de reação em cada cenário em todas as bases de dados utilizando teste estatístico.

Cenário	<i>Circle</i> <sup>1</sup>					<i>Line</i> <sup>1</sup>					<i>Hiperplane</i> <sup>1</sup>				
	EP	EI	EQ	EO	$\rho$ -v <sup>2</sup>	EP	EI	EQ	EO	$\rho$ -valor	EP	EI	EQ	EO	$\rho$ -v <sup>2</sup>
1	0,19	0,17	<b><u>0,04</u></b>	0,11	0,00	0,06	0,20	<b><u>0,04</u></b>	0,08	0,00	0,05	0,13	<b><u>0,04</u></b>	0,08	0,00
2	0,06	0,08	<b><u>0,03</u></b>	0,04	0,00	<b><u>0,03</u></b>	0,11	0,05	0,04	0,00	0,04 <sup>a</sup>	0,09	<b><u>0,04</u></b> <sup>b</sup>	0,05	0,00
4	0,02	0,08	<b><u>0,02</u></b>	0,02	0,00	0,04	0,04	<b><u>0,03</u></b>	0,04	0,00	<b><u>0,03</u></b>	0,05	0,04	0,06	0,00
5	0,05	0,18	<b><u>0,03</u></b>	0,12	0,00	0,03	0,04	<b><u>0,03</u></b>	0,08	0,00	0,04	0,05	<b><u>0,04</u></b>	0,09	0,00
6	0,50	0,14	<b><u>0,03</u></b>	0,05	0,00	<b><u>0,04</u></b>	0,05	0,04	0,05	0,00	0,05	0,06	<b><u>0,04</u></b>	0,07	0,00
8	<b><u>0,02</u></b>	0,16	0,03	0,05	0,00	0,50	0,05	<b><u>0,03</u></b>	0,04	0,00	0,05	0,15	<b><u>0,04</u></b>	0,06	0,00
9	0,04	0,11	<b><u>0,03</u></b>	0,09	0,00	0,04	0,09	<b><u>0,03</u></b>	0,07	0,00	0,06	0,08	<b><u>0,05</u></b>	0,09	0,00
13	0,04 <sup>a</sup>	0,17	<b><u>0,03</u></b> <sup>b</sup>	0,10	0,00	0,04	0,05	<b><u>0,03</u></b>	0,08	0,00	0,05	0,05	<b><u>0,04</u></b>	0,09	0,00
14	0,02	0,12	<b><u>0,01</u></b>	0,03	0,00	0,04	0,05	<b><u>0,02</u></b>	0,06	0,00	<b><u>0,03</u></b>	0,04	0,04	0,05	0,00
16	0,03	0,18	<b><u>0,02</u></b>	0,03	0,00	<b><u>0,02</u></b>	0,07	0,02	0,03	0,00	0,05 <sup>a</sup>	0,03 <sup>b</sup>	<b><u>0,03</u></b>	0,04	0,00
17	<b><u>0,06</u></b>	0,17	0,06	0,12	0,00	0,04	0,11	<b><u>0,04</u></b>	0,08	0,00	0,05	0,07	<b><u>0,04</u></b>	0,09	0,00
18	<b><u>0,03</u></b>	0,10	0,05	0,05	0,00	0,04	0,10	<b><u>0,03</u></b>	0,06	0,00	0,06	0,09	<b><u>0,04</u></b>	0,08	0,00
19	<b><u>0,01</u></b>	0,03	0,03	0,04	0,00	<b><u>0,01</u></b>	0,04	0,02	0,12	0,00	<b><u>0,04</u></b>	0,07	0,04	0,09	0,00
20	0,02 <sup>a</sup>	0,09	<b><u>0,02</u></b>	0,05 <sup>b</sup>	0,00	0,13	0,05	<b><u>0,04</u></b>	0,08	0,00	0,05	0,08	<b><u>0,05</u></b>	0,06	0,00
21	0,08	0,19	<b><u>0,04</u></b>	0,11	0,00	0,04	0,05	<b><u>0,03</u></b>	0,07	0,00	<b><u>0,04</u></b>	0,06	0,04	0,09	0,00
22	0,03	0,17	<b><u>0,03</u></b>	0,06	0,00	0,04	0,05	<b><u>0,03</u></b>	0,05	0,00	0,05	0,06	<b><u>0,04</u></b>	0,05	0,00
24	<b><u>0,02</u></b>	0,30	0,04	0,07	0,00	0,04	0,06	<b><u>0,02</u></b>	0,03	0,00	<b><u>0,03</u></b>	0,04	0,03	0,07	0,00
25	0,03	0,10	<b><u>0,02</u></b>	0,09	0,00	0,04	0,11	<b><u>0,03</u></b>	0,08	0,00	<b><u>0,04</u></b>	0,15	0,05	0,09	0,00
26	0,03	0,06	<b><u>0,01</u></b>	0,03	0,00	0,04	0,04	<b><u>0,02</u></b>	0,05	0,00	0,86	0,06	<b><u>0,05</u></b>	0,08	0,00
29	0,04	0,15	<b><u>0,02</u></b>	0,10	0,00	<b><u>0,03</u></b>	0,04	0,03	0,07	0,00	<b><u>0,03</u></b>	0,04	0,03	0,09	0,00
30	0,03	0,11	<b><u>0,01</u></b>	0,03	0,00	0,02	0,05	<b><u>0,02</u></b>	0,04	0,00	0,04	0,08	<b><u>0,04</u></b>	0,06	0,00
32	0,03	0,07	<b><u>0,01</u></b>	0,03	0,00	0,03	0,04	<b><u>0,02</u></b>	0,05	0,00	0,03	0,08	<b><u>0,03</u></b>	0,04	0,00

<sup>1</sup> Valores em negrito indicam o menor valor da mediana do erro prequencial dos cenários na segunda etapa dos experimentos. Valores sublinhados indicam diferença estatística significativa segundo os testes estatísticos aplicados para todos os descritores. Testes com nível de significância de 0,05.

<sup>2</sup>  $\rho$ -valor.

oito comparações; marcadas com as letras minúsculas "a" e "b" nas Tabelas 8 e 9. Para evitar excesso de informação, apenas as comparações sem diferença significativa no teste *posthoc* foram marcadas com letras minúsculas. Cabe esclarecer que valores iguais de medianas com interpretações distintas nas Tabelas 8 e 9 devem-se à limitação da representação com duas casas decimais.

Ao detalhar os resultados, observa-se que, na base de dados *Circle*, a estratégia EQ apresentou o melhor desempenho em 17 dos 22 cenários analisados, sendo superada pela estratégia EP apenas nos cenários 8, 17, 18, 19 e 24. As demais estratégias não apresentaram o melhor desempenho em nenhum dos cenários. A EQ também foi o melhor método na base de dados *Line* em 17 cenários, enquanto a EP apresentou melhor

Tabela 9 – Comparando as estratégias de reação em cada cenário em todas as bases de dados utilizando teste estatístico.

Cenário	<i>SineH</i> <sup>1</sup>					<i>SineV</i> <sup>1</sup>				
	EP	EI	EQ	EO	$\rho$ -valor	EP	EI	EQ	EO	$\rho$ -valor
1	0,18	0,24 <sup>a</sup>	<b><u>0,17</u></b>	0,24 <sup>b</sup>	0,00	<b><u>0,03</u></b>	0,13	0,04	0,07	0,00
2	0,18 <sup>a</sup>	0,31	<b><u>0,18</u></b> <sup>b</sup>	0,19	0,00	0,03	0,12	<b><u>0,02</u></b>	0,05	0,00
4	0,22	0,21	<b><u>0,19</u></b>	0,20	0,00	0,02 <sup>a</sup>	0,12	<b><u>0,02</u></b> <sup>b</sup>	0,03	0,00
5	0,25	0,28	<b><u>0,18</u></b>	0,22	0,00	0,02	0,03	<b><u>0,02</u></b>	0,08	0,00
6	0,22	<b><u>0,18</u></b>	0,19	0,19	0,00	0,03	0,06	<b><u>0,02</u></b>	0,04	0,00
8	0,19	0,19	<b><u>0,18</u></b>	0,21	0,00	0,03	0,12	<b><u>0,02</u></b>	0,02	0,00
9	<b><u>0,19</u></b>	0,19	0,20	0,24	0,00	0,04	0,06	<b><u>0,03</u></b>	0,07	0,00
13	0,20	0,39	<b><u>0,19</u></b>	0,23	0,00	0,02	0,05	<b><u>0,02</u></b>	0,07	0,00
14	0,19	0,39	<b><u>0,18</u></b>	0,19	0,00	0,03	0,06	<b><u>0,03</u></b>	0,04	0,00
16	<b><u>0,18</u></b>	0,34	0,18	0,18	0,00	<b><u>0,02</u></b>	0,07	0,03	0,05	0,00
17	0,29	0,20	<b><u>0,19</u></b>	0,24	0,00	0,04	0,10	<b><u>0,03</u></b>	0,06	0,00
18	0,50	0,20	<b><u>0,19</u></b>	0,20	0,00	0,03	0,09	<b><u>0,03</u></b>	0,09	0,00
19	<b><u>0,14</u></b>	0,23	0,17	0,16	0,00	<b><u>0,04</u></b>	0,43	0,04	0,24	0,00
20	0,19	0,33	<b><u>0,19</u></b>	0,23	0,00	0,03	0,06	<b><u>0,02</u></b>	0,04	0,00
21	0,20	<b><u>0,18</u></b>	0,19	0,22	0,00	<b><u>0,02</u></b>	0,06	0,02	0,07	0,00
22	0,20	0,39	<b><u>0,19</u></b>	0,21	0,00	0,03	0,05	<b><u>0,02</u></b>	0,06	0,00
24	0,44	<b><u>0,18</u></b>	0,19	0,21	0,00	<b><u>0,02</u></b>	0,02	0,02	0,06	0,00
25	0,19	<b><u>0,19</u></b>	0,19	0,23	0,00	0,03	0,11	<b><u>0,03</u></b>	0,08	0,00
26	0,19	0,33	<b><u>0,18</u></b>	0,21	0,00	<b><u>0,05</u></b>	0,07	0,05	0,09	0,00
29	0,19	<b><u>0,19</u></b> <sup>a</sup>	0,19 <sup>b</sup>	0,23	0,00	<b><u>0,02</u></b>	0,03	0,02	0,07	0,00
30	0,20	0,20	<b><u>0,19</u></b>	0,21	0,00	0,03	0,05	<b><u>0,03</u></b>	0,06	0,00
32	<b><u>0,18</u></b>	0,29	0,19	0,21	0,00	0,03	0,15	<b><u>0,02</u></b>	0,05	0,00

<sup>1</sup> Valores em negrito indicam a menor mediana do erro frequencial nos cenários da segunda etapa dos experimentos. Valores sublinhados indicam diferença estatística significativa segundo os testes aplicados para todos os descritores. Os testes foram realizados com nível de significância de 0,05.

desempenho nos cinco cenários restantes (2, 6, 16, 19 e 29). O mesmo comportamento foi observado na base de dados *Hiperplane*: a EQ obteve o melhor desempenho em 15 cenários; a EP superou a EQ em sete cenários (4, 14, 19, 21, 24, 25 e 29). Para a base *SineV*, a superioridade da EQ também foi observada em 15 cenários. Novamente, a EP foi a melhor opção nos sete cenários restantes (1, 16, 19, 21, 24, 26 e 29). Por fim, na base de dados *SineH*, a EQ também superou os demais métodos, embora essa superioridade tenha ocorrido em apenas 13 cenários. Além disso, diferentemente das outras bases de dados, a estratégia EI apresentou o segundo melhor desempenho em cinco cenários (6, 21, 24, 25 e 29), enquanto a EP foi superior nos cenários 9, 16, 19 e 32.

A Figura 34 resume os resultados obtidos no experimento. Conforme ilustrado,

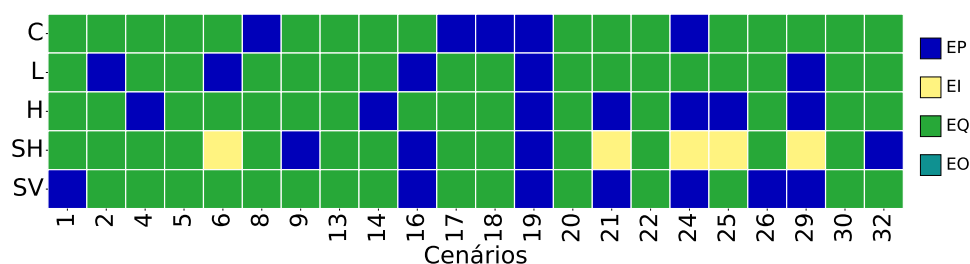


Figura 34 – O mapa de calor ilustra os cenários em que cada estratégia alcançou o melhor resultado.

a estratégia EQ apresentou o melhor desempenho em 18 cenários, enquanto a EP foi a melhor nos cenários 16, 19, 24 e 29. Assim, a EQ pode ser considerada a estratégia mais adequada para reagir à mudança de conceito. Algumas observações podem ser feitas a partir desses resultados. Primeiramente, destaca-se que, em 100% dos cenários caracterizados por mudanças repentinas, de alta severidade, com ou sem recorrência (cenários 1, 2, 5 e 6) a EQ obteve o melhor desempenho. Isso indica que mudanças com essas características específicas são mais bem tratadas por estratégias que incorporam mecanismos de esquecimento.

A segunda observação também diz respeito aos mecanismos de esquecimento. Considerando que o objetivo da estratégia EQ foi analisar o impacto do esquecimento parcial do classificador na estratégia de reação, seu alto desempenho indica que, ao menos em ambientes supervisionados, uma estratégia de reação eficaz deve incorporar mecanismos de esquecimento. Isso se insere no chamado dilema estabilidade–plasticidade (GROSSBERG, 1988; BAYRAM; AHMED; KASSLER, 2022), no qual a estabilidade se refere à preservação do conhecimento relevante, enquanto a plasticidade implica na substituição do conhecimento obsoleto. A necessidade de incorporar mecanismos de esquecimento ao reagir a mudanças de conceito constitui a segunda principal conclusão deste trabalho.

## 5.5 Teste de Ablação

Estudo de ablação é um método de pesquisa utilizado mais comumente nas áreas da Medicina e Psicologia, no qual o desempenho de um organismo é analisado após a

remoção de determinadas partes (SHEIKHOESLAMI, 2019). No contexto de aprendizado de máquina, um estudo de ablação é definido como um experimento em que componentes individuais de uma solução são removidos ou substituídos, com o objetivo de avaliar sua contribuição para o desempenho geral da solução (COHEN; HOWE, 1988; NEWELL, 1975).

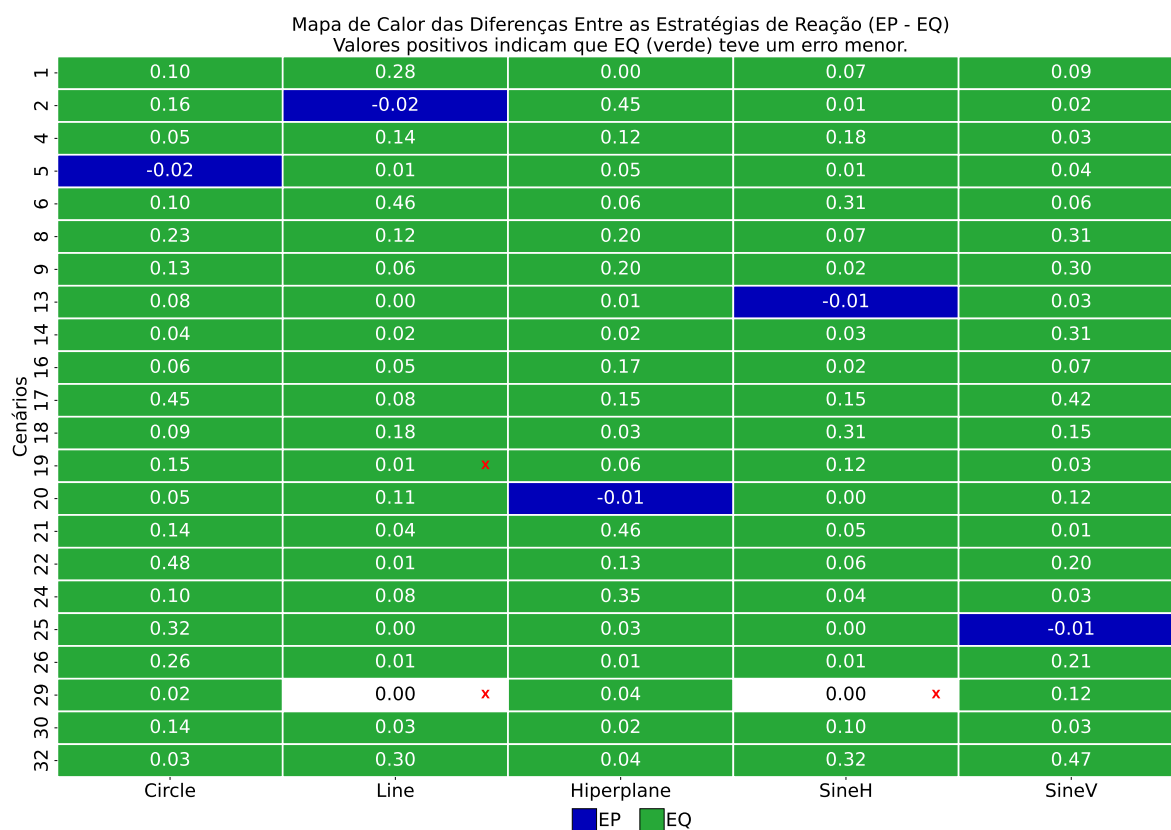


Figura 35 – Valores positivos indicam que EQ alcançou a menor taxa de erro. X indica que não há diferença significativa entre EQ e EP, no nível de significância 0,05.

Com base nos resultados obtidos no segundo estágio do experimento, o objetivo do estudo de ablação conduzido neste trabalho foi avaliar o impacto do EDDM como detector de mudanças no desempenho observado. Para isso, o EDDM foi substituído pelo DDM (GAMA et al., 2004), e os cenários da segunda etapa dos experimentos foram reanalisados, com foco apenas nas estratégias EP e EQ. Os resultados, apresentados na Figura 35, indicam que o alto desempenho da estratégia EQ não se deve a características particulares do detector EDDM. Mesmo após a substituição do EDDM pelo DDM, a EQ continuou apresentando os melhores resultados na maioria dos cenários. Para comparar

as diferenças de desempenho, foi aplicado o teste estatístico de Mann-Whitney com nível de significância de 0,05.

## 5.6 Considerações Finais

No estudo apresentado neste capítulo, investigou-se o impacto dos descritores de mudança de conceito no desempenho dos classificadores e nas estratégias de reação. Para isso, foram realizadas análises experimentais em cinco bases de dados sintéticas, abrangendo 32 cenários distintos. Aplicaram-se cinco diferentes estratégias de reação, incluindo abordagens tradicionais e incrementais. Além disso, foram utilizados testes estatísticos, análise do tamanho de efeito e realizado um teste de ablação para validar a robustez das conclusões obtidas. Os experimentos foram divididos em duas etapas e um teste de ablação, totalizando 1.000 testes individuais conduzidos.

Os resultados demonstram que a escolha da estratégia de reação à mudança de conceito está fortemente associada às características da própria mudança. A adoção de uma estratégia padrão, frequentemente relatada na literatura, pode levar à redução do desempenho do classificador em determinados cenários. Em particular, optar por não reagir à mudança é a estratégia mais adequada para lidar com mudanças de baixa frequência e periódicas. Nesses casos, os custos de retreinamento do classificador superam os possíveis benefícios da adaptação ao novo conceito. Esta constitui a nossa primeira conclusão principal.

Em cenários nos quais o retreinamento se mostrou necessário, a estratégia mais comumente utilizada também não foi a mais efetiva. Em vez disso, a abordagem de aprendizado incremental que incorporou o esquecimento de instâncias antigas apresentou melhores resultados. Esse resultado destaca um elemento crucial, porém pouco explorado na literatura sobre mudança de conceito: o mecanismo de esquecimento. Assim, conclui-se que abordagens voltadas ao tratamento da mudança de conceito devem incorporar mecanismos eficazes de esquecimento. Esta é a nossa segunda conclusão principal.

Além disso, os resultados dos experimentos também indicaram que caracterizar

a mudança de conceito com base em seus descritores ajuda a evitar custos desnecessários de retreinamento e melhora o desempenho do classificador em cenários específicos, especialmente quando essa caracterização é incorporada às estratégias de reação. Esta é a nossa terceira conclusão principal.

