



CHARLES CARVALHO DE AGUIAR

**UNSUPERVISED FUZZY EIX:
CLUSTERIZAÇÃO INTERNA-EXTERNA FUZZY EVOLUTIVA
DE FLUXOS DE DADOS NÃO-ESTACIONÁRIOS**

LAVRAS – MG

2022

CHARLES CARVALHO DE AGUIAR

UNSUPERVISED FUZZY EIX:

**CLUSTERIZAÇÃO INTERNA-EXTERNA FUZZY EVOLUTIVA DE FLUXOS DE DADOS
NÃO-ESTACIONÁRIOS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia de Sistemas e Automação, área de concentração em Inteligência Computacional, para a obtenção do título de Mestre.

Prof. DSc. Daniel Leite

Orientador

LAVRAS – MG

2022

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

de Aguiar, Charles Carvalho

Unsupervised Fuzzy eIX : Clusterização Interna-Externa
Fuzzy Evolutiva de Fluxos de Dados Não-Estacionários /
Charles Carvalho de Aguiar. – Lavras : UFLA, 2022.

62 p. : il.

Dissertação (Mestrado Acadêmico)–Universidade Federal
de Lavras, 2022.

Orientador: Prof. DSc. Daniel Leite.

Bibliografia.

1. Aprendizado não supervisionado. 2. Sistema Fuzzy
Evolutivo. 3. Computação Granular. I. Leite, Daniel Furtado.
II. Título.

CHARLES CARVALHO DE AGUIAR

**UNSUPERVISED FUZZY EIX: CLUSTERIZAÇÃO INTERNA-EXTERNA FUZZY
EVOLUTIVA DE FLUXOS DE DADOS NÃO-ESTACIONÁRIOS
UNSUPERVISED FUZZY EIX: EVOLVING INTERNAL-EXTERNAL FUZZY
CLUSTERING FOR NON-STATIONARY ONLINE DATA STREAMS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia de Sistemas e Automação, área de concentração em Inteligência Computacional, para a obtenção do título de Mestre.

APROVADA em 31 de Janeiro de 2022.

DSc. Daniel Furtado Leite UFLA
DSc. Luiz Alberto Queiroz Cordovil Junior UFMG
DSc. Murilo Cesar Osorio Camargos Filho LANCASTER

Prof. DSc. Daniel Leite
Orientador

**LAVRAS – MG
2022**

RESUMO

Modelos classificadores com fronteiras de decisão variantes no tempo, isto é, classificadores evolutivos, desempenham um papel importante em um cenário em que as informações só podem ser descobertas a partir de um fluxo de dados online. Este texto apresenta um novo método de aprendizado não supervisionado para dados numéricos chamado *evolving Internal-eXternal Fuzzy Clustering* (Fuzzy eIX). Desenvolve-se a noção de grânulos fuzzy de dupla fronteira e elabora-se sobre suas implicações. Sistemas de inferência fuzzy evolutivos tipo 1 e tipo 2 podem ser obtidos a partir de projeções de grânulos fuzzy eIX em eixos ortogonais correspondentes às dimensões do problema. O algoritmo de aprendizado de modelos Fuzzy eIX é guiado pelo princípio da Granularidade Balanceada de Pedrycz para obter um nível mais alto de compreensibilidade e interpretabilidade do modelo em um domínio. Grânulos internos e externos são atualizados a partir de um fluxo de dados numérico. Concomitantemente, a estrutura granular global do classificador é autonomamente construída e adaptada. Um problema sintético preliminar chamado Rotação das Gaussianas Gêmeas mostra o comportamento do classificador frente a um fluxo de dados não estacionário. Além disso, o desempenho do método é comparado àqueles de outros métodos já estabelecidos na literatura quanto à classificação de bases de dados *benchmark* em aprendizado de máquina online. As comparações levam em conta a qualidade das partições geradas por meio de índices de validação incrementais, a acurácia e a compactação da estrutura de regras resultante.

Palavras-chave: Aprendizado não supervisionado. Sistema Fuzzy Evolutivo. Computação Granular. Fluxos de Dados Online.

ABSTRACT

Classifiers with time-varying decision boundaries, namely, evolving classifiers, play an important role in a scenario in which information is available as an online data stream. This text presents a new unsupervised learning method for numerical data called evolving Internal-eXternal Fuzzy clustering method (Fuzzy eIX). The notion of double-boundary fuzzy granules and some of its implications are developed and explored. It will be shown how type 1 and type 2 fuzzy inference systems can be obtained from the projection of Fuzzy eIX granules on orthogonal axes corresponding to the dimensions of a problem. Fuzzy eIX learning algorithm performs Pedrycz Balanced Information Granularity principle within fuzzy eIX classifiers to achieve a higher level of model understandability in a given problem domain. Internal and external granules are updated from a numerical data stream at the same time that the global granular structure of the classifier is autonomously evolved. A synthetic preliminary problem called Rotation of Twin Gaussians shows the behavior of the classifier for a nonstationary data stream input. Additionally, the performance of the Fuzzy eIX method will be compared to other two evolving methods already established in the literature when it comes to the classification of *benchmark* data sets usually employed in online machine learning models assessments. Comparisons will also be conducted in terms of partition quality through incremental cluster validation indexes, the accuracy and compactness of the resulting rules structure.

Keywords: Unsupervised Learning. Evolving Fuzzy System. Granular Computing. Online Data Stream.

LISTA DE FIGURAS

Figura 2.1 – Mapeamento granular em modelos IBeM: (a) Limites inferiores e superiores do mapeamento; (b) Grânulos intervalares.	18
Figura 2.2 – Funções de pertinência com incertezas associadas e representação gráfica da FOU	20
Figura 3.1 – Derivação de funções de pertinência evolutivas do tipo 1 e de um sistema de inferência Fuzzy evolutivo do tipo 1 a partir de grânulos Fuzzy eIX . . .	28
Figura 3.2 – Derivação de funções de pertinência evolutivas do tipo 2 e de um sistema de inferência Fuzzy evolutivo do tipo 2 a partir de grânulos Fuzzy eIX . . .	29
Figura 3.3 – Informação sobre incertezas locais e inter-granulares em grânulos Fuzzy eIX	30
Figura 4.1 – Representação de um grânulo Fuzzy eIX bidimensional genérico	32
Figura 4.2 – Encolhendo e arrastando um grânulo Fuzzy eIX após processamento de instância pertencente à região interna	34
Figura 4.3 – Expandindo um grânulo Fuzzy eIX após processamento de instância pertencente à região externa	36
Figura 4.4 – Balanceamento de granularidade em grânulos Fuzzy eIX bidimensionais . .	37
Figura 4.5 – Exemplo do efeito da mesclagem de grânulos Fuzzy eIX	39
Figura 4.6 – Abordagem de mesclagem de grânulos Fuzzy eIX por envoltório convexo .	40
Figura 4.7 – Algoritmo de aprendizagem: Fuzzy eIX	41
Figura 5.1 – Experimento das Gaussianas Rotativas em torno do ponto (5,5)	42
Figura 6.1 – Posição final dos grânulos gerados no estágio estacionário e respectiva fronteira de decisão.	48
Figura 6.2 – Posição final dos grânulos gerados no estágio não-estacionário e respectiva fronteira de decisão.	48
Figura 6.3 – Evolução estrutural do modelo classificador Fuzzy eIX..	49

LISTA DE TABELAS

Tabela 6.1 – Fuzzy eIX - Rotação de Gaussianas Gêmeas	47
---	----

LISTA DE QUADROS

Quadro 8.1 – Cronograma de Atividades	55
---	----

SUMÁRIO

1	INTRODUÇÃO	9
1.1	Objetivos	11
1.2	Contribuições	12
1.3	Publicações Diretas e Indiretas	12
1.4	Estrutura do Trabalho	13
2	SISTEMAS INTELIGENTES EVOLUTIVOS	14
2.1	Computação Granular	16
2.2	Sistemas Granulares Evolutivos	17
2.3	Modelagem Evolutiva Granular Baseada em Intervalos	17
2.4	Sistemas Evolutivos Fuzzy Tipo 2	19
2.5	Sistemas de Inferência Fuzzy Tipo 2	19
2.6	Classificador Fuzzy Tipo 2 Evolutivo eT2Class	20
2.7	Tipicidade e Excentricidade	21
2.8	TEDA-Cluster	22
2.9	Validação de Modelos para Classificação Não Supervisionada	23
2.10	Índice de Validação Xie-Beni	24
2.11	Índice de Validação Davies-Bouldin	24
2.12	Formulação de Índices de Validação Incrementais	25
2.13	Índice Xie-Beni Incremental	26
2.14	Índice Davies-Bouldin Incremental	26
3	DECORRÊNCIAS DE MODELOS FUZZY eIX	28
3.1	Derivação de Sistemas de Inferência Fuzzy Evolutivos do Tipo 1	28
3.2	Derivação de Sistemas de Inferência Fuzzy Evolutivos do Tipo 2	28
3.3	Análise de Incertezas Inter-Granulares Local e Global	29
4	ALGORITMO DE APRENDIZADO FUZZY eIX	31
4.1	Descrição do método Fuzzy eIX	31
4.2	Inicialização de Grânulos	31
4.3	Instância em Região Interna	32
4.4	Instância em Região Externa	34
4.5	Instância Excêntrica	36
4.6	Princípio da Granularidade Balanceada	36

4.7	Mesclagem de Grânulos	38
4.8	Mesclagem por Média Ponderada	39
4.9	Mesclagem por Envoltório Convexo	40
4.10	Sumário	41
5	METODOLOGIA	42
5.1	Rotação de Gaussianas Gêmeas	42
5.2	Clusterização de Conjuntos de Dados Reais	43
5.3	Software, Hardware e Avaliação de Desempenho	45
6	RESULTADOS	46
6.1	Rotação de Gaussianas Gêmeas	46
6.2	Clusterização de Conjuntos de Dados Reais	49
7	CONCLUSÕES	51
7.1	Alternativas em ambiente eIX	52
8	CRONOGRAMA DE ATIVIDADES	54
9	ORÇAMENTO	56
	REFERÊNCIAS	57

1 INTRODUÇÃO

Fluxos de dados provenientes de processos reais mudam ao longo do tempo, o que reflete a evolução do ambiente em que estão inseridos, a mudança dos parâmetros inerentes a um sistema, e/ou a variação da intensidade da interação mútua entre sistemas adjacentes. Para conceber, representar e lidar com as informações contidas em tais fluxos de dados, um modelo computacional deve ser capaz de se adaptar em resposta às mudanças. A capacidade de auto-ajuste consiste na atualização dos parâmetros e da estrutura de um modelo com a finalidade de responder a eventos ou padrões comportamentais desconhecidos e à alterações graduais ou abruptas das condições operacionais do sistema. Algoritmos de inferência e aprendizado de máquina capazes de perceber mudanças no fluxos de dados de entrada e são úteis para rearranjar ou adaptar os parâmetros e estrutura de um modelo para manter uma representação apropriada do perfil dos fluxos de dados (ŠKRJANC et al., 2019; ANGELOV, 2013; LEITE; ŠKRJANC; GOMIDE, 2020; LEITE, 2012).

Modelos fuzzy podem ser construídos e adaptados por meio de algoritmos incrementais que atuam de maneira online. Como discutido em (ŠKRJANC et al., 2019; ANGELOV, 2013; LEITE; ŠKRJANC; GOMIDE, 2020; LEITE, 2012; BOUCHACHIA; GABRYS; SAHEL, 2007; CORDOVIL et al., 2019), os modelos evolutivos são caracterizados por:

1. habilidade de aprendizado online, i.e., o modelo é adaptado a cada nova instância do fluxo de dados;
2. habilidade de sintonizar seus parâmetros e estrutura frente a mudanças graduais ou abruptas na distribuição de probabilidade que descreve a fonte de dados (*concept drift* e *concept shift*, respectivamente);
3. não requerer conhecimento *a priori* sobre os dados ou suas propriedades estatísticas, nem sobre a distribuição das classes.

O aprendizado online deve equilibrar a plasticidade estrutural e a estabilidade do modelo. Em outras palavras, os procedimentos de aprendizagem devem, de maneira parcimoniosa, decidir entre criar novos grânulos ou adaptar aquelas que já existem. Plasticidade estrutural significa criar novos grânulos para memorizar novos conceitos. A plasticidade evita que grânulos já construídos se alterem exageradamente, e que conhecimento adquirido no passado seja esquecido pelo modelo. A estabilidade estrutural preserva a estrutura do modelo, e, ao mesmo

tempo, permite a adaptação dos grânulos existentes para variações suaves e graduais do fluxo de dados. Geralmente, os algoritmos de aprendizado de máquina online não possuem mecanismos eficazes na determinação deste equilíbrio, pois não consideram métodos de (a) verificação da relevância atual dos grânulos de informação; (b) balanço de granularidade; e (c) mesclagem de grânulos, com o objetivo de diminuir a redundância do conhecimento representado pelo modelo. Os quesitos (a) e (b) são raramente tratados na literatura. O quesito (c), sobre mescla de elementos locais, é diverso. Não há unanimidade a respeito de um método de mescla melhor.

O armazenamento de grandes volumes de dados provenientes de ambientes não estacionários é muitas vezes inviável ou ineficaz (PRATAMA; PEDRYCZ; LUGHOFER, 2018; SOUZA et al., 2019; SILVA et al., 2019). Os sistemas inteligentes evolutivos provêm uma abordagem de análise e aprendizado autônoma a partir de fluxos de dados. Esses sistemas geralmente implementam métodos de aprendizado de varredura única dos dados, (*single-pass learning* ou, ainda, aprendizado incremental) que é de extrema importância em aplicações de tempo crítico, alta frequência (dados gerados muito rapidamente), de caráter não estacionário, e bases de dados com muitos atributos e amostras (*Big data*). A aplicação de métodos *single-pass* implica que cada instância de dados deve ser processada apenas uma vez pelo modelo e, em seguida, pode ser descartada. Isso elimina a necessidade de armazenamento de grandes volumes de dados (YANG, 2003). Adicionalmente, os sistemas evolutivos oferecem uma plataforma de desenvolvimento, ou *framework*, na qual componentes locais, e.g., grânulos, neurônios, folhas, clusters ou regras, podem ser criados, adaptados, mesclados, divididos ou reconsiderados com base unicamente no comportamento do fluxo de dados (ANGELOV, 2013; TUNG; QUEK; GUAN, 2013; LAÑA et al., 2019). Modelos evolutivos também se mostram auto-organizáveis, uma vez que nenhum conhecimento prévio sobre os dados ou suas propriedades é necessário (LUGHOFER, 2011; SILVA et al., 2014; ŠKRJANC et al., 2019; CORDOVIL et al., 2019).

Além de variações ao longo do tempo, os fluxos de dados produzidos a partir de, por exemplo, dispositivos eletrônicos, sensores ou por qualquer outra forma de se perceber o mundo real consistem em conjuntos de valores inerentemente incertos. A incerteza é uma propriedade que indica a imperfeição do valor medido com relação ao valor verdadeiro. A incerteza pode se originar de flutuações de natureza aleatória nos fluxos de dados, imprecisões numérica devido ao cálculo binário durante o processo de aquisição dos dados, fusão de informações de diferentes fontes, instrumentos de medição não ideais, métodos de pré-processamento de dados e outros (LEITE, 2012), (LIU, 2007; LIO; LIU, 2018; GARCIA et al., 2019). Modelos granula-

res evolutivos que são particularmente capazes de aprender a partir de fluxos de dados fuzzy ou intervalares, e incorporar incertezas de dados em suas representações são propostos em (LEITE; COSTA; GOMIDE, 2013; LEITE et al., 2014; LEITE; ŠKRJANC, 2019). A maioria dos modelos evolutivos, no entanto, aprende com fluxos de dados numéricos, enquanto a representação dos dados em modelos locais é dada por objetos fuzzy ou intervalares (ŠKRJANC et al., 2019).

