



Universidade Federal da Bahia
Instituto de Matemática

Programa de Pós-Graduação em Ciência da Computação

**APLICANDO REDES DE FUNÇÃO DE BASE
RADIAL PARA DETECÇÃO DE MUDANÇAS
DE CONCEITO EM FLUXOS CONTÍNUOS
DE DADOS**

Ruivaldo Azevedo Lobão Neto

QUALIFICAÇÃO DE MESTRADO

Salvador
03 de Abril de 2019

RUIVALDO AZEVEDO LOBÃO NETO

**APLICANDO REDES DE FUNÇÃO DE BASE RADIAL PARA
DETECÇÃO DE MUDANÇAS DE CONCEITO EM FLUXOS
CONTÍNUOS DE DADOS**

Esta Qualificação de Mestrado foi apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal da Bahia, como requisito parcial para obtenção do grau de Mestre em Ciência da Computação.

Orientador: Ricardo Araújo Rios

Salvador
03 de Abril de 2019

TERMO DE APROVAÇÃO

RUIVALDO AZEVEDO LOBÃO NETO

APLICANDO REDES DE FUNÇÃO DE BASE RADIAL PARA DETECÇÃO DE MUDANÇAS DE CONCEITO EM FLUXOS CONTÍNUOS DE DADOS

Esta Qualificação de Mestrado foi julgada adequada à obtenção do título de Mestre em Ciência da Computação e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciência da Computação da Universidade Federal da Bahia.

Salvador, 03 de Abril de 2019

Prof. Dr. Ricardo Araújo Rios
UFBA

RESUMO

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Palavras-chave: Aprendizado de Máquina, Fluxos Contínuos de Dados, Mudanças de Conceito, Redes de Função de Base Radial, Não supervisionado

ABSTRACT

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Keywords: Machine Learning, Data Streams, Concept Drift, Radial Basis Function Networks, RBF Network, Unlabeled

SUMÁRIO

Capítulo 1—Introdução	1
1.1 Contexto e Motivação	1
1.2 Hipótese e Objetivo	2
Capítulo 2—Revisão Bibliográfica	5
2.1 Considerações Iniciais	5
2.2 Fluxos Contínuos de Dados e Aprendizado de Máquina	5
2.3 Mudança de Conceito	6
2.3.1 Taxonomia	9
2.3.2 Algoritmos para Detecção de Mudança de Conceito	10
2.3.3 Ferramentas	12
2.3.4 MOA	12
2.3.5 Tornado	15
2.4 Redes de Função de Base Radial	16
2.5 Trabalhos Relacionados	18
2.6 Considerações Finais	18
Capítulo 3—Plano de Pesquisa	21
3.1 Considerações Iniciais	21
3.2 Descrição do Problema	21
3.3 Atividades de Pesquisa	22
3.4 Considerações Finais	22
Capítulo 4—Experimentos Iniciais	23
4.1 Considerações Iniciais	23
4.2 Configuração dos Experimentos	23
4.3 Método de Pettitt	23
4.4 Redes de Função de Base Radial	24
4.5 Considerações Finais	24

LISTA DE FIGURAS

2.1	Mudança de Conceito Virtual vs. Mudança de Conceito Real	8
2.2	Padrões de ocorrência de Mudanças de Conceito	8
2.3	MOA - Tela Inicial	13
2.4	MOA - Configuração detector	14
2.5	Framework Tornado (PESARANGHADER, 2018).	15
2.6	Tornado - Resultado para múltiplos pares (PESARANGHADER, 2018) .	16
2.7	Arquitetura RBF	17

LISTA DE TABELAS

2.1	Terminologia - Mudança de Conceito (ZLIOBAITE, 2010)	9
2.2	Sumário - Abordagens de detecção (SETHI; KANTARDZIC, 2017) . . .	12

INTRODUÇÃO

1.1 CONTEXTO E MOTIVAÇÃO

Nos últimos anos, o volume de dados produzido por sistemas computacionais tem crescido de forma significativa. Segundo (ZWOLENSKI; WEATHERILL, 2014), estima-se que entre os anos de 2014 e 2020 a quantidade de dados produzida anualmente aumentará de 4,4 zettabytes (trilhões de gigabytes) para 44 zettabytes.

Nos principais domínios de aplicação de sistemas computacionais, as informações são produzidas na forma de Fluxos Contínuos de Dados (FCDs): sequências ininterruptas, ordenadas, de alta frequência e potencialmente infinitas (AGGARWAL, 2006). Exemplos desses domínios, incluem: redes de sensores, redes de computadores, comércio eletrônico, análise de sentimentos, entre outros.

A área de Aprendizado de Máquina (AM) estuda algoritmos que melhoram o seu desempenho na realização de determinada tarefa conforme ganham experiência (MIT-CHELL, 1997). Na última década, tem havido um grande interesse de pesquisa na aplicação dessas técnicas em diferentes domínios. Contudo, muitos desses domínios lidam com dados na forma de fluxos contínuos, sendo necessário adaptar os algoritmos para atenderem a severas restrições de tempo de execução e de uso dos recursos computacionais (BIFET, 2009).

Além dessas restrições, as técnicas de aprendizado também devem considerar que os domínios de aplicação com fluxos contínuos de dados são, em sua maioria, dinâmicos. Isto é, nesses cenários, o contexto do processo gerador ou a distribuição dos dados podem mudar ao longo do tempo. Este problema é denominado mudança de conceito (GAMA, 2010) e a sua ocorrência pode impactar a acurácia do modelo de decisão criado.

A fim de minimizar o impacto das mudanças de conceito, foram desenvolvidas diferentes técnicas de detecção (GAMA et al., 2014). As primeiras técnicas propostas atuavam em conjunto com os classificadores, detectando alterações através do monitoramento da taxa de erro. São exemplos desta abordagem: DDM (GAMA et al., 2004), EDDM (BAENA-GARCÍA et al., 2006), ADWIN (BIFET; GAVALDÀ, 2007), ECDD (ROSS

et al., 2012), PL (Bach; Maloof, 2008), FCWM (SEBASTIÃO et al., 2010), STEPD (NISHIDA; YAMAUCHI, 2007), DOF (SOBHANI; BEIGY, 2011), dentre outros.

Todavia, as técnicas mencionadas dependem que o rótulo correto de cada exemplo esteja disponível. Em muitos contextos, o tempo ou o custo para obter esse rótulo é proibitivo (AGGARWAL, 2006). Diante disso, foram desenvolvidos algoritmos para detecção de mudanças de conceito que independem dessa informação.

As técnicas independentes de rótulos baseiam-se na identificação de eventos que não se enquadram na atual estrutura dos dados (SPINOSA; CARVALHO; GAMA, 2007). Para tanto, técnicas de agrupamento, detecção de *outliers*, sumarização dos dados e aplicação de medidas de dissimilaridade são utilizadas nas implementações (RYU et al., 2012). São exemplos desta categoria: OLINDDA (SPINOSA; CARVALHO; GAMA, 2007), MINAS (FARIA; GAMA; CARVALHO, 2013), ECSMiner (MASUD et al., 2011), GC3 (SETHI; KANTARDZIC; HU, 2016), dentre outros.

Entretanto, segundo (AGGARWAL, 2006; GAMA et al., 2014), as técnicas para identificação de mudanças de conceito propostas ainda possuem deficiências. Os algoritmos dependentes de rótulos, por exemplo, são impraticáveis para muitos problemas do mundo real. De forma similar, as técnicas independentes têm dificuldade para atender às restrições de tempo de execução e de uso dos recursos computacionais em cenários com fluxos contínuos de dados.

Visando resolver essas limitações, este projeto de mestrado propõe um algoritmo baseado em redes de função de base radial capaz de detectar mudanças de conceito em fluxos contínuos de dados, de forma online, precisa e independente de exemplos prévios ou rótulos.

1.2 HIPÓTESE E OBJETIVO

Com base nas observações citadas anteriormente, a seguinte hipótese foi formulada:

“A análise de fluxos contínuos de dados através de redes de função de base radial permite a detecção de mudanças de conceito de forma online, precisa e independente de exemplos prévios ou rótulos.”

O objetivo deste trabalho será o desenvolvimento e comprovação desta hipótese. Para atingir este objetivo, será implementado um algoritmo para detecção de mudanças de conceito baseado em redes de função de base radial. Os principais diferenciais deste método serão a escolha dinâmica dos centros, conforme novos dados são recepcionados, e a aplicação de um raio variável. A ativação de novos centros será utilizada como indicador de possíveis mudanças de conceito.

O algoritmo implementado será comparado com o estado da arte. O primeiro conjunto de testes será formado por dados sintéticos, que permitirão uma análise detalhada da abordagem, uma vez que as características e os comportamentos dos dados são conhecidos. O segundo conjunto será composto por dados obtidos a partir de sistemas computacionais utilizados na indústria, visando apresentar uma aplicação prática para a solução proposta.