O presente estudo contribui de forma significativa para a compreensão do papel dos descritores na definição de estratégias de reação à mudança de conceito. No entanto, algumas limitações devem ser consideradas. Primeiramente, este trabalho não abordou explicitamente mudanças de conceito causadas pelo surgimento ou desaparecimento de novos atributos ou classes, conforme proposto por [Bayram, Ahmed e Kassler \(2022\)](#). Isso pode limitar a aplicabilidade dos resultados em ambientes que apresentam essas características, embora tais mudanças sejam relativamente raras. Ainda assim, é importante observar que, entre os 32 cenários testados, mudanças com essas características são implicitamente contempladas, o que minimiza o impacto dessa limitação.

Outra limitação está relacionada à ausência de testes envolvendo o descritor zona de influência. Essa limitação se deve à dificuldade de simular com precisão o local onde as mudanças ocorrem. No entanto, os potenciais efeitos da zona de influência sobre a generalização são conhecidos e afetam exclusivamente mudanças virtuais com alta severidade. Nesse caso, a zona de influência não impacta a fronteira de decisão, pois o descritor severidade superestima a severidade real da mudança, fazendo com que uma mudança virtual seja tratada como uma mudança real.

O próximo capítulo discute as conclusões gerais desta tese e apresenta sugestões para trabalhos futuros.

Esta tese abordou o fenômeno da mudança de conceito, que pode acarretar em perda de desempenho dos classificadores. Esse problema é amplamente explorado na literatura a partir da perspectiva de métodos de detecção supervisionados. Esses métodos assumem como premissa que os rótulos verdadeiros estarão sempre disponíveis, o que não é uma condição realística em ambientes reais. Além disso, as estratégias de reação à mudança amplamente utilizadas não levam em consideração as características das mudanças, que podem ser insumos úteis para a tomada de decisão baseada em evidências.

Dado esse contexto, esta tese aborda o fenômeno da mudança de conceito considerando os pontos destacados no parágrafo anterior, resultando em três principais contribuições, organizadas nos três estudos realizados. O primeiro estudo abordou uma revisão sistemática da literatura sobre métodos de detecção não supervisionados considerados estado-da-arte, com a proposição de uma nova classificação taxonômica.

A revisão sistemática da literatura identificou dois grupos de publicações: i) artigos de revisão e análise de diferentes abordagens de detecção de mudança de conceito; e ii) artigos que apresentam métodos não supervisionados. O primeiro grupo evidencia a escassez de trabalhos que sintetizem tais abordagens, enquanto o segundo permitiu agrupar métodos não supervisionados e semissupervisionados em duas categorias principais: baseados em lote e *online*.

Os métodos em lote e *online* referem-se exclusivamente ao componente de detecção. No primeiro, a mudança é identificada mediante o processamento de conjuntos de instâncias, ao passo que, no segundo, a detecção ocorre instância a instância ao longo do fluxo de dados.

O segundo trabalho conduzido analisou o impacto de quatro descritores da

mudança de conceito no desempenho do classificador, com o objetivo de identificar os descritores mais relevantes para o processo de reação à mudança. Os resultados desse estudo indicam que a reação à mudança de conceito não deve ser uniforme, mas sim ajustada às características específicas de cada cenário, sendo que em determinadas situações a melhor resposta pode ser não reagir. Observou-se que velocidade, severidade e recorrência apresentaram valores bem definidos em cenários de melhor e pior desempenho, enquanto a frequência mostrou comportamento menos consistente. No entanto, apenas a severidade se destacou de forma predominante, seguida pela frequência e pela recorrência, ao passo que a velocidade não demonstrou impacto significativo. Dessa forma, severidade, recorrência e frequência configuram os descritores mais relevantes para orientar estratégias de reação, com destaque para a severidade como o fator isolado de maior influência. Por fim, ressalta-se a importância de incorporar mecanismos de descrição da mudança de conceito, executados de forma concorrente à detecção, de modo que as propriedades descritas subsidiem a escolha da estratégia de reação mais adequada.

O último estudo realizado analisou o impacto dos descritores da mudança de conceito no desempenho do classificador e na estratégia de reação. O estudo demonstrou que a escolha da estratégia de reação à mudança de conceito depende fortemente das características do fenômeno. Em cenários de mudanças de baixa frequência e periódicas, a melhor resposta pode ser simplesmente não reagir, evitando custos desnecessários de retreinamento. Quando o retreinamento se mostrou necessário, a estratégia incremental com esquecimento de instâncias antigas superou a abordagem padrão, o que evidencia a relevância do mecanismo de esquecimento, ainda pouco explorado na literatura, como componente essencial das estratégias de reação. Além disso, verificou-se que a caracterização da mudança por meio de descritores potencializa a definição de reações mais adequadas, reduz custos e melhora o desempenho dos classificadores.

Ao correlacionar os resultados dos Capítulos 4 e 5, observa-se que o descritor com maior influência no desempenho do classificador é a severidade. Os cenários de alta severidade requerem a adoção de uma estratégia de retreino total ou de aprendizado incremental, indicando a necessidade de um custo maior de rotulagem. Esses resultados



reforçam tanto o impacto do descritor de severidade, como a definição de uma estratégia de acordo com as características da mudança. Outra correlação relevante diz respeito ao descritor de recorrência. No Capítulo 4, todos os experimentos foram conduzidos com a estratégia padrão de reação, e verificou-se que as mudanças recorrentes impactaram mais o desempenho do que as não recorrentes. Por sua vez, no Capítulo 5, os resultados demonstraram que, em cenários de recorrência, a melhor estratégia foi não reagir, reforçando a hipótese da tese de que a definição da estratégia deve estar fundamentada nas evidências fornecidas pelos descritores. Essa correlação também se verifica nas mudanças de baixa frequência, em que a melhor estratégia é não reagir, e nas de alta frequência, em que estratégias adaptativas demonstraram resultados superiores. Dessa forma, adotar a estratégia padrão resultaria em desempenho inferior em comparação às demais estratégias.

## 6.1 Limitações e Trabalhos Futuros

A tese contribui de forma significativa para o entendimento da relevância das características da mudança de conceito no desempenho dos classificadores, assim como de sua influência na escolha da estratégia de reação mais adequada. Entretanto, algumas limitações devem ser consideradas.

Primeiramente, a tese não abordou explicitamente a mudança de conceito causada pelo surgimento ou desaparecimento de novos atributos ou classes, conforme proposto por [Bayram, Ahmed e Kassler \(2022\)](#). Isso pode limitar a aplicabilidade de nossos resultados em ambientes que apresentem essas características, embora tais mudanças sejam relativamente raras. Ainda assim, é importante destacar que, entre os 32 cenários testados nos experimentos do Capítulo 5, mudanças com essas características estão implicitamente contempladas, o que minimiza o impacto dessa limitação.

Outra limitação está relacionada à ausência de testes envolvendo o descritor de zona de influência. Essa limitação decorreu da dificuldade em simular a localização exata onde as mudanças ocorrem. Contudo, os potenciais efeitos da zona de influência sobre a generalização são conhecidos e afetam exclusivamente mudanças virtuais de

alta severidade. Nesse caso, a zona de influência não impacta a fronteira de decisão, pois o descritor de severidade superestima a severidade real da mudança, fazendo com que uma mudança virtual seja tratada como real.

Para pesquisas futuras, em relação à revisão sistemática da literatura e à classificação taxonômica proposta, recomenda-se expandir esse estudo por meio da avaliação e da comparação dos métodos revisados em experimentos computacionais, como realizado por [Barros e Santos \(2018\)](#) para detectores supervisionados. Também sugere-se uma sistematização dos detectores de mudança de conceito de acordo com as estratégias de reação adotadas.

Em relação à compreensão da natureza da mudança de conceito, sugere-se: (i) analisar o impacto mútuo entre os descritores; (ii) aprofundar a investigação das relações existentes entre os descritores; (iii) desenvolver mecanismos para caracterizar a mudança em termos desses descritores em tempo execução; (iv) avaliar os impactos da mudança em modelos de aprendizado profundo e verificar se diferem daqueles observados em modelos rasos; (v) investigar de que forma características específicas de problemas do mundo real podem influenciar a ação dos descritores sobre as estratégias de reação; e (vi) projetar detectores que considerem tais características.

Por fim, em relação aos descritores e à definição de estratégias de reação, sugere-se também alguns trabalhos futuros. Primeiramente, analisar explicitamente mudanças geradas pelo surgimento ou desaparecimento de atributos ou classes. Em segundo lugar, é importante investigar se mudanças no número de atributos ou classes representam um novo descritor espacial para a mudança de conceito, dado que essas alterações modificam o espaço de entrada, em vez de constituírem apenas outro tipo de mudança. No que se refere a novos descritores, também se sugere a exploração de características adicionais que possam descrever a mudança de conceito, como a Entropia da Informação ([SHANNON, 1948](#); [SUN; MI; JIN, 2024](#)), os metaatributos discutidas em ([KOMORNICZAK; KSIENIEWICZ, 2024](#)) e as relações causais entre atributos abordadas em ([YANG et al., 2025](#)). Recomenda-se ainda o desenvolvimento ou aprimoramento de métodos para mensurar as características das mudanças de conceito. Por fim, o estudo de mecanismos de esquecimento, com especial atenção ao dilema entre estabilidade e

plasticidade, pode ser abordado em trabalhos futuros.

---

## Referências

---

- ABDALLAH, A.; MAAROF, M.; ZAINAL, A. Fraud detection system: A survey. *Journal of Network and Computer Applications*, v. 68, p. 90–113, 2016. [113](#)
- AGRAHARI, S.; SINGH, A. K. Concept drift detection in data stream mining : A literature review. *Journal of King Saud University - Computer and Information Sciences*, 2021. ISSN 1319-1578. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1319157821003062>. [29](#)
- AGRAHARI, S.; SINGH, A. K. Disposition-based concept drift detection and adaptation in data stream. *Arabian Journal for Science and Engineering*, v. 47, n. 8, p. 10605–10621, Aug 2022. ISSN 2191-4281. Disponível em: <https://doi.org/10.1007/s13369-022-06653-4>. [56](#), [60](#)
- AGUIAR, G. J.; CANO, A. A comprehensive analysis of concept drift locality in data streams. *Knowledge-Based Systems*, v. 289, p. 111535, 2024. ISSN 0950-7051. [57](#), [59](#), [60](#), [62](#), [64](#), [65](#), [72](#)
- ALIPPI, C.; ROVERI, M. Just-in-time adaptive classifiers—part i: Detecting nonstationary changes. *IEEE Transactions on Neural Networks*, v. 19, n. 7, p. 1145–1153, 2008. [113](#)
- BAENA-GARCÍA, M. et al. Early drift detection method. *ECML PKDD Workshop on Knowledge Discovery from Data Streams*, 01 2006. [19](#), [54](#), [66](#), [80](#), [83](#), [111](#), [112](#), [136](#), [138](#)
- BARROS, R.; SANTOS, S. A large-scale comparison of concept drift detectors. *Information Sciences*, v. 451-452, p. 348–370, 2018. [97](#), [114](#), [142](#)
- BARROS, R.; SANTOS, S. An overview and comprehensive comparison of ensembles for concept drift. *Information Fusion*, v. 52, p. 213–244, 2019. [114](#)
- BASHIR, S.; PETROVSKI, A.; DOOLAN, D. A framework for unsupervised change detection in activity recognition. *International Journal of Pervasive Computing and Communications*, v. 13, n. 2, p. 157–175, 2017. [120](#), [123](#), [140](#)
- BAYRAM, F.; AHMED, B. S.; KASSLER, A. From concept drift to model degradation: An overview on performance-aware drift detectors. *Knowledge-Based Systems*, v. 245, p. 108632, 2022. ISSN 0950-7051. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0950705122002854>. [18](#), [25](#), [28](#), [29](#), [30](#), [31](#), [32](#), [33](#), [34](#), [35](#), [36](#), [37](#), [38](#), [44](#), [90](#), [93](#), [96](#)
- BEYENE, A. et al. Improved concept drift handling in surgery prediction and other applications. *Knowledge and Information Systems*, v. 44, n. 1, p. 177–196, 2015. [113](#)

- BIFET, A.; GAVALDÀ, R. Adaptive learning from evolving data streams. In: ADAMS, N. M. et al. (Ed.). *Advances in Intelligent Data Analysis VIII*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. p. 249–260. ISBN 978-3-642-03915-7. 80
- BIFET, A.; GAVALDÀ, R. Learning from time-changing data with adaptive windowing. In: *7th SIAM International Conference on Data Mining (SDM'07)*. Mineapolis, Minnesota: Society for Industrial and Applied Mathematics, 2007. v. 7. 63, 132
- BISHOP, C. M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN 0387310738. 28
- BOBOWSKA, B.; CHORÁS, M.; WOŹNIAK, M. Advanced analysis of data streams for critical infrastructures protection and cybersecurity. *Journal of Universal Computer Science*, v. 24, n. 5, p. 622–633, 2018. 113
- BOUSQUET, O.; ELISSEFF, A. Stability and generalization. *Journal of Machine Learning Research*, JMLR.org, v. 2, p. 499–526, mar. 2002. ISSN 1532-4435. 134
- BRZEZINSKI, D. *Mining Data Streams with Concept Drift*. Dissertação (Mestrado) — Poznan University of Technology, 2010. Disponível em: <<http://www.cs.put.poznan.pl/dbrzezinski/publications/ConceptDrift.pdf>>. 46
- CASTELLANI, A.; SCHMITT, S.; HAMMER, B. Task-sensitive concept drift detector with constraint embedding. In: *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*. [S.l.: s.n.], 2021. p. 01–08. 26, 44
- CERQUEIRA, V. et al. Studd: a student–teacher method for unsupervised concept drift detection. *Machine Learning*, Jun 2022. Disponível em: <<https://doi.org/10.1007/s10994-022-06188-7>>. 55, 66
- CHANDRA, S. et al. An adaptive framework for multistream classification. In: *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. New York, NY, USA: Association for Computing Machinery, 2016. (CIKM '16), p. 1181–1190. ISBN 9781450340731. Disponível em: <<https://doi.org/10.1145/2983323.2983842>>. 111
- CHIANG, H.-S.; WU, Z.-W. Online incremental learning for sleep quality assessment using associative petri net. *Applied Soft Computing Journal*, v. 68, p. 774–783, 2018. 113
- COHEN, J. *Statistical Power Analysis for the Behavioral Sciences*. [S.l.]: Lawrence Erlbaum Associates, 1988. 82
- COHEN, P. R.; HOWE, A. E. How evaluation guides ai research: The message still counts more than the medium. *AI Magazine*, v. 9, n. 4, p. 35, Dec. 1988. Disponível em: <<https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/952>>. 91
- COSTA, A.; GIUSTI, R.; SANTOS, E. M. d. Analysis of descriptors of concept drift and their impacts. *Informatics*, v. 12, n. 1, 2025. ISSN 2227-9709. Disponível em: <<https://www.mdpi.com/2227-9709/12/1/13>>. 19, 26, 40, 41, 42, 44, 46, 49, 51
- COSTA, A. F. J.; ALBUQUERQUE, R. A. S.; SANTOS, E. M. d. A drift detection method based on active learning. In: *2018 International Joint Conference on Neural Networks (IJCNN)*. [S.l.: s.n.], 2018. p. 1–8. 55, 66, 120, 125, 128

- DASU, T. et al. An information-theoretic approach to detecting changes in multi-dimensional data streams. In: *In Proc. Symp. on the Interface of Statistics, Computing Science, and Applications*. [S.l.: s.n.], 2006. 42, 56, 59, 60, 62, 63
- DAWID, A.; VOVK, V. Prequential probability: Principles and properties. *Bernoulli*, v. 5, 02 1999. 81
- DONG, F. et al. Concept drift region identification via competence-based discrepancy distribution estimation. In: *2017 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*. [S.l.: s.n.], 2017. p. 1–7. 56, 59, 60
- DONG, F. et al. A drift region-based data sample filtering method. *IEEE Transactions on Cybernetics*, p. 1–14, 2021. 56, 59, 60
- DONMEZ, P.; LEBANON, G.; BALASUBRAMANIAN, K. Unsupervised supervised learning i: Estimating classification and regression errors without labels. *Journal of Machine Learning Research*, v. 11, p. 1323–1351, 04 2010. 133
- DUDA, R. O.; HART, P. E. et al. *Pattern classification*. [S.l.]: John Wiley & Sons, 2006. 28
- ELWELL, R.; POLIKAR, R. Incremental learning of concept drift in nonstationary environments. *IEEE Transactions on Neural Networks*, v. 22, n. 10, p. 1517–1531, 2011. 32
- FAHY, C.; YANG, S.; GONGORA, M. Scarcity of labels in non-stationary data streams: A survey. Association for Computing Machinery, New York, NY, USA, v. 55, n. 2, jan 2022. ISSN 0360-0300. Disponível em: <<https://doi.org/10.1145/3494832>>. 26, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 44
- FANGYU, W. et al. Research on imbalanced data set preprocessing based on deep learning. In: *2021 Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS)*. [S.l.: s.n.], 2021. p. 75–79. 32
- FARRUGIA, D. et al. A real-time prescriptive solution for explainable cyber-fraud detection within the igaming industry. *SN Computer Science*, v. 2, n. 3, p. 215, Apr 2021. ISSN 2661-8907. Disponível em: <<https://doi.org/10.1007/s42979-021-00623-7>>. 20
- FORMAN, G. Tackling concept drift by temporal inductive transfer. In: *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. New York, NY, USA: Association for Computing Machinery, 2006. (SIGIR '06), p. 252–259. ISBN 1595933697. Disponível em: <<https://doi.org/10.1145/1148170.1148216>>. 35, 36
- FRIEDMAN, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, Taylor & Francis, v. 32, n. 200, p. 675–701, 1937. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/01621459.1937.10503522>>. 68
- FRÍAS-BLANCO, I. et al. Online and non-parametric drift detection methods based on hoeffding's bounds. *IEEE Transactions on Knowledge and Data Engineering*, v. 27, n. 3, p. 810–823, 2015. 83
- GAMA, J. et al. Learning with drift detection. In: BAZZAN, A. L. C.; LABIDI, S. (Ed.). *Advances in Artificial Intelligence – SBIA 2004*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004. p. 286–295. ISBN 978-3-540-28645-5. 19, 37, 54, 66, 80, 83, 91, 136, 138



- GAMA, J. a. et al. A survey on concept drift adaptation. *ACM Comput. Surv.*, Association for Computing Machinery, New York, NY, USA, v. 46, n. 4, mar 2014. ISSN 0360-0300. Disponível em: <<https://doi.org/10.1145/2523813>>. 18, 25, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 44, 54, 111, 112, 114
- GEMAQUE, R. N. et al. An overview of unsupervised drift detection methods. *WIREs Data Mining and Knowledge Discovery*, v. 10, n. 6, p. e1381, 2020. Disponível em: <<https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.1381>>. 18, 25, 60
- GILLELAND, M. *Levenshtein Distance*, in *Three Flavors*. 2020. Accessed = 2020-04-14. Disponível em: <<http://www.merriampark.com/ld.htm>>. 133
- GOLDENBERG, I.; WEBB, G. I. Survey of distance measures for quantifying concept drift and shift in numeric data. *Knowl. Inf. Syst.*, Springer-Verlag, Berlin, Heidelberg, v. 60, n. 2, p. 591–615, aug 2019. ISSN 0219-1377. Disponível em: <<https://doi.org/10.1007/s10115-018-1257-z>>. 41
- GOMES, H. M. et al. Correction to: Adaptive random forests for evolving data stream classification. *Machine Learning*, Springer Science and Business Media LLC, v. 108, n. 10, p. 1877–1878, mar. 2019. Disponível em: <<https://doi.org/10.1007/s10994-019-05793-3>>. 111
- GROSSBERG, S. Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks*, v. 1, n. 1, p. 17–61, 1988. ISSN 0893-6080. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0893608088900214>>. 90
- GUO, H. et al. Concept drift type identification based on multi-sliding windows. *Information Sciences*, v. 585, p. 1–23, 2022. ISSN 0020-0255. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0020025521011439>>. 25, 56, 59, 60
- GUZELLA, T.; CAMINHAS, W. A review of machine learning approaches to spam filtering. *Expert Systems with Applications*, v. 36, n. 7, p. 10206–10222, 2009. 113
- GÂLMEANU, H.; ANDONIE, R. Concept drift adaptation with incremental-decremental svm. *Applied Sciences*, v. 11, n. 20, 2021. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/11/20/9644>>. 37
- HAMMER, B.; VAQUET, V.; HINDER, F. One or two things we know about concept drift—a survey on monitoring in evolving environments. part b: locating and explaining concept drift. *Frontiers in Artificial Intelligence*, v. 7, 2024. ISSN 2624-8212. 58, 59, 60, 63, 64, 65
- HAQUE, A.; KHAN, L.; BARON, M. Sand: Semi-supervised adaptive novel class detection and classification over data stream. In: *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*. [S.l.: s.n.], 2016. (AAAI'16), p. 1652–1658. 111, 130, 134, 135, 140, 141
- HEDGES, L.; VEVEA, J. Meta-analysis: Tools. In: SMELSER, N. J.; BALTES, P. B. (Ed.). *International Encyclopedia of the Social & Behavioral Sciences*. Oxford: Pergamon, 2001. p. 9724–9730. ISBN 978-0-08-043076-8. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B0080430767004629>>. 82