Este texto apresenta um ambiente de aprendizado online chamado *evolving Internal-external Fuzzy Clustering*, abreviado como Fuzzy eIX. Um modelo Fuzzy eIX é utilizado para classificação não supervisionada. Diferentemente de qualquer outra abordagem evolutiva, um modelo Fuzzy eIX é formado por grânulos com dupla fronteira. Os grânulos contêm uma estrutura interna que resume estatisticamente e localmente as instâncias de dados mais recentes. Portanto, informação granular local é usada para tomada de decisões autônoma ao longo do processo de aprendizado online. A propriedade de fronteira dupla dos grânulos Fuzzy eIX pode ser usada para converter o resultado do agrupamento em um sistema de inferência fuzzy evolutivo tipo 1 ou tipo 2 a qualquer instante. Além disso, o algoritmo de aprendizado associado a um modelo Fuzzy eIX implementa o princípio da granularidade balanceada da informação (BARGIELA; PEDRYCZ, 2016). Este princípio diz que os grânulos devem possuir tamanhos equilibrados em todas as dimensões para que uma melhor compreensibilidade e interpretabilidade dos resultados no domínio de aplicação sejam alcançadas. Deste modo, o Fuzzy eIX está fortemente alinhado com o conceito de *eXplainable Artificial Intelligence* (XAI) (GUNNING, 2017; TJOA; GUAN, 2019), que afirma que as decisões tomadas e as soluções fornecidas por algoritmos e modelos de aprendizado de máquina devem ser passíveis de compreensão e interpretação por seres humanos.

1.1 Objetivos

Os objetivos gerais desta pesquisa são:

1. Rever métodos evolutivos importantes e recentes de classificação não supervisionada existentes na literatura e compreender pontos fortes e fraquezas dos mesmos;
2. Propor um novo método de classificação não supervisionada que potencialmente supere os métodos estado da arte da área em termos de acurácia e interpretabilidade;

3. Comparar o desempenho do método proposto com àqueles alcançados por métodos alternativos em conjuntos de dados *benchmark* sintéticos e reais, que exploram a capacidade de classificação em ambientes não estacionários e de grandes volumes de dados.

Os objetivos específicos são:

1. Formalizar matematicamente o algoritmo de aprendizado de modelos Fuzzy eIX;
2. Apresentar índices de validação incrementais, que serão utilizados na comparação de qualidade das partições geradas por diferentes algoritmos;
3. Descrever procedimentos experimentais que avaliem a robustez do método Fuzzy eIX para classificação.

1.2 Contribuições

A principal contribuição deste projeto de pesquisa é a proposição de um novo método evolutivo de classificação não-supervisionada, denominado *evolving Internal-eXternal Fuzzy Clustering* (Fuzzy eIX). Este método tem como principal diferencial a construção de grânulos com fronteiras duplas, que são convenientes na representação de incertezas presentes em fluxos de dados online. Adicionalmente, o método Fuzzy eIX realiza o princípio da granularidade balanceada, pode produzir sistemas de inferência fuzzy tipo 2 e está alinhado às ideias de XAI.

1.3 Publicações Diretas e Indiretas

1. **AGUIAR, C. C.** ; LEITE, D. F. . Unsupervised Fuzzy eIX: Evolving Internal-External Fuzzy Clustering. In: 2020 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS), 2020, Bari. 2020 IEEE Conference on Evolving and Adaptive Intelligent Systems, 2020.
2. **AGUIAR, C. C.** ; LEITE, D. . Controle Fuzzy PDC Baseado em Inequações Matriciais Lineares de um Sistema de Guindaste Overhead. In: 14^o Simpósio Brasileiro de Automação Inteligente (SBAI). SBA, 2019.
3. LEITE, D. ; **AGUIAR, C. C.** ; SOUZA, G. ; Škrjanc, I. . Nonlinear Fuzzy State-Space Modeling and LMI Fuzzy Control of Overhead Cranes. In: 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) (pp. 1-6). IEEE, 2019.

1.4 Estrutura do Trabalho

Este texto é dividido em 9 capítulos como segue.

Neste capítulo, foram introduzidos conceitos sobre sistemas inteligentes evolutivos e suas principais características. Apontou-se as contribuições a serem obtidas a partir do método proposto, Fuzzy eIX. No capítulo 2 é realizada a revisão de alguns sistemas inteligentes evolutivos aplicados à classificação não supervisionada de fluxos de dados online. Estes apresentam, em certos aspectos, pontos em comum com a metodologia do aprendizado Fuzzy eIX. O capítulo 3 traz decorrências de um modelo Fuzzy eIX, i.e., o que pode ser derivado a partir do resultado produzido por um modelo Fuzzy eIX. No capítulo 4, os procedimentos de desenvolvimento e adaptação de um modelo Fuzzy eIX são formalmente descritos. No capítulo 5, detalha-se a metodologia para a condução dos experimentos computacionais e comparação de desempenho de algoritmos e modelos evolutivos. O capítulo 6 traz os resultados experimentais alcançados. O capítulo 7 traz as conclusões, sumário das principais características do algoritmo apresentado e breves discussões sobre o *framework* Fuzzy eIX. Finalmente, os capítulos 8 e 9 apresentam, respectivamente, o cronograma de execução e o orçamento.

2 SISTEMAS INTELIGENTES EVOLUTIVOS

Neste capítulo, apresentam-se conceitos e características de sistemas inteligentes evolutivos e como estes são empregados em aplicações de classificação não supervisionada. Três modelos evolutivos são discutidos neste capítulo: Modelagem Evolutiva baseada em Intervalos (IBeM), classificador evolutivo Fuzzy Tipo 2 (eT2Class) e Análise de Dados baseada em tipicidade e excentricidade (TEDA). Também serão apresentadas comparações conceituais destes modelos com o modelo Fuzzy eIX proposto. No Capítulo 6, as performances dos modelos TEDA-Cluster e Fuzzy eIX serão comparadas frente ao problema de classificação de conjuntos de dados *benchmark*.

Sistemas inteligentes são entendidos como sistemas que possuem a habilidade de compreender e responder ao ambiente no qual estão inseridos. São capazes de aprender e descrever padrões presentes nestes ambientes (YAGER; ZADEH, 2012). Tomadas de decisões e regras de aprendizagem de sistemas inteligentes são baseadas em métodos de inteligência computacional. Alguns exemplos de métodos de inteligência computacional são: redes neurais artificiais, sistemas fuzzy, sistemas neuro-fuzzy, máquinas de vetores suporte, algoritmos genéticos, entre outros.

De maneira geral, sistemas inteligentes estão inseridos em ambientes complexos, onde grandes quantidades de dados são produzidas de diferentes fontes, de maneira contínua e independente. Além disso, a quantidade e as características dos dados varia constantemente ao longo do tempo (GAROFALAKIS; GEHRKE; RASTOGI, 2016). Tais conjuntos de dados são denominados *data streams*, ou, fluxos de dados. Fluxos de dados são ainda caracterizados pelos seguintes aspectos (GAMA, 2010):

1. as instâncias se fazem disponíveis de maneira online, i.e., não são totalmente conhecidas em um dado instante de tempo e se tornam conhecidas ao longo do tempo, uma observação por vez;
2. os dados não possuem limitações de magnitude e tamanho, pode ser consideradas infinitas;
3. o sistema não possui conhecimento nem controle sobre a ordem de chegada das instâncias, e;
4. as características e comportamentos dos dados podem mudar de maneiras imprevisíveis, impossibilitando a aproximação de perfis estatísticos dos mesmos. Estas mudanças são

classificadas como *concept drifts* e *concept shifts*, que consistem em variações graduais e abruptas, respectivamente, das distribuições dos dados (LUGHOFER; ANGELOV, 2011).

Em decorrência destes aspectos, o processamento e extração de conhecimento dos fluxos de dados se tornam desafiadores para métodos de inteligência computacional e aprendizado de máquina que operam em batelada, de maneira offline, i.e., processando volumes contidos e limitados de dados por vez. O processamento de fluxos de dados online exige que modelos com capacidade de adaptação sejam empregados, de modo a garantir boa performance em previsões, classificações e aproximações destes dados. Neste contexto, sistemas inteligentes baseados em modelos evolutivos vêm sendo estudados e propostos, como em (SILVA et al., 2014; LUGHOFER et al., 2015; LUGHOFER; PRATAMA; SKRJANC, 2017) para lidar com processamento e aprendizado a partir de fluxos de dados online.

Segundo (KASABOV, 2007), um modelo evolutivo pode ser entendido como aquele que se adapta e se desenvolve de modo contínuo. A adaptação se dá em função das características do ambiente em que o sistema está inserido. Como discutido anteriormente, variações temporais nas distribuições de dados podem ter impactos significativos nos processos de aprendizado e, assim, modelos evolutivos devem ser capazes de observar tais variações e auto ajustar adequadamente.

Ainda de acordo com (KASABOV, 2007), o termo "evolutivo" possui maior abrangência quando comparado ao termo "evolucionário". Enquanto modelos evolucionários constituem populações de indivíduos e como suas gerações são formadas e sofrem mutações, métodos que são evolutivos compreendem, de forma mais geral, o desenvolvimento das funcionalidades e estruturas de modelos. Sistemas evolutivos conseguem aprender e se adaptar baseados em experiências e na percepção de variações do comportamento do ambiente (ANGELOV; FILEV; KASABOV, 2010).

Em comparação à sistemas adaptativos, ou incrementais, os sistemas evolutivos são capazes de responder à quaisquer variações no comportamento dos dados a partir da alterações em seus parâmetros e também de suas estruturas internas. Por outro lado, sistemas adaptativos ou incrementais são, tipicamente, capazes apenas de ajustar seus parâmetros e supõe modelos com estruturas fixas (ANGELOV, 2013).

2.1 Computação Granular

De acordo com (PEDRYCZ, 2001; ZADEH, 1997), o termo computação granular se refere à formas de representar e processar volumes básicos de informação – denotados como *grânulos* de informação. Estes grânulos de informação podem também ser entendidos como coleções de entidades que são agrupadas juntas devido ao grau de similaridade que possuem entre si (PEDRYCZ et al., 2007). Esta similaridade pode se dar espacial, temporal ou funcionalmente (PEDRYCZ; SKOWRON; KREINOVICH, 2008; YAO; VASILAKOS; PEDRYCZ, 2013). No contexto da inteligência computacional, a abordagem granular pode ser utilizada na decomposição de problemas em partes que podem ser solucionadas independentemente e depois combinadas utilizando medidas de similaridade entre grânulos (LEITE, 2012).

Nos último anos, um tópico em particular na área de inteligência computacional vem ganhando extrema notoriedade: *eXplainable Artificial Intelligence*, ou XAI. Este tópico reúne uma série de métodos e técnicas que visam conferir às decisões tomadas pelos algoritmos aspectos de interpretabilidade e compreensibilidade por humanos, i.e., tornar mais fácil aos humanos a interpretação e o entendimento dos motivos e fatos que levaram o algoritmo a tomar determinadas decisões (GUNNING, 2017; TJOA; GUAN, 2019). É um contraste ao paradigma da "caixa preta", onde não se possui completa ciência das motivações que levaram os modelos a tomar suas decisões (ADADI; BERRADA, 2018). Vários trabalhos abordam métodos de XAI e discutem a importância da aferição de compreensibilidade aos modelos de inteligência computacional e aprendizado de máquina, dentre eles (TJOA; GUAN, 2019; BYRNE, 2019; WOLF, 2019).

Os grânulos possuem importante papel na construção de características cognitivas e tomada de decisões em algoritmos evolutivos, aferindo-os, principalmente, interpretabilidade às decisões tomadas (PEDRYCZ, 2018). No algoritmo de aprendizado Fuzzy eIX, métodos de balanceamento de granularidade são considerados de modo a aumentar a interpretabilidade do modelo, segundo o princípio da granularidade balanceada (BARGIELA; PEDRYCZ, 2016). Portanto, o Fuzzy eIX está fortemente alinhado com os conceitos de XAI. É importante notar que a medida de granularidade adotada neste trabalho é o tamanho de determinado grânulo em cada uma das dimensões do problema, conforme será discutido no capítulo 4.

2.2 Sistemas Granulares Evolutivos

Os sistemas granulares evolutivos são abordagens de modelagem de fluxos de dados incertos ou numéricos que consideram grânulos de informação para representar o conhecimento contido no fluxo. Os grânulos sofrem adaptações a medida que instâncias do fluxo de dados são processadas. Grânulos são criados, adaptados (expandidos, reduzidos ou arrastados), mesclados e deletados baseando-se exclusivamente no comportamento dos dados (LEITE et al., 2011). Nesta seção, apresentam-se a abordagem de modelagem evolutiva granular baseada em intervalos (IBeM), proposta em (LEITE; COSTA; GOMIDE, 2009) e os pontos de similaridade e divergência em relação ao método Fuzzy eIX proposto neste trabalho.

2.3 Modelagem Evolutiva Granular Baseada em Intervalos

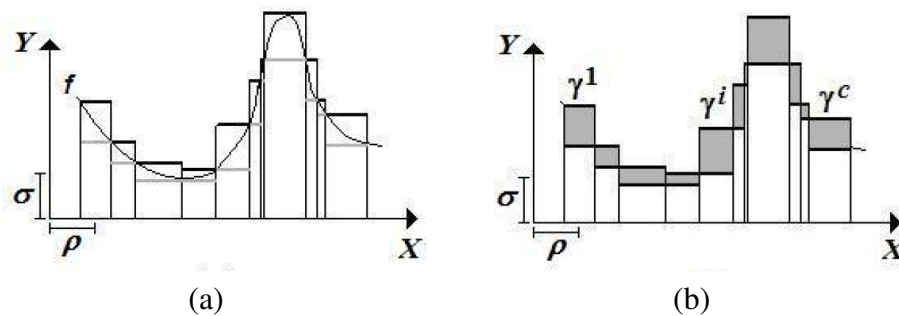
Modelagem Evolutiva Granular baseada em Intervalos (IBeM) é uma abordagem de modelagem evolutiva baseada em regras SE-ENTÃO que constitui uma representação granular do fluxo de dados, proposta em (LEITE; COSTA; GOMIDE, 2009). Cada uma das regras de um modelo intervalar provê uma descrição granular do comportamento do sistema em conjunto a uma função local das variáveis de entrada, que é associada a cada grânulo. Semelhante a modelos Fuzzy Takagi-Sugeno (TAKAGI; SUGENO, 1985), as funções locais são parte do consequente das regras, porém, a noção de conjuntos Fuzzy é substituída por intervalos dos espaço das entradas e saídas. O conteúdo de cada grânulo é sumarizado por uma regra SE-ENTÃO. Um conjunto destas regras fornece uma estimativa granular de uma função f , simultaneamente com uma aproximação ponto-a-ponto desta função por meio das funções e intervalos locais.

Novo conhecimento é acomodado em um modelo IBeM a partir da criação ou adaptação de grânulos, conseqüentemente, criando ou adaptando o banco de regras. A adaptação de grânulos já existentes é feita atualizando os valores dos limites inferior e superior dos intervalos associados aos antecedentes e consequentes das respectivas regras, bem como os coeficientes das funções locais. Estas adaptações ocorrem a cada nova instância que chega do fluxo de dados se dão de acordo com posição desta instância em relação aos intervalos associados aos grânulos. Em modelos Fuzzy eIX, os grânulos também são adaptados por meio de atualizações dos tamanhos de intervalos de acordo com a posição de uma nova instância, com a diferença da existência de um segundo intervalo (externo) que é empregado na representação de incertezas

de pertinência. Portanto, a regra de adaptação de modelos Fuzzy eIX pode ser entendida como uma extensão a modelos IBeM, com a ressalva de que não são empregadas funções locais.

Em modelos IBeM, os hiper parâmetros $\rho \in X \subseteq \mathbb{R}^n$ e $\sigma \in Y \subseteq \mathbb{R}^n$, em que X e Y são os conjuntos de entrada e saída, respectivamente, determinam os máximos tamanhos dos grânulos associadas às regras no espaço das entradas e saídas. Estes parâmetros possuem impacto direto na acurácia do modelo, uma vez que eles ditam a granularidade do mapeamento granular e, possivelmente, no formato dos grânulos. Quando ρ e σ tendem a zero, a granularidade é alta, i.e., o número de grânulos gerados é maior. Por outro lado, quando ρ e σ crescem, a granularidade diminui e grânulos maiores tendem a cobrir o universo de discurso, o que pode prejudicar a acurácia do modelo, e.g., um grânulo maior pode compreender instâncias que idealmente estariam em grânulos separados. A Figura 2.1 ilustra um exemplo de um mapeamento granular em um modelo IBeM denotado por $\gamma^1, \dots, \gamma^c$ a partir de um um fluxo de dados para aproximar uma função f . Em Fuzzy eIX, apenas o tamanho mínimo dos grânulos é controlado por um meta parâmetro $\varepsilon \in [0, 1]$ e a granularidade pode ser indiretamente controlada ajustando-se ε e o limiar de distância para mesclagem de grânulos, dado pelo meta parâmetro $\rho \in [0, 1]$.