Este projeto está organizado conforme a seguinte estrutura: O **Capítulo 2** apresenta uma revisão bibliográfica dos principais conceitos utilizados neste trabalho: fluxos

contínuos de dados e aprendizado de máquina, mudança de conceito e redes de função de base radial; No **Capítulo 3** o plano de pesquisa é detalhado, identificando a metodologia que será aplicada na pesquisa e o cronograma de atividades. Por fim, o **Capítulo 4** apresenta um conjunto de experimentos preliminares e a análise dos resultados obtidos.

REVISÃO BIBLIOGRÁFICA

2.1 CONSIDERAÇÕES INICIAIS

Este capítulo discute os principais conceitos utilizados neste trabalho de pesquisa. Em primeiro momento, a aplicação de técnicas de aprendizado de máquina a fluxos contínuos de dados será detalhada. Em seguida, abordaremos a ideia de mudança de conceito, seus tipos e características, taxonomia, principais algoritmos de detecção e ferramentas. Logo após, as redes de função de base radial serão explanadas. Ao fim do capítulo, serão elencados os trabalhos relacionados encontrados na literatura.

2.2 FLUXOS CONTÍNUOS DE DADOS E APRENDIZADO DE MÁQUINA

O volume de dados produzidos por sistemas computacionais tem crescido de forma significativa (COHEN et al., 2009). Parte relevante desses dados é produzida em sequências contínuas, ordenadas, de alta frequência e potencialmente infinitas. Estas sequências são denominadas Fluxos Contínuos de Dados (FCDs) e estão presentes em muitos domínios de aplicação (AGGARWAL, 2006).

A área de Aprendizado de Máquina (AM) estuda algoritmos que melhoram o seu desempenho na realização de determinada tarefa conforme ganham experiência (MITCHELL, 1997). Nos últimos anos, há um crescente interesse de pesquisa na aplicação dessas técnicas em diferentes contextos. Entretanto, segundo (GAMA, 2010), a área teve como foco, durante muitos anos, os problemas de aprendizado em lote. Nestes cenários, os algoritmos consideram que a distribuição dos dados e o contexto do processo não sofrem alterações ao longo do tempo. Ou seja, não há o surgimento de novos conceitos-alvo ou alteração dos existentes.

Contudo, os principais domínios de aplicação com fluxos contínuos de dados são dinâmicos. Isto é, a distribuição dos dados ou o contexto do processo gerador podem sofrer alterações que afetam os conceitos-alvo esperados. Estas alterações, denominadas mudanças de conceito (GAMA, 2010), podem impactar a acurácia do modelo de decisão.

As técnicas de aprendizado de máquina dividem-se em duas categorias principais: não supervisionadas (agrupamento ou *clustering*) e supervisionadas (classificação ou regressão) (MITCHELL, 1997). Algoritmos de ambas as categorias foram adaptados para que pudessem ser aplicados à cenários com fluxos contínuos de dados. Os detalhes de cada categoria e as adaptações propostas na literatura serão discutidas a seguir.

As técnicas de agrupamento associam objetos conforme suas características. Estas técnicas buscam formar grupos com alta similaridade interna e baixa similaridade com elementos externos (JAIN; DUBES, 1988). Os principais métodos de agrupamento para cenários em lote são: K-Means (LLOYD, 2006), DBSCAN (ESTER et al., 1996), PAM (KAUFMAN; ROUSSEUW, 1990), OPTICS (ANKERST et al., 1999), dentre outros.

Segundo (GAMA, 2010), a principal dificuldade na aplicação de técnicas de agrupamento em cenários com fluxos contínuos de dados é a manutenção da qualidade e consistência dos grupos formados em relação à sequência de dados observados. Portanto, é necessário que os algoritmos atuem de forma incremental, evoluindo os grupos formados ao longo do tempo (BARBARÁ, 2002). Sendo assim, foram desenvolvidos algoritmos de agrupamento especializados para fluxos contínuos de dados (AGGARWAL et al., 2003): CluStream (AGGARWAL et al., 2003), StreamKM++ (ACKERMANN et al., 2012), DenStream (CAO et al., 2006), D-Stream (LING; LING-JUN; LI, 2009) e ClusTree (KRANEN et al., 2011), dentre outros.

Os algoritmos supervisionados realizam previsões para novos exemplos utilizando um modelo de decisão criado a partir de uma base de treinamento previamente rotulada (KOTSIANTIS, 2007). Se a previsão é categórica, entende-se como um problema de classificação. Se a previsão resulta em um valor numérico, trata-se de uma tarefa de regressão. Os principais algoritmos de classificação e regressão para cenários em lote são: árvores de decisão (BREIMAN et al., 1984), métodos baseados em regras, redes neurais e máquinas de vetores suporte (SVM) (VAPNIK, 1998), dentre outros.

Contudo, esses métodos não atualizam o modelo conforme novos dados são observados. Além disso, também assumem que todos os dados estão disponíveis em memória e que é possível realizar múltiplas varreduras sobre eles. Dessa forma, novos algoritmos supervisionados foram propostos para os cenários com fluxos contínuos de dados (DOMINGOS; HULTEN, 2000; BIFET et al., 2013; WANG et al., 2003; AGGARWAL et al., 2004; GAMA; ROCHA; MEDAS, 2003).

No contexto deste trabalho de mestrado, assume-se que os dados são obtidos a partir de fluxos contínuos de dados com mudanças de conceito. O algoritmo para detecção de mudanças proposto é independente de técnicas de aprendizado, pois identifica as mudanças a partir de uma análise direta e contínua do fluxo de dados. Na próxima seção, o fenômeno mudança de conceito será discutido.

2.3 MUDANÇA DE CONCEITO

Ao aplicar técnicas de aprendizado de máquina em cenários dinâmicos, é comum que a distribuição dos dados ou o contexto do processo mudem ao longo do tempo, levando ao fenômeno de mudança de conceito (SCHLIMMER; GRANGER, 1986). Como exemplo, considere um modelo criado a partir do histórico de compras de um cliente em uma

livraria. Este cliente tem demonstrado, ao longo do tempo, uma absoluta preferência por livros de ficção científica. Entretanto, ao ingressar na universidade, passa a consumir apenas livros técnicos. Esta alteração de comportamento é considerada uma mudança de conceito e pode ter impacto na acurácia do modelo.

A Teoria Bayesiana de Decisão (DUDA; HART; STORK, 2000) é comumente aplicada para descrever as técnicas de classificação. Utilizaremos esta descrição como base teórica para formalizar o problema de mudança de conceito: Sendo $X \in \mathbb{R}^p$ uma instância em um espaço p -dimensional de atributos e $X \in c_i$ onde c_1, c_2, \dots, c_k é o conjunto de classes, o classificador ótimo para classificar $x \rightarrow c_i$ é determinado pelas probabilidades a priori das classes $P(c_i)$ e a função de densidade de probabilidade condicionada às classes $p(X|c_i), i = 1, \dots, k$.

Dessa forma, um conceito pode ser definido como um conjunto de probabilidades a priori e condicionais das classes, como mostra a Equação 2.1:

$$S = \{(P(c_1), P(X|c_1)), (P(c_2), P(X|c_2)), \dots, (P(c_k), P(X|c_k))\} \quad (2.1)$$

Ainda segundo a Teoria Bayesiana, a classificação de uma instância X baseada na máxima probabilidade a posteriori pode ser obtida através da Equação 2.2:

$$p(c_i|X) = \frac{p(c_i) * p(X|c_i)}{p(X)} \quad (2.2)$$

Sendo assim, é possível formalmente definir que há mudança de conceito entre os instantes t_0 e t_1 se:

$$\exists X : p_{t_0}(X, c) \neq p_{t_1}(X, c) \quad (2.3)$$

onde, p_{t_0} e p_{t_1} denotam as distribuições de probabilidades conjuntas nos instantes t_0 e t_1 , respectivamente, para X e c (GAMA et al., 2014). Isto é, um conjunto de exemplos possui rótulos de classe legítimos em t_0 , mas passa a ter rótulos diferentes, também legítimos, em t_1 (KOLTER; MALOOF, 2007).

Com base na formalização exposta, segundo (ZLIOBAITE, 2010), mudanças de conceito podem ocorrer devido a:

1. Alterações na probabilidade a priori das classes $P(c)$;
2. Alterações na distribuição de uma ou mais classes $p(X|c)$;
3. Alterações nas distribuições a posteriori das classes $p(c|X)$.

De acordo com (ZLIOBAITE, 2010; GAMA et al., 2014), as mudanças de conceito podem ser categorizadas como virtuais ou reais. As virtuais indicam mudanças na probabilidade a priori das classes, $P(c)$, e não têm impacto nos conceitos-alvo (representadas na Figura 2.1 (b)). Enquanto as mudanças de conceito reais referem-se a mudanças na probabilidade a posteriori, $p(c|X)$, e afetam os conceitos-alvo (representadas na Figura 2.1 (c)).