- HINDER, F.; HAMMER, B. Concept drift segmentation via kolmogorov-trees. In: *29th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2021, Online event (Bruges, Belgium), October 6-8, 2021*. [S.l.: s.n.], 2021. 56, 59, 60
- HINDER, F.; KUMMERT, J.; HAMMER, B. Explaining concept drift by mean of direction. In: FARKAŠ, I.; MASULLI, P.; WERMTER, S. (Ed.). *Artificial Neural Networks and Machine Learning – ICANN 2020*. Cham: Springer International Publishing, 2020. p. 379–390. ISBN 978-3-030-61609-0. 57, 59, 60, 62, 64, 65
- HINDER, F. et al. Localization of concept drift: Identifying the drifting datapoints. In: *2022 International Joint Conference on Neural Networks (IJCNN)*. [S.l.: s.n.], 2022. p. 1–9. 62, 64, 65
- HINDER, F.; VAQUET, V.; HAMMER, B. One or two things we know about concept drift—a survey on monitoring in evolving environments. part a: detecting concept drift. *Frontiers in Artificial Intelligence*, v. 7, 2024. ISSN 2624-8212. 60
- HOENS, T. R.; POLIKAR, R.; CHAWLA, N. V. Learning from streaming data with concept drift and imbalance: an overview. *Progress in Artificial Intelligence*, v. 1, n. 1, p. 89–101, Apr 2012. ISSN 2192-6360. Disponível em: <<https://doi.org/10.1007/s13748-011-0008-0>>. 24
- HOSSEINI, M. J.; GHOLIPOUR, A.; BEIGY, H. An ensemble of cluster-based classifiers for semi-supervised classification of non-stationary data streams. *Knowledge and Information Systems*, Springer Science and Business Media LLC, v. 46, n. 3, p. 567–597, abr. 2015. Disponível em: <<https://doi.org/10.1007/s10115-015-0837-4>>. 111
- HU, H.; KANTARDZIC, M.; SETHI, T. No free lunch theorem for concept drift detection in streaming data classification: A review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, v. 10, p. 1–25, 2019. 111, 114
- IWASHITA, A. S.; PAPA, J. P. An overview on concept drift learning. *IEEE Access*, v. 7, p. 1532–1547, 2019. 26, 37, 44, 60, 114
- JIA, R.; KOH, Y. S.; DOBBIE, G. Predicting concept drift severity. In: *Workshop on learning in the presence of class imbalance and concept drift (LPCICD'17)*. [S.l.: s.n.], 2017. 41
- JR, P. M. G.; BARROS, R. S. M. D. Rcd: A recurring concept drift framework. *Pattern Recogn. Lett.*, Elsevier Science Inc., USA, v. 34, n. 9, p. 1018–1025, jul. 2013. ISSN 0167-8655. Disponível em: <<https://doi.org/10.1016/j.patrec.2013.02.005>>. 113
- KADLEC, P.; GRBIĆ, R.; GABRYS, B. Review of adaptation mechanisms for data-driven soft sensors. *Computers and Chemical Engineering*, v. 35, n. 1, p. 1–24, 2011. 113
- KHAMASSI, I. et al. Discussion and review on evolving data streams and concept drift adapting. *Evolving Systems*, v. 9, n. 1, p. 1–23, Mar 2018. ISSN 1868-6486. Disponível em: <<https://doi.org/10.1007/s12530-016-9168-2>>. 26, 50, 51, 111, 112, 113, 118, 138
- KIM, Y.; PARK, C. H. An efficient concept drift detection method for streaming data under limited labeling. *IEICE Transactions on Information and Systems*, E100.D, n. 10, p. 2537–2546, 2017. 130, 136, 137



- KOH, Y. S. CD-TDS: change detection in transactional data streams for frequent pattern mining. In: *2016 International Joint Conference on Neural Networks, IJCNN 2016, Vancouver, BC, Canada, July 24-29, 2016*. [S.l.]: IEEE, 2016. p. 1554–1561. [130](#), [132](#), [133](#), [141](#)
- KOMORNICZAK, J.; KSIENIEWICZ, P. On metafeatures' ability of implicit concept identification. *Machine Learning*, v. 113, n. 10, p. 7931–7966, Oct 2024. ISSN 1573-0565. Disponível em: <https://doi.org/10.1007/s10994-024-06612-0>. [63](#), [64](#), [65](#), [68](#), [97](#)
- KORE, A. et al. Empirical data drift detection experiments on real-world medical imaging data. *Nature Communications*, v. 15, n. 1, p. 1887, Feb 2024. ISSN 2041-1723. Disponível em: <https://doi.org/10.1038/s41467-024-46142-w>. [19](#)
- KOSINA, P.; GAMA, J.; SEBASTIÃO, R. Drift severity metric. In: . [S.l.: s.n.], 2010. v. 215, p. 1119–1120. [41](#)
- KRAWCZYK, B. et al. Ensemble learning for data stream analysis: A survey. *Information Fusion*, v. 37, p. 132–156, 2017. [54](#), [114](#)
- KRUSKAL, W. H.; WALLIS, W. A. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association, ASA Website*, v. 47, n. 260, p. 583–621, 1952. [82](#), [87](#)
- KUH, A.; PETSCHKE, T.; RIVEST, R. L. Learning time-varying concepts. In: *Proceedings of the 3rd International Conference on Neural Information Processing Systems*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1990. (NIPS'90), p. 183–189. ISBN 1558601848. [49](#)
- KULLBACK, S.; LEIBLER, R. A. On Information and Sufficiency. *The Annals of Mathematical Statistics*, Institute of Mathematical Statistics, v. 22, n. 1, p. 79 – 86, 1951. Disponível em: <https://doi.org/10.1214/aoms/1177729694>. [42](#), [56](#)
- KUNCHEVA, L. I. Classifier ensembles for changing environments. In: ROLI, F.; KITTNER, J.; WINDEATT, T. (Ed.). *Multiple Classifier Systems*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004. p. 1–15. ISBN 978-3-540-25966-4. [24](#)
- KUNCHEVA, L. I.; ZLIOBAIT, I. On the window size for classification in changing environments. *Intell. Data Anal.*, IOS Press, NLD, v. 13, n. 6, p. 861–872, dez. 2009. ISSN 1088-467X. [113](#)
- LAZARESCU, M. M.; VENKATESH, S.; BUI, H. H. Using multiple windows to track concept drift. *Intell. Data Anal.*, IOS Press, NLD, v. 8, n. 1, p. 29–59, jan 2004. ISSN 1088-467X. [41](#)
- LEO, J.; KALITA, J. Survey of continuous deep learning methods and techniques used for incremental learning. *Neurocomputing*, v. 582, p. 127545, 2024. ISSN 0925-2312. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0925231224003163>. [18](#)
- LI, B. et al. Faad: an unsupervised fast and accurate anomaly detection method for a multi-dimensional sequence over data stream. *Frontiers of Information Technology and Electronic Engineering*, v. 20, n. 3, p. 388–404, 2019. [120](#), [122](#), [138](#), [140](#)
- LIU, A. et al. Accumulating regional density dissimilarity for concept drift detection in data streams. *Pattern Recognition*, v. 76, p. 256–272, 2018. [120](#), [121](#), [138](#)

- LIU, A. et al. Regional concept drift detection and density synchronized drift adaptation. In: *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*. [s.n.], 2017. p. 2280–2286. Disponível em: <<https://doi.org/10.24963/ijcai.2017/317>>. 58, 62, 63
- LOPES, R. H. C. Kolmogorov-smirnov test. In: \_\_\_\_\_. *International Encyclopedia of Statistical Science*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. p. 718–720. ISBN 978-3-642-04898-2. 56
- LOSING, V.; HAMMER, B.; WERSING, H. Knn classifier with self adjusting memory for heterogeneous concept drift. In: *2016 IEEE 16th International Conference on Data Mining (ICDM)*. [S.l.: s.n.], 2016. p. 291–300. 25
- LOVRIC, M. International encyclopedia of statistical science. In: \_\_\_\_\_. [S.l.]: Springer, 2011. p. 718–720. 125
- LU, J. et al. Learning under concept drift: A review. *IEEE Transactions on Knowledge and Data Engineering*, v. 31, n. 12, p. 2346–2363, 2019. 25, 26, 29, 39, 40, 41, 42, 43, 45, 46, 47, 54, 68, 111, 112, 114, 120
- LU, N. et al. A concept drift-tolerant case-base editing technique. *Artificial Intelligence*, v. 230, p. 108–133, 2016. ISSN 0004-3702. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0004370215001459>>. 25, 46
- LUGHOFFER, E. et al. Recognizing input space and target concept drifts in data streams with scarcely labeled and unlabelled instances. *Information Sciences*, v. 355–356, p. 127–151, 2016. 130, 133
- MALETZKE, A.; REIS, D. dos; BATISTA, G. Combining instance selection and self-training to improve data stream quantification. *Journal of the Brazilian Computer Society*, v. 24, n. 1, p. 12, 2018. 120, 124, 125
- MALETZKE, A. G.; REIS, D. M. dos; BATISTA, G. E. Quantification in data streams: Initial results. In: *2017 Brazilian Conference on Intelligent Systems (BRACIS)*. [S.l.: s.n.], 2017. p. 43–48. 124
- MALETZKE, A. G. et al. On the need of class ratio insensitive drift tests for data streams. In: *Second International Workshop on Learning with Imbalanced Domains: Theory and Applications*. [S.l.: s.n.], 2018. (Proceedings of Machine Learning Research, v. 94), p. 110–124. 125
- MANN, H. B.; WHITNEY, D. R. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, Institute of Mathematical Statistics, v. 18, n. 1, p. 50 – 60, 1947. Disponível em: <<https://doi.org/10.1214/aoms/1177730491>>. 68, 81, 82, 87
- MASTELINI, S. M. et al. Benchmarking multi-target regression methods. In: *2018 7th Brazilian Conference on Intelligent Systems (BRACIS)*. [S.l.: s.n.], 2018. p. 396–401. 69
- MASUD, M. M. et al. Classification and novel class detection of data streams in a dynamic feature space. In: BALCÁZAR, J. L. et al. (Ed.). *Machine Learning and Knowledge Discovery in Databases*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. p. 337–352. ISBN 978-3-642-15883-4. 31, 32

- MATTOS, J. G. et al. Interpretable concept drift. In: TAVARES, J. M. R. S.; PAPA, J. P.; HIDALGO, M. G. (Ed.). *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*. Cham: Springer International Publishing, 2021. p. 271–280. ISBN 978-3-030-93420-0. 56, 59, 60
- MBALLO, C.; DIDAY, E. Kolmogorov-smirnov for decision trees on interval and histogram variables. In: BANKS, D. et al. (Ed.). *Classification, Clustering, and Data Mining Applications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004. p. 341–350. ISBN 978-3-642-17103-1. 56
- MELLO, R. de et al. On learning guarantees to unsupervised concept drift detection on data streams. *Expert Systems with Applications*, v. 117, p. 90–102, 2019. 130, 134
- MINKU, L. L.; WHITE, A. P.; YAO, X. The impact of diversity on online ensemble learning in the presence of concept drift. *IEEE Transactions on Knowledge and Data Engineering*, v. 22, n. 5, p. 730–742, May 2010. ISSN 2326-3865. 26, 35, 37, 41, 42, 45, 46, 47, 48, 50, 69, 135
- MORENO-TORRES, J. G. et al. A unifying view on dataset shift in classification. *Pattern Recognition*, v. 45, n. 1, p. 521–530, 2012. ISSN 0031-3203. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0031320311002901>>. 25
- MOUCHAWEH, M. S. *Learning from Data Streams in Dynamic Environments*. 1. ed. Cham, Switzerland: Springer International Publishing, 2016. (SpringerBriefs in applied sciences and technology). 19, 26, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41, 42, 43, 45, 46, 47, 49, 50, 51
- MOUSS, H. et al. Test of page-hinckley, an approach for fault detection in an agro-alimentary production system. In: *2004 5th Asian Control Conference (IEEE Cat. No.04EX904)*. [S.l.: s.n.], 2004. v. 2, p. 815–818 Vol.2. 133
- MUSTAFA, A. M. et al. Unsupervised deep embedding for novel class detection over data stream. In: *2017 IEEE International Conference on Big Data, BigData 2017*. [S.l.: s.n.], 2017. p. 1830–1839. 130, 133, 140
- NARASIMHAMURTHY, A.; KUNCHEVA, L. I. A framework for generating data to simulate changing environments. In: *Proceedings of the 25th Conference on Proceedings of the 25th IASTED International Multi-Conference: Artificial Intelligence and Applications*. USA: ACTA Press, 2007. (AIAP'07), p. 384–389. 50
- NEWELL, A. A tutorial on speech understanding systems. In: ACADEMIC PRESS. *Speech recognition: Invited papers presented at the 1974 IEEE Symposium*. [S.l.], 1975. p. 3–54. 91
- NISHIDA, K.; YAMAUCHI, K. Detecting concept drift using statistical testing. In: CORRUBLE, V.; TAKEDA, M.; SUZUKI, E. (Ed.). *Discovery Science*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. p. 264–269. ISBN 978-3-540-75488-6. 63
- OKOLI, C. A guide to conducting a standalone systematic literature review. *Communications of the Association for Information Systems*, v. 37, n. 43, p. 879–910, 2015. 115, 141

- PAGE, E. S. Continuous inspection schemes. *Biometrika*, v. 41, n. 1-2, p. 100–115, 06 1954. ISSN 0006-3444. Disponível em: <<https://doi.org/10.1093/biomet/41.1-2.100>>. 63, 112
- PARDO, L.; ANTIVILO-BRUNA, A.; MIGUEZ, G. Effect size. In: \_\_\_\_\_. *Encyclopedia of Animal Cognition and Behavior*. Cham: Springer International Publishing, 2020. p. 1–5. ISBN 978-3-319-47829-6. Disponível em: <[https://doi.org/10.1007/978-3-319-47829-6\\_242-1](https://doi.org/10.1007/978-3-319-47829-6_242-1)>. 82
- PESARANGHADER, A.; VIKTOR, H.; PAQUET, E. Mcdiarmid drift detection methods for evolving data streams. In: . [S.l.: s.n.], 2018. p. 1–9. 112
- PESARANGHADER, A.; VIKTOR, H. L. Fast hoeffding drift detection method for evolving data streams. In: FRASCONI, P. et al. (Ed.). *Machine Learning and Knowledge Discovery in Databases*. Cham: Springer International Publishing, 2016. p. 96–111. ISBN 978-3-319-46227-1. 63
- PINAGÉ, F.; SANTOS, E. M. dos; GAMA, J. A drift detection method based on dynamic classifier selection. *Data Mining and Knowledge Discovery*, v. 34, n. 1, p. 50–74, Jan 2020. Disponível em: <<https://doi.org/10.1007/s10618-019-00656-w>>. 54, 130, 135, 140
- PINAGE, F. A.; SANTOS, E. M. dos. A dissimilarity-based drift detection method. In: 2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI). [S.l.: s.n.], 2015. p. 1069–1076. 37, 66, 84
- QUIÑONERO-CANDELA, J. et al. When training and test sets are different: Characterizing learning transfer. In: \_\_\_\_\_. *Dataset Shift in Machine Learning*. [S.l.: s.n.], 2009. 25
- RAMÍREZ-GALLEGO, S. et al. A survey on data preprocessing for data stream mining: Current status and future directions. *Neurocomputing*, v. 239, p. 39–57, 2017. ISSN 0925-2312. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925231217302631>>. 50
- REIS, D. M. dos et al. Fast unsupervised online drift detection using incremental kolmogorov-smirnov test. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. [S.l.: s.n.], 2016. (KDD '16), p. 1545–1554. 130, 131, 139, 141
- RICHARDSON, J. T. Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, v. 6, n. 2, p. 135–147, 2011. ISSN 1747-938X. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1747938X11000029>>. 81
- RIOS, R. A. et al. Applying concept drift to understand hepatitis evolution in brazil. *Cybernetics and Systems*, Taylor & Francis, v. 51, n. 6, p. 631–645, 2020. Disponível em: <<https://doi.org/10.1080/01969722.2020.1724640>>. 19
- ROTALINTI, Y. et al. Detecting drift in healthcare ai models based on data availability. In: KOPRINSKA, I. et al. (Ed.). *Machine Learning and Principles and Practice of Knowledge Discovery in Databases*. Cham: Springer Nature Switzerland, 2023. p. 243–258. ISBN 978-3-031-23633-4. 19

- SAADALLAH, A. et al. Bright—drift-aware demand predictions for taxi networks. *IEEE Transactions on Knowledge and Data Engineering*, v. 32, n. 2, p. 234–245, 2020. 139
- SCHLIMMER, J. C.; GRANGER, R. H. Incremental learning from noisy data. *Machine Learning*, v. 1, n. 3, p. 317–354, Sep 1986. ISSN 1573-0565. Disponível em: <<https://doi.org/10.1007/BF00116895>>. 18, 25
- SETHI, T.; KANTARDZIC, M. Don't pay for validation: Detecting drifts from unlabeled data using margin density. *Procedia Computer Science*, v. 53, n. 1, p. 103–112, 2015. 120, 125, 126
- SETHI, T.; KANTARDZIC, M. Handling adversarial concept drift in streaming data. *Expert Systems with Applications*, v. 97, p. 18–40, 2018. 120, 127, 139, 140
- SETHI, T. S.; KANTARDZIC, M. On the reliable detection of concept drift from streaming unlabeled data. *Expert Syst. Appl.*, Pergamon Press, Inc., USA, v. 82, n. C, p. 77–99, oct 2017. ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957417417302439>>. 54, 66, 120, 126, 128, 140
- SHAKER, A.; LUGHOFER, E. Self-adaptive and local strategies for a smooth treatment of drifts in data streams. *Evolving Systems*, Springer Science and Business Media LLC, v. 5, n. 4, p. 239–257, maio 2014. Disponível em: <<https://doi.org/10.1007/s12530-014-9108-y>>. 113
- SHAMITHA, S. K.; ILANGO, V. Importance of self-learning algorithms for fraud detection under concept drift. In: SANYAL, G. et al. (Ed.). *International Conference on Artificial Intelligence and Sustainable Engineering*. Singapore: Springer Singapore, 2022. p. 343–354. ISBN 978-981-16-8546-0. 20
- SHANNON, C. E. A mathematical theory of communication. *The Bell System Technical Journal*, v. 27, p. 379–423, 1948. Disponível em: <<http://plan9.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf>>. 97
- SHEIKHOLESLAMI, S. *Ablation Programming for Machine Learning*. 52 p. Dissertação (Mestrado) — KTH, School of Electrical Engineering and Computer Science (EECS), 2019. 91
- SOUZA, V. M. A. et al. Challenges in benchmarking stream learning algorithms with real-world data. *Data Mining and Knowledge Discovery*, v. 34, n. 6, p. 1805–1858, Nov 2020. ISSN 1573-756X. Disponível em: <<https://doi.org/10.1007/s10618-020-00698-5>>. 68
- STEVANOSKI, B. et al. Change detection and adaptation in multi-target regression on data streams. *Machine Learning*, v. 113, n. 11, p. 8585–8622, Dec 2024. ISSN 1573-0565. Disponível em: <<https://doi.org/10.1007/s10994-024-06621-z>>. 69
- SUN, Y.; MI, J.; JIN, C. Entropy-based concept drift detection in information systems. *Knowledge-Based Systems*, v. 290, p. 111596, 2024. ISSN 0950-7051. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0950705124002314>>. 97
- SUSNJAK, T.; MADDIGAN, P. Forecasting patient flows with pandemic induced concept drift using explainable machine learning. *EPJ Data Science*, v. 12, n. 1, p. 11, Apr 2023. ISSN 2193-1127. Disponível em: <<https://doi.org/10.1140/epjds/s13688-023-00387-5>>. 20