Figura 2.1 – Mapeamento granular em modelos IBeM: (a) Limites inferiores e superiores do mapeamento; (b) Grânulos intervalares.



Fonte: (LEITE; COSTA; GOMIDE, 2009)

Mesclagem é executada em modelos IBeM quando dois grânulos quaisquer estão suficientemente próximos de acordo com uma medida de *dissimilaridade*. Esta medida é uma função das distâncias entre os limites dos intervalos de entrada (antecedentes) e saída (consequentes) associados aos grânulos e é atualizada sempre que um novo grânulo é criado. Os grânulos produzidos a partir de procedimentos de mesclagem possuem, também, tamanhos limitados pelos valores de ρ e σ . Em Fuzzy eIX, a mesclagem é feita com base nas distâncias infinito entre centros de dois grânulos e ocorre sempre que esta distância for menor que o meta parâmetro ρ . Ainda em Fuzzy eIX, a mesclagem pode ocorrer via um procedimento de média, que conserva o tamanho médio dos grânulos anteriores, mantendo a granularidade maior. Ou, ainda,

pode ocorrer pela formação de um envoltório convexo, que tente a produzir um grânulo maior, que cobre totalmente a região dos grânulos anteriores e os espaços entre eles. Desta forma, a granularidade é menor.

De maneira geral, as regras de criação e adaptação de grânulos Fuzzy eIX podem ser entendidas como extensões das regras de análogas em modelos IBeM. A mesclagem difere principalmente em relação às propriedades granulares utilizadas na definição das medidas de distâncias entre grânulos.

2.4 Sistemas Evolutivos Fuzzy Tipo 2

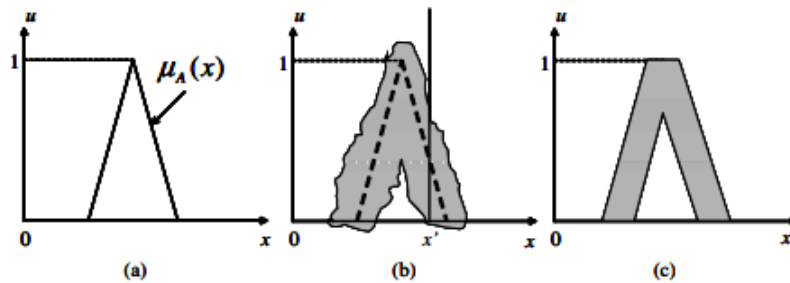
Nesta seção, apresentam-se conceitos dos sistemas de inferência Fuzzy do tipo 2 e como este difere dos sistemas do tipo 1. Adicionalmente, apresenta-se um classificador Fuzzy do tipo 2 evolutivo, o eT2Class, proposto em (PRATAMA; LU; ZHANG, 2015) e analisam-se pontos comuns e divergentes ao Fuzzy eIX.

2.5 Sistemas de Inferência Fuzzy Tipo 2

Sistemas de inferência Fuzzy do tipo 2 generalizam os sistemas Fuzzy do tipo 1 de modo que incertezas possam ser melhor tratadas e explicadas (MENDEL, 2007b; AISBETT; RICKARD; MORGENTHALER, 2010). Nos sistemas Fuzzy tipo 1 associamos incerteza aos dados atribuindo funções de pertinência aos mesmos. Porém, não admite-se que as próprias funções de pertinência possuam incertezas associadas. Nos sistemas do tipo 2, propostos também por L. Zadeh em (ZADEH, 1975), as funções de pertinência são tridimensionais e associadas, em cada ponto de seus domínios a outras funções de pertinência que descrevem a incerteza naquele ponto, i.e., as funções de pertinência são Fuzzy. Considere a função de pertinência do tipo 1 ilustrada na Figura 2.2-a, onde um valor x do universo de discurso possua uma pertinência bem definida $\mu(x)$. A seguir, considere que para cada ponto de μ é associada uma incerteza, de modo que μ possa assumir quaisquer valores na região hachurada na Figura 2.2-b. Assim, um valor x' não terá mais apenas um valor de pertinência e, sim, uma gama de valores que são definidos nos pontos onde a linha vertical intersecta a região hachurada (de incerteza). Nos sistemas fuzzy do tipo 2, então, cada valor é associado a um grau de pertinência primário (por meio de uma função de pertinência do tipo 1) e secundário (por meio de uma função de pertinência do tipo 2). Ao conjunto de todas funções de pertinência do tipo 1 que podem estar presentes em um

sistema Fuzzy do tipo 2 dá-se o nome de *Footprint Of Uncertainty* (FOU) (MENDEL, 2007a; MENDEL, 2007b). Isto é, o FOU compreende a função de pertinência mostrada na Figura 2.2-a e a segunda função de pertinência que descreve a incerteza sobre seus valores, na Figura 2.2-b. Estas duas funções de pertinência são mostradas na Figura 2.2-c. Em linhas gerais, a FOU descreve a área hachurada na Figura 2.2-b que é completamente descrita pelas duas funções de pertinência em suas bordas: a função superior e a função inferior, ambas sendo conjuntos Fuzzy do tipo 1 (BILGIN et al., 2012). Isso implica que a descrição de sistemas Fuzzy tipo 2 é feita por agregação de sistemas Fuzzy do tipo 1.

Figura 2.2 – Funções de pertinência com incertezas associadas e representação gráfica da FOU



Fonte: Adaptado de (MENDEL, 2007b)

2.6 Classificador Fuzzy Tipo 2 Evolutivo eT2Class

Os sistemas evolutivos baseados em lógica Fuzzy são geralmente construídos a partir de um sistema de inferência Fuzzy do tipo 1, que processam dados *crisp*, i.e., precisamente definidos, associando-os funções de pertinência também *crisp* (ANGELOV; ZHOU, 2008; YANG et al., 2010). Contudo, sistemas de inferência Fuzzy do tipo 1 podem não ser robustos o suficiente para lidar com incertezas presentes nos fluxos de dados comumente observados no mundo real. Neste cenário, os sistemas de inferência Fuzzy do tipo 2 provêm mais subsídios ao projeto de sistemas evolutivos, no sentido de melhor representar tais incertezas (KARNIK; MENDEL; LIANG, 1999; MENDEL; JOHN, 2002).

Um modelo classificador baseado em sistema de inferência Fuzzy do tipo 2, o eT2Class, é proposto em (PRATAMA; LU; ZHANG, 2015). O eT2Class é evolutivo no sentido de criação e adaptação de regras Fuzzy e coeficientes das funções consequentes, a partir unicamente do comportamento do fluxo de dados. Cada regra Fuzzy em um modelo eT2Class está vinculada a uma única partição dos dados (cluster) e é composta de uma função de pertinência Gaussiana multivariada intervalar e uma função polinomial não linear de Chebyshev no consequente.

Quando a primeira instância chega do fluxo de dados, a primeira regra é criada por meio do método do potencial por classe (PRATAMA et al., 2014b). Adicionalmente, regras Fuzzy são criadas a partir de uma medida proximidade espacial entre uma nova instância e todas os outros padrões de treinamento já observados, sem que haja a necessidade de armazenar internamente todas as instâncias anteriores. Isto é feito atribuindo a cada cluster uma noção de densidade de instâncias (PRATAMA et al., 2014a). Uma nova regra é criada para acomodar uma instância do fluxo de dados quando esta ocupa uma região Fuzzy muito densa ou muito isolada da zona de influência das demais regras Fuzzy. Também é notável que a adaptação dos coeficientes da função consequente de determinada regra é feita por meio de um algoritmo de mínimos quadrados denominado *Fuzzily Weighted Generalized Recursive Least Square* (FWGRLS), que é uma generalização Fuzzy do método GRLS (XU; WONG; LEUNG, 2006). É notável que o classificador eT2Class utiliza perfis estatísticos em seus algoritmos de criação e adaptação de regras. Modelos Fuzzy eIX simplificam estes procedimentos, tornando o processo de criação e adaptação de grânulos dependente majoritariamente da distância entre a instância e o centro e limites inferior e superior dos grânulos. Além disso, as funções de pertinência do tipo 2 são formuladas e adaptadas de forma trivial em modelos Fuzzy eIX, enquanto que em eT2Class estas também dependem cálculos recursivos de medidas estatísticas. Estes fatores contribuem para o menor custo computacional em modelos Fuzzy eIX.

Em modelos eT2Class, clusters são mescladas segundo o método de similaridade vetorial, proposto em (WU; MENDEL, 2008). Neste método, a similaridade entre dois cluster é calculada como uma função das distâncias entre os centros dos clusters e suas características geométricas, e.g., tamanho, forma, volume, tornando a quantificação da similaridade mais precisa. Dois clusters são mesclados quando suas medidas de similaridade ultrapassam um parâmetro pré-definido, a saber, ρ_3 . Em Fuzzy eIX, a similaridade entre dois grânulos é calculada apenas como uma função da distância entre os centro. Não são levadas em consideração fatores geométricos ou estatísticos. esta medida simplificada de similaridade pode ocasionar em perda de acurácia de modelos Fuzzy eIX, ao ganho de uma menor complexidade computacional.

2.7 Tipicidade e Excentricidade

O método de análise de dados baseada em tipicidade e excentricidade (TEDA) é proposto por (ANGELOV, 2014b) e foi primeiramente aplicado em (ANGELOV, 2014a). Este *framework* tem por objetivo prover uma ferramenta de análise e extração de informações de

dados sem que sejam necessárias suposições restritivas sobre distribuições, independência entre instâncias, perfis estatísticos ou especificações de parâmetros. Suposições essas herdadas de métodos estatísticos tradicionais. Estas características são de extrema importância quando trabalha-se com fluxos de dados online. O TEDA é uma metodologia sistemática que não necessita de dados anteriores e pode ser utilizado para vários métodos de processamento de imagens, clusterização, classificação, previsão, controle, filtragem, regressão etc (ANGELOV, 2014b).

O termo "tipicidade" é uma medida que quantifica o quanto um determinado objeto é um bom exemplo de um conceito (OSHERSON; SMITH, 1997). A "excentricidade" é uma medida complementar à tipicidade, i.e., indicando o quanto um objeto se afasta de um conceito. No *framework* TEDA, ambos conceitos são formalmente descritos, sob uma óptica de densidade e proximidade no espaço dos dados. A cada nova amostra do fluxo de dados, a excentricidade e tipicidade são calculadas recursivamente, no sentido de determinar se a amostra é típica aos dados já existentes ou se trata de uma anomalia. Estabelecem-se níveis de densidade de dados que permitem a identificação de anomalias baseado em distâncias. Dessa forma, o TEDA se apresenta como uma abordagem estatística alternativa, capaz de trabalhar com qualquer tipo de processos com dados aleatórios, onde cada observação é completamente independente da outra (SOARES et al., 2017).

As medidas de tipicidade no TEDA remetem às funções de pertinência presentes em sistemas de inferência Fuzzy, com a diferença de que, no TEDA, estas são formulada e adaptadas diretamente a partir do fluxo de dados. Esta característica também encontra-se presente no processo de aprendizagem de modelos Fuzzy eIX.

2.8 TEDA-Cluster

O algoritmo TEDA-Cluster é proposto em (KANGIN; ANGELOV, 2015a) e utiliza conceitos do *framework* TEDA na realização de clusterização de fluxos de dados online, i.e., sem que sejam necessárias informações sobre a distribuição de classes (não-supervisionada). Este algoritmo é baseado também no conceito de "nuvem de dados" (KOUTRIKA; ZADEH; GARCIA-MOLINA, 2009), onde cada instância do fluxo de dados é descrita em termos de pertinências Fuzzy. As nuvens de dados, assim como clusters, são agrupamentos de instâncias similares, mas não são limitadas ou possuem um formato geométrico específico (KANGIN; ANGELOV, 2015a; SOARES et al., 2018).

Dada uma instância do fluxo de dados, a regra Fuzzy associada é disparada em um nível proporcional à tipicidade normalizada daquela amostra à nuvem. Esta característica é denotada pelo predicado *like* (\sim) presente nos antecedentes das regras (KANGIN; ANGELOV, 2015b). Este predicado, então, reflete uma noção de proximidade da instância à regra Fuzzy.

Novas nuvens de dados são criadas quando a tipicidade de uma dada instância em cada uma das nuvens existentes (tipicidades locais) é menor do que um limite $T \in [0, 1]$ pré definido. Além disso, todas as regras que são mais típicas que T à nova regra devem ser deletadas.

2.9 Validação de Modelos para Classificação Não Supervisionada

Os processos de validação de clusterização são utilizados na verificação da qualidade das partições, e.g., grânulos de fronteira dupla, produzidas por algoritmos de classificação não-supervisionada. Diferentes algoritmos de classificação produzem diferentes partições para um mesmo conjunto de dados. Alternativamente, um mesmo algoritmo pode produzir diferentes partições para combinações distintas de meta parâmetros (SILVA; WUNSCH, 2018; XU; XU; WUNSCH, 2012). Nestes casos, os índices de validação podem ser úteis na comparação entre os diversos resultados obtidos. As medidas de qualidade das partições geradas são baseadas (i) no grau de concordância entre duas partições, onde uma delas é conhecida *a priori* sobre o conjunto de dados, e a segunda é uma partição produzida pelo algoritmo considerado (índices de validação *externos*), e; (ii) no grau de concordância apenas entre as partições geradas pelo algoritmo, sem a necessidade de informações externas (índices de validação *internos*) (RENDÓN et al., 2011b; RENDÓN et al., 2011a).

Recentemente, índices de validação incrementais vem sendo desenvolvidos para métodos de classificação online de fluxos de dados, i.e., índices que são computados a cada nova instância de dados (MOSHTAGHI et al., 2018; MOSHTAGHI et al., 2019; IBRAHIM; KELLER; BEZDEK, 2018; IBRAHIM; WANG; KELLER, 2018). Estes consistem de versões online de índices internos baseados em soma de quadrados (SS), que geralmente trabalham sobre medidas de custo-benefício entre compactação intra-grânulos (o quão distantes estão as instâncias dentro de um mesmo grânulo) e separação entre-grânulos (XU; XU; WUNSCH, 2012; ZHAO; FRÄNTI, 2014).

2.10 Índice de Validação Xie-Beni

O índice de validação Xie-Beni (XB), proposto em (XIE; BENI, 1991), quantifica a qualidade das partições baseando-se em medidas de variância intra-partições e separação entre-partições. A variância intra-partições define o quão distante as instâncias que possuem algum grau de pertinência à partição estão de seu centro. A medida de separação é a distância entre os centros de diferentes partições. Seja um sistema classificador composto por k partições. O índice de Xie-Beni pode ser computado como:

$$XB(k) = \frac{\sum_{i=1}^k \sum_{h=1}^N \mu_{ih}^2 \|\mathbf{c}^i - \mathbf{x}^{[h]}\|^2}{N \cdot \min_{i,h} \|\mathbf{c}^i - \mathbf{x}^{[h]}\|^2} \quad (2.1)$$

em que N é o número total de instâncias de dados, $\mathbf{x}^{[h]}, h = 1, \dots, N$ é uma instância particular e $\mathbf{c}^i, i = 1, \dots, k$ são os centros ou protótipos das partições geradas. A matriz μ é denominada *matriz de partição*, e suas entradas μ_{ih} relacionam os graus de pertinência da h -ésima instância à i -ésima partição. É importante notar que $\mu_{ih} = 0$ ou $\mu_{ih} = 1$ nos casos de particionamentos rígidos, e $0 < \mu_{ih} < 1$ nos casos de particionamentos Fuzzy. O melhor conjunto de partições minimiza (2.1).