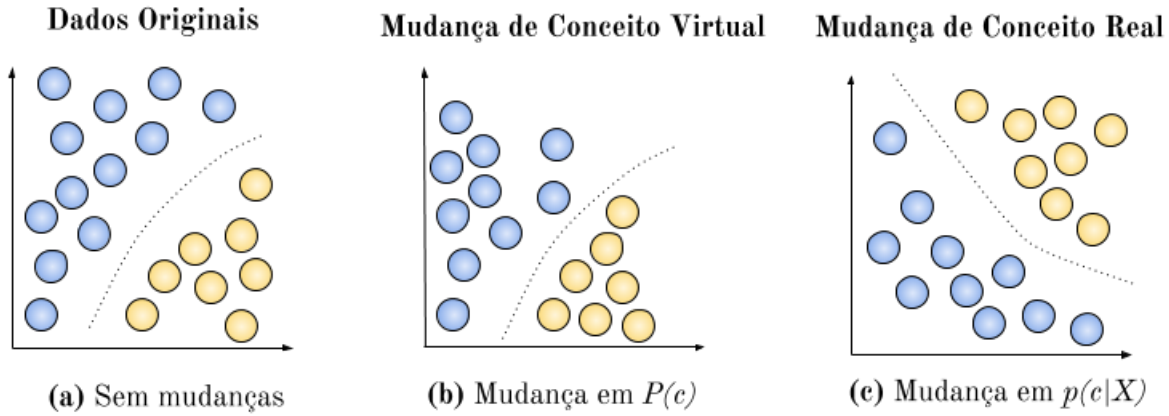


Figura 2.1 Mudança de Conceito Virtual vs. Mudança de Conceito Real

Conforme (GAMA et al., 2014), as mudanças de conceito podem ocorrer de forma abrupta, gradual, incremental ou recorrente. A Figura 2.2 ilustra estes padrões:

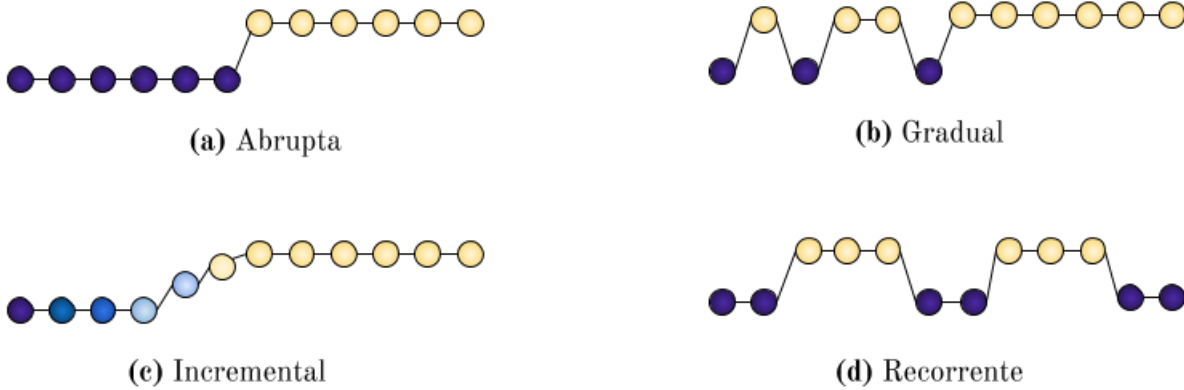


Figura 2.2 Padrões de ocorrência de Mudanças de Conceito

Mudanças de conceito abruptas são alterações repentinas entre conceitos, como descrito na Figura 2.2 (a). Por exemplo, o gênero literário preferido de um cliente pode mudar repentinamente de ficção científica para auto-ajuda. Outro exemplo, é a troca de um sensor por outro com diferente calibração, em uma planta industrial química (GAMA et al., 2014).

A mudança de conceito gradual representa uma transição mais lenta de um conceito para outro. Isto é, durante a transição, são percebidos eventos de ambos os conceitos. Pode-se considerar que o cliente, ao invés de subitamente passar a preferir livros de auto-ajuda, adquire interesse nestes ao longo do tempo. Isto significa que ainda existe certo interesse em livros de ficção científica, mas, em paralelo, desenvolve-se uma crescente pre-

ferência por livros de auto-ajuda, que se tornará predominante. A mudança de conceito gradual é representada na Figura 2.2 (b).

Mudanças de conceito incrementais apresentam conceitos intermediários durante a transição de um conceito para outro. Para o exemplo utilizado, é como se o interesse literário variasse de ficção científica para ficção científica e aventura, de ficção científica e aventura para aventura e suspense, de aventura e suspense para suspense e auto-ajuda, e finalmente de suspense e auto-ajuda para apenas auto-ajuda. Considerando a planta industrial química e os sensores, seria como se um sensor perdesse precisão ao longo do tempo. A Figura 2.2 (c) demonstra a mudança de conceito incremental.

A mudança de conceito recorrente surge quando um conceito, previamente ativo, reaparece após determinado período de tempo. Em nosso exemplo, considere que o cliente está atualmente interessado em livros de ficção científica. Contudo, ele é presenteado com uma série de livros de auto-ajuda de um renomado autor. Após ler toda a série, o cliente é premiado com descontos para livros de ficção científica e volta a lê-los. A mudança de conceito recorrente está representada na Figura 2.2 (d).

O método proposto neste trabalho de mestrado busca identificar mudanças de conceito reais em fluxos contínuos de dados, independente do padrão de ocorrência destas. Na subseção seguinte, será discutida a taxonomia do termo mudança de conceito e diferenciada a técnica de detecção de mudança de conceito de outras relacionadas.

2.3.1 Taxonomia

O fenômeno mudança de conceito tem sido estudado em diferentes comunidades de pesquisa, incluindo mineração de dados, aprendizado de máquina, estatística e recuperação de informação (ZLIOBAITE, 2010). Contudo, o mesmo conceito pode ter diferentes nomenclaturas em cada comunidade. Na Tabela 2.3.1 são listados os termos correspondentes a mudança de conceito para cada área de pesquisa.

Área	Termos
Mineração de Dados	Mudança de Conceito
Aprendizado de Máquina	Mudança de Conceito, Mudança de Covariável
Computação Evolucionária	Ambiente Evolutivo, Ambiente em Mudança
IA e Robótica	Ambiente Dinâmico
Estatísticas, Séries Temporais	Não Estacionário
Recuperação de Informação	Evolução Temporal

Tabela 2.1 Terminologia - Mudança de Conceito (ZLIOBAITE, 2010)

Outra fonte comum de equívocos são os termos Detecção de *Outliers*, Detecção de Novidade, Detecção de *Change Points* e Detecção de Mudança de Conceito. Estes termos são, em muitos cenários, utilizados de forma indistinta. Mas, para o contexto deste trabalho, é importante distingui-los.

De acordo com (CHANDOLA; BANERJEE; KUMAR, 2009), a expressão Detecção de *Outliers* refere-se a tarefa de encontrar padrões nos dados em desacordo com o comportamento esperado: anomalias, ruídos, etc.

Ainda conforme (CHANDOLA; BANERJEE; KUMAR, 2009), as técnicas para Detecção de Novidade têm como objetivo detectar padrões ainda não observados nos dados. Diferentemente das técnicas para Detecção de *Outliers*, as técnicas para Detecção de Novidade incorporam os novos padrões ao modelo de decisão.

Por fim, as técnicas para Detecção de *Change Points* identificam variações abruptas de valor, que podem representar transições entre estados, em séries temporais unidimensionais estacionárias (AMINIKHANGHAHI; COOK, 2017).

Na próxima subseção, detalharemos as principais técnicas para detecção de mudanças de conceito encontradas na literatura.

2.3.2 Algoritmos para Detecção de Mudança de Conceito

Os algoritmos para detecção de mudança de conceito caracterizam e quantificam as mudanças de conceito através da delimitação dos instantes ou intervalos de tempo em que as mudanças ocorrem (BASSEVILLE; NIKIFOROV, 1993).

Essas técnicas podem ser divididas em duas categorias principais, conforme a dependência ou não da rotulação dos dados:

Algoritmos Explícitos/Supervisionados Dependem da rotulação dos dados por um especialista. Estes rótulos são utilizados no cálculo de medidas de performance como taxa de erro e acurácia, que são monitoradas ao longo do tempo. Mudanças de conceito são sinalizadas quando essas medidas atingem um limite previamente definido.

Algoritmos Implícitos/Não Supervisionados Independem da rotulação por especialistas, baseando-se em características dos próprios dados ou indicadores das técnicas de aprendizagem aplicadas, para sinalizar desvios. São mais propensos a alarmes falsos, mas a independência de rótulos os torna interessantes para contextos onde a obtenção desses é dispendiosa, demorada ou inviável.