- SUÁREZ-CETRULO, A. L.; QUINTANA, D.; CERVANTES, A. A survey on machine learning for recurring concept drifting data streams. *Expert Systems with Applications*, v. 213, p. 118934, 2023. ISSN 0957-4174. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0957417422019522>>. 50
- TAN, C. H.; LEE, V. C.; SALEHI, M. Information resources estimation for accurate distribution-based concept drift detection. *Information Processing e Management*, v. 59, n. 3, p. 102911, 2022. ISSN 0306-4573. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0306457322000383>>. 26, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 44
- TOBIAS, P. Statistical process control and failure mode analysis. In: BUSCHOW, K. J. et al. (Ed.). *Encyclopedia of Materials: Science and Technology*. Oxford: Elsevier, 2001. p. 8816 – 8824. ISBN 978-0-08-043152-9. 124
- TRAN, D. Automated change detection and reactive clustering in multivariate streaming data. In: *2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF)*. [S.l.: s.n.], 2019. p. 1–6. 113
- TSYMBAL, A. et al. Dynamic integration of classifiers for handling concept drift. *Information Fusion*, v. 9, n. 1, p. 56–68, 2008. ISSN 1566-2535. Special Issue on Applications of Ensemble Methods. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1566253506001138>>. 30
- TSYMBAL, A.; PUURONEN, S. Bagging and boosting with dynamic integration of classifiers. In: ZIGHED, D. A.; KOMOROWSKI, J.; ŻYTKOW, J. (Ed.). *Principles of Data Mining and Knowledge Discovery*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000. p. 116–125. ISBN 978-3-540-45372-7. 113
- WANG, P. et al. Quadcdd: A quadruple-based approach for understanding concept drift in data streams. *Expert Systems with Applications*, v. 238, p. 122114, 2024. ISSN 0957-4174. 58, 59, 60, 61, 64, 65
- WARES, S.; ISAACS, J.; ELYAN, E. Data stream mining: methods and challenges for handling concept drift. *SN Applied Sciences*, v. 1, n. 11, p. 1412, 2019. 111, 112, 113, 139
- WEBB, G. I. et al. Characterizing concept drift. *Data Mining and Knowledge Discovery*, v. 30, n. 4, p. 964–994, Jul 2016. ISSN 1573-756X. Disponível em: <<https://doi.org/10.1007/s10618-015-0448-4>>. 24, 25, 35, 37, 41, 46, 47, 49, 50, 51, 54
- WIDMER, G.; KUBAT, M. Learning in the presence of concept drift and hidden contexts. *Machine Learning*, v. 23, n. 1, p. 69–101, Apr 1996. ISSN 1573-0565. Disponível em: <<https://doi.org/10.1007/BF00116900>>. 25, 49
- WILCOX, R. R. Chapter 5 - comparing two groups. In: WILCOX, R. R. (Ed.). *Introduction to Robust Estimation and Hypothesis Testing (Fifth Edition)*. Fifth edition. Academic Press, 2022. p. 153–251. ISBN 978-0-12-820098-8. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780128200988000117>>. 81
- XU, S.; WANG, J. Dynamic extreme learning machine for data stream classification. *Neurocomput.*, Elsevier Science Publishers B. V., NLD, v. 238, n. C, p. 433–449, maio 2017. ISSN 0925-2312. Disponível em: <<https://doi.org/10.1016/j.neucom.2016.12.078>>. 111

- YAN, M. M. W. Accurate detecting concept drift in evolving data streams. *ICT Express*, v. 6, n. 4, p. 332–338, 2020. ISSN 2405-9595. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2405959520300655>. 26, 44
- YANG, L. et al. Detecting and rationalizing concept drift: A feature-level approach for understanding cause–effect relationships in dynamic environments. *Expert Systems with Applications*, v. 260, p. 125365, 2025. ISSN 0957-4174. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0957417424022322>. 61, 64, 65, 97
- ŽLIOBAITĖ, I. Controlled permutations for testing adaptive learning models. *Knowledge and Information Systems*, v. 39, n. 3, p. 565–578, Jun 2014. ISSN 0219-3116. 129
- ŽLIOBAITĖ, I.; HOLLMÉN, J. Optimizing regression models for data streams with missing values. *Machine Learning*, v. 99, n. 1, p. 47–73, Apr 2015. ISSN 1573-0565. Disponível em: <https://doi.org/10.1007/s10994-014-5450-3>. 25
- ŽLIOBAITĖ, I. *Learning under Concept Drift: an Overview*. arXiv, 2010. Disponível em: <https://arxiv.org/abs/1010.4784>. 41

---

## **Uma Visão Geral dos Métodos Não Supervisionados de Detecção de Mudança de Conceito**

---

**H**á na literatura revisões sobre o tema da detecção de mudanças, como as de [Lu et al. \(2019\)](#), [Hu, Kantardzic e Sethi \(2019\)](#), [Wares, Isaacs e Elyan \(2019\)](#) e [Khamassi et al. \(2018\)](#). No entanto, observa-se que nenhum desses trabalhos é especificamente voltado à detecção de mudanças não supervisionada, embora [Khamassi et al. \(2018\)](#) considerem uma categoria que inclui métodos desse tipo. Nesse cenário, o trabalho apresentado neste Apêndice concentra-se exclusivamente em métodos de detecção não supervisionada. Assim, a taxonomia aqui proposta é inédita e distingue-se das anteriores por tratar especificamente das propriedades de detectores projetados para operar em ambientes não supervisionados.

[Lu et al. \(2019\)](#) propuseram uma categorização dos métodos de detecção de mudança em três grandes classes. A primeira classe inclui métodos que monitoram taxas de erro, como o DDM ([GAMA et al., 2014](#)), o EDDM ([BAENA-GARCÍA et al., 2006](#)) e outros ([XU; WANG, 2017](#); [GOMES et al., 2019](#)). A segunda classe é composta por métodos que utilizam medidas de distância para estimar a similaridade entre as distribuições dos dados anteriores e atuais. A última classe inclui métodos que fazem uso de testes de hipóteses múltiplas para detectar a mudança de conceito. Nenhuma dessas classes é específica para métodos não supervisionados; entretanto, os trabalhos de [Haque, Khan e Baron \(2016\)](#), [Chandra et al. \(2016\)](#) e [Hosseini, Gholipour e Beigy \(2015\)](#) são apresentados como métodos não supervisionados ou semissupervisionados que pertencem às duas últimas classes. Portanto, se considerada a taxonomia proposta em [Lu et al. \(2019\)](#), detectores não supervisionados podem ser enquadrados nas duas últimas categorias.

[Hu, Kantardzic e Sethi \(2019\)](#) apresentaram uma categorização dos métodos de detecção de mudança de conceito em dois grandes grupos: abordagens baseadas em desempenho e abordagens baseadas na distribuição dos dados. Os métodos ba-



seados em desempenho monitoram continuamente alguma métrica relacionada ao erro, como acurácia, F-score, precisão e revocação. Em geral, uma mudança é sinalizada quando se observa uma queda significativa nessa métrica. Como esses métodos requerem rótulos verdadeiros para estimar os erros, eles não são aplicáveis a tarefas não supervisionadas. Por sua vez, as abordagens baseadas na distribuição utilizam medidas de monitoramento da distribuição dos dados, como localização, densidade e amplitude. Os métodos pertencentes a essa categoria podem ser supervisionados ou não supervisionados. Ressalta-se que esse grupo abrange os métodos das duas últimas classes da categorização apresentada por [Lu et al. \(2019\)](#).

[Wares, Isaacs e Elyan \(2019\)](#) descrevem uma categorização para métodos supervisionados, em que os métodos são classificados em quatro grupos: i) métodos estatísticos, ii) métodos baseados em janelas, iii) comitês baseados em blocos e iv) comitês baseados em aprendizado incremental. O primeiro grupo abrange detectores que utilizam testes estatísticos, como a Soma Cumulativa (*Cumulative Sum*) e o teste de Page-Hinckley ([PAGE, 1954](#)). Esse grupo inclui o DDM ([GAMA et al., 2014](#)), o EDDM ([BAENA-GARCÍA et al., 2006](#)) e o método de detecção de mudança de McDiarmid ([PESARANGHADER; VIKTOR; PAQUET, 2018](#)). O segundo grupo compreende métodos baseados em janelas, que, em geral, monitoram a acurácia do classificador sobre as instâncias contidas em uma janela deslizando. Os dois últimos grupos englobam métodos que monitoram a acurácia de comitês de classificadores, diferenciando-se pela forma como os dados são processados durante a reação à mudança: métodos do terceiro realizam o retreinamento dos classificadores em blocos de instâncias, enquanto os do último grupo o fazem de forma incremental, a cada nova instância observada.

[Khamassi et al. \(2018\)](#) introduzem uma taxonomia guiada pelas seguintes perguntas: "Como os dados são processados?", "Como o aprendizado é realizado?", "Como a mudança de conceito é monitorada?", "Como a mudança de conceito é tratada?" e "Quais são os critérios de desempenho?". Com relação à terceira pergunta, os métodos são agrupados de acordo com o tipo de detecção, especificamente se ela é supervisionada, não supervisionada ou semi-supervisionada. Os autores refinam ainda mais a classificação dos métodos não supervisionados em: similaridade no tempo, similaridade no espaço e medidas de complexidade do modelo.

O primeiro grupo refere-se a como a distribuição evolui entre dois instantes de tempo, sendo que, geralmente, as diferenças são detectadas por meio de testes de

hipóteses ([KUNCHEVA; ZLIOBAIT, 2009](#); [ALIPPI; ROVERI, 2008](#); [SHAKER; LUGHO-FER, 2014](#)). O segundo grupo inclui métodos que monitoram a evolução da distribuição dos dados no espaço utilizando funções de distância, como as distâncias Euclidiana, Euclidiana-Heterogênea com sobreposição e de Mahalanobis; conforme observado em [TRAN \(2019\)](#), [Tsymbal e Puuronen \(2000\)](#) e [Jr e Barros \(2013\)](#), respectivamente. O último grupo foca em mudanças na estrutura ou nos parâmetros do modelo.

Na taxonomia proposta por [Khamassi et al. \(2018\)](#), o grupo de similaridade no tempo é composto por trabalhos publicados antes do limite temporal deste trabalho (2015). De fato, conforme mencionado por [Wares, Isaacs e Elyan \(2019\)](#), há poucos estudos que consideram a dependência temporal na detecção de mudança de conceito. Como o foco deste trabalho está na análise de características mais recentes dos métodos de detecção não supervisionada de mudanças, esse critério não é considerado em nossa taxonomia. Ao observar a categoria de similaridade no espaço, é possível notar que todos os métodos discutidos neste artigo podem ser incluídos nesse grupo. Assim, a taxonomia proposta expande essa categoria de métodos. Por fim, o último grupo é dependente do algoritmo, enquanto os trabalhos descritos neste artigo podem ser gerados utilizando qualquer algoritmo de aprendizado de máquina.

Em Aprendizado de Máquina, a forma como os dados devem ser processados depende de características específicas determinadas pelo modo de acesso e pela disponibilidade dos dados. No caso de dados em fluxo, eles se diferenciam de outras formas de dados pelo fato de que as instâncias chegam de forma contínua e sequencial. Com o tempo, a distribuição subjacente dos dados pode mudar dinamicamente. Esse fenômeno é conhecido como mudança de conceito, e representa um desafio relevante que pode impactar diversos domínios de aplicação. Esse fato pode explicar o elevado número de estudos de revisão publicados na última década, conforme resumido na Figura 36, os quais se concentram em investigar abordagens para lidar com a mudança de conceito em diferentes áreas de aplicação, incluindo saúde; como predição de cirurgias ([BEYENE et al., 2015](#)) e avaliação da qualidade do sono ([CHIANG; WU, 2018](#)); segurança da informação; como filtragem de spam ([GUZELLA; CAMINHAS, 2009](#)), detecção de fraudes ([ABDALLAH; MAAROF; ZAINAL, 2016](#)) e cibersegurança ([BOBOWSKA; CHORÁS; WOŹNIAK, 2018](#)); além de sensores inteligentes baseados em dados ([KADLEC; GRBIĆ; GABRYS, 2011](#)). De modo geral, os métodos descritos nesses trabalhos não se baseiam exclusivamente em técnicas de aprendizado de máquina, uma vez que são métodos

dependentes do problema.

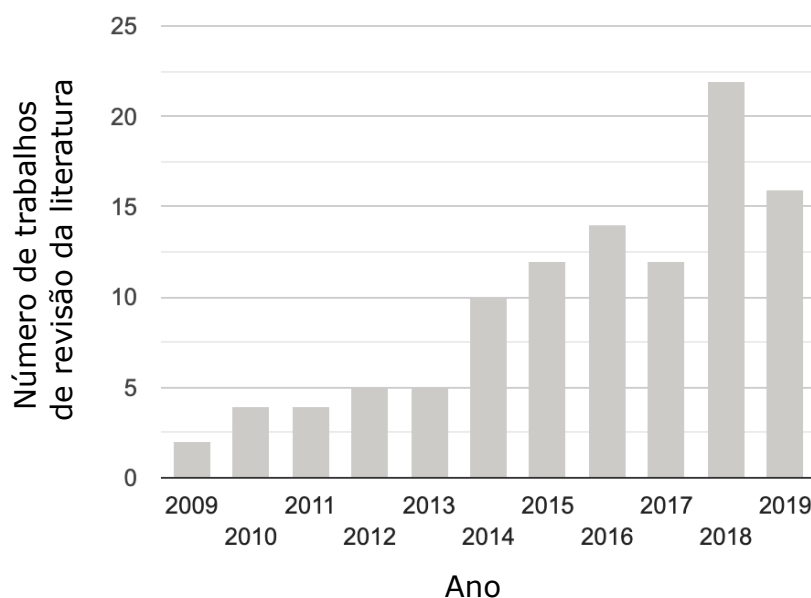


Figura 36 – O crescente número de publicações entre os anos de 2009 e 2019 em formato de revisão da literatura sobre detecção de mudança de conceito.

Estudos de revisão sobre soluções baseadas em aprendizado de máquina também têm atraído o interesse da comunidade científica (GAMA et al., 2014; LU et al., 2019; HU; KANTARDZIC; SETHI, 2019). No entanto, apesar do crescente volume de publicações sobre mudança de conceito nesse contexto, a ampla maioria dos trabalhos concentra-se em abordagens supervisionadas. De fato, diversos métodos supervisionados para detecção de mudança de conceito foram compilados em revisões como as de Lu et al. (2019) e Barros e Santos (2018). Por exemplo, Barros e Santos (2018) apresentam uma ampla comparação de 14 detectores supervisionados de mudança de conceito. Os autores indicam que abordagens ou semi-supervisionadas ou não supervisionadas poderiam ser exploradas em estudos futuros. Já Krawczyk et al. (2017) e Barros e Santos (2019) discutem a mudança de conceito em dados em fluxos com ênfase em abordagens baseadas em comitês, não sendo, portanto, voltadas principalmente para métodos não supervisionados. Iwashita e Papa (2019) destacam a expressiva concentração de publicações sobre abordagens supervisionadas, representando aproximadamente 85% dos estudos, em comparação com 12% de semi-supervisionados e apenas 5% de não supervisionados. Por fim, de acordo com Lu et al. (2019), que analisaram mais de 130 publicações sobre mudança de conceito, a detecção e adaptação de mudanças em cenários não supervisionados ou semi-supervisionados constitui uma área de pesquisa

ainda pouco explorada.

Por outro lado, ao considerar problemas do mundo real, é mais realista tentar resolver o problema da mudança de conceito de forma não supervisionada, uma vez que dispor dos rótulos das instâncias que chegam imediatamente após sua classificação, na maioria das vezes, não é viável. Até onde se tem conhecimento, não há revisões sistemáticas que descrevam o estado da arte em abordagens não supervisionadas para detecção de mudança de conceito e que proponham uma taxonomia específica para esses métodos. Esta revisão tem como objetivo preencher essa lacuna na literatura.

A presente revisão foi elaborada para responder às seguintes perguntas de pesquisa:

- P1) Quais são as abordagens não supervisionadas para detecção de mudança de conceito propostas na literatura?;
- P2) Essas abordagens foram propostas para dados em fluxo *online* ou em lotes?;
- P3) Como categorizar essas abordagens em uma taxonomia significativa?; e
- P4) Qual é a contribuição da taxonomia proposta em relação às taxonomias existentes na literatura?

Para responder às perguntas de pesquisa, foi adotada uma metodologia de revisão sistemática (OKOLI, 2015). A seleção das referências utilizadas nesta revisão foi realizada por meio de uma *string* de busca no mecanismo de meta-pesquisa Scopus (<[www.scopus.com](http://www.scopus.com)>). A primeira *string* de busca foi elaborada para recuperar estudos de revisão da literatura (*overviews*, *surveys* e *reviews*) que abordassem o fenômeno da mudança de conceito no período de 2009 a 2019, resultando em 106 documentos. Em seguida, a busca foi refinada para incluir apenas artigos publicados em inglês e disponíveis em periódicos ou anais de conferências. Esse novo filtro resultou em 78 documentos. Uma verificação preliminar deste grupo revelou uma maior concentração de publicações após 2014. Assim, com o objetivo de revisar os avanços mais recentes sobre detecção de mudança de conceito por meio de estratégias não supervisionadas, a busca foi limitada ao período de 2015 a 2019. Essa etapa retornou 55 documentos.

Como resultado, o presente trabalho descreve 16 métodos de detecção não supervisionada de mudança de conceito, conforme apresentados nas Tabelas 10 e 11.

Esses métodos são detalhados nas seções sobre os métodos baseados em lote e sobre os *online*. As principais etapas da metodologia estão descritas a seguir:

- Etapa 1: Mecanismo de busca. Ambas as buscas foram realizadas utilizando o mecanismo de meta-pesquisa Scopus ([<www.scopus.com>](http://www.scopus.com)), que realiza pesquisas nas bases de dados Science Direct ([<www.sciencedirect.com>](http://www.sciencedirect.com)), ACM Digital Library ([<dl.acm.org>](http://dl.acm.org)), IEEE Xplore ([<ieeexplore.ieee.org>](http://ieeexplore.ieee.org)) e Springer Link ([<link.springer.com>](http://link.springer.com));
- Etapa 2: Seleção dos documentos. Triagem preliminar dos artigos: A primeira busca foi baseada em palavras-chave. Os artigos foram então selecionados como referências caso: 1) apresentassem nova teoria, algoritmo ou metodologia na área de mudança de conceito; ou 2) relatassem uma aplicação relacionada à mudança de conceito;
- Etapa 3: Filtragem dos resultados para os artigos. Os artigos selecionados na Etapa 2 foram divididos em dois grupos: detecção de mudança de conceito em lote e detecção de mudança de conceito em dados em fluxo *online*. As referências de cada grupo foram novamente filtradas com base no período de publicação (2016-2020); e
- Etapa 4: Seleção dos métodos. A última etapa analisa todos os documentos para extrair apenas os artigos que propõem um componente de detecção de mudança semi-supervisionado ou não supervisionado.

A principal contribuição desta revisão é propor uma taxonomia dos estado-da-arte das abordagens para detecção de mudança de conceito baseadas em estratégias não supervisionadas, bem como apresentar uma visão abrangente das abordagens que lidam com a mudança de conceito em problemas de classificação de forma não supervisionada.

Portanto, as principais diferenças entre a taxonomia proposta e as atualmente existentes são:

- Uma taxonomia específica para métodos de detecção de mudança de conceito não supervisionada, e

- Categorização dos métodos com base em como a detecção da mudança é realizada, ao contrário de outras taxonomias que se concentram em como os classificadores são treinados ou em como reagem à mudança.

### Classificação Taxonômica

Na revisão sistemática da literatura realizada, foram identificados atributos comuns entre os métodos de detecção de mudança de conceito não supervisionados e semissupervisionados, com o objetivo de propor uma taxonomia. As categorias dessa taxonomia são apresentadas de forma resumida na Figura 37, e serão detalhadas nas subseções seguintes.

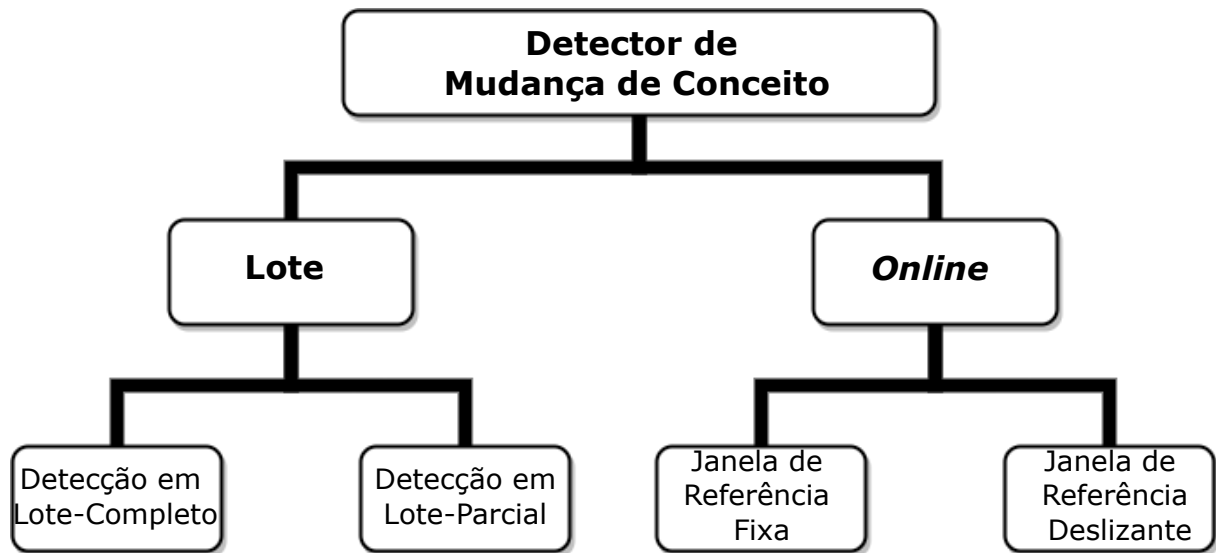


Figura 37 – Taxonomia proposta de métodos não supervisionados de detecção de mudança de conceito.