2.11 Índice de Validação Davies-Bouldin

O índice de validação de Davies-Bouldin (DB) (DAVIES; BOULDIN, 1979) para cada partição i baseia-se na média da similaridade R entre i e uma segunda partição l com $i \neq l$, onde a similaridade entre as duas partições é máxima, i.e.,

$$DB(k) = \frac{1}{k} \sum_{i=1}^k R_i, \quad (2.2)$$

em que:

$$R_i = \max_{i \neq l} \left(\frac{S_i + S_l}{M_{i,l}} \right), \quad (2.3)$$

$$S_i = \left[\frac{1}{N^i} \sum_{h=1}^{N^i} \|\mathbf{x}^{[h]} - \mathbf{c}^i\|_2^q \right]^{\frac{1}{q}}, \quad q \geq 1, i = 1, \dots, k, \quad (2.4)$$

$$M_{i,l} = \left[\sum_{j=1}^n |c_j^i - c_j^l|^p \right]^{\frac{1}{p}}, \quad p \geq 1, \quad (2.5)$$

onde c_j^i é a j -ésima componente do centro n -dimensional \mathbf{c}^i da i -ésima partição. O parâmetro N^i corresponde ao número de instâncias pertencentes á i -ésima partição. As variáveis (p, q) em (2.5) são parâmetros definidos pelo usuário, e $(S_i, M_{i,l})$ mensuram, respectivamente, a compactação e separação das partições. O melhor conjunto de partições minimiza (2.2).

2.12 Formulação de Índices de Validação Incrementais

As medidas de compactação e separação comumente empregada nas formulações dos índices de validação utilizam informações sobre os centros das partições e as instâncias de dados. O último é geralmente um informação conhecida *a priori*, i.e., um conjunto de dados estático cujas características estão todas disponíveis antes da classificação (MOSHTAGHI et al., 2018; IBRAHIM; KELLER; BEZDEK, 2018). Para situações onde a classificação é feita de maneira online, uma formulação de índices de validação incremental que estima compactação e separação de forma recursiva é introduzida em (MOSHTAGHI et al., 2018; MOSHTAGHI et al., 2019). Nesta formulação, a medida de compactação da i -ésima partição, CP_i é dada por

$$CP_i = \sum_{h=1}^{N_i} \mu_{ih}^2 \|\mathbf{x}^{[h]} - \mathbf{c}^i\|_2^2. \quad (2.6)$$

Quando uma nova instância chega do fluxo de dados e é apresentada e processada pela i -ésima partição, seu novo valor de compactação é computado de acordo com a regra

$$CP_i(\text{new}) = \sum_{h=1}^{N_i(\text{new})} \mu_{ih}^2 \|\mathbf{x}^{[h]} - \mathbf{c}^i(\text{new})\|_2^2, \quad (2.7)$$

em que $N_i(\text{new}) = N_i(\text{old}) + 1$ e $\mathbf{c}^i(\text{new})$ é obtido de $\mathbf{c}^i(\text{old})$ a partir da aplicação de uma regra de adaptação adequada, que pode variar de acordo com o algoritmo classificador empregado. O índice de compactação em (2.6) é adaptado incrementalmente de acordo com

$$CP_i(\text{new}) = CP_i(\text{old}) + \|\mathbf{z}^i\|_2^2 + N_i(\text{old})\|\Delta\mathbf{c}^i\|_2^2 + 2\Delta(\mathbf{c}^i)^T \mathbf{g}^i(\text{old}), \quad (2.8)$$

onde $\mathbf{z}^i = \mathbf{x}^{[h]} - \mathbf{c}^i(\text{new})$ e $\Delta\mathbf{c}^i = \mathbf{c}^i(\text{old}) - \mathbf{c}^i(\text{new})$. O vetor \mathbf{g}^i , que é formalmente definido por

$$\mathbf{g}^i = \sum_{h=1}^{N_i} (\mathbf{x}^{[h]} - \mathbf{c}^i), \quad (2.9)$$

é adaptado incrementalmente a cada iteração utilizando a regra

$$\mathbf{g}^i(\text{new}) = \mathbf{g}^i(\text{old}) + \mathbf{z}^i + N_i(\text{old})\Delta\mathbf{c}^i. \quad (2.10)$$

A partir desta formulação, é possível, então, derivar versões incrementais dos índices de validação de Xie-Beni (iXB) e Davies-Bouldin (iDB) (SILVA; MELTON; WUNSCH, 2020; MOSHTAGHI et al., 2018; IBRAHIM; KELLER; BEZDEK, 2018). Estas formulações são apresentadas a seguir.

2.13 Índice Xie-Beni Incremental

Referências para a formulação e utilização do índice iXB podem ser encontradas em (SILVA; MELTON; WUNSCH, 2020; MOSHTAGHI et al., 2018; IBRAHIM; KELLER; BEZDEK, 2018). O cálculo recursivo de iXB é feito por meio de

$$XB(\text{new}) = \left(\frac{1}{N_i(\text{new})} \right) \frac{\sum_{i=1}^{k(\text{new})} CP_i(\text{new})}{\min_{i \neq l} (\|\mathbf{c}^i(\text{new}) - \mathbf{c}^l(\text{new})\|_2^2)}. \quad (2.11)$$

O valor $k(\text{new}) = k(\text{old}) + 1$ representa o número atualizado de partições após o processamento da nova instância.

2.14 Índice Davies-Bouldin Incremental

Referências para a formulação e utilização do índice iDB podem ser encontradas em (SILVA; MELTON; WUNSCH, 2020; MOSHTAGHI et al., 2018; IBRAHIM; KELLER; BEZDEK, 2018). O cálculo recursivo de iDB é feito por meio de

$$DB(\text{new}) = \frac{1}{k(\text{new})} \sum_{i=1}^{k(\text{new})} \max_{l, i \neq l} \left(\frac{\frac{CP_i(\text{new})}{N^i(\text{new})} + \frac{CP_l(\text{new})}{N^l(\text{new})}}{\|\mathbf{c}^i(\text{new}) - \mathbf{c}^l(\text{new})\|_2^2} \right). \quad (2.12)$$

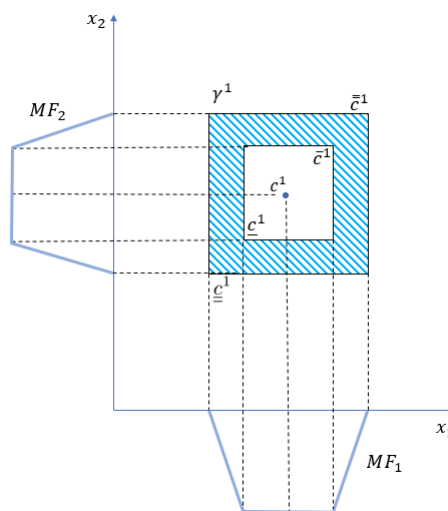
3 DECORRÊNCIAS DE MODELOS FUZZY EIX

Neste capítulo são apresentadas possíveis implicações em se considerar grânulos Fuzzy eIX, i.e., evolutivos com fronteiras duplas, internas e externas, na representação de conhecimento contido em um fluxo ativo de dados.

3.1 Derivação de Sistemas de Inferência Fuzzy Evolutivos do Tipo 1

Ao se projetar as fronteiras interna e externa de um grânulo Fuzzy eIX em eixos ortogonais que representam os atributos de um problema, funções de pertinência trapezoidais podem ser formuladas. Conseqüentemente, um sistema de inferência Fuzzy evolutivo do tipo 1 é obtido. Grânulos são criados e adaptados no espaço Cartesiano, de modo que o núcleo e o suporte das funções de pertinência associadas evoluem a partir do fluxo de dados. A Figura 3.1 ilustra um exemplo de como as funções de pertinências MF_1 e MF_2 podem ser extraídas a partir de um grânulo Fuzzy eIX. Naturalmente, um modelo Fuzzy eIX pode ser linguisticamente interpretado a partir de um conjunto de regras SE-ENTÃO, associando uma regra por grânulo.

Figura 3.1 – Derivação de funções de pertinência evolutivas do tipo 1 e de um sistema de inferência Fuzzy evolutivo do tipo 1 a partir de grânulos Fuzzy eIX



Fonte: Arquivo Pessoal

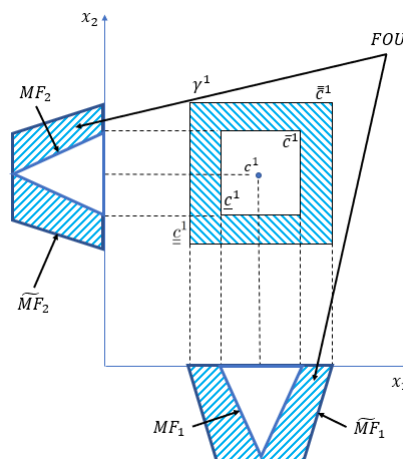
3.2 Derivação de Sistemas de Inferência Fuzzy Evolutivos do Tipo 2

Uma maneira mais rigorosa de se projetar grânulos Fuzzy eIX em eixos ortogonais consiste em considerar o centro do grânulo como seu protótipo e a informação descrita por este

como apenas aquela que existe dentro do domínio de seu intervalo interno. Funções de pertinência rigorosas que descrevem a incerteza associada à proximidade de uma instância (amostra) ao protótipo de um determinado grânulo são definidas a partir das fronteiras internas deste grânulo.

Um sistema de inferência Fuzzy evolutivo do tipo 2 pode ser obtido a partir das projeções das fronteiras interna e externa, resultando em funções de pertinência MF_i e $\tilde{M}F_i$, respectivamente, como mostrado na Figura 3.2. A área hachurada compreendida entre as funções de pertinência é chamada de *footprint of uncertainty*, ou FOU. O quão maior for a área da FOU, maior será a incerteza associada à pertinência, e vice-versa. Como os parâmetros, i.e., núcleo e suporte, das funções MF_i e $\tilde{M}F_i$ são obtidos diretamente de grânulos Fuzzy eIX, estes são assimilados pelo modelo autonomamente a partir do fluxo de dados.

Figura 3.2 – Derivação de funções de pertinência evolutivas do tipo 2 e de um sistema de inferência Fuzzy evolutivo do tipo 2 a partir de grânulos Fuzzy eIX



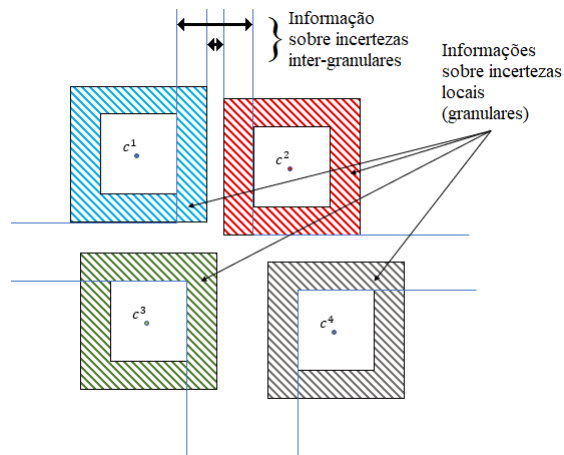
Fonte: Arquivo Pessoal

3.3 Análise de Incertezas Inter-Granulares Local e Global

Um grânulo Fuzzy eIX hiper retangular de dupla fronteira é caracterizado por um conjunto de quatro parâmetros por dimensão do problema, que definem suas fronteiras dos intervalos interno e externo. Adaptar tais intervalos de acordo com apenas um atributo do fluxo de dados é apenas uma maneira rápida e direta de promover o aprendizado do modelo em um ambiente dinâmico. Além da representação local de incertezas associadas à pertinência de uma dada instância a um grânulo, i.e., incertezas locais – dadas pelas áreas hachuradas na Figura 3.3 também é possível representar e calcular informações sobre incertezas globais, inter-granulares. Estas informações e medidas podem ser empregadas na tomada de decisões do algoritmo de

aprendizado associado ao modelo Fuzzy eIX - que será descrito no capítulo 4. Procedimentos de mesclagem entre grânulos e resolução de conflitos, e.g., dada instância pertencer a mais de um grânulo, podem ser derivados a partir da comparação de incertezas associadas a cada um dos grânulos.

Figura 3.3 – Informação sobre incertezas locais e inter-granulares em grânulos Fuzzy eIX



Fonte: Arquivo Pessoal

4 ALGORITMO DE APRENDIZADO FUZZY EIX

A descrição formal do *framework* de aprendizado online proposto neste trabalho, denominado *Fuzzy evolving Internal-eXternal* (Fuzzy eIX) é apresentada neste capítulo. Especificadamente, o método Fuzzy eIX descrito será aplicado à problemas de classificação não supervisionada de dados, ou *clustering*. Os processos de inicialização e mesclagem de grânulos serão detalhados, bem como regras de adaptação dos mesmos: expansão, encolhimento e balanceamento de granularidade.

4.1 Descrição do método Fuzzy eIX

Grânulos Fuzzy eIX são espacialmente delimitados por hiper retângulos internos e externos. O grau de pertinência de uma instância do fluxo de dados que se localiza na região interna de um grânulo é unitário. Uma instância localizada na área compreendida entre os hiper retângulos interno e externo é parcialmente considerada um membro do respectivo grânulo. O centro dos grânulos, bem como suas fronteiras interna e externa são recursivamente adaptados ao longo do tempo, conforme novas instâncias do fluxo de dados são processadas pelo modelo.

4.2 Inicialização de Grânulos

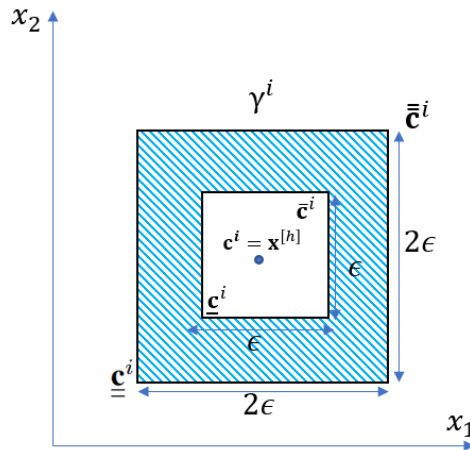
Seja um fluxo de dados denotado por $\mathbf{x}^{[h]}$, $h = 1, \dots$. Dada a primeira instância deste fluxo, i.e., $\mathbf{x}^{[1]} \in \mathbb{R}^n$, a saber, $\mathbf{x}^{[1]} = [x_1^{[1]} \dots x_j^{[1]} \dots x_n^{[1]}]$ – sendo n o número de atributos – o primeiro grânulo Fuzzy eIX, γ^1 é criado. O centro deste grânulo, denotado por \mathbf{c}^1 é igual a $\mathbf{x}^{[1]}$. As larguras iniciais dos intervalos interno e externo do grânulo γ^1 são iguais a ε e 2ε , respectivamente, em qualquer dimensão j , sendo $\varepsilon \in [0, 0.5]$ um meta parâmetro. As fronteiras interna e externa de γ^1 são caracterizadas pelos vetores limítrofes $[\underline{\mathbf{c}}^1, \bar{\mathbf{c}}^1]$ e $[\underline{\underline{\mathbf{c}}}^1, \bar{\bar{\mathbf{c}}}^1]$, respectivamente. Os vetores que definem o centro e os limites das bordas interna e externa o grânulo Fuzzy eIX são adaptativos ao longo do tempo, conforme novas instâncias são processadas pelo modelo.

Um grânulo Fuzzy eIX genérico γ^i de uma coleção $\gamma = \{\gamma^1, \dots, \gamma^k\}$ é caracterizado por:

1. um ponto prototípico, ou centro, $\mathbf{c}^i \in \mathbb{R}^n$; e
2. vetores limítrofes internos superior e inferior, $\underline{\mathbf{c}}^1, \bar{\mathbf{c}}^1 \in \mathbb{R}^n$, externos superior e inferior, $\underline{\underline{\mathbf{c}}}^1, \bar{\bar{\mathbf{c}}}^1 \in \mathbb{R}^n$.