Em sua pesquisa (GAMA et al., 2014) categorizou os algoritmos supervisionados em três grupos principais:

Métodos Baseados em Análise Sequencial Avaliam, de forma sequenciada, os resultados das predições conforme tornam-se disponíveis (performance). Indicam a ocorrência de mudança de conceito quando um limite pré-definido é atingido. Os algoritmos *Cumulative Sum (CUSUM)*, *PageHinkley (PH)* (PAGE, 1954) e *Geometric Moving Average (GMA)* (ROBERTS, 2000) são representantes deste grupo.

Abordagens baseadas em Estatística Identificam mudanças de conceito através da análise de parâmetros estatísticos como média e desvio padrão para os resultados das predições. Os métodos *Drift Detection Method (DDM)* (GAMA et al., 2004), *Early Drift Detection Method (EDDM)* (BAENA-GARCÍA et al., 2006), *Exponentially*

Weighted Moving Average (EWMA) (ROSS et al., 2012) e *Reactive Drift Detection Method (RDDM)* (BARROS et al., 2017) são exemplos deste grupo.

Métodos baseados em Janelas Utilizam uma janela de tamanho fixo para sumarizar informações passadas e uma janela deslizante para sumarizar dados mais recentes. A mudança de conceito é detectada quando há uma diferença significativa entre as distribuições das janelas. Esta diferença é detectada a partir de testes estatísticos ou desigualdades matemáticas, considerando como hipótese nula a igualdade entre as distribuições. Os algoritmos *Adaptive Windowing (ADWIN)* (BIFET; GAVALDÀ, 2007), *SeqDrift* (PEARS; SAKTHITHASAN; KOH, 2014), *HDDMA* e *HDDMW* (BLANCO et al., 2015) pertencem a este grupo.

Os algoritmos implícitos também foram divididos em três grupos (GONÇALVES et al., 2014):

Detecção de Novidade / Métodos de Agrupamento Utilizam estratégias oriundas das técnicas de agrupamento e detecção de *outliers* para identificar novos padrões (RYU et al., 2012). Os métodos *OLINDDA* (SPINOSA; CARVALHO; GAMA, 2007), *MINAS* (FARIA; GAMA; CARVALHO, 2013), *Woo* (RYU et al., 2012), *DETECTNOD* (Hayat; Hashemi, 2010), *ECSMiner* (MASUD et al., 2011) e *GC3* (SETHI; KANTARDZIC; HU, 2016) fazem parte deste grupo.

Monitoramento de distribuição multivariada Monitoram diretamente a distribuição dos dados através de subconjuntos. Um subconjunto de treinamento tem sua distribuição sumarizada e armazenada, para ser utilizada como referência. A distribuição referência é comparada à distribuição do subconjunto atual e, havendo diferenças significativas, a mudança de conceito é sinalizada. Os algoritmos *CoC* (Lee; Magoulès, 2012), *HDDDM* (Ditzler; Polikar, 2011), *PCA-detect* (KUNCHEVA, 2008) são representantes deste grupo.

Monitoramento dependente de modelo São aplicadas em conjunto com classificadores probabilísticos, o que permite realizar a detecção de mudanças de conceito através do monitoramento das probabilidades a posteriori (ZLIOBAITE, 2010). A utilização destas probabilidades diminui a incidência de falsos positivos e torna o processo computacionalmente eficiente, pois apenas um único fluxo univariado de valores é observado. Os métodos *A-distance* (DREDZE; OATES; PIATKO, 2010), *CDBD* (LINDSTROM; NAMEE; DELANY, 2013) e *Margin* (DRIES; RÜCKERT, 2009) integram este grupo.

Por fim, a Tabela 2.2 sumariza as categorias, os grupos e as respectivas técnicas mencionadas nesta seção.

O algoritmo desenvolvido neste projeto de mestrado enquadra-se na categoria de algoritmos implícitos ou não supervisionados, mais especificamente no grupo *Detecção de Novidades / Métodos de Agrupamento*. Na próxima seção, abordaremos as ferramentas utilizadas para implementação da técnica e sua validação.

Algoritmos Explícitos/Supervisionados	Métodos Baseados em Análise Sequencial	Cumulative Sum (CUSUM) PageHinkley (PH) (PAGE, 1954) Geometric Moving Average (GMA) (ROBERTS, 2000)
	Abordagens baseadas em Estatística	Drift Detection Method (DDM) (GAMA et al., 2004) Early Drift Detection Method (EDDM) (BAENA-GARCÍA et al., 2006) Exponentially Weighted Moving Average (EWMA) (ROSS et al., 2012) Reactive Drift Detection Method (RDDM) (BARROS et al., 2017)
	Métodos baseados em Janelas	Adaptive Windowing (ADWIN) (BIFET; GAVALDÀ, 2007) SeqDrift (PEARS; SAKTHITHASAN; KOH, 2014) HDDMA/HDDMW (BLANCO et al., 2015)
Algoritmos Implícitos/Não Supervisionados	Detecção de Novidade / Métodos de Agrupamento	OLINDDA (SPINOSA; CARVALHO; GAMA, 2007) MINAS (FARIA; GAMA; CARVALHO, 2013) Woo (RYU et al., 2012) DETECTNOD (Hayat; Hashemi, 2010) ECSMiner (MASUD et al., 2011) GC3 (SETHI; KANTARDZIC; HU, 2016)
	Monitoramento de distribuição multivariada	CoC (Lee; Magoulès, 2012) HDDDM (Ditzler; Polikar, 2011) PCA-detect (KUNCHEVA, 2008)
	Monitoramento dependente de modelo	A-distance (DREDZE; OATES; PIATKO, 2010) CDBD (LINDSTROM; NAMEE; DELANY, 2013) Margin (DRIES; RÜCKERT, 2009)

Tabela 2.2 Sumário - Abordagens de detecção (SETHI; KANTARDZIC, 2017)

2.3.3 Ferramentas

Nesta seção, os frameworks *Massive Online Analysis* (MOA) e *Tornado* serão apresentados. Estes frameworks possibilitam a implementação e a análise - perante o estado da arte - de algoritmos para detecção de mudanças de conceito. Ambos permitem a avaliação entre técnicas e a utilização de diferentes fontes de dados.

2.3.4 MOA

O *MOA – Massive Online Analysis*¹ é, atualmente, o principal framework para mineração de dados em fluxos contínuos. O projeto é de código-aberto² e apresenta uma comunidade bastante ativa e crescente (BIFET et al., 2010). A aplicação é composta por uma ampla coleção de algoritmos da área de aprendizado de máquina: classificação, regressão, agrupamento, busca por padrões, detecção de *outliers*, detecção de mudanças de conceito e sistemas de recomendação. Além das implementações, também estão disponíveis métodos para avaliação dessas técnicas. A aplicação é desenvolvida em Java, o que permite a sua execução nos principais sistemas operacionais, e dispõe de integração com o projeto WEKA (HALL et al., 2009).

O MOA divide suas funcionalidades em tarefas (*tasks*). As principais tarefas incluem: produção de fluxos de dados, treino de classificadores, aplicação de algoritmos de agrupamento, análise de algoritmos para detecção de *outliers* e de *concept drift*, dentre outras. As tarefas podem ser executadas a partir da interface gráfica (GUI) ou por linha de comando. A tela principal da aplicação é demonstrada na Figura 2.3. Através da interface gráfica é possível executar múltiplas tarefas de forma concorrente, controlar suas execuções e visualizar os resultados parciais.

¹<https://moa.cms.waikato.ac.nz/>

²<https://github.com/Waikato/moa>



Figura 2.3 MOA - Tela Inicial

A aplicação é capaz de ler arquivos em formato *ARFF*, populares por serem utilizados no projeto WEKA (HALL et al., 2009). A ferramenta também permite a produção de fluxos de dados dinamicamente, através de geradores. Alguns dos geradores de fluxo disponíveis no MOA são: *Random Trees* (DOMINGOS; HULTEN, 2000) *SEA* (STREET; KIM, 2001), *STAGGER* (SCHLIMMER; GRANGER, 1986), *Rotating Hyperplane* (WANG et al., 2003), *Random RBF*, *LED* (GAMA; ROCHA; MEDAS, 2003), *Waveform* (GAMA; ROCHA; MEDAS, 2003), e *Function* (JIN; AGRAWAL, 2003).

Outra funcionalidade importante do framework é a possibilidade de adicionar mudanças de conceito a fluxos estacionários existentes. Isto é feito através de uma função sigmóide, que modela o evento de mudança de conceito como uma combinação balanceada de duas distribuições homogêneas, que caracterizam os conceitos-alvo antes e depois da mudança (BIFET; KIRKBY, 2009). Além destes conceitos, o usuário também pode definir o momento da mudança e a sua duração (BIFET et al., 2010).