Embora a mudança de conceito não seja exclusiva de dados em fluxo, todos os métodos identificados nesta revisão lidam com as instâncias por meio de alguma forma de janela aplicada a um fluxo de instâncias. Em geral, pelo menos uma das janelas contém as instâncias consideradas como pertencentes ao conceito mais recente conhecido, que foram utilizadas para treinar ou atualizar o classificador mais recente. E, pelo menos uma janela contém as instâncias que podem ter sofrido uma mudança de conceito. Embora cada autor adote sua própria terminologia, neste trabalho essas janelas são denominadas de janela de referência e janela de detecção. Todas as categorias da taxonomia proposta podem ser explicadas com base em como essas janelas e as instâncias contidas nelas são tratadas pelos algoritmos.

No primeiro nível da categorização, os métodos são diferenciados de acordo com a forma como a janela de detecção é construída. Observa-se facilmente que os detectores de mudança não supervisionados baseiam-se ou no acúmulo de um lote de instâncias ou na verificação de mudança a cada nova instância, o que reforça os achados apresentados em [Khamassi et al. \(2018\)](#). Assim, os métodos do primeiro grupo são descritos como métodos de detecção de mudança em lote, enquanto os do segundo grupo são categorizados como métodos de detecção de mudança *online*. É importante destacar que essa classificação considera apenas como os métodos diferem no momento de *detectar* a mudança, e não como treinam, atualizam ou reprocessam seus modelos, ou realizam qualquer outra tarefa específica após a detecção da mudança.

No segundo nível da taxonomia, os detectores de mudança baseados em lote são categorizados de acordo com a forma como manipulam os dados na janela de detecção. Embora os métodos se diferenciem em vários aspectos; incluindo se o tamanho do lote é fixo ou dinâmico,, a diferença mais notável entre eles é se utilizam todas as instâncias do lote ou apenas uma amostra para detectar mudanças. Assim, a divisão no segundo nível contempla as categorias detecção com lote-completo e detecção com lote-parcial.

Por sua vez, os detectores *online* foram categorizados de acordo com a forma como manipulam a janela de referência. Nesses métodos, a janela de detecção é sempre deslizante, avançando conforme novas instâncias chegam, e refere-se ao conceito atual no fluxo de dados. Por outro lado, a janela de referência pode ser fixa ou deslizante. Portanto, esses métodos são categorizados como detectores com janela de referência fixa e detectores com janela de referência deslizante.

Não se faz distinção entre os métodos quanto à estratégia de reação à mudança de conceito. Em geral, todos treinam um novo modelo (ou comitê de modelos), ou atualizam o modelo (ou comitê), utilizando as instâncias contidas na janela de detecção no momento em que a mudança é identificada. Em seguida, a janela de detecção passa a ser a próxima janela de referência. Além disso, essas janelas podem ou não se sobrepor e, no caso de detectores *online*, pode haver um período de inatividade até que a nova janela de detecção atinja o número mínimo de amostras necessário para a detecção da mudança.

### **Métodos Baseados em Lote**

Na taxonomia proposta, os detectores baseados em lote diferenciam-se pela forma como as instâncias no lote são utilizadas. Os termos "lote" e "janela" são utilizados

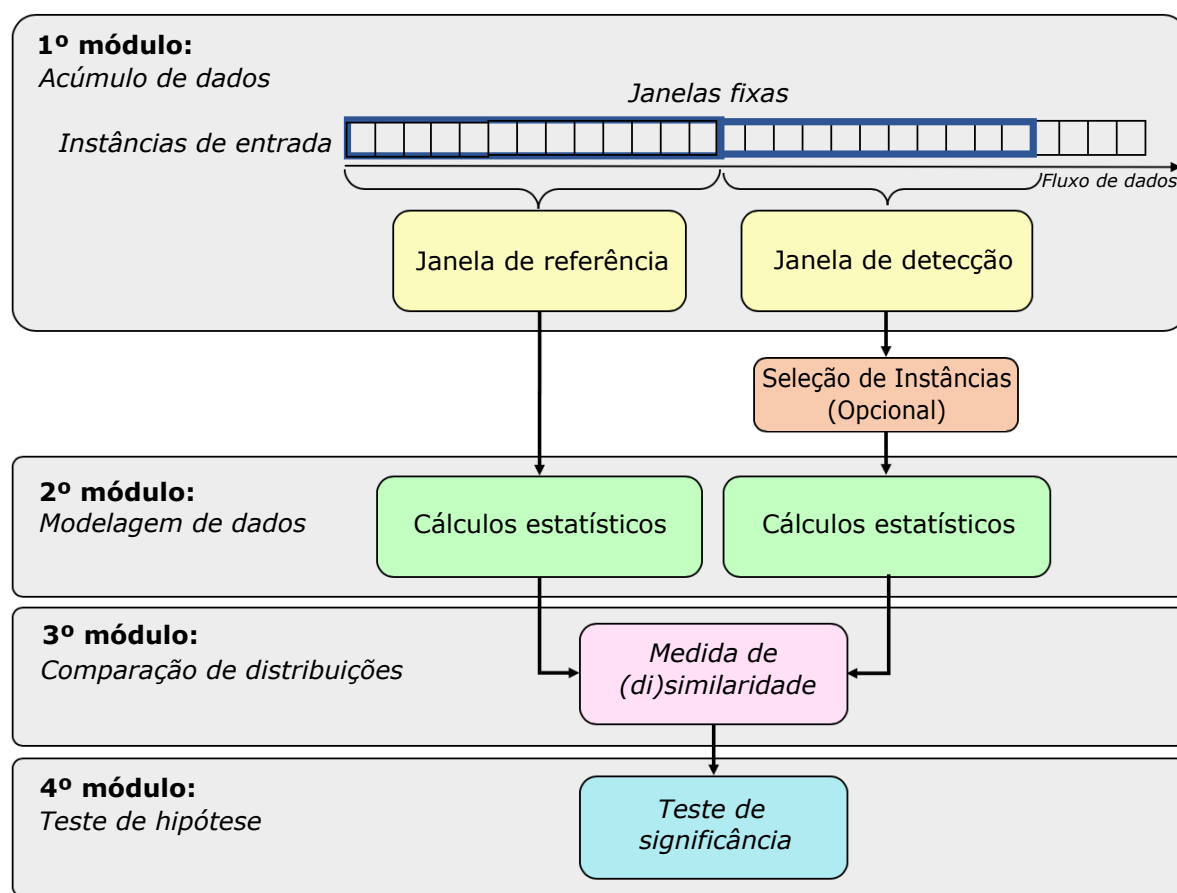


Figura 38 – Estrutura geral de métodos não supervisionados de detecção de mudança de conceito em lote.

de forma intercambiável. No primeiro grupo, denominado detecção em lote-completo, todo o lote é utilizado no processo de detecção. No segundo grupo, chamado detecção em lote-parcial, apenas um subconjunto do lote é utilizado.

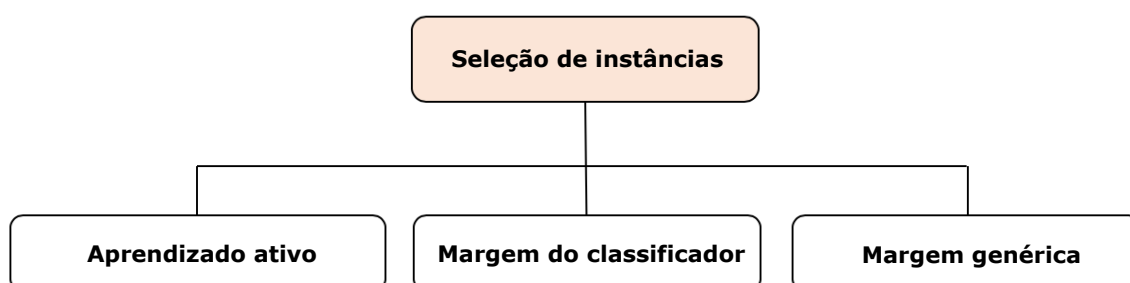


Figura 39 – Estratégias de seleção de instâncias empregadas nos métodos de detecção por lote parcial analisados.



Um arcabouço geral para métodos de detecção baseados em lote, adaptado de [Lu et al. \(2019\)](#), é apresentado na Figura 38. Esse arcabouço combina quatro módulos: acumulação de dados; modelagem dos dados; comparação de distribuições; e teste de hipótese. O primeiro módulo tem como objetivo acumular as instâncias iniciais em uma janela de referência de tamanho fixo, e as instâncias subsequentes em uma janela de detecção também de tamanho fixo, uma vez que os dados chegam de forma contínua e sequencial. Como etapa opcional, um subconjunto de instâncias da janela de detecção pode ser selecionado por meio de uma abordagem de seleção de instâncias, conforme resumido na Figura 39, para gerar uma janela de detecção menor nos métodos de detecção com lote parcial. Em seguida, o segundo módulo visa representar os dados de ambas as janelas por meio de estatísticas e outras informações, como resumos ou quantificações, a fim de expressar a distribuição dos dados. Depois, esses valores são utilizados como entrada para o terceiro módulo, que aplica medidas de similaridade ou dissimilaridade para comparar as janelas. Por fim, o último módulo utiliza testes de hipótese para avaliar limiares e decidir quando uma mudança de conceito deve ser sinalizada.

Uma lista dos trabalhos considerados para a definição da taxonomia dos métodos de detecção de mudança de conceito baseados em lote é apresentada na Tabela 10.

Tabela 10 – Categorização dos métodos de detecção de mudança de conceito baseados em lote.

Categoria	Subcategoria	Método	Referência
Lote	Detecção Lote-Completo	NN-DVI	<a href="#">Liu et al. (2018)</a>
		FAAD	<a href="#">Li et al. (2019)</a>
		UDetect	<a href="#">Bashir, Petrovski e Doolan (2017)</a>
		SQSI-IS	<a href="#">Maletzke, Reis e Batista (2018)</a>
	Detecção Lote-Parcial	MD3	<a href="#">Sethi e Kantardzic (2015)</a>
		MD3-EGM*	<a href="#">Sethi e Kantardzic (2017)</a>
		PDetect*	<a href="#">Sethi e Kantardzic (2018)</a>
		DDAL	<a href="#">Costa, Albuquerque e Santos (2018)</a>

\* Os nomes marcados com asterisco foram introduzidos neste trabalho porque os métodos não foram nomeados em seus trabalhos originais.

### Detecção em Lote-Completo

O primeiro método descrito nesta subcategoria é um dos detectores baseados em lote mais proeminentes. O algoritmo *Nearest Neighbor-based Density Variation Identification*

(NN-DVI) foi proposto por [Liu et al. \(2018\)](#) para lidar com mudanças de conceito regionais. Segundo os autores, detectores de mudança baseados em distribuição são orientados a mudanças globais e frequentemente não detectam mudanças causadas por variações na densidade regional. O NN-DVI é um método de detecção de mudança baseado em distribuição, fundamentado na estimativa de densidade regional, composto por três módulos: 1) particionamento do espaço baseado em k-NN para modelagem dos dados; 2) função de distância para acumular as discrepâncias de densidade; e 3) teste de significância estatística.

Esse método funciona comparando a janela de referência com a janela de detecção da seguinte forma. Em seu primeiro módulo, o NN-DVI concentra-se na criação de partições de dados com base na vizinhança de cada exemplo. Essa estratégia resulta em cada instância sendo agrupada com sua partícula de instância para formar uma partição. A partícula de uma instância  $x_i$  é composta por  $x_i$  e seus vizinhos. Em seguida, grupos de partículas de instância são criados levando-se em conta a vizinhança compartilhada, ou seja, cada instância de dados é representada por um conjunto de partículas de instância compartilhadas, e uma função de multiplicidade é utilizada para fornecer um multiconjunto de partículas para cada instância, de forma a representar instâncias com pesos uniformes. O objetivo desse primeiro módulo é permitir que o NN-DVI seja mais sensível a pequenas discrepâncias na distribuição dos dados. As diferenças de densidade entre as partições, calculadas considerando as partículas de instância, são utilizadas na medição de similaridade no segundo módulo do NN-DVI. Uma medida de dissimilaridade baseada em conjuntos, também chamada de medida de distância, é aplicada para quantificar as diferenças no número de partículas de instância entre pares de multiconjuntos de instâncias. Ao comparar os conjuntos amostrais, a distância entre dois conjuntos é calculada em termos de suas diferenças acumuladas no número de partículas de instância. Por fim, o último módulo envolve a aplicação de um teste estatístico de significância adaptado para indicar a ocorrência de mudança. Os autores assumem que as instâncias nas janelas de tempo são independentes e demonstram que os valores de distância dos conjuntos amostrais seguem uma distribuição normal. Como resultado, utilizam o método da máxima verossimilhança para estimar a média e a variância das distribuições de distância dos conjuntos amostrais. Em seguida, a função de distribuição acumulada é comparada com um limiar de mudança definido pelo usuário para decidir se uma mudança foi observada ou não. Um classificador *Naive*

*Bayes* é utilizado como aprendiz base. Quando uma mudança é detectada, a janela de detecção atual passa a representar o novo conceito.

O segundo é denominado de *Fast and Accurate Anomaly Detection* (FAAD), que foi proposto por Li et al. (2019) para lidar com a detecção não supervisionada de anomalias em sequências, no contexto de problemas envolvendo sequências multidimensionais em fluxos de dados sujeitos à mudança de conceito. Esse tipo de aplicação refere-se à análise das relações de ordenação em sequências com o objetivo de detectar anomalias. Para alcançar esse objetivo desafiador, o FAAD é composto por três algoritmos distintos. O primeiro é um método de seleção de atributos baseado em informação mútua e incerteza simétrica, com a finalidade de reduzir a redundância entre atributos e lidar melhor com sequências multidimensionais. Em seguida, o segundo algoritmo realiza a detecção de anomalias por meio da construção de modelos utilizando amostragem aleatória de atributos e cálculo de escores de anomalia com esses modelos. Os escores de anomalia são comparados a um limiar definido pelo usuário para determinar se uma sequência é considerada anômala. Por fim, o algoritmo *Anomaly Buffer based on Model Dynamic Adjustment* (ABMDA) é proposto para detectar mudanças de conceito. Esse terceiro componente é detalhado a seguir.

A principal questão no FAAD é distinguir entre anomalias verdadeiras e sequências normais que são incorretamente detectadas como anomalias devido à mudança de conceito. Consequentemente, o algoritmo ABMDA é dividido em duas etapas de detecção de anomalias. Somente as anomalias identificadas em ambas as etapas são consideradas como verdadeiras anomalias. A segunda etapa é executada utilizando o segundo componente do FAAD, enquanto o detector de mudança é utilizado na primeira etapa. Esse detector opera com base em um escore obtido a partir de duas considerações: 1) distribuição das anomalias; e 2) estatísticas utilizadas para expressar a distribuição dos dados. Em ambos os procedimentos de modelagem de dados, o ABMDA compara a janela de referência com a janela de detecção. O primeiro procedimento é projetado para calcular a proporção de anomalias na janela de detecção, enquanto se assume que essa proporção na janela de referência segue uma distribuição normal; a média e a variância das anomalias de referência são calculadas previamente. Por fim, calcula-se a diferença entre ambas as proporções. Parte-se do pressuposto de que a proporção de anomalias é baixa quando o fluxo de dados está estável. Assim, quanto maior for o desvio da proporção de anomalias em cada período de tempo em relação à proporção

de anomalias de referência, maior será a probabilidade de que anomalias falsas estejam sendo detectadas devido à mudança de conceito.

No segundo procedimento, o algoritmo compara duas matrizes de frequência de palavras  $M$ : uma proveniente dos dados de referência e outra dos dados de detecção. Cada matriz  $M$  é calculada da seguinte forma: dado um atributo  $x_i$ , calcula-se a frequência de ocorrência do valor  $v_j$  para  $x_i$  ao longo de seu respectivo lote, armazenando esse valor em  $M_{i,j}$ . A diferença entre as duas matrizes está correlacionada com uma maior probabilidade de que tenha ocorrido uma mudança de conceito. Em seguida, é calculada uma soma ponderada entre a diferença de proporção de anomalias mencionada anteriormente e a 1-norma da diferença entre as matrizes de frequência. Se essa soma for superior a um limite definido pelo usuário, então as anomalias atuais são armazenadas em um *buffer* de anomalias. Quando esse *buffer* atinge seu tamanho máximo, um novo modelo é gerado e adicionado ao conjunto de modelos iniciais; os quais estão sujeitos a atualizações de pesos e à remoção de modelos para manter o número total sob controle. Posteriormente, o segundo algoritmo do FAAD é utilizado para detectar sequências anômalas no *buffer*. Quando as sequências são classificadas como anomalias nesse segundo processo de detecção, elas são consideradas anomalias verdadeiras. Por fim, a distribuição das anomalias detectadas e a matriz de frequência são utilizadas para atualizar os dados de referência. É importante destacar que o FAAD é uma abordagem dependente da aplicação, cujo foco está voltado para a detecção de anomalias nos dados, e não especificamente para a detecção de mudanças de conceito.

No terceiro método, Bashir, Petrovski e Doolan (2017) propuseram uma estrutura em dois níveis baseada em distribuição, chamada *Unsupervised Change Detection for Activity Recognition* (UDetect). Essa estrutura baseia-se na premissa de que uma mudança de conceito é indicada por uma alta variância nas instâncias classificadas com um determinado rótulo, em relação às instâncias de treinamento desse mesmo rótulo. No primeiro nível, o UDetect treina um classificador e coleta dados para caracterizar cada classe. O segundo nível é não supervisionado e coleta lotes de dados para cada classe, verificando se as instâncias dessas classes diferem significativamente dos dados previamente coletados.

Mais especificamente, durante o treinamento, os dados são divididos em blocos de tamanho fixo, de forma que nenhum bloco contenha instâncias pertencentes a mais de uma classe. Cada bloco é resumido em um único parâmetro, que é a distância média

de todas as suas instâncias até o seu centróide. Em seguida, os parâmetros de cada classe são utilizados para calcular limiares por meio de três heurísticas diferentes. A estrutura não impõe nenhuma restrição ao classificador, e a totalidade dos dados pode ser utilizada para treinar um modelo. O resultado dessa fase é um classificador e um conjunto de limiares para cada classe.

A segunda fase está relacionada ao uso do modelo com o lote de dados que chega. Para cada classe conhecida na primeira fase, a estrutura designa um *buffer* com capacidade idêntica ao tamanho dos blocos da fase anterior. O objetivo desse *buffer* é acumular um lote de instâncias que foram atribuídas a um rótulo específico. Quando uma quantidade suficiente de dados é coletada para essa classe, o bloco é resumido utilizando as mesmas heurísticas empregadas na primeira fase. O parâmetro resultante é comparado aos limiares utilizando gráficos de Shewhart (TOBIAS, 2001), e picos são identificados para sinalizar que ocorreu uma mudança de conceito. A estrutura proposta foi avaliada com dois conjuntos de dados de reconhecimento de atividades, nos quais os participantes realizavam atividades como caminhar, sentar, ficar em pé, e etc. Apesar do nome, o artigo não fornece uma indicação clara de que o método seja específico para reconhecimento de atividades.

Por fim, o último método Maletzke, Reis e Batista (2018) é uma versão aprimorada do *Stream Quantification by Score Inspection* (SQSI) (MALETZKE; REIS; BATISTA, 2017), uma metodologia de quantificação de dados em fluxos que emprega um detector de mudança de conceito baseado em distribuição. O problema da quantificação está intimamente relacionado à classificação, na medida em que a tarefa do quantificador é estimar o número de instâncias de cada classe em uma amostra. De fato, uma abordagem comum, conhecida como classificar e depois contar (*classify-then-count*), consiste em primeiro classificar um conjunto de instâncias e, em seguida, contar os rótulos.