Adicionalmente, o meta parâmetro ε define as larguras inicial e mínima dos atributos de um grânulo, i.e., um grânulo Fuzzy eIX terá sua largura em qualquer dimensão j limitada inferiormente pelo valor ε , para o intervalo interno, e 2ε , para o intervalo externo. Um segundo meta parâmetro do algoritmo, ρ é utilizado nos procedimentos de mesclagem, como será descrito na seção 4.1.6. Portanto, $\gamma^i := \{\underline{\mathbf{c}}^i, \underline{\mathbf{c}}^i, \bar{\mathbf{c}}^i, \bar{\mathbf{c}}^i\}$ é definido como uma coleção de cinco vetores n -dimensionais. A Figura 4.1 exemplifica um grânulo Fuzzy eIX bidimensional.

Figura 4.1 – Representação de um grânulo Fuzzy eIX bidimensional genérico



Fonte: Arquivo Pessoal

4.3 Instância em Região Interna

Se uma dada instância do fluxo de dados $\mathbf{x}^{[h]}$ localiza-se entre os limites da região interna de um grânulo γ^i , isto é, se

$$\begin{aligned} x_j^{[h]} &> \underline{c}_j^i \\ x_j^{[h]} &< \bar{c}_j^i \end{aligned}, \quad \forall j, j = 1, \dots, n, \quad (4.1)$$

então, o grau de pertinência de $\mathbf{x}^{[h]}$ em γ^i é 1 considerando-se qualquer operador \mathbb{T} -norma. Em decorrência do aumento de certeza sobre o correto posicionamento do grânulo γ^i a partir da inclusão de $\mathbf{x}^{[h]}$ na região interna, as larguras interna e externa de γ^i sofrem reduções em cada uma das dimensões $j = 1, \dots, n$. Portanto, o componente limítrofe interno inferior é aumentado de acordo com a regra de adaptação

$$\underline{c}_j^i(\text{new}) = (1 + d_j^{i[h]}) \underline{c}_j^i(\text{old}), \quad (4.2)$$

com $d_j^{i[h]} \in [0, \beta]$ dado por

$$d_j^{i[h]} := \beta - \beta \left(\frac{|c_j^i - x_j^{[h]}|}{c_j^i - \underline{c}_j^i} \right). \quad (4.3)$$

O valor padrão de β é 0.3. O componente limítrofe externo inferior, \underline{c}_j^i , é aumentando de forma análoga à (4.2)-(4.3), substituindo c_j^i por \underline{c}_j^i . O componente limítrofe interno superior, \bar{c}_j^i , é, por sua vez, reduzido proporcionalmente a partir da regra de adaptação

$$\bar{c}_j^i(\text{new}) = \bar{c}_j^i(\text{old}) - (\underline{c}_j^i(\text{new}) - \underline{c}_j^i(\text{old})), \quad (4.4)$$

$j = 1, \dots, n$. Similarmente, o componente limítrofe externo superior, \bar{c}_j^i , é obtido de (4.4) utilizando a componente limítrofe externa inferior no último termo. Nota-se que as reduções observadas nas larguras das regiões interna e externa de γ^i são mais expressivas quando $\mathbf{x}^{[h]}$ for mais próxima \mathbf{c}^i .

O tamanho dos grânulos Fuzzy eIX é limitado de modo que a largura das regiões interna e externa, em qualquer dimensão, sejam maiores que ε e 2ε , respectivamente, em qualquer instante de tempo.

Após o encolhimento, o i -ésimo grânulo γ^i é arrastado em direção à nova instância $\mathbf{x}^{[h]}$. Este arraste é feito sobre a caminho definido pelos pontos $\mathbf{x}^{[h]}$ e \mathbf{c}^i em um módulo inversamente proporcional à densidade de γ_i , i.e., ao número de instância N^i que pertencem à região interna de γ^i , desconsiderando $\mathbf{x}^{[h]}$. Este processo é chamado de *arraste ponderado*.

O centro \mathbf{c}^i é arrastado utilizando a regra

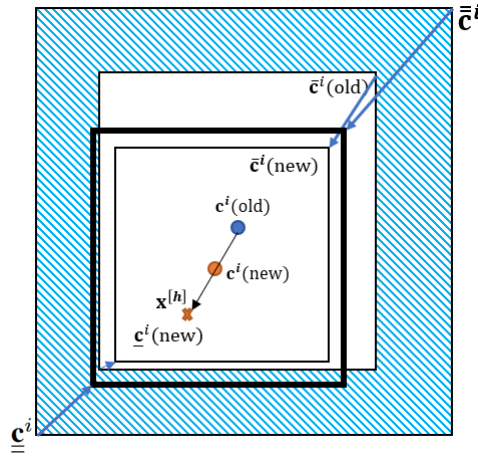
$$\mathbf{c}^i(\text{new}) = \mathbf{c}^i(\text{old}) + \frac{1}{N^i + 1} (\mathbf{x}^{[h]} - \mathbf{c}^i(\text{old})). \quad (4.5)$$

Considere que \bullet^i denote qualquer um dos vetores limítrofes $\underline{\underline{\mathbf{c}}^i}$, $\underline{\mathbf{c}}^i$, $\bar{\mathbf{c}}^i$, $\bar{\bar{\mathbf{c}}^i}$. De modo a preservar a simetria geométrica dos hiper retângulos que definem o grânulo, os vetores limítrofes interno e externo são arrastados de acordo com

$$\bullet^i(\text{new}) = \bullet^i(\text{old}) + \frac{1}{N^i + 1}(\mathbf{x}^{[h]} - \mathbf{c}^i(\text{old})). \quad (4.6)$$

A Figura 4.2 ilustra um exemplo dos procedimentos de encolhimento e arraste devido ao processamento de uma nova instância $\mathbf{x}^{[h]}$ que pertence à região interna de γ^i .

Figura 4.2 – Encolhendo e arrastando um grânulo Fuzzy eIX após processamento de instância pertencente à região interna



Fonte: Arquivo Pessoal

4.4 Instância em Região Externa

Uma instância $\mathbf{x}^{[h]}$ pode estar localizada nos limites da região externa de um grânulo Fuzzy eIX, i.e., com um ou mais de seus atributos localizados no intervalo externo. Neste caso:

$$\begin{aligned} x_j^{[h]} &> \underline{\underline{c}}_j^i \\ x_j^{[h]} &< \bar{\bar{c}}_j^i \end{aligned}, \quad \forall j, j = 1, \dots, n. \quad (4.7)$$

Além da factibilidade de (4.7), a equação (4.1) deve ser falsa para pelo menos uma dimensão j . Para $\mathbf{x}^{[h]}$ pertencente à região externa, a distância L_∞ de $\mathbf{x}^{[h]}$ à \mathbf{c}^i é maior que a distância entre \mathbf{c}^i e qualquer uma das instâncias que se localizam na região interna do grânulo γ^i .

Uma nova instância $\mathbf{x}^{[h]}$ pertencente à região externa de um grânulo γ^i implica que seu centro \mathbf{c}^i não deve ser adaptado, uma vez que a inclusão de $\mathbf{x}^{[h]}$ traz incertezas sobre o correto posicionamento e conteúdo de γ^i . Além disso, a incerteza associada ao grau de pertinência de $\mathbf{x}^{[h]}$ em γ^i sugere uma expansão granular no sentido de incorporação da incerteza ao modelo Fuzzy eIX.

As larguras interna e externa de γ^i se tornam maiores de acordo com as seguintes regras de expansão. O componente limítrofe interno inferior é reduzido a partir de

$$\underline{c}_j^i(\text{new}) = (1 - f_j^{i[h]}) \underline{c}_j^i(\text{old}), \quad (4.8)$$

com

$$f_j^{i[h]} := \beta \left(\frac{\underline{c}_j^i - x_j^{[h]}}{\underline{c}_j^i - \underline{\underline{c}}_j^i} \right), \quad (4.9)$$

considerando apenas as dimensões j para as quais $\underline{\underline{c}}_j^i \leq x_j^{[h]} \leq \underline{c}_j^i$ seja verdadeiro. Caso contrário, \underline{c}_j^i não é adaptado. O valor padrão de β é 0.3 e $f_j^{i[h]} \in [0, \beta]$. O componente limítrofe externo inferior, $\underline{\underline{c}}_j^i$ é reduzido de forma análoga, utilizando (4.8) e o mesmo valor de $f_j^{i[h]}$ obtido em (4.9).

O componente limítrofe interno superior, \bar{c}_j^i , é aumentado proporcionalmente a partir da regra

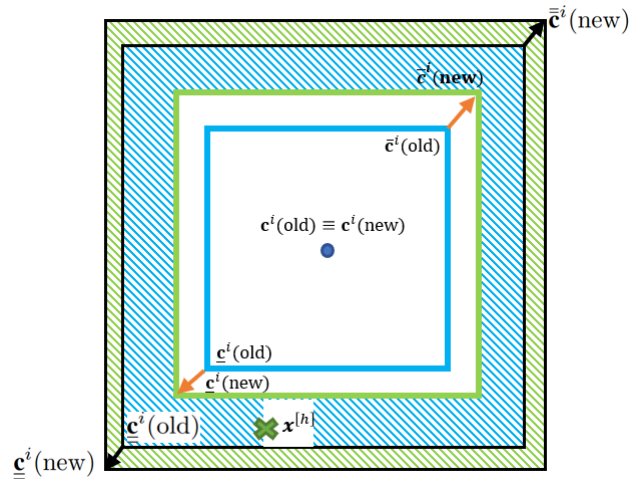
$$\bar{c}_j^i(\text{new}) = \bar{c}_j^i(\text{old}) + (\underline{c}_j^i(\text{old}) - \underline{c}_j^i(\text{new})) \quad (4.10)$$

$j = 1, \dots, n$. Similarmente, o componente limítrofe externo superior, $\bar{\bar{c}}_j^i$, é obtido a partir de (4.10) utilizando $\bar{\bar{c}}_j^i$ no último termo. Nota-se que a expansão do grânulo é maior para instâncias $\mathbf{x}^{[h]}$ mais próximas das bordas da região externa e, conseqüentemente, mais distantes do centro \mathbf{c}^i .

O tamanho dos grânulos Fuzzy eIX é limitado, de modo que a largura das regiões interna e externa, em qualquer dimensão, não sejam menores que ε e 2ε , respectivamente. Figura

4.3 ilustra um exemplo da expansão granular devido a nova instância $\mathbf{x}^{[h]}$ pertencente à região externa de um grânulo γ^i .

Figura 4.3 – Expandindo um grânulo Fuzzy eIX após processamento de instância pertencente à região externa



Fonte: Arquivo Pessoal

4.5 Instância Excêntrica

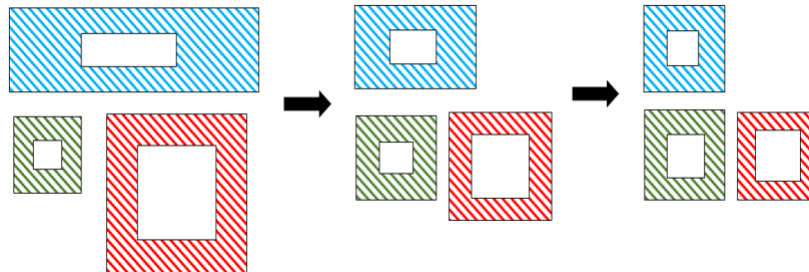
Se uma nova instância $\mathbf{x}^{[h]}$ não pertencer à nenhuma região interna ou externa dos grânulos existentes, um novo grânulo é criado para incluir $\mathbf{x}^{[h]}$. O procedimento de criação deste grânulo é análogo àquele descrito no processo de inicialização do modelo, quando a primeira instância é processada. O grânulo criado incorpora ao modelo Fuzzy eIX o novo conhecimento associado à $\mathbf{x}^{[h]}$.

4.6 Princípio da Granularidade Balanceada

O princípio da granularidade balanceada (BARGIELA; PEDRYCZ, 2016) afirma que a concepção de grânulos com tamanhos balanceados ao longo de todas as dimensões deve ser dada preferência frente a grânulos com tamanhos desbalanceados. Isto é, um processo de balanceamento de granularidade deve ser adotado para que, a cada nova instância processada, as larguras das regiões interna e externa dos grânulos sejam gradualmente equalizadas de acordo com alguma medida global de granularidade. Em suma, as larguras das retas que delineiam um hiper retângulo devem ser idealmente similares. Uma melhor compreensibilidade dos resultados no domínio de aplicação do algoritmo pode ser alcançada por meio de um modelo cujos grânulos sejam balanceados em termos de suas dimensões (PEDRYCZ; BARGIELA, 2002). A

Figura 4.4 ilustra um possível processo de balanceamento de granularidade em grânulos Fuzzy eIX bidimensionais ao longo das iterações do algoritmo.

Figura 4.4 – Balanceamento de granularidade em grânulos Fuzzy eIX bidimensionais



Fonte: Arquivo Pessoal

Os tamanhos das regiões interna e externa de grânulos Fuzzy eIX são balanceadas a cada iteração a partir de um regra baseada em larguras médias. Sejam

$$w_j^{i(\text{int})} := \bar{c}_j^i - \underline{c}_j^i, \quad w_j^{i(\text{ext})} := \bar{c}_j^i - \underline{c}_j^i, \quad (4.11)$$

com $j = 1, \dots, n$, as larguras interna e externa, respectivamente, de um grânulo γ^i .

Seja k a quantidade total de grânulos criados. Para atributos individuais j , $j = 1, \dots, n$, as médias das regiões interna e externa, respectivamente, de todos os grânulos a longo do j -ésimo eixo são dadas por

$$\hat{w}_j^{(\text{int})} = \frac{1}{k} \sum_{i=1}^k w_j^{i(\text{int})}, \quad \hat{w}_j^{(\text{ext})} = \frac{1}{k} \sum_{i=1}^k w_j^{i(\text{ext})}. \quad (4.12)$$

As larguras individuais (4.11) são adaptadas na direção da largura média global, (4.12) considerando um atributo j por vez. Isto é, os lados relativamente menores dos hiper retângulos vão sendo gradualmente aumentados, enquanto os maiores, reduzidos. Estas adaptação são feitas em torno da média global (4.12). Formalmente, esta regra pode ser escrita como

$$\begin{aligned}
\underline{\underline{c}}_j^i(\text{new}) &= \underline{\underline{c}}_j^i(\text{old}) - \alpha(\hat{w}_j^{(\text{ext})} - w_j^{i(\text{ext})}) \\
\underline{\underline{c}}_j^i(\text{new}) &= \underline{\underline{c}}_j^i(\text{old}) - \alpha(\hat{w}_j^{(\text{int})} - w_j^{i(\text{int})}) \\
\bar{\bar{c}}_j^i(\text{new}) &= \bar{\bar{c}}_j^i(\text{old}) + \alpha(\hat{w}_j^{(\text{int})} - w_j^{i(\text{int})}) \\
\bar{\bar{c}}_j^i(\text{new}) &= \bar{\bar{c}}_j^i(\text{old}) + \alpha(\hat{w}_j^{(\text{ext})} - w_j^{i(\text{ext})})
\end{aligned} \tag{4.13}$$

$\forall i, j; i = 1, \dots, k; j = 1, \dots, n$, onde $\alpha \in [0, 1]$ é a taxa de balanceamento. Neste trabalho, assume-se $\alpha = 0.3$. Valores mais elevados de α aumentam a velocidade da convergência dos tamanhos dos grânulos Fuzzy eIX para um valor similar médio e, conseqüentemente, aumenta a interpretabilidade do modelo. Porém, o aumento da velocidade de balanceamento vem com o custo de potencial perda na acurácia do modelo decorrente de rápidas e significativas flutuações nos tamanhos dos grânulos, portanto, incluindo ou excluindo instâncias já processadas. Se a acurácia do modelo é um aspecto importante no domínio da aplicação do método, então, valores menores de α devem ser considerados de modo a melhor conservar a geometria dos hiper retângulos ao longo das iterações.