Os principais métodos para detecção de mudança de conceito propostos na literatura estão disponíveis no MOA. O framework também permite a utilização de classificadores do WEKA (HALL et al., 2009) combinados aos detectores. A janela para configuração de um detector é demonstrada na Figura 2.4.



Figura 2.4 MOA - Configuração detector

A arquitetura do framework é modular, o que permite aos usuários implementar novas tarefas com pouco esforço. Por exemplo, para criar um novo detector, basta estender a classe abstrata `moa.classifiers.core.driftdetection.AbstractChangeDetector` e implementar o algoritmo desejado. A janela de configuração para o detector (similar a 2.4) é criada dinamicamente, a partir dos atributos da classe.

O MOA dispõe de diversas classes para avaliação de técnicas de aprendizado de máquina. Para este trabalho, destacam-se as classes `DriftDetectionMethodClassifier` e `BasicConceptDriftPerformanceEvaluator`, que realizam a análise de algoritmos para detecção de mudança de conceito. A classe `DriftDetectionMethodClassifier` permite avaliar técnicas de detecção que encapsulam um classificador. Por sua vez, a classe `BasicConceptDriftPerformanceEvaluator` avalia a performance das técnicas de detecção diretamente, sem a necessidade de um classificador. Estes avaliadores e os seus indicadores serão detalhados juntamente com os resultados dos experimentos iniciais, na Seção 4.

2.3.5 Tornado

O Tornado é um framework para avaliação de algoritmos de detecção de mudança de conceito (PESARANGHADER, 2018). O projeto é desenvolvido na linguagem Python e seu código está disponível³. O framework diferencia-se do MOA por apresentar um cenário de avaliação específico: analisar a execução, em paralelo, de pares (classificador, detector de mudança de conceito), para identificar o par ótimo ao longo do tempo, em relação ao fluxo de dados.

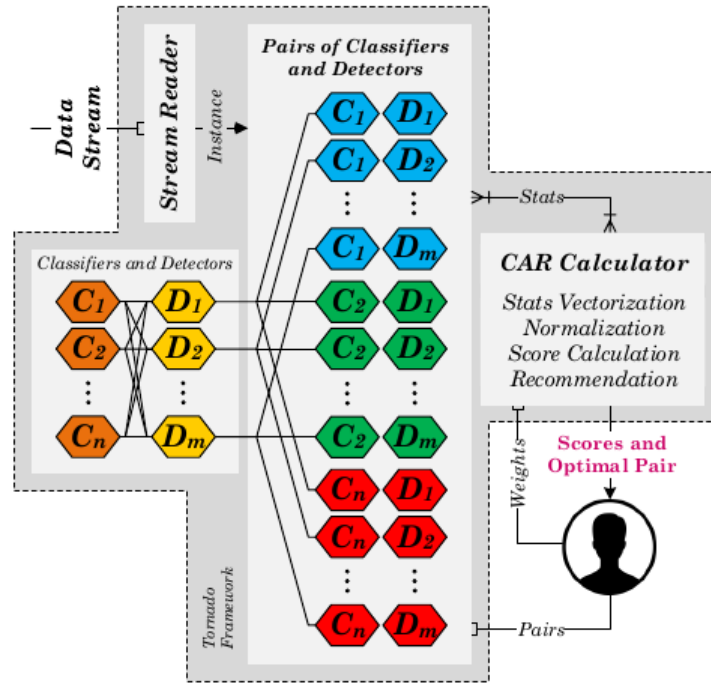


Figura 2.5 Framework Tornado (PESARANGHADER, 2018).

Conforme apresentado na Figura 2.5, os principais componentes do framework são: *Stream Reader*, *Classifiers*, *Detectors*, *Classifier-Detector Pairs* e *CAR Calculator*. A entrada de dados é composta por um fluxo (*Stream*), uma lista de pares (classificador, detector) e um vetor com pesos.

O componente *Stream Reader* lê instâncias e as encaminha para construção do modelo. Os modelos são construídos de forma incremental. Por seguir a abordagem *prequential*, cada instância é primeiramente utilizada para testes e depois como treinamento. Simultaneamente, os classificadores enviam suas estatísticas (taxas de erro ou resultados das predições) aos detectores, para que a mudança de conceito possa ser sinalizada. Por fim, o componente *CAR Calculator* calcula uma pontuação para cada par (classificador e detector), considerando taxas de erro, atraso para detecção da mudança de conceito, falsos positivos, falsos negativos, quantidade de memória utilizada e tempo de execução

³<https://github.com/alipsgh/tornado>

(PESARANGHADER, 2018). O framework apresenta ao usuário o par com maior pontuação em cada instante de execução. A avaliação do framework é necessária pois este par pode mudar ao longo do tempo, devido ao aprendizado incremental ou as mudanças de conceito. A Figura 2.6 apresenta um resultado obtido através do framework.

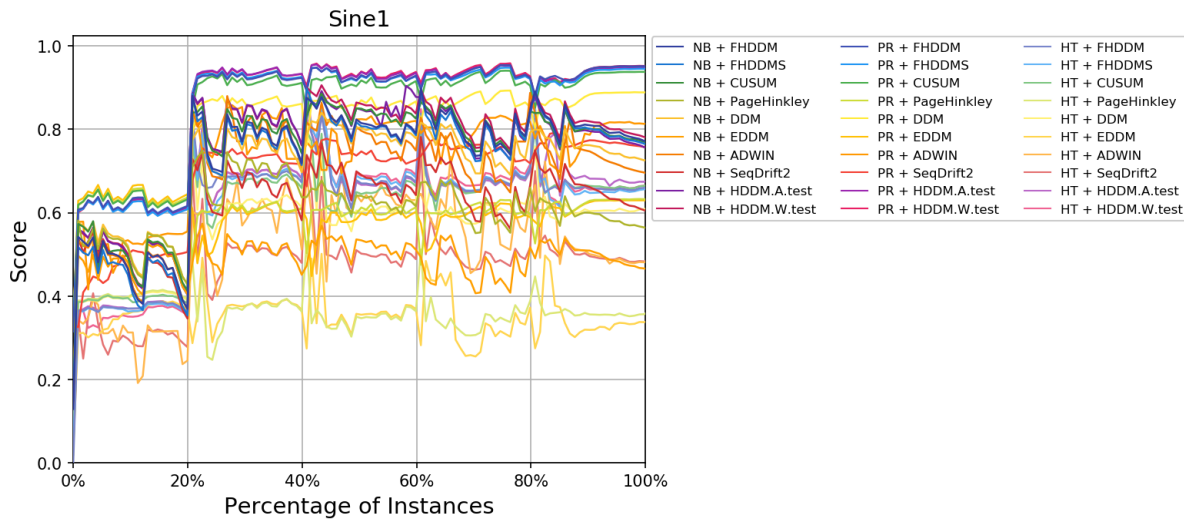


Figura 2.6 Tornado - Resultado para múltiplos pares (PESARANGHADER, 2018)

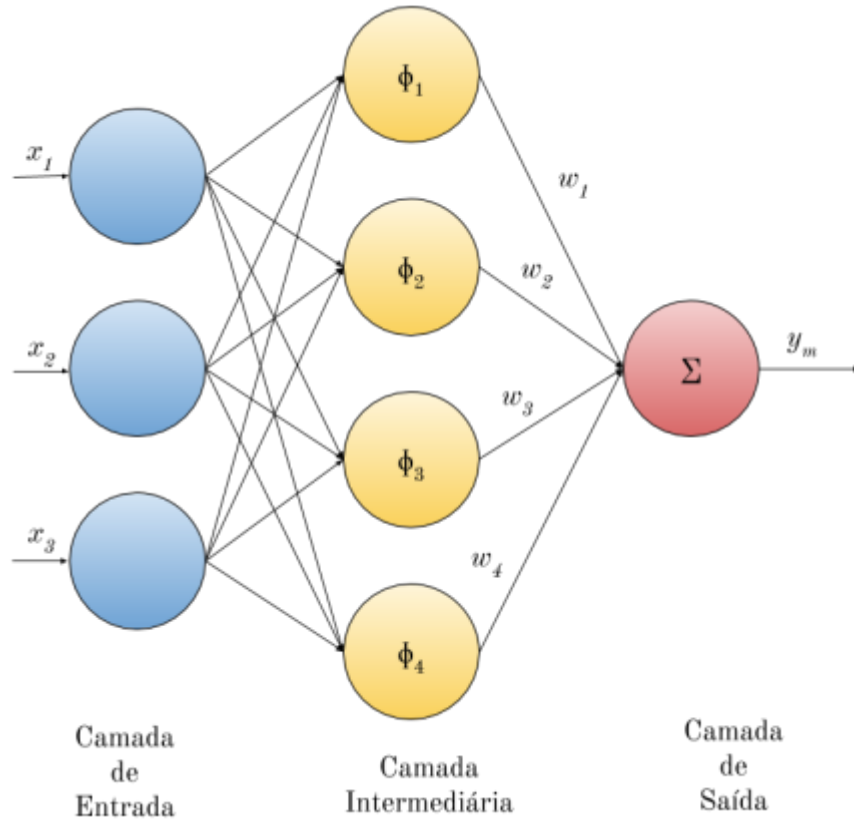
Neste trabalho, o algoritmo proposto foi implementado e testado nas duas ferramentas apresentadas. Os detalhes de implementação e os resultados desses testes são discutidos na Seção 4. A seguir, as redes de função de base radial serão detalhadas.