O SQSI detecta mudança de conceito por meio de um teste em duas etapas, com o objetivo de verificar se a distribuição dos erros do classificador foi alterada (MALETZKE; REIS; BATISTA, 2017). À medida que novas instâncias chegam, elas são agrupadas em uma janela de detecção com tamanho definido pelo usuário. Quando essa janela está completa, um classificador fornece scores para cada instância de forma não supervisionada, os quais podem ser quaisquer estatísticas que representem a confiança do classificador naquela instância; como, por exemplo, a distância da instância até o hiperplano em uma Máquina de Vetores de Suporte (MVS). Em seguida, um teste

estatístico é aplicado para verificar se os scores atuais provêm da mesma distribuição dos scores estimados a partir dos dados de referência; que, inicialmente, correspondem aos dados de treinamento. Caso a hipótese nula seja rejeitada, uma transformação linear é aplicada aos dados de referência, de modo que as instâncias da janela tenham a mesma média e desvio padrão dos dados de detecção (o conceito atual). Novos scores são então calculados e o teste estatístico é reaplicado. Se a hipótese nula ainda for rejeitada, o SQSI sinaliza a ocorrência de uma mudança de conceito. Após detectar a mudança, o SQSI solicita os rótulos verdadeiros para todas as instâncias presentes na janela de detecção. No entanto, se algum dos testes não rejeitar a hipótese nula, uma nova janela de detecção é iniciada, sem sobreposição de instâncias com a anterior; ou seja, um novo lote.

O SQSI com Seleção de Instâncias (SQSI-IS) (MALETZKE; REIS; BATISTA, 2018) aprimora a versão inicial do SQSI (MALETZKE; REIS; BATISTA, 2018). Tanto o SQSI quanto o SQSI-IS realizam a verificação em duas etapas mencionada anteriormente e treinam um novo classificador após a detecção de uma mudança de conceito, utilizando a janela de detecção como dados de referência para identificar a próxima mudança. No entanto, enquanto o SQSI solicita os rótulos verdadeiros de todas as instâncias, o SQSI-IS emprega técnicas de seleção de instâncias e auto-aprendizado para requisitar os rótulos verdadeiros de apenas uma fração das instâncias no lote.

O teste utilizado em ambos os algoritmos foi o Kolmogorov-Smirnov (KS), que é um teste não paramétrico utilizado para identificar se duas amostras provêm da mesma distribuição (LOVRIC, 2011). Como o teste KS é muito sensível a mudanças na proporção das classes (MALETZKE et al., 2018), os autores sugerem o uso de um nível de confiança bastante rigoroso (no máximo 0,001), caso contrário, uma quantidade significativa de falsos alarmes tende a ser gerada. Para problemas em que a proporção entre as classes não muda de forma tão significativa; por exemplo, em classificação; um teste ou nível de confiança diferente pode ser empregado.

### **Detecção em Lote-Parcial**

Nesta subcategoria, os detectores de mudança envolvem a seleção de um subconjunto de instâncias da janela de detecção para verificar se uma mudança de conceito ocorreu. Em geral, esse módulo pode ser executado utilizando qualquer estratégia de seleção de instâncias, variando desde a margem do MVS (SETHI; KANTARDZIC, 2015) até Aprendizado Ativo (COSTA; ALBUQUERQUE; SANTOS, 2018). Todos os

métodos discutidos nesta seção empregam uma das estratégias de seleção de instâncias resumidas na Figura 39.

Em [Sethi e Kantardzic \(2015\)](#), o principal objetivo do método proposto, *Margin Density Drift Detection* (MD3), é monitorar mudanças na margem do classificador; região do espaço de decisão onde as predições são altamente incertas. Parte-se do pressuposto de que variações na densidade acima de um limiar definido pelo usuário indicam a ocorrência de uma mudança de conceito. O número de instâncias localizadas na região delimitada pela margem é utilizado para calcular a densidade da margem. Esta primeira versão do MD3 é um detector de mudança não supervisionado que emprega MVS para determinar a margem do classificador. Um classificador MVS é gerado utilizando o conjunto de dados de treinamento inicial. Em seguida, duas medidas de referência são obtidas: a densidade máxima e a densidade mínima histórica. A densidade é então calculada para o novo lote de dados que chega. Quando o número de amostras não rotuladas dentro da margem aumenta, isso indica uma mudança gradual. Por outro lado, quando esse número diminui, sinaliza-se uma mudança repentina. O MD3 utiliza uma janela deslizante para processar o fluxo de dados. Após a detecção de uma mudança, o método utiliza dados rotulados para substituir o classificador MVS e também para atualizar os valores de densidade de referência. O principal problema desse método é que ele funciona apenas com o MVS como algoritmo de aprendizado.

Essa limitação motivou [Sethi e Kantardzic \(2017\)](#) a proporem uma versão aprimorada do MD3 para contornar a desvantagem mencionada da primeira versão. Nesta nova versão, denominada neste levantamento como MD3 Ensemble Generic Margin (MD3-EGM), utiliza-se um ensemble de classificadores para produzir uma margem genérica, definida como o conjunto de instâncias desconhecidas com maior incerteza. A incerteza é medida pelo grau de discordância entre os classificadores do conjunto ao atribuírem um rótulo a cada instância. O ensemble é construído por meio da abordagem de subespaços aleatórios (*random subspace*), ou seja, selecionam-se aleatoriamente  $n$  subespaços do espaço original de atributos para treinar  $n$  classificadores diferentes. A detecção de mudança é realizada monitorando a diversidade entre os classificadores, de modo que se espera a ocorrência de uma mudança sempre que houver aumento da discordância do ensemble ao classificar amostras desconhecidas contidas na janela de detecção. No entanto, como essa nova versão do MD3 pode gerar altas taxas de falso positivo, foi adicionado ao método um módulo de confirmação supervisionada da mudança. Assim,



este método é considerado um detector de mudança semi-supervisionado.

O monitoramento da discordância é realizado da seguinte forma. Quatro medidas de referência são obtidas a partir da janela de referência utilizando um procedimento de validação cruzada *10-fold* para treinamento e teste: discordância e desempenho preditivo, juntamente com seus respectivos desvios padrão. Em seguida, a discordância entre os classificadores do comitê é calculada com base nas instâncias da janela de detecção. Considerando as instâncias situadas na margem, o detector de mudança é dividido em dois níveis principais: indicadores de mudança não supervisionados e supervisionados. Ambos os níveis utilizam um limiar ponderado com base no desvio padrão das medidas de referência. No primeiro nível de detecção, a discordância acumulada atual é comparada à discordância de referência. Se a diferença entre esses valores for superior ao limiar estabelecido, o segundo indicador de mudança é acionado. Esse segundo nível funciona selecionando um grupo de instâncias do lote atual para serem rotuladas e incluídas em um conjunto de dados, que será utilizado para calcular o desempenho preditivo atual. Esse desempenho é então comparado ao desempenho de referência. Assume-se que ocorreu uma mudança de conceito quando a diferença entre os valores excede o limiar. A seleção das instâncias consiste em rotular uma quantidade fixa de instâncias que chegam após a primeira indicação de mudança. Por fim, o comitê de classificadores é retreinado utilizando o conjunto de dados criado após a confirmação da mudança no segundo nível. Vale ressaltar que esse método pode ser aplicado a classificadores que não possuam margens explícitas, porém exige uma confirmação supervisionada da mudança para tomar uma decisão final.

A estrutura de classificação em fluxo *Predict-Detect*, proposto por [Sethi e Kantardzic \(2018\)](#), e que será referido nesta revisão como PDetect, foi desenvolvido para lidar com uma aplicação específica: ataques adversariais. Parte-se da premissa de que tais ataques levam a alterações na distribuição dos dados e à redução da capacidade preditiva dos modelos. Quando esse tipo de problema é tratado no contexto de mineração de dados em fluxo, essas alterações são denominadas mudanças adversariais de conceito. Segundo os autores, propor um método dependente da aplicação é essencial nesse tipo de problema, pois a natureza independente de aplicação dos detectores de mudança baseados em distribuição pode representar uma desvantagem frente à mudanças adversariais. O problema surge porque esses detectores podem ser vulneráveis à manipulação adversarial em tempo de teste. Para superar essa limitação, os autores propõem um



método de detecção de mudança consciente da presença de adversários, desenvolvido como uma estrutura baseada em comitês e independente do tipo de classificador.

A estrutura proposta divide o espaço de atributos do conjunto de dados de treinamento em dois subconjuntos disjuntos para treinar dois modelos de classificação: (1) predict (predizer) e (2) detect (detectar). A divisão do espaço de atributos é realizada classificando todas as variáveis disponíveis de acordo com seu valor-F obtido via ANOVA, distribuindo-as entre os dois classificadores por meio de uma estratégia de *round robin*, com o objetivo de obter classificadores com alto poder preditivo. O classificador de previsão é responsável por atribuir rótulos às instâncias desconhecidas da janela de detecção, enquanto o classificador de detecção é encarregado de indicar atividade adversarial. O primeiro tende a ser alvo de ataques em algum momento. Já o segundo é considerado invulnerável à manipulação adversarial em tempo de teste, uma vez que é um componente oculto no processo de predição. Esses dois classificadores compõem um esquema de auto monitoramento. Quando o desacordo entre suas predições diminui significativamente ao longo do tempo, isso pode indicar uma possível mudança de conceito.

O detector de mudança funciona conforme descrito em [Sethi e Kantardzic \(2017\)](#). Aqui, no entanto, apenas dois classificadores são considerados. Consequentemente, este método emprega um detector de mudança semissupervisionado. Uma vez detectada a mudança, os dois classificadores são retreinados utilizando o conjunto de dados criado após a primeira indicação de mudança. Em relação à partição dos atributos, a divisão pode ser mantida ou uma nova pode ser gerada.

[Costa, Albuquerque e Santos \(2018\)](#) também propuseram um método de detecção de mudança baseado na hipótese de que a variação de densidade das instâncias mais significativas pode indicar uma mudança. No trabalho, o procedimento de seleção de instâncias é baseado em Aprendizado Ativo. O método proposto, denominado *Drift Detection Method Based on Active Learning* (DDAL), é dividido em duas fases. A primeira fase envolve a geração de um classificador utilizando instâncias contidas na janela de referência (primeiro lote de dados), enquanto a segunda é subdividida em três módulos: detecção de mudança, reação e classificação. Na segunda fase, o DDAL monitora a ocorrência de mudança de conceito para cada novo lote de dados não rotulados e, dependendo do resultado da detecção de mudança, também executa os módulos de reação e classificação.

O módulo de detecção funciona calculando a densidade das instâncias mais significativas selecionadas da janela de detecção, utilizando a estratégia de Incerteza Fixa (ŽLIOBAITĖ, 2014), da seguinte forma: margens virtuais são criadas como projeções de hiperplanos, definidos com base em um limiar de incerteza determinado pelo usuário, que são equidistantes ao hiperplano separador. Para cada instância na janela de detecção, calcula-se a probabilidade posterior máxima fornecida pelo classificador e a compara-se ao limiar de incerteza. Todas as instâncias cujo valor de confiança for inferior ao limiar são consideradas dentro do subespaço delimitado pelas margens virtuais e, conseqüentemente, são selecionadas para o cálculo da densidade. O valor da densidade é obtido dividindo-se o número de instâncias que caem dentro do subespaço delimitado pelas margens virtuais pelo número total de instâncias na janela de detecção. Em seguida, esse valor de densidade é comparado com os valores máximos e mínimos de densidade de referência. Os novos valores substituem os valores de referência quando são superiores ao máximo de referência ou inferiores ao mínimo de referência. Por fim, se a diferença entre esses dois valores for maior que um limiar de mudança definido pelo usuário, então uma mudança de conceito é sinalizada e o módulo de reação é ativado. No módulo de reação, um novo classificador é gerado utilizando os dados da janela de detecção em substituição ao classificador atual. Somente nesse ponto as verdadeiras classes das instâncias são requeridas, uma vez que um novo conjunto de treinamento deve ser fornecido. Por outro lado, se nenhuma mudança for detectada, o DDAL avança para o seu módulo de classificação, no qual as amostras da janela de detecção são classificadas. Em seguida, o DDAL retorna ao módulo de detecção para monitorar o próximo lote de instâncias não rotuladas.

### **Métodos Online**

Conforme discutido anteriormente, os métodos baseados em lotes funcionam acumulando instâncias em um lote de dados e, quando o lote está completo, um ou mais algoritmos são acionados para detectar se ocorreu uma mudança de conceito desde que o último classificador foi treinado. Em contraste, os métodos baseados em fluxo contínuo (*online*) verificam a ocorrência de mudança a cada nova instância que chega; possivelmente após um período de inicialização para acumular instâncias antes que a primeira verificação possa ser realizada. Todos os métodos *online* analisados neste estudo empregam uma janela deslizante de detecção. No entanto, alguns deles diferem pelo fato de a janela de referência ser fixa sobre as instâncias de treinamento, enquanto

outros fazem a janela de referência deslizar para incluir as amostras não rotuladas que chegam após a mudança; possivelmente sobrepondo-se à janela de detecção. Portanto, a taxonomia proposta refina a categoria *online* em métodos que utilizam janelas de referência fixas e deslizantes. Os trabalhos analisados em cada categoria são discutidos nas próximas subseções, e [Kim e Park \(2017\)](#) é tratado na seção sobre múltiplas abordagens, uma vez que o método proposto pelos autores possui variantes tanto para janelas de referência fixas quanto deslizantes.

Uma estrutura geral de métodos não supervisionados de detecção de mudança de conceito *online* é apresentado na Figura 40. A estrutura é dividida em três módulos: recuperação de dados; medidas de (dis)similaridade; e teste de significância. O primeiro módulo é responsável por receber as instâncias de dados que chegam, a fim de gerar a janela de referência, que, como mencionado anteriormente, pode ser fixa ou deslizante. Este primeiro módulo também é responsável por manipular a janela de detecção, que é sempre uma janela deslizante. O segundo e o terceiro módulos são semelhantes, respectivamente, ao terceiro e ao quarto módulos da estrutura dos métodos baseados em lotes. Uma lista dos trabalhos considerados para a definição da taxonomia de métodos online de detecção de mudança de conceito é apresentada na Tabela 11.

Tabela 11 – Categorização dos métodos de detecção de mudança de conceito online.

Categoria	Subcategoria	Método	Referência
<i>Online</i>	Janela de Referência Fixa	IKS-bdd*	<a href="#">Reis et al. (2016)</a>
		CD-TDS	<a href="#">Koh (2016)</a>
	Janela de Referência Deslizante	OMV-PHT*	<a href="#">Lughofer et al. (2016)</a>
		NM-DDM*	<a href="#">Mustafa et al. (2017)</a>
		Plover*	<a href="#">Mello et al. (2019)</a>
		SAND	<a href="#">Haque, Khan e Baron (2016)</a>
		DSDD	<a href="#">Pinagé, Santos e Gama (2020)</a>
	Múltiplas Abordagens	DbDDA*	<a href="#">Kim e Park (2017)</a>

\* Os nomes marcados com asterisco foram introduzidos neste trabalho porque os métodos não foram nomeados em seus trabalhos originais.

### Janela de Referência Fixa

O teste Incremental de Kolmogorov-Smirnov é uma variante *online* do teste KS proposta em [Reis et al. \(2016\)](#). Ele utiliza uma estrutura de dados *treap* para inserir e remover observações em um conjunto dinâmico, permitindo atualizar o valor de  $p$  do

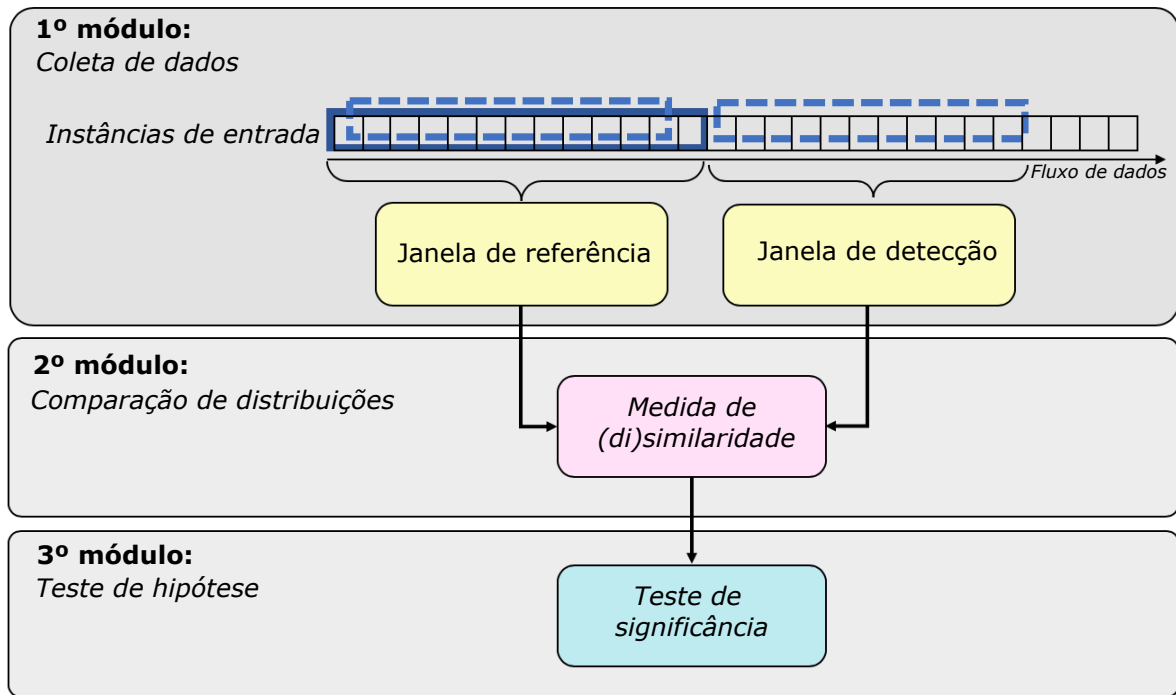


Figura 40 – Estrutura geral de métodos não supervisionados de detecção de mudança de conceito *online*.

teste KS sem a necessidade de revisitar lotes inteiros de dados. A vantagem do IKS é que ele se comporta exatamente como o teste KS tradicional, mas com complexidade temporal total muito menor quando aplicado repetidas vezes à medida que uma amostra é atualizada com novos elementos.

Reis et al. (2016) também propuseram um método *online* não supervisionado de detecção de mudança utilizando o IKS, *IKS-based Drift Detector* (IKS-bdd). O princípio por trás do IKS-bdd é aplicar o teste IKS individualmente a cada atributo. Isso elimina a necessidade de utilizar testes multivariados ou funções de mapeamento, reduzindo assim a quantidade de dados necessária. A detecção de uma mudança em um único atributo pode ser suficiente para acionar a detecção de mudança de conceito. A desvantagem dessa abordagem é que algumas mudanças podem não ser detectáveis quando as variáveis são analisadas individualmente, embora um detector multivariado possa funcionar para os mesmos dados.

Para cada atributo, o IKS-bdd mantém duas janelas de tamanho  $W$ . A janela de referência é fixa e contém os dados utilizados para treinar o modelo mais recente. A janela de detecção é deslizante e contém dados provenientes do fluxo. À medida

que cada instância chega, ela é classificada e o teste IKS é realizado, comparando a distribuição dos dados na janela de referência com a da janela de detecção. Se a hipótese nula de que as distribuições são idênticas for rejeitada, uma mudança é sinalizada. Caso contrário, a janela é deslocada e uma nova instância é retirada do fluxo.

Os autores propõem três reações à mudança. A primeira é a Substituição de Modelo (*Model Replacement* – MR), que consiste em solicitar os rótulos verdadeiros para todas as instâncias na janela de detecção e, em seguida, treinar um novo classificador. A segunda abordagem é a transformação AB, que envolve: (1) aplicar uma transformação linear aos dados de referência, de modo que os atributos envolvidos na mudança passem a ter média e desvio padrão idênticos aos das instâncias nos dados de detecção; e (2) realizar novamente o teste KS, recorrendo à MR caso a hipótese nula ainda seja rejeitada. A terceira abordagem é uma árvore de decisão modificada: quando uma mudança é detectada, os rótulos verdadeiros são solicitados apenas para as instâncias que alcançam folhas com decisão baseada no atributo que desencadeou a mudança.

Koh (2016) propuseram um método não supervisionado de detecção de mudança de conceito especificamente para fluxos de dados transacionais; problemas que envolvem representar interações entre entidades ou itens. O método é denominado *Change Detection in Transactional Data Stream for Frequent Pattern Mining* (CD-TDS) e pode ser dividido em duas partes: i) Detecção Local de Mudança; e ii) Detecção Global de Mudança. Quando há mudanças em alguns itens, mas não são observadas alterações nas conexões entre eles, esse comportamento pode indicar uma mudança local, mais precisamente uma mudança de conceito virtual. No caso da detecção global de mudança, as conexões entre os itens precisam mudar para indicar uma mudança devido à geração de um novo conceito ou ao desaparecimento de um conceito antigo.