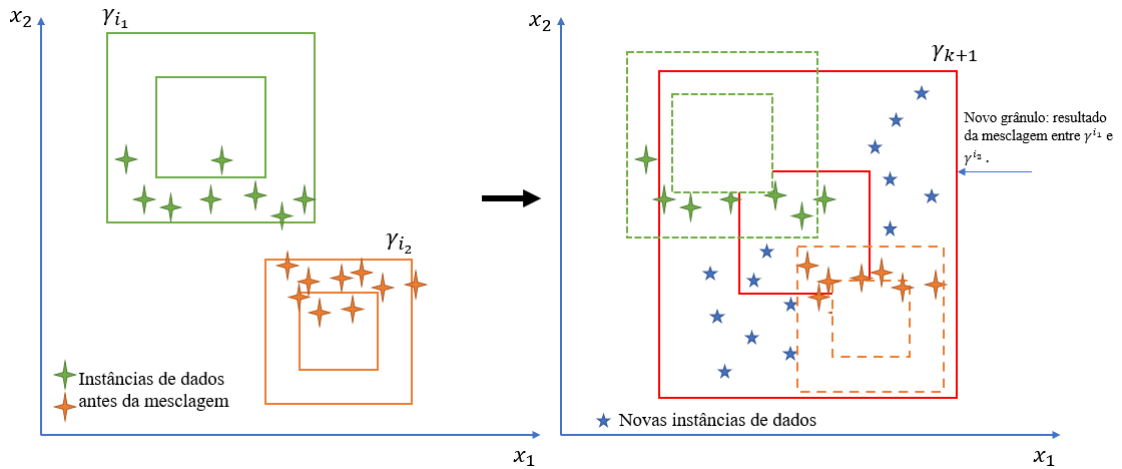
4.7 Mesclagem de Grânulos

A mesclagem é um procedimento aplicado a um par de grânulos que apresentam significativa sobreposição. Frequentemente, uma série de instâncias podem pertencer à lacuna existente entre dois grânulos próximos, ou ainda, pertencer à regiões de interseção entre grânulos. Esta situação implica diretamente que existem conhecimentos similares sendo representados por dois (ou mais) grânulos. Portanto, a redundância é minimizada, ou até mesmo evitada, a partir da mesclagens destes grânulos (ŠKRJANC et al., 2019). A título de visualização, um exemplo para grânulos Fuzzy eIX bidimensionais é ilustrado na Figura 4.5. No caso mostrado na Figura 4.5, novas instâncias preenchem a lacuna existentes entre dois grânulos distintos que foram mesclados e, assim, reduzindo a redundância existente na representação das informações contidas em ambos grânulos e na lacuna entre eles.

Quando dois grânulos, e.g., γ^{i1} e γ^{i2} estão suficiente próximos um do outro, isto é, quando seus centros \mathbf{c}^{i1} e \mathbf{c}^{i2} são tais que

$$\|\mathbf{c}^{i1} - \mathbf{c}^{i2}\|_\infty \leq \rho, \tag{4.14}$$

Figura 4.5 – Exemplo do efeito da mesclagem de grânulos Fuzzy eIX



Fonte: Adaptado de (ŠKRJANC et al., 2019)

com $\rho \in [0, 1]$, então, eles são mesclados gerando o grânulo γ^{k+1} como resultado. O meta parâmetro ρ define, então, o quão próximo dois grânulos devem estar para que sejam mesclados. Conseqüentemente, o valor de ρ controla a quantidade de grânulos que são formados, bem como a fuzzcidade do modelo, i.e., o quão próximo os grânulos estarão um dos outros e o grau de pertinência de dada instância a cada um dos grânulos gerados.

Neste trabalho, propõe-se dois métodos de mesclagem para grânulos Fuzzy eIX: baseado em média ponderada e em envoltório convexo. Ambos métodos serão detalhados a seguir.

4.8 Mesclagem por Média Ponderada

No método baseado em média ponderada, o centro do novo grânulo, γ^{k+1} , é

$$\mathbf{c}^{k+1} = \mathbf{c}^{i_1} - \frac{N^{i_2}}{N^{i_1} + N^{i_2}} (\mathbf{c}^{i_1} - \mathbf{c}^{i_2}), \quad (4.15)$$

onde N^i é o número de vezes que o i -ésimo grânulo foi adaptado durante o processamento de dada instância de dados, i.e., tal instância pertence à região interna ou externa de γ^i . Esta propriedade é denominada **densidade de dados** representada pelo grânulo γ^i . O novo grânulo é posicionado no segmento de reta $\mathbf{c}^{i_1}\bar{\mathbf{c}}^{i_2}$. Por meio de (4.15), nota-se que este centro \mathbf{c}^{k+1} será posicionado mais próximo de \mathbf{c}^{i_1} caso $N^{i_1} > N^{i_2}$, e vice-versa. Caso $N^{i_1} = N^{i_2}$, este será posicionado no ponto médio do segmento de reta $\mathbf{c}^{i_1}\bar{\mathbf{c}}^{i_2}$. Os vetores limítrofes internos, $\underline{\mathbf{c}}^{k+1}$, $\bar{\mathbf{c}}^{k+1}$ e externos, $\underline{\underline{\mathbf{c}}}^{k+1}$, $\bar{\bar{\mathbf{c}}}^{k+1}$ do novo grânulo são obtidos a partir de (4.15) por analogia.

4.9 Mesclagem por Envoltório Convexo

Como uma forma alternativa de mesclagem, o processo por envoltório convexo consiste em produzir um novo grânulo γ^{k+1} menos refinado, que encapsula todas a informação inerente aos grânulos anteriores. Ou seja, o novo grânulo herdará todas as instâncias que faziam parte dos grânulos mesclados. Neste caso, o centro do novo grânulo é concebido de forma natural, a partir das adaptações aplicadas aos vetores limítrofes internos e externos, a saber

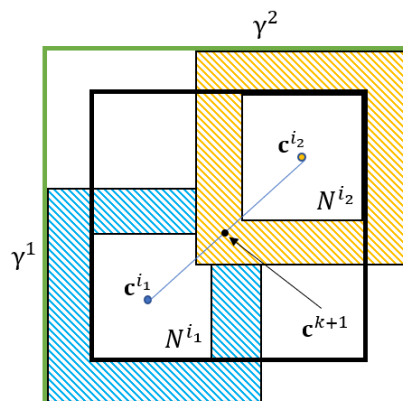
$$\begin{aligned}\underline{c}_j^{k+1} &= \min\{\underline{c}_j^{i_1}, \underline{c}_j^{i_2}\} \\ \underline{c}_j^{k+1} &= \min\{\underline{c}_j^{i_1}, \underline{c}_j^{i_2}\} \\ \bar{c}_j^{k+1} &= \max\{\bar{c}_j^{i_1}, \bar{c}_j^{i_2}\} \\ \bar{c}_j^{k+1} &= \max\{\bar{c}_j^{i_1}, \bar{c}_j^{i_2}\}\end{aligned}\tag{4.16}$$

$j = 1, \dots, n$. Então, o novo centro

$$\mathbf{c}^{k+1} = \frac{1}{2} (\bar{\mathbf{c}}^{k+1} - \underline{\mathbf{c}}^{k+1})\tag{4.17}$$

é o ponto médio do grânulo. À primeira vista, esta abordagem produz grânulos maiores, aumentando a área de cobertura do modelo e, simultaneamente, preservando as informações adquiridas em iterações passadas. Tais grânulos tendem a ser reduzidos nas iterações futuras, por meio dos procedimentos de balanceamento granular. Figura 4.6 exemplifica o procedimento no espaço bidimensional.

Figura 4.6 – Abordagem de mesclagem de grânulos Fuzzy eIX por envoltório convexo



Fonte: Arquivo Pessoal

4.10 Sumário

O algoritmo não-supervisionado Fuzzy eIX é sumarizado na listagem 4.7. O classificador resultante é parametricamente e estruturalmente adaptável. Portanto, consegue ser aplicado inclusive à problemas que apresentam fluxos de dados não estacionários.

Figura 4.7 – Algoritmo de aprendizagem: Fuzzy eIX

Listagem 1: FuzzyEIX (ϵ, ρ)

Result: $\{\gamma\}$: um conjunto de grânulos
 $\gamma \leftarrow []$, $c = 0$, $\alpha = \beta = 0.3$
foreach $x^{[h]}$, $h = 1, \dots$ (*fluxo de dados ativo*) **do**
 if $h = 1$ **then**
 $\gamma^{c+1} \leftarrow$ CriarGranulo($x^{[h]}$, ϵ)
 else
 foreach $\gamma^j \in \gamma$ **do**
 if $x^{[h]} \in \gamma^j.$ *Interna* **then**
 EncolherInterna($x^{[h]}$, γ^j , ϵ)
 EncolherExterna($x^{[h]}$, γ^j , ϵ)
 ArrastarCentro($x^{[h]}$, γ^j)
 else if $x^{[h]} \in \gamma^j.$ *Externa* **then**
 ExpandirInterna($x^{[h]}$, γ^j)
 ExpandirExterna($x^{[h]}$, γ^j)
 else
 $\gamma^{c+1} \leftarrow$ CriarGranulo($x^{[h]}$, ϵ)
 end
 end
 end
 MesclarGranulos(γ , ρ)
 BalancearGranulos(γ , α)
end

Fonte: Arquivo Pessoal

5 METODOLOGIA

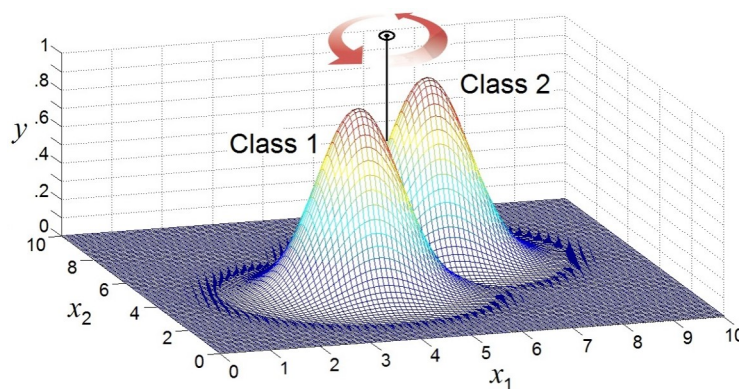
Neste capítulo será apresentada a metodologia adotada na realização de experimentos. Inicialmente, será apresentado o experimento das gaussianas rotativas, utilizado para averiguação de forma isolada da performance do modelo Fuzzy eIX frente a dados não estacionários. A seguir, será fornecido um detalhamento dos conjuntos de dados *benchmark* empregados na averiguação e comparação entre a performance dos modelos evolutivos em cenários mais realísticos (dados não sintéticos) e onde há variação temporal das distribuição dos dados (dados sintéticos). Finalmente, será apresentada a metodologia de comparação dos resultados, que envolve a utilização dos índices de validação incrementais apresentados no capítulo 2.

5.1 Rotação de Gaussianas Gêmeas

O experimento denominado "Rotação de Gaussianas Gêmeas" (LEITE, 2012) consiste em um problema de classificação não-estacionário proposto para avaliar a efetividade do método Fuzzy eIX. A não-estacionariedade dos dados implica diretamente que a média das distribuições Gaussianas não será constante ao longo do tempo.

Um fluxo de dados é gerado a partir de duas funções Gaussianas que se sobrepõem parcialmente. Estas função têm suas médias rotacionadas em relação a um ponto fixo, a partir de um determinado número de amostras geradas. A Figura 5.1 ilustra o processo.

Figura 5.1 – Experimento das Gaussianas Rotativas em torno do ponto (5,5)



Fonte: (LEITE, 2012)

Para o experimento discutido neste trabalho, as funções Gaussianas são inicialmente posicionadas com seus centros (médias) em $g_1^{[0]} = (4,4)$ e $g_2^{[0]} = (6,6)$, e possuem um desvio padrão de 0.8. Elas rotacionam em torno do ponto (5,5) de acordo com

$$\begin{aligned}
\theta_i^{[h]} &= \theta_i^{[h-1]} + \phi \\
g_{i(1)}^{[h]} &= 5 + \sqrt{2} \cos(\theta_i^{[h]}) \\
g_{i(2)}^{[h]} &= 5 + \sqrt{2} \sin(\theta_i^{[h]}),
\end{aligned} \tag{5.1}$$

onde θ_i é um ângulo medido na direção anti-horária entre o ponto fixo (5,5) e o centro da i -ésima Gaussiana. Inicialmente, os centros são (4,4) e (6,6), portanto, $\theta_1^{[0]} = 45$ graus para a ‘Classe 1’, e $\theta_2^{[0]} = 225$ graus para a ‘Classe 2’. Considera-se, então, dois estágios. No primeiro, a taxa de rotação, ϕ , é nula, i.e., as Gaussianas estão paradas (estacionárias). Este estágio dura $h = 200$ passos. A seguir, a taxa de rotação, ϕ , vale 0.45 de $h = 201$ até $h = 400$. Uma instância $\mathbf{x}^{[h]}$ é produzida por uma das Gaussianas por passo h .

Os dados são normalizados no intervalo $[0, 1]$. Assume-se que a primeira classe é a classe positiva (P). Para a quantificação da performance do modelo Fuzzy eIX neste experimento, considera-se uma matriz de confusão de duas colunas e duas linhas, que representam o número de "verdadeiros positivos" (TP), "falsos positivos" (FP), "verdadeiros negativos" (TN) e "falsos negativos" (FN) (FAWCETT, 2006). A acurácia do classificador é obtida a partir de

$$Acc(\%) = \frac{TP + TN}{TP + FP + TN + FN} \cdot 100\%. \tag{5.2}$$

5.2 Clusterização de Conjuntos de Dados Reais

O objetivo deste trabalho é propor um novo *framework* de aprendizado *online* denominado Fuzzy eIX e demonstrar aplicações do método a problemas de classificação não-supervisionada. Para isto, três conjuntos de dados para *benchmark* são utilizados para verificar e comparar as performances de modelos Fuzzy eIX e TEDA-Cluster. Estes conjuntos de dados possuem grande número de instâncias, atributos e classes e, além disso, apresentam *concept drifts* abruptos, graduais e incrementais. Estas bases são amplamente utilizados na literatura para verificação de performance de classificação de dados multi-classes e não estacionários.

1. *LED*: O conjunto LED foi introduzido em Breiman et al. (1984). O conjunto possui um total de 1.000.000 de instâncias. Cada instância gerada possui 24 atributos, sendo que destes apenas 7 são relevantes e correspondem a cada segmento de um *display* LED de sete segmentos, indicando se o segmento está energizado (1) ou não (0). O objetivo

é, dada uma instância, classificar o dígito mostrado pelo *display*, i.e., designando um problema de 10 classes (dígitos de 0 a 9). Cada atributo tem uma chance de 10% de estar invertido. Este gerador simula *concept drifts* intercalando os atributos relevantes e irrelevantes entre si. Os atributos são intercalados para 50.000 instâncias e são dispostos da seguinte maneira: a partir da amostra 250.000, três atributos são intercalados; a partir da amostra 500.000, cinco atributos são intercalados, e; a partir da amostra 750.000, sete atributos são intercalados. Existem duas variações deste conjunto de dados: i) LED_a , que possui *drifts* abruptos, e; ii) LED_g , que possui *drifts* graduais. Neste trabalho, considera-se o subconjunto LED_a ;

2. *HYPER*: O conjunto de dados *HYPER* simula *concept drifts* incrementais e é concebido a partir do gerador hiperplano introduzido em (HULTEN; SPENCER; DOMINGOS, 2001). Um hiperplano é um subconjunto de pontos que divide o espaço $n - dimensional$ onde está inserido em duas partes independentes. É possível alterar a orientação e posição do hiperplano ajustando-se suavemente seus parâmetros (pesos), $w_i, i = 0, \dots, n$, onde n é o número de atributos. Este gerador pode ser utilizado na simulação de *concept drifts* a partir da variação de seus parâmetros ao longo do tempo (BIFET et al., 2011). O conjunto *HYPER* possui 1.000.000 instâncias e é parametrizado com $n = 10$ atributos w_i e uma taxa de variação incremental em cada atributo de 0,001.