2.4 REDES DE FUNÇÃO DE BASE RADIAL

As Redes de Função de Base Radial (RBF *networks*) são aproximadoras universais de funções e têm como principal diferencial a sua forma de ativação, a qual é feita através do cálculo da distância entre os vetores de entrada e os centros estabelecidos (BRAGA; CARVALHO; LUDERMIR, 2007).

Em sua forma básica, a arquitetura de uma rede do tipo RBF é composta por três camadas: i) uma camada de entrada; ii) uma intermediária; e iii) uma de saída (ROJAS, 1996). A camada de entrada recebe os dados, a camada intermediária (oculta) utiliza funções de base radial para organizar os dados de entrada em grupos, transformando-os em um conjunto de valores linearmente separáveis. Por fim, a camada de saída realiza uma combinação linear para classificar os padrões recebidos a partir da camada intermediária (BRAGA; CARVALHO; LUDERMIR, 2007).

A Figura 2.7 demonstra essa arquitetura.

**Figura 2.7** Arquitetura RBF

Conforme a literatura, a função de base radial Gaussiana (Eq. 2.4) é a mais utilizada na camada intermediária:

$$\phi(v_i) = \exp\left(-\frac{\|v_i - c_i\|^2}{2\sigma^2}\right) \quad (2.4)$$

onde c_i representa o centro e σ é o parâmetro limitador do raio.

Na equação 2.4, $v = \|x - t_i\|$ é dado pela distância euclidiana, onde x é o valor de entrada da rede, enquanto t_i e σ correspondem respectivamente ao centro e a largura da função radial. Dessa maneira, a resolução de um determinado problema por uma rede do tipo RBF consiste na resolução das funções 2.5 e 2.6, obtendo o sistema 2.7.

$$f(x) = \sum_{i=1}^N w_{ij} \phi(\|\mathbf{x} - \mathbf{t}_i\|) \quad (2.5)$$

$$y_i = \sum_{i=1}^N w_{ij} \phi(\|\mathbf{x} - \mathbf{t}_i\|) + w_{j0} \quad (2.6)$$

$$\begin{bmatrix} \varphi(\|x_1 - t_1\|) & \varphi(\|x_1 - t_2\|) & \dots & \varphi(\|x_1 - t_N\|) \\ \varphi(\|x_2 - t_1\|) & \varphi(\|x_2 - t_2\|) & \dots & \varphi(\|x_2 - t_N\|) \\ \dots & \dots & \dots & \dots \\ \varphi(\|x_N - t_1\|) & \varphi(\|x_N - t_2\|) & \dots & \varphi(\|x_N - t_N\|) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \quad (2.7)$$

onde w_{ij} são os pesos de cada conexão, ϕ é a matriz de interpolação originada do conjunto de N funções de base radial aplicadas nas entradas x e dos seus respectivos centros t_i , w_{j_0} representa o bias, $\varphi(\|x - t_i\|)$ é o conjunto de N funções de base radial, $\|\dots\|$ é a norma euclidiana e y são as saídas geradas pela rede.

As camadas inicial e intermediária e suas propriedades de agrupamento são utilizadas como base para o algoritmo de detecção de mudança de conceito proposto neste projeto de mestrado.

2.5 TRABALHOS RELACIONADOS

Além das referências básicas apresentadas neste capítulo, foi realizada uma pesquisa na literatura visando identificar trabalhos que também propõem a identificação de mudanças de conceito em fluxos contínuos de dados através da aplicação de redes de função de base radial ou que apliquem técnicas similares.

Em (JIANPING; VENKATESWARLU, 2002) redes de função de base radial, com funções gaussianas, são utilizadas para detecção de novidades. A técnica proposta atua sobre cenários estacionários, mais especificamente o problema de identificação de falas. Durante a preparação da rede, o algoritmo *k-means* é utilizado para definir os centros e as matrizes de covariância.

Roberts e Penny (Roberts; Penny; Pillot, 1996) propõem um método para detecção de novidade baseado no monitoramento das taxas de erro e confiança, utilizando um comitê de redes de função de base radial. Cada rede é inicializada com um vetor de pesos diferente. A taxa de erro final é calculada a partir da matriz de covariância de erro do comitê criado. Esta abordagem foi testada na classificação de pacientes com problemas de tremor muscular.

As redes de função de base radial também foram aplicadas para detecção de anomalias (BAZARGANI; NAMEE, 2018). Neste trabalho, às funções de perda (*loss*) das redes são modificadas, tornando-as classificadores de uma única classe. Estas modificações permitem às redes identificar exemplos divergentes dos padrões conhecidos.

Este projeto de mestrado se diferencia dos trabalhos mencionados por utilizar apenas as camadas de entrada e intermediária das redes de função de base radial para detecção de mudanças de conceito. Além disso, etapas como a escolha dos centros e o cálculo do tamanho do raio são realizadas de forma dinâmica, tornando possível a sua aplicação em cenários com fluxos contínuos de dados.

2.6 CONSIDERAÇÕES FINAIS

Neste capítulo foram apresentados os principais conceitos utilizados neste projeto de mestrado. Discutiu-se a aplicação de técnicas de aprendizado de máquina à fluxos contínuos

de dados, o fenômeno mudança de conceito, suas técnicas de detecção e ferramentas, redes de função de base radial e os principais trabalhos relacionados encontrados na literatura. No próximo capítulo, o plano de pesquisa será detalhado.

PLANO DE PESQUISA

3.1 CONSIDERAÇÕES INICIAIS

Este capítulo descreve como a pesquisa proposta neste mestrado será desenvolvida. O objetivo principal é a implementação de um algoritmo para detecção de mudanças de conceito em fluxos de dados contínuos utilizando as camadas inicial e intermediária das redes da função de base radial. A seguir, são apresentados detalhes sobre cada etapa do desenvolvimento do projeto.

3.2 DESCRIÇÃO DO PROBLEMA

A partir da Teoria Bayesiana de Decisão, apresentada na Seção 2.3, é possível afirmar que existe mudança de conceito entre os instantes t_0 e t_1 se:

$$\exists X : p_{t_0}(X, c) \neq p_{t_1}(X, c) \quad (3.1)$$

onde, p_{t_0} e p_{t_1} denotam as distribuições de probabilidades conjuntas nos instantes t_0 e t_1 , respectivamente, para X e c (GAMA et al., 2014). Isto é, um conjunto de exemplos possui rótulos de classe legítimos em t_0 , mas passa a ter rótulos diferentes, também legítimos, em t_1 (KOLTER; MALOOF, 2007). Esta mudança pode ocorrer devido a alterações no contexto do processo gerador ou na distribuição dos dados, e pode impactar a acurácia de modelos de decisão utilizados.

Considerando que a camada intermediária (oculta) de uma rede de função de base radial realiza o agrupamento dos dados de entrada em clusters, transformando padrões de entrada não linearmente separáveis em um conjunto de valores linearmente separáveis (JIANPING; VENKATESWARLU, 2002), este projeto de mestrado tem como objetivo comprovar a hipótese que mudanças de conceito em fluxos contínuos de dados podem ser detectadas através de redes de função de base radial, de forma online e independente de exemplos prévios ou rótulos.

Para exemplificar a execução desta proposta de mestrado, considere o conjunto $D = \{0.1, 0.13, 0.14, 0.4, 0.5, 0.6, 0.16, 0.14\}$, recorte do fluxo contínuo utilizado como entrada.