O procedimento principal da detecção local da mudança envolve a comparação de duas janelas  $W_0$  e  $W_1$  de um fluxo  $S$  utilizando o limite de *Hoeffding*. A diferença entre as médias amostrais das duas janelas é comparada a um valor  $\epsilon$  definido pelo limite de *Hoeffding*. Uma mudança é sinalizada quando essa diferença é maior que  $\epsilon$ . É importante destacar que esse processo de monitoramento de mudança local é semelhante ao utilizado no ADWIN (BIFET; GAVALDÀ, 2007).

Para detectar a mudança global, é empregada uma estrutura em árvore para representar as conexões entre itens em  $S$ . Dessa forma, as duas janelas  $S_0$  e  $S_1$  são representadas por grafos. Para determinar se as duas janelas apresentam ou não alta

dissimilaridade, é utilizado um teste estatístico pareado para avaliar a discordância entre as duas estruturas de árvore. Diversas métricas de discordância reportadas na literatura podem ser empregadas para medir essa diversidade, como distância de edição, maior subárvore comum e menor super árvore comum. Koh (2016) trabalharam com a distância de edição de Levenshtein (GILLELAND, 2020).

### Janela de Referência Deslizante

Lughofer et al. (2016) propuseram duas técnicas para lidar com mudanças de conceito com poucas instâncias rotuladas e até mesmo sem instâncias rotuladas: a primeira técnica emprega filtros de aprendizado ativo de passagem única para selecionar as amostras mais relevantes para classificação supervisionada, e a segunda considera a sobreposição entre duas classes na distribuição da saída de um classificador. Em ambos os casos, é executada uma versão *online* modificada do teste de Page-Hinkley (MOUSS et al., 2004) para detectar mudanças; denominada de Versão Online Modificada do Teste de Page-Hinkley (OMV-PHT). O PHT é utilizado para detectar mudanças repentinas em um sinal Gaussiano  $x$ , empregando a diferença entre  $x(t)$  e a média anterior  $\bar{x}_{t-1}$ .

Esse sinal corresponde ao indicador de desempenho do classificador. No cenário semissupervisionado, o aprendizado ativo é utilizado para selecionar as instâncias mais significativas a fim de estimar o erro do classificador. A técnica não supervisionada para detecção de mudança baseia-se na ideia proposta por Donmez, Lebanon e Balasubramanian (2010) para estimar erros de classificação sem rótulos. Assume-se que as estimativas de incerteza das duas classes mais prováveis correspondem a uma distribuição bimodal e que essa distribuição pode ser considerada como duas distribuições Gaussianas distintas. A sobreposição dessas distribuições pode ser utilizada como uma estimativa do erro. Assim, o grau de variação dessa sobreposição é usado como sinal para o PHT.

Mustafa et al. (2017) propuseram um detector de novas classes, *Non-parametric Multidimensional Drift Detection Method* (NM-DDM), que é baseado em *denoising autoencoders* e em um método não paramétrico multidimensional de detecção de mudança baseado em logaritmo da verossimilhança de caminho aleatório. O interesse recai sobre o primeiro. Esse método de detecção de mudança baseia-se no cálculo da razão de verossimilhança entre dados de duas janelas: uma anterior ao ponto potencial de mudança e a outra posterior. A verossimilhança é estimada para cada dimensão (atributo) e, se a maior razão entre as duas janelas for superior a um limiar, uma mudança é sinalizada.

Após uma mudança ser identificada, as amostras da última janela são utilizadas para treinar um novo classificador. As instâncias com valores de confiança baixos são rotuladas por um componente externo, enquanto as amostras com valores de confiança altos são rotuladas usando o rótulo previsto.

Plover ([MELLO et al., 2019](#)) é um algoritmo de prova de conceito. Os autores realizaram uma análise dos fundamentos teóricos e das garantias de aprendizado na detecção não supervisionada de mudança de conceito. Eles concluem que, enquanto os métodos supervisionados são amparados pela teoria do aprendizado estatístico, as abordagens não supervisionadas dependem de métricas internas ou externas que não contam com garantias semelhantes. Métricas internas, como a compacidade de agrupamento e a distância entre centróides, são calculadas a partir da estrutura dos dados, enquanto métricas externas dependem de rótulos supervisionados; os quais não podem ser realisticamente considerados disponíveis em tarefas verdadeiramente não supervisionadas.

O conceito de Estabilidade de Algoritmo ([BOUSQUET; ELISSEEFF, 2002](#)) estabelece condições para que qualquer função arbitrária convirja para o seu valor esperado. O Plover é um algoritmo de detecção de mudança de conceito uniformemente estável derivado desse conceito. De forma semelhante a outros métodos *online* discutidos nesta seção, ele calcula a divergência entre uma janela deslizante de detecção e uma janela de referência e, se essa diferença for maior que um limiar, uma mudança é sinalizada. Segundo os autores, duas premissas são necessárias para empregar a teoria proposta: i) a função deve ser escolhida de forma independente dos dados de entrada; e ii) os dados de entrada devem ser independentes e identicamente distribuídos. Tal função deve ser adequada ao problema em questão; por exemplo, média, variância, potência espectral, etc.

[Haque, Khan e Baron \(2016\)](#) propuseram um método para detecção de mudança de conceito e de novas classes em dados em fluxos denominado *Semi-supervised Adaptive Novel Class Detection and Classification over Data Stream* (SAND). O SAND utiliza um conjunto de classificadores de k-vizinhos mais próximos e é dividido em dois módulos: detecção de *outliers* e detecção de mudança de conceito. Além disso, o método também é voltado para a autoanotação de dados. À medida que o sistema observa um número crescente de *outliers*, um algoritmo baseado em agrupamento verifica se esses *outliers* são instâncias de uma nova classe, a fim de treinar posteriormente o conjunto com



amostras dessa nova classe.

A cada novo exemplo desconhecido  $x_i$ , o SAND prevê seu rótulo por meio de votação majoritária e estima a confiança de cada classificador utilizando duas heurísticas: associação e pureza. A heurística de associação estima a confiança de um classificador medindo a distância entre  $x_i$  e seu pseudoponto mais próximo; definido pelos autores como uma estrutura que contém o centróide, o raio e o número de pontos de dados pertencentes a cada classe. A segunda heurística concentra-se principalmente em monitorar a classe mais frequente no pseudoponto associado a  $x_i$ . Em seguida, uma janela de tamanho variável armazena  $x_i$ , sua previsão de classe e os escores de confiança.

Após a instância ser classificada, o SAND monitora os escores de confiança para detectar mudanças. O processo de detecção é realizado assumindo que os valores de confiança seguem uma distribuição *beta*. Os valores armazenados na janela de tamanho variável são usados para estimar os parâmetros dessa distribuição, e o resultado é comparado a um limiar. Se uma mudança for detectada, o SAND solicita apenas os rótulos verdadeiros para as instâncias com escores de confiança baixos. Essas instâncias são incorporadas ao novo conjunto de treinamento, juntamente com aquelas autoanotadas, e o conjunto é então atualizado utilizando o novo conjunto de treinamento.

Em linha com a ideia desenvolvida em [Haque, Khan e Baron \(2016\)](#); [Pinagé, Santos e Gama \(2020\)](#) propuseram um método não supervisionado de detecção de mudança que também permite autoanotação. Esse método, denominado de *Dynamic Selection Drift Detector* (DSDD), realiza a seleção dinâmica de classificadores a partir de um conjunto inicial de classificadores e utiliza um detector de mudança auxiliar configurável, guiado por uma taxa de pseudoerro como métrica para detectar mudanças. Ele é dividido em três módulos: 1) criação do conjunto; 2) seleção dinâmica; e 3) detecção. O primeiro módulo funciona particionando um lote inicial de instâncias rotuladas; composto pelas primeiras  $m$  instâncias recebidas; em conjuntos de dados de treinamento e validação.

O primeiro conjunto de dados é utilizado para gerar um conjunto de classificadores por meio de uma versão modificada do algoritmo *Online Bagging* ([MINKU; WHITE; YAO, 2010](#)), ou seja, cada classificador membro é treinado com  $n$  cópias de cada instância, onde  $n$  é definido pela distribuição de *Poisson*. Os autores destacam que valores baixos de  $n$  proporcionam alta diversidade entre os membros do conjunto. Eles definem valores suficientemente altos para resultar em diversidade significativa, de



modo a gerar diferentes classificadores membros. As instâncias do segundo conjunto de dados são usadas como regiões de competência de referência no módulo de seleção dinâmica, que parte do pressuposto de que cada classificador membro é um especialista em uma região de competência. Esse segundo módulo funciona da seguinte forma: quando uma instância desconhecida  $x_i$  chega, define-se uma região em termos de sua vizinhança no conjunto de validação. Em seguida, calcula-se a competência de cada classificador membro. O classificador com maior competência é escolhido para atribuir um rótulo a  $x_i$ . Essa decisão é considerada o pseudo-rótulo verdadeiro de  $x_i$ .

Por fim, o módulo de detecção aplica um detector de mudança para cada classificador membro. Como a proposta não é utilizar os rótulos verdadeiros das instâncias, a questão central nesse método é comparar a previsão fornecida por cada classificador membro para  $x_i$  com seu pseudo-rótulo verdadeiro, em um processo de monitoramento de pseudoerro. Esse pseudoerro pode ser usado como métrica em diversos detectores de mudança supervisionados disponíveis na literatura. Os autores utilizaram tanto o DDM (GAMA et al., 2004) quanto o EDDM (BAENA-GARCÍA et al., 2006). Quando um número fixo de classificadores membros indica uma mudança, este é sinalizado e o conjunto de validação é atualizado com novas amostras rotuladas. Caso contrário, o aprendizado online do conjunto de classificadores prossegue utilizando os rótulos autoanotados.

### Múltiplas Abordagens

Kim e Park (2017) propuseram uma abordagem de detecção de mudança baseada em distribuição, *Distribution-based Drift Detection Approach* (DbDDA). O DbDDA sinaliza mudanças quando a incerteza do classificador se torna muito alta. Inicialmente, um classificador é treinado utilizando instâncias rotuladas. Em seguida, para cada nova instância  $x_i$ , é estimado um vetor de probabilidade a posteriori  $f(x_i)$ . Um segundo vetor  $\bar{y}(x_i)$  também é determinado. Esse vetor possui valor 1 apenas no componente correspondente à classe prevista para  $x_i$ , enquanto os demais recebem valor 0. A variável aleatória  $X$  é obtida a partir das distâncias  $X(x_i) = \|f(x_i) - \bar{y}(x_i)\|^2$  e é utilizada para estimar a incerteza do classificador.

O método proposto monitora as diferenças  $X(x_i)$  para as instâncias na janela de referência e na janela de detecção. Como o classificador é treinado com dados da janela de referência, as instâncias nessa janela provêm do conceito atual, enquanto as instâncias na janela de detecção podem ter sido geradas a partir de um novo conceito.

Antes de iniciar a classificação, o algoritmo determina um intervalo de confiança para  $X$  a partir dos dados da janela de referência. À medida que cada novo exemplo chega, o valor médio de  $X(x_i)$  é calculado a partir das instâncias na janela de detecção. Se essa diferença for maior que o limite superior do intervalo de confiança, assume-se que ocorreu uma mudança. Nesse caso, os rótulos verdadeiros são solicitados para as instâncias na janela de detecção, e essas instâncias são utilizadas para treinar um novo classificador.

A janela de detecção é sempre uma janela de tempo deslizante. No entanto, os autores propõem três estratégias diferentes para a janela de referência. As duas primeiras são as clássicas janelas fixa e deslizante. Em ambos os casos, a janela de referência não se sobrepõe à janela de detecção. A terceira abordagem, Conjunto de Janelas de Referência, do inglês *Ensemble of Reference Windows* (ERW), baseia-se no uso de cartas de controle de médias e desvios-padrão. Essas cartas monitoram mudanças na distribuição usando a média das médias das janelas e a média dos desvios; padrão das janelas. Enquanto nos métodos anteriores é utilizada apenas uma janela de referência, no ERW um novo par de janelas; referência e detecção; é introduzido à medida que cada instância chega, até que uma mudança seja detectada.

Além disso, [Kim e Park \(2017\)](#) propõem uma abordagem para a situação em que não há dados rotulados disponíveis. Nesse caso, não é possível construir um modelo inicial. Os autores, então, sugerem aplicar o *k-means* para obter uma estrutura de agrupamento, resultando em um classificador virtual. Em seguida, o procedimento de classificação e detecção de mudança é realizado conforme explicado anteriormente.

### Discussão

Mudança de conceito é um problema desafiador para o aprendizado em cenários de dados em fluxo. Uma forma de lidar com essa questão é monitorar os fluxos e atualizar os classificadores atuais com novos conceitos quando mudanças significativas ocorrem. Por um lado, é amplamente aceito que, devido ao enorme volume de dados recebidos dos fluxos, os rótulos verdadeiros geralmente não estão disponíveis de forma imediata. Além disso, dependendo da tarefa a ser resolvida, rotular esses dados pode ser um processo custoso que, em alguns casos, também pode envolver trabalho manual de vários especialistas no domínio. Por outro lado, é facilmente observável a predominância de métodos que abordam o problema de mudança de conceito monitorando a evolução de indicadores ou medidas baseadas na taxa de erro do(s) modelo(s) de

predição atual(is). Por exemplo, o DDM (GAMA et al., 2004) que funciona contando o número de erros. Outro exemplo é o EDDM (BAENA-GARCÍA et al., 2006), baseado no monitoramento da distância entre dois erros de classificação.

Em vista das dificuldades das abordagens supervisionadas, foram propostos detectores de mudança capazes de operar em modo não supervisionado. Espera-se que essas abordagens reduzam significativamente a dependência de rótulos, ao mesmo tempo em que possam fazer melhor uso dos dados disponíveis, já que dados não rotulados são mais frequentemente acessíveis. Além disso, elas podem auxiliar no tratamento da mudança de conceito virtual, ou seja, quando os limites de decisão não são afetados pelas mudanças ocorridas (KHAMASSI et al., 2018). De acordo com o levantamento apresentado neste artigo, os métodos não supervisionados descritos se enquadram em duas categorias básicas: abordagens baseadas em lote e abordagens *online*. No entanto, ambos os grupos tratam o problema de mudança de conceito sob o mesmo ponto de vista: a distribuição subjacente dos dados deve ser monitorada para identificar os pontos em que essa distribuição sofre uma mudança significativa. Ainda assim, cada abordagem emprega estratégias para lidar com essa questão que diferem em seus detalhes e tendem a apresentar limitações distintas.

A descrição apresentada neste artigo permite observar várias propriedades discutidas nesta seção e resumidas na Tabela 12. Primeiramente, a maioria dos métodos não supervisionados é voltada para a detecção de mudanças globais. Consequentemente, uma grande desvantagem é a possibilidade de não detectar mudanças regionais, o que pode levar a problemas como: baixa ou alta sensibilidade e alta taxa de falsos alarmes, especialmente na detecção de conceitos que mudam de forma muito gradual. Considerando que os métodos baseados em abordagem *online* monitoram cada instância individualmente à medida que ela chega, esses métodos são menos propensos a deixar de detectar mudanças locais, mesmo quando se concentram apenas na detecção de mudanças globais. As abordagens baseadas em lote, no entanto, podem ser mais severamente impactadas, já que monitoram mudanças considerando todo o conjunto amostral, mesmo nos casos de lotes-parciais. Como resultado das tentativas de superar essa limitação, algumas estratégias foram desenvolvidas. Por exemplo, o NN-DVI (LIU et al., 2018) concentra-se em definir vizinhanças para lidar com mudanças regionais, enquanto o FAAD (LI et al., 2019) realiza seleção de subconjuntos de atributos e amostragem de atributos.

Tabela 12 – Compilação de algumas características de detectores de mudança não supervisionados encontrados na literatura, destacando o tipo de mudança tratada; se há dependência da aplicação; se é realizada autoanotação; e a dependência de dados rotulados.

Método	Tipo de Mudança	Aplicação	Autoanotação	Requer Rótulos
NN-DVI	Local	Independente		
FAAD	Local/Global	Detecção de Anomalias em Sequência		
UDetect	Global	Independente		
SQSI-IS	Global	Independente	✓	
MD3	Global	Independente		
MD3-EGM	Global	Independente		✓
PDetect	Global	Ataque Adversarial		✓
DDAL	Global	Independente		✓
IKS-bdd	Global	Independente		
CD-TDS	Local/Global	Dados Transacionais		
OMV-PHT	Global	Independente		
NM-DDM	Global	Independente	✓	
Plover	Global	Independente		
SAND	Global	Independente	✓	✓
DSDD	Global	Independente	✓	✓
DbDDA	Global	Independente		✓

A partir da Tabela 12, também é possível observar que os métodos estudados são, em sua maioria, independentes de aplicação. Eles geralmente são testados em dados sintéticos e em poucos conjuntos de dados reais. Quando conjuntos de dados do mundo real são investigados, estes são, em geral, de pequena escala e se assemelham mais a bases de dados de referência simples (*toy benchmarks*) quando comparados aos conjuntos de dados reais recentes de grande escala. Há, entretanto, trabalhos que utilizam conjuntos de dados reais de maior escala e mais desafiadores, como o IKS-bdd (REIS et al., 2016), e trabalhos que investigam problemas de aplicação recentes, como ataques adversariais, estudados no PDetect (SETHI; KANTARDZIC, 2018). Vale destacar que a falta de bases de dados de referência para avaliação é uma limitação da literatura existente sobre mudança de conceito em geral (WARES; ISAACS; ELYAN, 2019). Ainda assim, um cenário diferente é observado para os detectores de mudança supervisionados, cuja avaliação tem sido realizada com conjuntos de dados mais realistas e recentes; por exemplo, demanda por redes de táxi (SAADALLAH et al., 2020).

Embora esta pesquisa esteja principalmente voltada para métodos não super-

visionados de detecção de mudança, algumas técnicas revisadas envolvem instâncias rotuladas. No grupo de detecção lote-parcial, por exemplo, um problema central é que essas abordagens podem ser fortemente afetadas pela estratégia de seleção de instâncias empregada. Além disso, a redução no número de instâncias utilizadas para monitorar a distribuição dos dados pode resultar em altas taxas de detecção falsa. Métodos que buscam lidar com essa questão, como o MD3-EGM (SETHI; KANTARDZIC, 2017) e o PDetect (SETHI; KANTARDZIC, 2018), incluem confirmação supervisionada de mudança para reduzir as taxas de detecção falsa. No entanto, a confirmação supervisionada adiciona dependência de rótulos reais, em contraste com a independência de rótulos buscada pelos detectores não supervisionados. Alguns métodos *online* também podem sofrer dessa limitação, como o DSDD (PINAGÉ; SANTOS; GAMA, 2020), que retém um subconjunto de instâncias rotuladas para ser usado como conjunto de validação, permitindo a seleção dinâmica de classificadores.

Uma vantagem distinta de alguns métodos não supervisionados é a capacidade de encontrar amostras não rotuladas informativas para serem autoanotadas e adicionadas ao conjunto de treinamento, o que é especialmente importante para detectores empregados em cenários de aprendizado *online*. O critério utilizado para a seleção de amostras é basicamente o mesmo: os rótulos de classe só são solicitados para instâncias cujos valores de confiança sejam baixos, como observado no SAND (HAQUE; KHAN; BARON, 2016), no NM-DDM (MUSTAFA et al., 2017) e no DSDD (PINAGÉ; SANTOS; GAMA, 2020). Por fim, há métodos não supervisionados cujo objetivo não é apenas estabelecer um detector de mudança. Nesses casos, seu objetivo final é uma tarefa diferente, como a detecção de anomalias, no FAAD (LI et al., 2019), ou a detecção de novas classes, no NM-DDM (MUSTAFA et al., 2017).

Há propriedades adicionais que merecem ser mencionadas. Uma delas é a necessidade de acumular um lote de instâncias em uma janela de detecção antes de tentar identificar uma mudança nas abordagens baseadas em lote. Isso pode ser apontado como uma limitação pelos seguintes motivos: 1) mudanças podem ocorrer em intervalos que não são limitados pelo tamanho da janela; e 2) lidar com mudança de conceito exige a atualização contínua dos modelos. Esse problema é ainda mais evidente em abordagens que assumem premissas mais restritivas. Por exemplo, no UDetect (BASHIR; PETROVSKI; DOOLAN, 2017), como um tamanho específico de janela é fixado para cada classe, o detector de mudança precisa acumular um número fixo de instâncias para

cada classe antes de determinar se uma mudança de conceito ocorreu.