Os modelos evolutivos Fuzzy eIX e TEDA-Cluster serão utilizados na classificação dos conjuntos de dados apresentados. Estes conjuntos de dados consistem em "fotografias" de processos e não são ativos, isto é, dados não são produzidos ao longo do tempo. No entanto, os algoritmos considerados neste trabalho são projetados para trabalhar em cenários mais realísticos, onde os dados são apresentados aos modelos em tempo real e em grande volume, i.e., de maneira *online*. Naturalmente, estes cenários podem ser criados em experimentos computacionais, apresentando cada instância individualmente aos modelos e desenvolvendo as rotinas de adaptação para cada instância. Este procedimento será utilizado na condução dos experimentos com os conjuntos de dados.

Os dados foram redimensionados na escala $[0, 1]$ para que as variáveis sejam consideradas na mesma proporção, i.e., atributos com valores muito elevados ou muito pequenos não viesem os algoritmos de aprendizado. A normalização é realizada de maneira recursiva, de modo a proporcionar consistência em um ambiente online. Isto significa que os valores de má-

ximo e mínimo observados são atualizados a cada nova instância de dados (SOARES et al., 2017).

5.3 Software, Hardware e Avaliação de Desempenho

Além dos valores de diferentes índices de validação, os modelos também serão comparados quanto ao tempo de processamento médio. Esta média será calculada para 10 execuções de cada algoritmo, para cada um dos conjuntos de dados *benchmark*. Os algoritmos rodam em ambiente do *software Matlab 2014a*, em um computador com CPU de *2.50GHz* e memória RAM de *8Gb*.

Para cálculo das acurácias e apuração da capacidade de generalização de cada um dos modelos, os conjuntos de dados foram divididos em subconjuntos de treinamento, com 70% das amostras, e teste, com 30%. Adicionalmente, os resultados serão comparados em termos do número de grânulos gerados e, mais notadamente, equilíbrio das granularidades obtidas por cada modelo considerado.

A qualidade das partições geradas por cada um dos algoritmos será avaliada por meio da aplicação dos índices de validação incrementais de Xie-Beni e Davies-Bouldin discutidos no Capítulo 2, ver equações (2.11) e (2.12). Para obtenção destes índices, os graus de pertinência da h -ésima instância à i -ésima partição, μ_{ih} , serão calculados como

$$\mu_{ih} = \frac{1}{\sum_{l=1}^k \left(\frac{L_{ih}}{L_{lh}} \right)}, \quad (5.3)$$

onde L_{AB} é uma medida de distância entre a partição A e a instância B . Neste trabalho, será considerada a norma de Chebyshev, ou norma infinito, dada por

$$L_{\infty} = \|\mathbf{x}^{[h]} - \mathbf{c}^i\| = \lim_{p \rightarrow \infty} \left(\sum_{j=1}^n |x_j^{[h]} - c_j^i|^p \right)^{\frac{1}{p}} \quad (5.4)$$

6 RESULTADOS

Neste capítulo, a performance de modelos Fuzzy eIX será explorada a partir dos resultados obtidos da clusterização de fluxos de dados não estacionários. Primeiramente, serão explorados aspectos do modelo frente a dados sintéticos não-estacionários, por meio do experimento das Gaussianas rotativas (ver seção 5.1). A acurácia do modelo classificador será obtida considerando-se uma matriz de confusão (5.2) para diversas parametrizações (ε , ρ) distintas.

A seguir, índices de validação incrementais serão utilizados para verificar o desempenho e qualidade das partições geradas por modelos Fuzzy eIX e TEDA-Cluster ao clusterizar os conjuntos de dados reais e não-estacionários LED e HYPER, discutidos na seção 5.2.

Finalmente, algumas considerações finais serão abordadas sobre a performance de modelos Fuzzy eIX frente aos outros dois modelos abordados e serão discutidos alguns pontos potenciais em melhorar o desempenho do mesmo e que podem ser abordados no futuro.

6.1 Rotação de Gaussianas Gêmeas

Inicialmente, a acurácia do modelo Fuzzy eIX será apurada na clusterização das Gaussianas estacionárias, i.e., sem que haja rotação de seus centros. Este estágio acontece para as amostras $h = 0, \dots, 200$. Os centros iniciais são posicionados em $\theta_1^{[0]} = 45$ graus e $\theta_2^{[0]} = 225$ graus. A seguir, a acurácia será apurada de maneira análoga para as amostras $h = 201, \dots, 400$, onde os centros estarão sendo rotacionados a uma taxa de $\phi = 0.45$. Uma instância $x^{[h]} = (g_{i(1)}^{[h]}, g_{i(2)}^{[h]})$ é produzida por uma das Gaussianas por passo h , conforme (5.1). Durante ambos estágios, o método de mesclagem por envoltório convexo (seção 4.9) será considerado.

A tabela 6.1 ilustra os resultados obtidos para cada um dos estágios considerados, variando-se os valores dos parâmetros ε , ρ nos intervalos $[0.035, 0.065]$ e $[0.25, 0.65]$, respectivamente. Estes intervalos foram determinados por tentativa e erro. Os cenários onde a acurácia apurada foi a maior estão de negrito.

Nota-se que para os parâmetros $\varepsilon = 0.055$ e $\rho = 0.45$, uma acurácia de 94.5% e 90.5%, gerando uma média de 3.12 e 5.69 grânulos foi alcançada, respectivamente, para os estágios estacionário e não-estacionário. Para valores mais elevados de ε , a acurácia na clusterização é degradada, uma vez que grânulos mais largos são criados e acabam comportando instâncias de classes diferentes assim que as primeiras instâncias do fluxo de dados chegam (iterações iniciais). Analogamente, para valores menores de ε , são gerados grânulos menores que comportam uma quantidade muito inferior de instâncias e que podem ser interpretados como classes

Tabela 6.1 – Fuzzy eIX - Rotação de Gaussianas Gêmeas

Estágio estacionário (200 instâncias)			
Parâmetros $\{\varepsilon, \rho\}$	Acu (%)	Nro. Médio de Grânulos	Tempo (s)
$\{0.035, 0.25\}$	85.0	6.01	0.26
$\{0.045, 0.35\}$	88.0	5.78	0.10
$\{0.055, 0.45\}$	94.5	3.12	0.18
$\{0.065, 0.55\}$	84.5	4.41	0.21
$\{0.055, 0.55\}$	90.0	3.86	0.19
$\{0.065, 0.65\}$	81.0	3.07	0.24
Estágio não-estacionário (200 instâncias)			
Parâmetros $\{\varepsilon, \rho\}$	Acu (%)	Nro. Médio de Grânulos	Tempo (s)
$\{0.035, 0.25\}$	81.0	12.78	0.45
$\{0.045, 0.35\}$	83.0	9.32	0.41
$\{0.055, 0.45\}$	90.5	6.69	0.36
$\{0.065, 0.55\}$	79.5	4.12	0.31
$\{0.065, 0.45\}$	85.5	5.43	0.35
$\{0.065, 0.65\}$	77.5	4.25	0.29

Fonte: Arquivo Pessoal

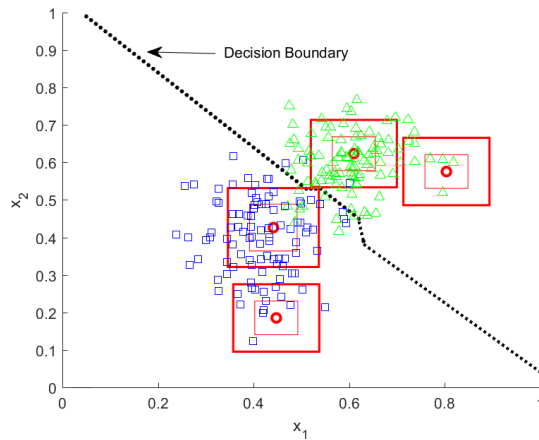
distintas pelo algoritmo de aprendizado associado. Além disso, nota-se que, em média, foram necessários 3.5 grânulos adicionais para cobrir o espaço dos dados e manter a acurácia durante o estágio não-estacionário. Este fato destaca o comportamento adaptativo e incremental do modelo em perceber variações na distribuição dos dados e se ajustar, de modo a manter um nível mínimo de acurácia.

As figuras 6.1 e 6.2 mostram os grânulos finais gerados com a parametrização $\varepsilon = 0.055$ e $\rho = 0.45$ para os estágios estacionário e não-estacionário, respectivamente. Também é mostrada a fronteira de decisão para cada caso.

É importante perceber os tamanhos similares de cada grânulo gerado ao final do processo de clusterização, mostrando a granularidade balanceada garantida pelo processo discutido na seção 4.6. Nota-se que alguns grânulos adicionais foram gerados para cobrir o espaço de dados durante a rotação das Gaussianas, porém não foram mesclados até o fim do estágio não-estacionário. Em especial, os três grânulos superiores na figura 6.2 poderiam ter sido mesclados e redimensionados, de modo a se ajustarem melhor à classe 2. Analogamente, os grânulos inferiores (classe 1) poderiam ter se ajustado da mesma maneira.

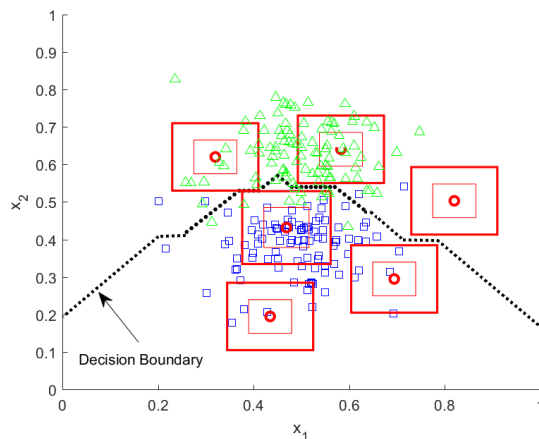
A figura 6.3 mostra a evolução do número de grânulos ao longo das instâncias do fluxo de dados, $[h]$.

Figura 6.1 – Posição final dos grânulos gerados no estágio estacionário e respectiva fronteira de decisão.



Fonte: Arquivo Pessoal

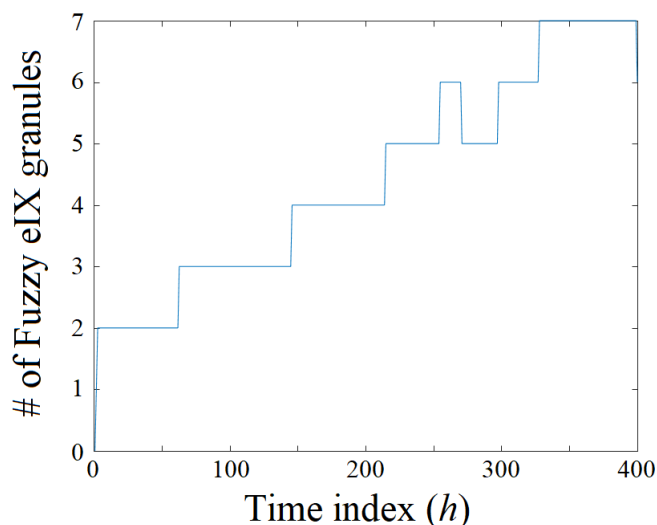
Figura 6.2 – Posição final dos grânulos gerados no estágio não-estacionário e respectiva fronteira de decisão.



Fonte: Arquivo Pessoal

Por meio da figura 6.3, nota-se um aumento expressivo e mais rápido do número de grânulos gerados a partir de $[h] = 200$, que é exatamente quando o estágio não-estacionário tem início. Estes grânulos adicionais são criados para endereçar o *drift* observado nas classes. Os grânulos são criados durante o processo *online* de clusterização (*on-the-fly*) para que as novas instâncias sejam cobertas e assinaladas a algum grau de pertinência a algum grânulo. É também notável que, de acordo com a tabela 6.1, a acurácia do modelo foi pouco impactada pelo *concept drift* simulado neste fluxo de dados, uma vez que adaptação paramétrica e estrutural são consideradas no algoritmo Fuzzy eIX.

Figura 6.3 – Evolução estrutural do modelo classificador Fuzzy eIX..



Fonte: Arquivo Pessoal

6.2 Clusterização de Conjuntos de Dados Reais

Para a clusterização de dados reais e não-estacionários, o algoritmo Fuzzy eIX não foi capaz de produzir resultados satisfatórios com a parametrização não adaptativa de ε , ρ , i.e., com valores de ε , ρ que não se alteram em função da estrutura atual do modelo durante as iterações do processo de aprendizado, ou a cada nova amostra que chega do fluxo de dados.

Foi notado que nos momentos onde ocorrem *concept drifts* existe um aumento relativamente excessivo dos grânulos eIX. Este aumento é feito para que toda a área de dados seja coberta por pelo menos um grânulo, mesmo que as distribuições dos dados esteja variando. Porém, o processo de mesclagem com ρ **fixado** não é eficiente ao executar a mesclagem destes grânulos gerados excessivamente, e portanto, redundantes, em iterações anteriores.

O valor fixado do parâmetro ε também pode ter prejudicado o desempenho do algoritmo em produzir um modelo satisfatório. O tamanho inicial e mínimo dos grânulos podem ter impacto direto no momento em que grânulos são mesclados, especialmente ao se utilizar o método de mesclagem por envoltório convexo, que leva em consideração a interseção entre-grânulos. Ao processar amostras durante os intervalos onde os *concept drifts* ocorrem poderia gerar grânulos maiores que seriam mesclados entre si posteriormente.

Além disso, um processo de deleção de grânulos não relevantes poderia ajudar a estabilizar o algoritmo. Os grânulos não relevantes seriam deletados e teriam impactos negativos no processo de balanceamento da granularidade, seja equalizando os demais grânulos (mais importantes) para dimensões maiores ou menores.

É importante notar alguns fatores que condicionaram bom desempenho do algoritmo no experimento das Gaussianas gêmeas rotativas, mas não para os conjuntos de dados reais *LED* e *HYPER*:

1. A velocidade dos *concept drifts* presentes nos conjuntos *LED* e *HYPER* é consideravelmente maior do que aquela utilizada no experimento das Gaussianas gêmeas rotativas, onde as instâncias eram rotacionadas a uma taxa de apenas 0.45 unidades/instância. Por exemplo, para o conjunto *LED*, os atributos são intercalados a uma taxa de 5 unidades/instâncias, já na primeira rodada em 50.000 instâncias. Em outras palavras, os *drifts* são muito mais abruptos nesses conjuntos de dados. A sensibilidade do algoritmo para valores fixados de ϵ , ρ não foi suficiente para acompanhar o intercalação dos atributos presentes nesses conjuntos de dados;
2. É sabido que tais conjuntos de dados possuem desbalanceamento de classes em suas instâncias, i.e., existem muito mais instâncias de determinadas classes do que outras. Este fato tem impacto negativo na generalização do algoritmo, uma vez que pode enviesar a determinadas classes majoritárias. No caso do experimento das Gaussianas gêmeas rotativas, os dados eram artificialmente gerados de forma perfeitamente balanceada, isto é, as duas classes possuíam o mesmo número de instâncias.

Na seção a seguir, serão discutidas alternativas que podem possibilitar um aumento do desempenho do algoritmo Fuzzy eIX frente a conjunto de dados reais, não estacionários como o *LED* e *HYPER* e que podem ser consideradas nas próximas fases de desenvolvimento do projeto.

7 CONCLUSÕES

Neste trabalho foi introduzido um novo ambiente de desenvolvimento, ou *framework* para clusterização não-supervisionada, denominado *Fuzzy Internal-External Clustering*, Fuzzy eIX. Nesta metodologia, atribuí-se a cada cluster regiões internas e externas no espaço dos dados. As regiões internas são construídas de modo a denotar regiões mais próximas ao centro dos clusters, implicando maior certeza de pertinência. Por outro lado, as regiões externas são construídas de modo a denotar incertezas de pertinência dos dados. Mais notavelmente, mostrou-se como as dimensões destas regiões são adaptadas de acordo com o perfil do fluxo de dados, i.e., de forma *online*. Não há necessidade de conhecimento dos padrões de classificação desejados em nenhum instante, atribuindo ao método a característica de ser *não supervisionado*. O eIX incorpora mecanismo de fusão de clusters similares, aferindo capacidade de compressão de conhecimento similar. Adicionalmente, a partir de um modelo eIX é possível derivar sistemas de inferência fuzzy dos tipos I ou II evolutivos, além de criar regiões de incerteza locais e globais. A capacidade de explicação de tomadas de decisão é melhorada.