3.3 ATIVIDADES DE PESQUISA

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

3.4 CONSIDERAÇÕES FINAIS

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

EXPERIMENTOS INICIAIS

4.1 CONSIDERAÇÕES INICIAIS

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

4.2 CONFIGURAÇÃO DOS EXPERIMENTOS

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

4.3 MÉTODO DE PETTITT

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut

porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

4.4 REDES DE FUNÇÃO DE BASE RADIAL

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

4.5 CONSIDERAÇÕES FINAIS

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

REFERÊNCIAS BIBLIOGRÁFICAS

- ACKERMANN, M. R. et al. Streamkm++: A clustering algorithm for data streams. *J. Exp. Algorithmics*, ACM, New York, NY, USA, v. 17, p. 2.4:2.1–2.4:2.30, maio 2012. ISSN 1084-6654. Disponível em: <http://doi.acm.org/10.1145/2133803.2184450>.
- AGGARWAL, C. C. *Data Streams: Models and Algorithms (Advances in Database Systems)*. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN 0387287590.
- AGGARWAL, C. C. et al. A framework for clustering evolving data streams. In: *Proceedings of the 29th International Conference on Very Large Data Bases - Volume 29*. VLDB Endowment, 2003. (VLDB '03), p. 81–92. ISBN 0-12-722442-4. Disponível em: <http://dl.acm.org/citation.cfm?id=1315451.1315460>.
- AGGARWAL, C. C. et al. On demand classification of data streams. In: *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, 2004. (KDD '04), p. 503–508. ISBN 1-58113-888-1. Disponível em: <http://doi.acm.org/10.1145/1014052.1014110>.
- AMINIKHANGHAHI, S.; COOK, D. J. A survey of methods for time series change point detection. *Knowl. Inf. Syst.*, Springer-Verlag New York, Inc., New York, NY, USA, v. 51, n. 2, p. 339–367, maio 2017. ISSN 0219-1377. Disponível em: <https://doi.org/10.1007/s10115-016-0987-z>.
- ANKERST, M. et al. Optics: Ordering points to identify the clustering structure. *SIGMOD Rec.*, ACM, New York, NY, USA, v. 28, n. 2, p. 49–60, jun. 1999. ISSN 0163-5808. Disponível em: <http://doi.acm.org/10.1145/304181.304187>.
- Bach, S. H.; Maloof, M. A. Paired learners for concept drift. In: *2008 Eighth IEEE International Conference on Data Mining*. [S.l.: s.n.], 2008. p. 23–32. ISSN 1550-4786.
- BAENA-GARCÍA, M. et al. Early drift detection method. In: *In Fourth International Workshop on Knowledge Discovery from Data Streams*. [S.l.: s.n.], 2006.
- BARBARÁ, D. Requirements for clustering data streams. *SIGKDD Explor. Newsl.*, ACM, New York, NY, USA, v. 3, n. 2, p. 23–27, jan. 2002. ISSN 1931-0145. Disponível em: <http://doi.acm.org/10.1145/507515.507519>.
- BARROS, R. S. M. de et al. RDDM: reactive drift detection method. *Expert Syst. Appl.*, v. 90, p. 344–355, 2017.
- BASSEVILLE, M.; NIKIFOROV, I. V. *Detection of Abrupt Changes: Theory and Application*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993. ISBN 0-13-126780-9.

BAZARGANI, M. H.; NAMEE, B. M. Radial basis function data descriptor (rbfdd) network : An anomaly detection approach. In: . [S.l.: s.n.], 2018.

BIFET, A. Adaptive learning and mining for data streams and frequent patterns. *SIGKDD Explor. Newsl.*, ACM, New York, NY, USA, v. 11, n. 1, p. 55–56, nov. 2009. ISSN 1931-0145. Disponível em: <http://doi.acm.org/10.1145/1656274.1656287>.

BIFET, A.; GAVALDÀ, R. Learning from time-changing data with adaptive windowing. In: *SDM*. SIAM, 2007. p. 443–448. ISBN 978-1-61197-277-1. Disponível em: <http://dblp.uni-trier.de/db/conf/sdm/sdm2007.html/#BifetG07>.

BIFET, A. et al. Moa: Massive online analysis. *J. Mach. Learn. Res.*, JMLR.org, v. 11, p. 1601–1604, ago. 2010. ISSN 1532-4435. Disponível em: <http://dl.acm.org/citation.cfm?id=1756006.1859903>.

BIFET, A.; KIRKBY, R. Data stream mining a practical approach. Citeseer, 2009.

BIFET, A. et al. Efficient data stream classification via probabilistic adaptive windows. In: *Proceedings of the 28th Annual ACM Symposium on Applied Computing*. New York, NY, USA: ACM, 2013. (SAC '13), p. 801–806. ISBN 978-1-4503-1656-9. Disponível em: <http://doi.acm.org/10.1145/2480362.2480516>.

BLANCO, I. I. F. et al. Online and non-parametric drift detection methods based on hoeffding's bounds. *IEEE Trans. Knowl. Data Eng.*, v. 27, n. 3, p. 810–823, 2015.

BRAGA, A.; CARVALHO, A. C.; LUDERMIR, T. B. *Redes Neurais Artificiais: Teoria e aplicações*. LTC Editora, 2007. ISBN 9788521615644. Disponível em: <http://www.worldcat.org/isbn/9788521615644>.

BREIMAN, L. et al. *Classification and Regression Trees*. Monterey, CA: Wadsworth and Brooks, 1984.

CAO, F. et al. Density-based clustering over an evolving data stream with noise. In: GHOSH, J. et al. (Ed.). *SDM*. SIAM, 2006. p. 328–339. ISBN 978-1-61197-276-4. Disponível em: <http://dblp.uni-trier.de/db/conf/sdm/sdm2006.html/#CaoEQZ06>.

CHANDOLA, V.; BANERJEE, A.; KUMAR, V. Anomaly detection: A survey. *ACM Comput. Surv.*, ACM, New York, NY, USA, v. 41, n. 3, p. 15:1–15:58, jul. 2009. ISSN 0360-0300. Disponível em: <http://doi.acm.org/10.1145/1541880.1541882>.

COHEN, J. et al. Mad skills: New analysis practices for big data. *Proc. VLDB Endow.*, VLDB Endowment, v. 2, n. 2, p. 1481–1492, ago. 2009. ISSN 2150-8097. Disponível em: <https://doi.org/10.14778/1687553.1687576>.

Ditzler, G.; Polikar, R. Hellinger distance based drift detection for nonstationary environments. In: *2011 IEEE Symposium on Computational Intelligence in Dynamic and Uncertain Environments (CIDUE)*. [S.l.: s.n.], 2011. p. 41–48.

DOMINGOS, P.; HULTEN, G. Mining high-speed data streams. In: *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, 2000. (KDD '00), p. 71–80. ISBN 1-58113-233-6. Disponível em: <http://doi.acm.org/10.1145/347090.347107>.

DREDZE, M.; OATES, T.; PIATKO, C. We're not in kansas anymore: Detecting domain changes in streams. In: . [s.n.], 2010. p. 585–595. Cited By 13. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-80053285630&partnerID=40&md5=2eea89f635e2cbc0920069028e9f7746>.

DRIES, A.; RÜCKERT, U. Adaptive concept drift detection. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, v. 2, n. 5-6, p. 311–327, 2009. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sam.10054>.

DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern Classification (2Nd Edition)*. New York, NY, USA: Wiley-Interscience, 2000. ISBN 0471056693.

ESTER, M. et al. A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise. In: *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. AAAI Press, 1996. (KDD'96), p. 226–231. Disponível em: <http://dl.acm.org/citation.cfm?id=3001460.3001507>.

FARIA, E. R.; GAMA, J. a.; CARVALHO, A. C. P. L. F. Novelty detection algorithm for data streams multi-class problems. In: *Proceedings of the 28th Annual ACM Symposium on Applied Computing*. New York, NY, USA: ACM, 2013. (SAC '13), p. 795–800. ISBN 978-1-4503-1656-9. Disponível em: <http://doi.acm.org/10.1145/2480362.2480515>.

GAMA, J. *Knowledge Discovery from Data Streams*. 1st. ed. [S.l.]: Chapman & Hall/-CRC, 2010. ISBN 1439826110, 9781439826119.

GAMA, J. et al. Learning with drift detection. In: BAZZAN, A. L. C.; LABIDI, S. (Ed.). *SBIA*. Springer, 2004. (Lecture Notes in Computer Science, v. 3171), p. 286–295. ISBN 3-540-23237-0. Disponível em: <http://dblp.uni-trier.de/db/conf/sbia/sbia2004.html/#GamaMCR04>.

GAMA, J. a.; ROCHA, R.; MEDAS, P. Accurate decision trees for mining high-speed data streams. In: *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, 2003. (KDD '03), p. 523–528. ISBN 1-58113-737-0. Disponível em: <http://doi.acm.org/10.1145/956750.956813>.

GAMA, J. a. et al. A survey on concept drift adaptation. *ACM Comput. Surv.*, ACM, New York, NY, USA, v. 46, n. 4, p. 44:1–44:37, mar. 2014. ISSN 0360-0300. Disponível em: <http://doi.acm.org/10.1145/2523813>.

GONÇALVES, P. M. et al. A comparative study on concept drift detectors. *Expert Systems with Applications*, v. 41, n. 18, p. 8144 – 8156, 2014. ISSN 0957-4174. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0957417414004175>.