Essa última limitação é evitada nos métodos baseados em abordagem *online*, pois eles monitoram cada instância individualmente à medida que ela chega. No entanto, embora se possa argumentar que as abordagens baseadas em *batch* têm baixo consumo de memória e baixa sobrecarga de tempo de execução; já que se concentram apenas no monitoramento das medidas extraídas da janela de referência, e não das instâncias presentes na própria janela; o oposto é observado no grupo *online*. Esse grupo de abordagens, especialmente os métodos com janela de referência fixa, como o IKS-bdd (REIS et al., 2016), geralmente precisa manter as instâncias da janela de referência na memória para conseguir detectar as mudanças na distribuição causadas pelas novas instâncias. Uma alternativa para enfrentar esse problema é o uso de representações dos dados, como as estruturas de árvore e de pseudopoints empregadas, respectivamente, em Koh (2016) e Haque, Khan e Baron (2016).

### Conclusão

Neste trabalho, foi proposta uma taxonomia para algumas abordagens de detecção de mudança de conceito não supervisionadas consideradas como estado da arte. Realizou-se uma revisão sistemática da literatura (OKOLI, 2015) para reunir dois grupos de publicações científicas recentes sobre detectores de mudança de conceito: i) Artigos que revisam e analisam diferentes abordagens de detecção de mudança de conceito, e ii) Artigos que propõem métodos não supervisionados para detecção de mudança de conceito. O primeiro grupo permitiu observar que pouquíssimos trabalhos têm se dedicado a resumir as abordagens não supervisionadas. O segundo grupo, por sua vez, possibilitou agrupar os métodos de detecção de mudança de conceito não supervisionados e semi-supervisionados em duas categorias principais: métodos baseados em lotes e métodos baseados em execução *online*. Os métodos *online* e em lote referem-se apenas ao componente de detecção da estratégia. Por um lado, os métodos *online* indicam a ocorrência de uma mudança no fluxo de dados uma instância por vez. Por outro, os métodos em lote precisam processar um conjunto de instâncias de forma conjunta para conseguir sinalizar uma mudança.

Neste estudo, foram analisados oito métodos baseados em lotes, observando-se que a principal diferença entre eles está em monitorar ou não mudanças significativas na distribuição dos dados em todo o lote de instâncias que chega ou apenas em um subconjunto selecionado de instâncias. Também foram analisados oito métodos baseados em

execução online. Esse grupo foi subdividido em duas categorias, de acordo com a forma como comparam as duas janelas envolvidas na detecção de mudança, denominadas janela de referência e janela de detecção. Um grupo difere do outro quanto ao uso de janelas de referência fixas ou deslizantes para detectar mudanças de conceito. Trabalhos futuros podem se concentrar em expandir este estudo preliminar para avaliar e comparar os métodos destacados por meio de experimentos computacionais, conforme realizado em [Barros e Santos \(2018\)](#) para detectores supervisionados de mudança de conceito.

## APÊNDICE B

## Valores dos parâmetros das bases de dados.

Tabela 13 – Valores dos parâmetros nas bases de dados *Line*, *Hiperplane*, *Circle*, *SineH* and *Sine* por cenário.

Cenário	<i>Line</i>				<i>Hiperplane</i>					<i>Circle</i>				
	$a_1$	$a_0$			$a_1$	$a_2$	$a_0$			a	b	r		
		Início	Fim	Passo			Início	Fim	Passo			Início	Fim	Passo
1 17	0,10	0,10	13,40	0,70	0,10	0,10	0,70	71,00	3,70	0,50	0,50	0,20	5,90	0,30
2 18	0,10	0,10	2,90	0,70	0,10	0,10	0,70	15,50	3,70	0,50	0,50	0,20	1,40	0,30
3 19	0,10	0,10	0,80	0,70	0,10	0,10	0,70	11,80	3,70	0,50	0,50	0,20	0,50	0,30
4 20	0,10	0,10	2,90	0,70	0,10	0,10	0,70	15,50	3,70	0,50	0,50	0,20	1,40	0,30
5 21	0,10	0,10	27,40	0,70	0,10	0,10	0,70	145,00	3,70	0,50	0,50	0,20	11,90	0,30
6 22	0,10	0,10	6,40	0,70	0,10	0,10	0,70	34,00	3,70	0,50	0,50	0,20	2,90	0,30
7 23	0,10	0,10	2,20	0,70	0,10	0,10	0,70	11,80	3,70	0,50	0,50	0,20	1,10	0,30
8 24	0,10	0,10	6,40	0,70	0,10	0,10	0,70	34,00	3,70	0,50	0,50	0,20	2,90	0,30
9 25	0,10	0,10	3,25	0,15	0,10	0,10	2,00	15,30	0,70	0,50	0,50	0,20	2,10	0,10
10 26	0,10	0,40	1,0	0,15	0,10	0,10	2,00	4,80	0,70	0,50	0,50	0,20	0,60	0,10
11 27	0,10	0,40	0,55	0,15	0,10	0,10	2,00	2,70	0,70	0,50	0,50	0,20	0,30	0,10
12 28	0,10	0,10	2,00	4,80	0,10	0,10	2,00	4,80	0,70	0,50	0,50	0,20	0,60	0,10
13 29	0,10	0,40	6,25	0,15	0,10	0,10	2,00	29,30	0,70	0,50	0,50	0,20	4,10	0,10
14 30	0,10	0,10	2,00	8,30	0,10	0,10	2,00	8,30	0,70	0,50	0,50	0,20	1,10	0,10
15 31	0,10	0,10	2,00	4,10	0,10	0,10	2,00	4,10	0,70	0,50	0,50	0,20	1,10	0,10
16 32	0,10	0,40	1,75	0,15	0,10	0,10	2,00	8,30	0,70	0,50	0,50	0,20	0,50	0,10

Cenário	<i>SineH</i>					<i>SineV</i>				
	a	c	b			a	b	c		
			Início	Fim	Passo			Início	Fim	Passo
1 17	5.00	5.00	0.00	59.66	3.14	1.00	1.00	-8.00	277.0	15.00
2 18	5.00	5.00	0.00	12.56	3.14	1.00	1.00	-8.00	52.00	15.00
3 19	5.00	5.00	0.00	3.14	3.14	1.00	1.00	-8.00	7.00	15.00
4 20	5.00	5.00	0.00	12.56	3.14	1.00	1.00	-8.00	52.00	15.00
5 21	5.00	5.00	0.00	122.46	3.14	1.00	1.00	-8.00	577.00	15.00
6 22	5.00	5.00	0.00	28.26	3.14	1.00	1.00	-8.00	127.00	15.00
7 23	5.00	5.00	0.00	9.42	3.14	1.00	1.00	-8.00	37.00	15.00
8 24	5.00	5.00	0.00	28.26	3.14	1.00	1.00	-8.00	127.00	15.00
9 25	5.00	5.00	0.00	14.82	0.78	1.00	1.00	-2.00	55.00	3.00
10 26	5.00	5.00	0.00	3.12	0.78	1.00	1.00	-2.00	10.00	3.00
11 27	5.00	5.00	0.00	0.78	0.78	1.00	1.00	-2.00	1.00	3.00
12 28	5.00	5.00	0.00	0.78	3.14	1.00	1.00	-2.00	10.00	3.00
13 29	5.00	5.00	0.00	30.42	0.78	1.00	1.00	-2.00	115.00	3.00
14 30	5.00	5.00	0.00	7.02	0.78	1.00	1.00	-2.00	25.00	3.00
15 31	5.00	5.00	0.00	2.34	0.78	1.00	1.00	-2.00	7.00	3.00
16 32	5.00	5.00	0.00	7.02	0.78	1.00	1.00	-2.00	25.00	3.00



## APÊNDICE C

## Métricas de desempenho dos detectores em cada cenário por contexto.

O eixo horizontal de cada mapa de calor representa cada cenário. O eixo vertical representa as bases de dados *Circle* (C), *Line* (L), *Hiperplane* (H), *SineV* (SV) e *SineH* (SH).

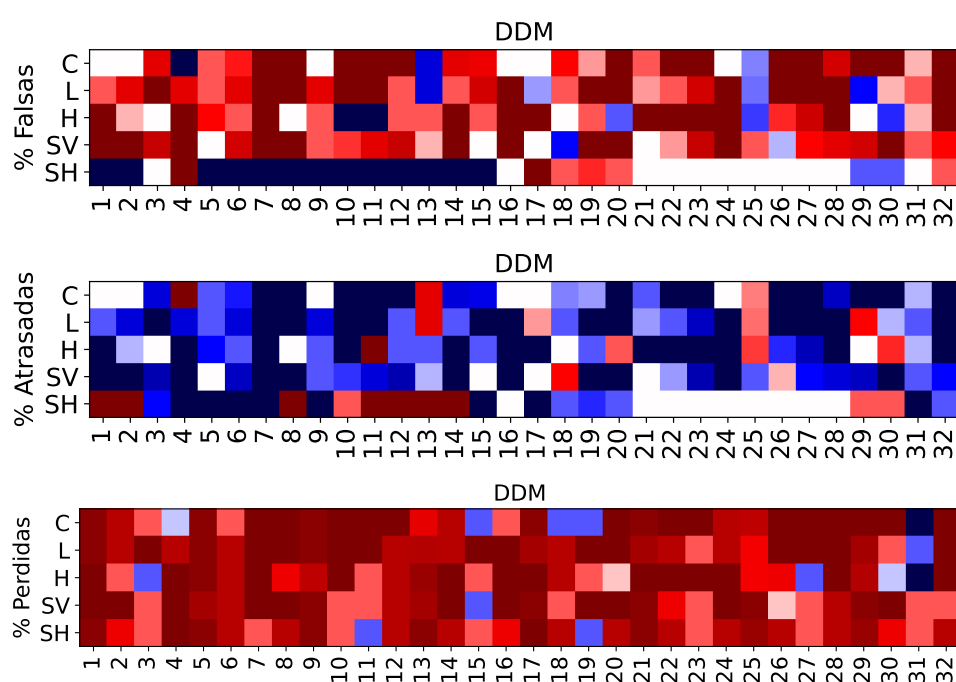


Figura 41 – Cenário DDM.

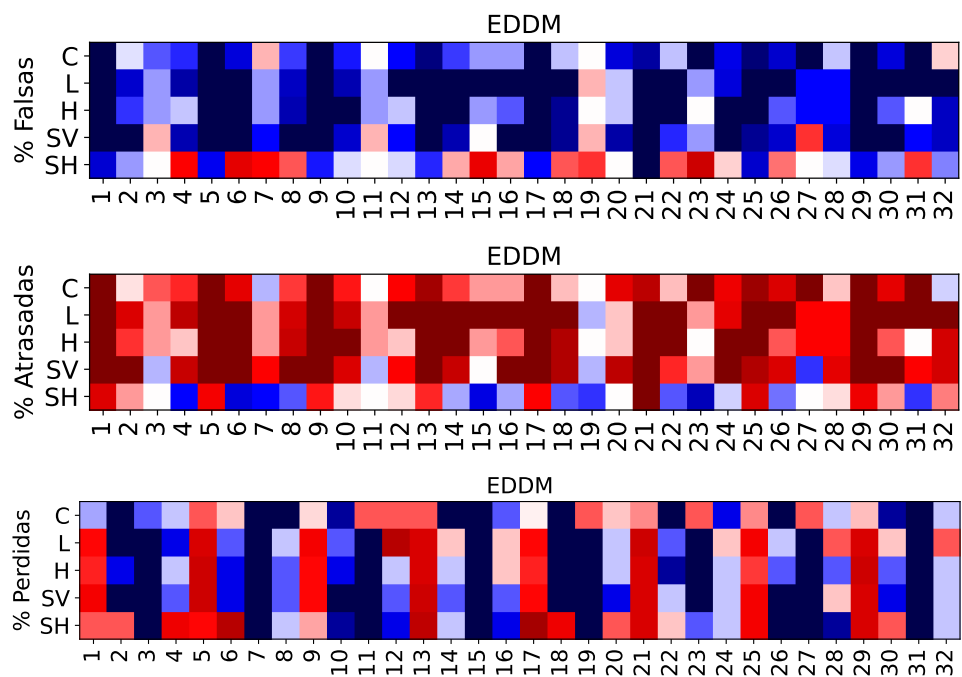


Figura 42 – Cenário EDDM.

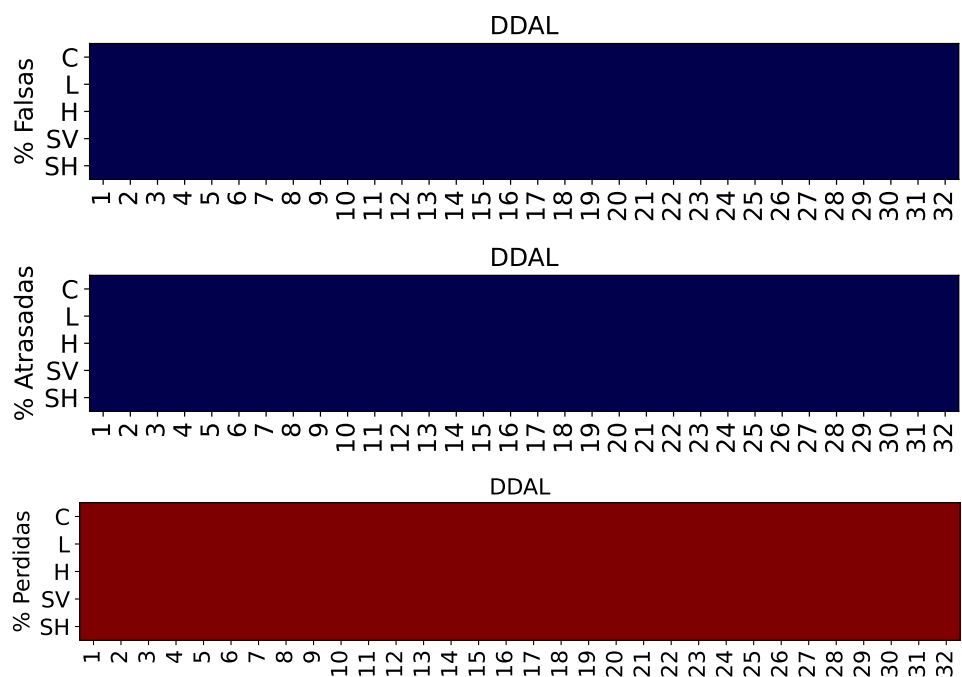


Figura 43 – Cenário DDAL.

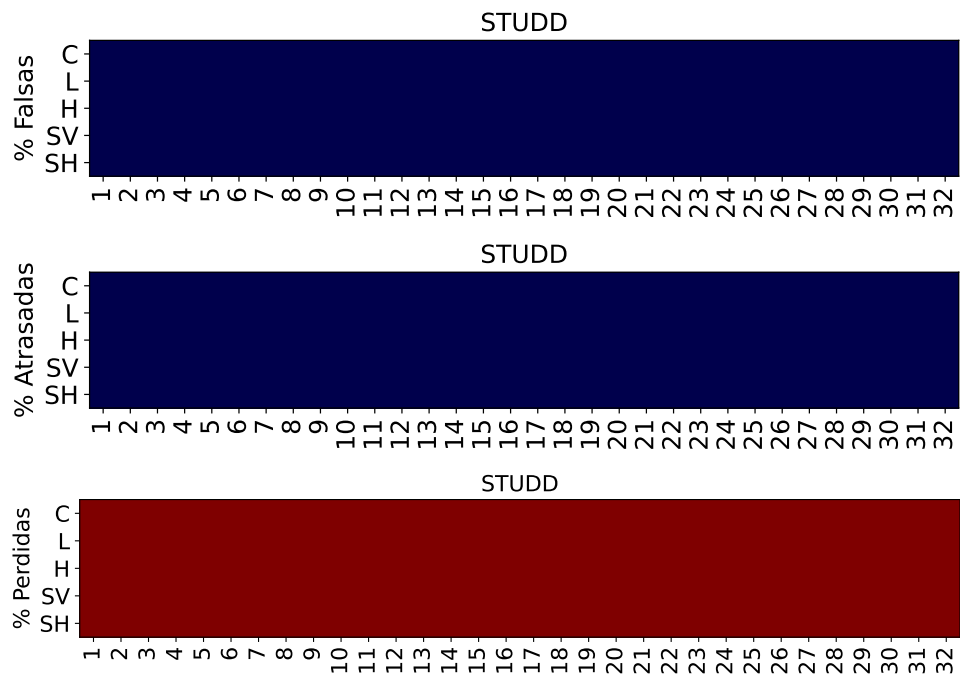


Figura 44 – Cenário STUDD.

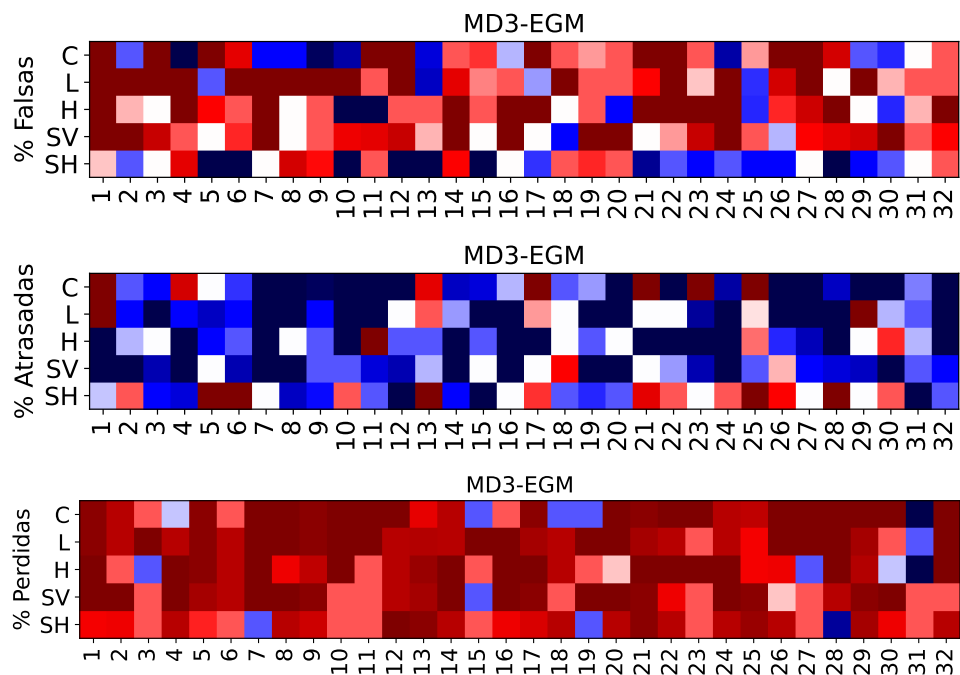


Figura 45 – Cenário MD3-EGM.

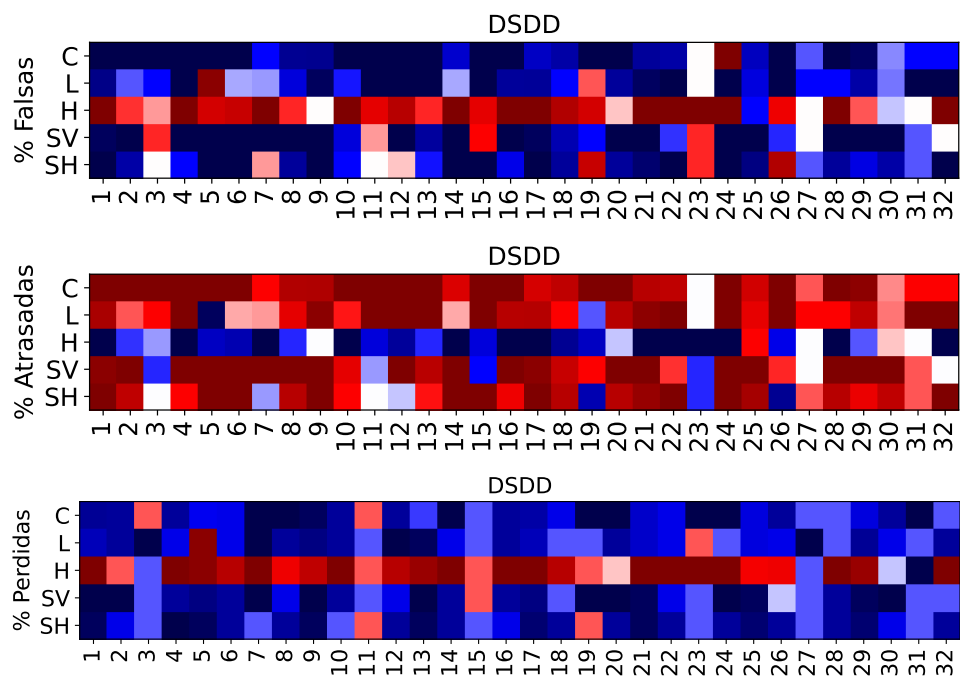


Figura 46 – Cenário DSDD.