HALL, M. et al. The weka data mining software: An update. *SIGKDD Explor. Newsl.*, ACM, New York, NY, USA, v. 11, n. 1, p. 10–18, nov. 2009. ISSN 1931-0145. Disponível em: <http://doi.acm.org/10.1145/1656274.1656278>.

Hayat, M. Z.; Hashemi, M. R. A dct based approach for detecting novelty and concept drift in data streams. In: *2010 International Conference of Soft Computing and Pattern Recognition*. [S.l.: s.n.], 2010. p. 373–378.

JAIN, A. K.; DUBES, R. C. *Algorithms for Clustering Data*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1988. ISBN 0-13-022278-X.

JIANPING, D.; VENKATESWARLU, R. Speaker recognition using radial basis function neural networks. In: ABRAHAM, A.; KÖPPEN, M. (Ed.). *Hybrid Information Systems*. Heidelberg: Physica-Verlag HD, 2002. p. 57–64. ISBN 978-3-7908-1782-9.

JIN, R.; AGRAWAL, G. Efficient decision tree construction on streaming data. In: *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, 2003. (KDD '03), p. 571–576. ISBN 1-58113-737-0. Disponível em: <http://doi.acm.org/10.1145/956750.956821>.

KAUFMAN, L.; ROUSSEEUW, P. *Finding Groups in Data: an introduction to cluster analysis*. [S.l.]: Wiley, 1990.

KOLTER, J. Z.; MALOOF, M. A. Dynamic weighted majority: An ensemble method for drifting concepts. *J. Mach. Learn. Res.*, JMLR.org, v. 8, p. 2755–2790, dez. 2007. ISSN 1532-4435. Disponível em: <http://dl.acm.org/citation.cfm?id=1314498.1390333>.

KOTSIANTIS, S. B. Supervised machine learning: A review of classification techniques. In: *Proceedings of the 2007 Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies*. Amsterdam, The Netherlands, The Netherlands: IOS Press, 2007. p. 3–24. ISBN 978-1-58603-780-2. Disponível em: <http://dl.acm.org/citation.cfm?id=1566770.1566773>.

KRANEN, P. et al. The clustree: Indexing micro-clusters for anytime stream mining. *Knowl. Inf. Syst.*, Springer-Verlag New York, Inc., New York, NY, USA, v. 29, n. 2, p. 249–272, nov. 2011. ISSN 0219-1377. Disponível em: <http://dx.doi.org/10.1007/s10115-010-0342-8>.

KUNCHEVA, L. Classifier ensembles for detecting concept change in streaming data: Overview and perspectives. *Proc. Eur. Conf. Artif. Intell.*, p. 5–10, 2008. Cited By 70.

Lee, J.; Magoulès, F. Detection of concept drift for learning from stream data. In: *2012 IEEE 14th International Conference on High Performance Computing and Communication 2012 IEEE 9th International Conference on Embedded Software and Systems*. [S.l.: s.n.], 2012. p. 241–245.

LINDSTROM, P.; NAMEE, B. M.; DELANY, S. J. Drift detection using uncertainty distribution divergence. *Evolving Systems*, v. 4, n. 1, p. 13–25, Mar 2013. ISSN 1868-6486. Disponível em: <https://doi.org/10.1007/s12530-012-9061-6>.

LING, C.; LING-JUN, Z.; LI, T. Stream data classification using improved fisher discriminate analysis. *Journal of Computers*, 01 2009.

LLOYD, S. Least squares quantization in pcm. *IEEE Trans. Inf. Theor.*, IEEE Press, Piscataway, NJ, USA, v. 28, n. 2, p. 129–137, set. 2006. ISSN 0018-9448. Disponível em: <http://dx.doi.org/10.1109/TIT.1982.1056489>.

MASUD, M. et al. Classification and novel class detection in concept-drifting data streams under time constraints. *IEEE Trans. on Knowl. and Data Eng.*, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 23, n. 6, p. 859–874, jun. 2011. ISSN 1041-4347. Disponível em: <http://dx.doi.org/10.1109/TKDE.2010.61>.

MITCHELL, T. M. *Machine Learning*. 1. ed. New York, NY, USA: McGraw-Hill, Inc., 1997. ISBN 0070428077, 9780070428072.

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.

PAGE, E. S. Continuous Inspection Schemes. *Biometrika*, Biometrika Trust, v. 41, n. 1/2, p. 100–115, 1954. ISSN 00063444. Disponível em: <http://dx.doi.org/10.2307/2333009>.

PEARS, R.; SAKTHITHASAN, S.; KOH, Y. S. Detecting concept change in dynamic data streams - A sequential approach based on reservoir sampling. *Machine Learning*, v. 97, n. 3, p. 259–293, 2014.

PESARANGHADER, A. *A Reservoir of Adaptive Algorithms for Online Learning from Evolving Data Streams*. Université d'Ottawa / University of Ottawa, 2018. Disponível em: <http://ruor.uottawa.ca/handle/10393/38190>.

Roberts, S. J.; Penny, W.; Pillot, D. Novelty, confidence and errors in connectionist systems. In: *IEE Colloquium on Intelligent Sensors (Digest No: 1996/261)*. [S.l.: s.n.], 1996. p. 10/1–10/6.

ROBERTS, S. W. Control chart tests based on geometric moving averages. *Technometrics*, American Society for Quality Control and American Statistical Association, Alexandria, Va, USA, v. 42, n. 1, p. 97–101, fev. 2000. ISSN 0040-1706. Disponível em: <http://dx.doi.org/10.2307/1271439>.

ROJAS, R. *Neural Networks: A Systematic Introduction*. Berlin, Heidelberg: Springer-Verlag, 1996. ISBN 3-540-60505-3.

ROSS, G. J. et al. Exponentially weighted moving average charts for detecting concept drift. *Pattern Recogn. Lett.*, Elsevier Science Inc., New York, NY, USA, v. 33, n. 2, p. 191–198, jan. 2012. ISSN 0167-8655. Disponível em: <http://dx.doi.org/10.1016/j.patrec.2011.08.019>.

RYU, J. W. et al. An efficient method of building an ensemble of classifiers in streaming data. In: SRINIVASA, S.; BHATNAGAR, V. (Ed.). *Big Data Analytics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. p. 122–133. ISBN 978-3-642-35542-4.

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>.

SEBASTIÃO, R. et al. Monitoring incremental histogram distribution for change detection in data streams. In: GABER, M. M. et al. (Ed.). *Knowledge Discovery from Sensor Data*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. p. 25–42. ISBN 978-3-642-12519-5.

SETHI, T. S.; KANTARDZIC, M. On the reliable detection of concept drift from streaming unlabeled data. *Expert Syst. Appl.*, Pergamon Press, Inc., Tarrytown, NY, USA, v. 82, n. C, p. 77–99, out. 2017. ISSN 0957-4174. Disponível em: <https://doi.org/10.1016/j.eswa.2017.04.008>.

SETHI, T. S.; KANTARDZIC, M.; HU, H. A grid density based framework for classifying streaming data in the presence of concept drift. *Journal of Intelligent Information Systems*, v. 46, n. 1, p. 179–211, Feb 2016.

SOBHANI, P.; BEIGY, H. New drift detection method for data streams. In: *Proceedings of the Second International Conference on Adaptive and Intelligent Systems*. Berlin, Heidelberg: Springer-Verlag, 2011. (ICAIS'11), p. 88–97. ISBN 978-3-642-23856-7. Disponível em: <http://dl.acm.org/citation.cfm?id=2045295.2045309>.

SPINOSA, E. J.; CARVALHO, A. P. de Leon F. de; GAMA, J. a. Olindda: A cluster-based approach for detecting novelty and concept drift in data streams. In: *Proceedings of the 2007 ACM Symposium on Applied Computing*. New York, NY, USA: ACM, 2007. (SAC '07), p. 448–452. ISBN 1-59593-480-4. Disponível em: <http://doi.acm.org/10.1145/1244002.1244107>.

STREET, W. N.; KIM, Y. A streaming ensemble algorithm (sea) for large-scale classification. In: *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, 2001. (KDD '01), p. 377–382. ISBN 1-58113-391-X. Disponível em: <http://doi.acm.org/10.1145/502512.502568>.

VAPNIK, V. N. *Statistical Learning Theory*. [S.l.]: Wiley-Interscience, 1998.

WANG, H. et al. Mining concept-drifting data streams using ensemble classifiers. In: *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, 2003. (KDD '03), p. 226–235. ISBN 1-58113-737-0. Disponível em: <http://doi.acm.org/10.1145/956750.956778>.

ZLIOBAITE, I. Learning under concept drift: an overview. *CoRR*, abs/1010.4784, 2010. Disponível em: <http://arxiv.org/abs/1010.4784>.

ZWOLENSKI, M.; WEATHERILL, L. The digital universe rich data and the increasing value of the internet of things. *Australian Journal of Telecommunications and the Digital Economy*, v. 2, 10 2014.