Inicialmente, foi apresentada uma contextualização sobre algoritmos granulares evolutivos de classificação não supervisionada. Também foram apresentados conceitos importantes de algoritmos de classificação fuzzy do tipo 2, balanceamento de granularidade, índices de validação de partições incrementais (Xie-Beni e Davis-Bouldin) e *eXplainable Artificial Intelligence* (XAI).

As decorrências de um modelo Fuzzy eIX foram apresentadas no capítulo 3 e ajudam a mostrar e entender como um modelo fuzzy eIX pode ser utilizado junto a algoritmos de aprendizado associados. As decorrências apresentadas são:

1. derivação de sistemas de inferência Fuzzy evolutivos do tipo 1;
2. derivação de sistemas de inferência Fuzzy evolutivos do tipo 2;
3. análise de incertezas associadas ao fluxo de dados em níveis inter-granulares local e global;

O algoritmo Fuzzy eIX foi então descrito no capítulo 4, onde foram apresentados os passos executados ao se processar uma nova instância recém chegada do conjunto de dados. Neste capítulo, também discutiu-se sobre como o algoritmo realiza os processos de mesclagem e balanceamento da granularidade, essenciais para a interpretabilidade e explicabilidade das decisões tomadas pelo algoritmo de aprendizado.

No capítulo 6, foram apresentados os resultados dos experimentos conduzidos neste trabalho. O primeiro experimento consistia em avaliar o desempenho do algoritmo Fuzzy eIX na classificação de dados de duas classes provenientes de uma distribuição Gaussiana que sofrem rotações de seus centros a partir de um dado instante. Foi mostrado que o algoritmo Fuzzy eIX foi capaz de produzir um modelo classificador com bons índices de desempenhos baseados em uma matriz de confusão. O desempenho do classificador foi aferido para diversas parametrizações de ε e ρ , onde a melhor performance foi dada para $\varepsilon = 0.55$ e $\rho = 0.45$, para os estágios estacionário e não-estacionário, respectivamente. As acurácias obtidas foram de 94.5% e 90.5%, respectivamente. O segundo experimento envolveu a obtenção de modelo classificador de dados não estacionários provenientes dos datasets *LED* e *HYPER*. Para estes conjuntos, o desempenho do classificador não foi satisfatório. Notou-se que o número de grânulos gerados foi extremamente alto e que o processo de mesclagem com os parâmetros ε e ρ fixados não foi eficiente em mesclar grânulos redundantes. Isso se deve basicamente a dois fatores: a) os *concept drifts* nos conjuntos *LED* e *HYPER* acontecem de forma mais abrupta e; b) tais conjuntos de dados possuem desbalanceamento de classes. Uma possível solução seria adaptar os parâmetros ε e ρ de acordo com a estrutura atual do modelo, tornando-os adaptativos.

7.1 Alternativas em ambiente eIX

Na metodologia eIX proposta neste trabalho diversas combinações nas formas de parametrização são possíveis, de modo que a customização do modelo é muito abrangente. Dentre elas, destacam-se:

1. As dimensões iniciais dos clusters, bem como as dimensões mínimas podem são reguladas pelo parâmetro ε ; este pode ser variado ao longo do processo de aprendizado em função da estrutura atual do modelo, tais como: a) número de grânulos; b) tamanho médio dos grânulos, e; c) número de instâncias associadas a cada grânulo, ou seja, com o maior grau de pertinência. O mesmo processo pode ser aplicado ao parâmetro ρ , que controla as mesclagens, porém, este parâmetro deveria se adaptar em função de medidas entre-clusters, como, por exemplo, distâncias entre os centros;
2. Os arrastes realizados no centro do cluster no processo de adaptação descrito na seção 4.3 são ponderados apenas pela densidade de cada cluster. É possível inserir outras

medidas no cálculo deste arraste, i.e., dimensões das regiões interna e externa no intuito de arrastar o centro proporcionalmente à quantidade de incerteza atribuída ao cluster;

3. Ao cálculo dos graus de pertinência é possível a inclusão de parâmetros que definem a fuzzicidade dos clusters. Ao mesmo tempo, é possível estudar efeitos de variações do parâmetro ρ (limiar para mesclagem) na fuzzicidade dos clusters criados;
4. Ainda referente à mesclagem, há possibilidade de variação no cálculo de similaridade entre dois clusters. Pode-se, por exemplo, utilizar distâncias de Hausdorff¹ simples, ou modificadas, que considerem a densidade de cada cluster como critério de decisão de mesclagem.
5. A velocidade do balanceamento das dimensões dos clusters é manipulada pelo hiperparâmetro α ; variações neste parâmetro podem ser analisadas no sentido de se obter um valor ótimo. Além disso, este seria outro parâmetro candidato a sofrer adaptações ao longo do processo de aprendizado *online*. A regra de adaptação pode ser baseada nas razões entre as dimensões dos diferentes grânulos na iteração atual.

Em meio a tantas possibilidades de otimização e customização, o eIX configura como um *ambiente* de desenvolvimento.

¹ Distâncias de Hausdorff consideram distâncias otimistas (vértices superiores à vértices inferiores) e pessimistas (vértices superiores à vértices superiores) em sua formulação.

8 CRONOGRAMA DE ATIVIDADES

Considerando a proposta apresentada e as exigências do Programa de Pós-Graduação em Engenharia de Sistemas e Automação, são especificadas as atividades a fim de elaborar um cronograma para execução, avaliação e finalização da dissertação de mestrado:

1. Revisão bibliográfica sobre o tema do projeto, buscando estudar, reunir e organizar publicações a respeito do assunto; compreender algoritmos já estabelecidos na literatura; levantar principais autores da área;
2. Desenvolvimento da abordagem: levantamento de contribuições e desenvolvimento do método proposto, formulações matemáticas formais, descrições de implicações e conceitos do mesmo;
3. Implementação da abordagem: utilização da ferramenta de computação científica MATLAB® para implementação do algoritmo proposto e elaboração das diversas simulações computacionais;
4. Execução e análise dos resultados: comparação entre os principais métodos de classificação evolutiva consolidados na literatura, levantados na atividade 1, com o método proposto;
5. Redação da dissertação: escrita da dissertação e defesa perante banca examinadora;
6. Publicações: escrita e submissão de artigos para publicação em conferências nacionais e internacionais.

O Quadro 8.1 dispõe as atividades citadas em um cronograma desde o início das atividades (2019/2) até a data prevista para defesa da dissertação. Cada marcação (X) corresponde a um bimestre do ano correspondente.

Quadro 8.1 – Cronograma de Atividades

Atividades Propostas	2019		2020				2021	
	3 ^o	4 ^o	1 ^o	2 ^o	3 ^o	4 ^o	1 ^o	2 ^o
1	X	X	X	X	X	X		
2		X	X	X	X	X		
3			X	X	X	X		
4				X	X	X	X	X
5						X	X	X
6					X	X	X	X

Fonte: Arquivo Pessoal

9 ORÇAMENTO

Este projeto será executado com recursos, i.e., notebook e *softwares*, próprios do autor. Não serão utilizados fundos de apoio ou recursos externos. Os computadores e *softwares* de projeto externo podem ser úteis ocasionalmente.

REFERÊNCIAS

- ADADI, A.; BERRADA, M. Peeking inside the black-box: A survey on explainable artificial intelligence (xai). **IEEE Access**, IEEE, v. 6, p. 52138–52160, 2018.
- AISBETT, J.; RICKARD, J. T.; MORGENTHALER, D. G. Type-2 fuzzy sets as functions on spaces. **IEEE Transactions on Fuzzy Systems**, IEEE, v. 18, n. 4, p. 841–844, 2010.
- ANGELOV, P. Anomaly detection based on eccentricity analysis. In: IEEE. **2014 IEEE symposium on evolving and autonomous learning systems (EALS)**. [S.l.], 2014. p. 1–8.
- ANGELOV, P. Outside the box: an alternative data analytics framework. **Journal of Automation Mobile Robotics and Intelligent Systems**, v. 8, n. 2, p. 29–35, 2014.
- ANGELOV, P.; FILEV, D. P.; KASABOV, N. **Evolving intelligent systems: methodology and applications**. [S.l.]: John Wiley & Sons, 2010. v. 12.
- ANGELOV, P. P. **Evolving rule-based models: a tool for design of flexible adaptive systems**. [S.l.]: Physica, 2013. v. 92.
- ANGELOV, P. P.; ZHOU, X. Evolving fuzzy-rule-based classifiers from data streams. **Ieee transactions on fuzzy systems**, IEEE, v. 16, n. 6, p. 1462–1475, 2008.
- BARGIELA, A.; PEDRYCZ, W. Granular computing. In: **Handbook on Computational Intelligence**. [S.l.]: World Scientific, 2016. p. 43–66.
- BIFET, A. et al. Moa data stream mining-a practical approach. **The University of Waikato**, p. 107–139, 2011.
- BILGIN, A. et al. Towards a general type-2 fuzzy logic approach for computing with words using linear adjectives. In: IEEE. **2012 IEEE International Conference on Fuzzy Systems**. [S.l.], 2012. p. 1–8.
- BOUCHACHIA, A.; GABRYS, B.; SAHEL, Z. Overview of some incremental learning algorithms. In: IEEE. **2007 IEEE International Fuzzy Systems Conference**. [S.l.], 2007. p. 1–6.
- BREIMAN, L. et al. **Classification and regression trees**. [S.l.]: CRC press, 1984.
- BYRNE, R. M. Counterfactuals in explainable artificial intelligence (xai): Evidence from human reasoning. In: **IJCAI**. [S.l.: s.n.], 2019. p. 6276–6282.
- CORDOVIL, L. A. Q. et al. Uncertain data modeling based on evolving ellipsoidal fuzzy information granules. **IEEE Transactions on Fuzzy Systems**, IEEE, 2019.
- DAVIES, D. L.; BOULDIN, D. W. A cluster separation measure. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, n. 2, p. 224–227, 1979.
- FAWCETT, T. An introduction to roc analysis. **Pattern recognition letters**, Elsevier, v. 27, n. 8, p. 861–874, 2006.
- GAMA, J. **Knowledge discovery from data streams**. [S.l.]: CRC Press, 2010.

GARCIA, C. et al. Evolvable fuzzy systems from data streams with missing values: With application to temporal pattern recognition and cryptocurrency prediction. **Pattern Recognition Letters**, Elsevier, v. 128, p. 278–282, 2019.

GAROFALAKIS, M.; GEHRKE, J.; RASTOGI, R. **Data stream management: processing high-speed data streams**. [S.l.]: Springer, 2016.

GUNNING, D. Explainable artificial intelligence (xai). **Defense Advanced Research Projects Agency (DARPA), nd Web**, v. 2, p. 2, 2017.

HULTEN, G.; SPENCER, L.; DOMINGOS, P. Mining time-changing data streams. In: **Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining**. [S.l.: s.n.], 2001. p. 97–106.

IBRAHIM, O. A.; KELLER, J. M.; BEZDEK, J. C. Analysis of streaming clustering using an incremental validity index. In: IEEE. **2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)**. [S.l.], 2018. p. 1–8.

IBRAHIM, O. A.; WANG, Y.; KELLER, J. M. Analysis of incremental cluster validity for big data applications. **International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems**, World Scientific, v. 26, n. Suppl. 2, p. 47–62, 2018.

KANGIN, D.; ANGELOV, P. Evolving clustering, classification and regression with teda. In: IEEE. **2015 International Joint Conference on Neural Networks (IJCNN)**. [S.l.], 2015. p. 1–8.

KANGIN, D.; ANGELOV, P. New autonomously evolving classifier teda class. In: **IEEE International Joint Conference on Neural Networks, IJCNN-2015**. [S.l.: s.n.], 2015.

KARNIK, N. N.; MENDEL, J. M.; LIANG, Q. Type-2 fuzzy logic systems. **IEEE transactions on Fuzzy Systems**, IEEE, v. 7, n. 6, p. 643–658, 1999.

KASABOV, N. K. **Evolving connectionist systems: the knowledge engineering approach**. [S.l.]: Springer Science & Business Media, 2007.

KOUTRIKA, G.; ZADEH, Z. M.; GARCIA-MOLINA, H. Data clouds: summarizing keyword search results over structured data. In: **Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology**. [S.l.: s.n.], 2009. p. 391–402.

LAÑA, I. et al. Adaptive long-term traffic state estimation with evolving spiking neural networks. **Transportation Research Part C: Emerging Technologies**, Elsevier, v. 101, p. 126–144, 2019.

LEITE, D. **Evolving Granular Systems**. Tese (Doutorado) — State University of Campinas (UNICAMP), Sao Paulo - Brazil, 2012.

LEITE, D.; COSTA, P.; GOMIDE, F. Evolving granular neural networks from fuzzy data streams. **Neural Networks**, Elsevier, v. 38, p. 1–16, 2013.

LEITE, D. et al. Fuzzy granular evolving modeling for time series prediction. In: IEEE. **2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)**. [S.l.], 2011. p. 2794–2801.

- LEITE, D. et al. Evolving granular fuzzy model-based control of nonlinear dynamic systems. **IEEE Transactions on Fuzzy Systems**, IEEE, v. 23, n. 4, p. 923–938, 2014.
- LEITE, D.; ŠKRJANC, I. Ensemble of evolving optimal granular experts, owa aggregation, and time series prediction. **Information Sciences**, Elsevier, v. 504, p. 95–112, 2019.
- LEITE, D.; ŠKRJANC, I.; GOMIDE, F. An overview on evolving systems and learning from stream data. **Evolving Systems**, Springer, p. 1–18, 2020.
- LEITE, D. F.; COSTA, P.; GOMIDE, F. Interval-based evolving modeling. In: IEEE. **2009 IEEE Workshop on Evolving and Self-Developing Intelligent Systems**. [S.l.], 2009. p. 1–8.
- LIO, W.; LIU, B. Residual and confidence interval for uncertain regression model with imprecise observations. **Journal of Intelligent & Fuzzy Systems**, IOS Press, v. 35, n. 2, p. 2573–2583, 2018.
- LIU, B. Uncertainty theory. In: **Uncertainty theory**. [S.l.]: Springer, 2007. p. 205–234.
- LUGHOFER, E. **Evolving fuzzy systems-methodologies, advanced concepts and applications**. [S.l.]: Springer, 2011. v. 53.
- LUGHOFER, E.; ANGELOV, P. Handling drifts and shifts in on-line data streams with evolving fuzzy systems. **Applied Soft Computing**, Elsevier, v. 11, n. 2, p. 2057–2068, 2011.
- LUGHOFER, E. et al. Generalized smart evolving fuzzy systems. **Evolving systems**, Springer, v. 6, n. 4, p. 269–292, 2015.
- LUGHOFER, E.; PRATAMA, M.; ŠKRJANC, I. Incremental rule splitting in generalized evolving fuzzy systems for autonomous drift compensation. **IEEE Transactions on Fuzzy Systems**, IEEE, v. 26, n. 4, p. 1854–1865, 2017.
- MENDEL, J. M. Advances in type-2 fuzzy sets and systems. **Information sciences**, Elsevier, v. 177, n. 1, p. 84–110, 2007.
- MENDEL, J. M. Type-2 fuzzy sets and systems: an overview. **IEEE computational intelligence magazine**, IEEE, v. 2, n. 1, p. 20–29, 2007.
- MENDEL, J. M.; JOHN, R. B. Type-2 fuzzy sets made simple. **IEEE Transactions on fuzzy systems**, IEEE, v. 10, n. 2, p. 117–127, 2002.
- MOSHTAGHI, M. et al. Online cluster validity indices for streaming data. **arXiv preprint arXiv:1801.02937**, 2018.
- MOSHTAGHI, M. et al. Online cluster validity indices for performance monitoring of streaming data clustering. **International Journal of Intelligent Systems**, Wiley Online Library, v. 34, n. 4, p. 541–563, 2019.
- OSHERSON, D.; SMITH, E. E. On typicality and vagueness. **Cognition**, Elsevier, v. 64, n. 2, p. 189–206, 1997.
- PEDRYCZ, W. Granular computing: an introduction. In: IEEE. **Proceedings joint 9th IFSA world congress and 20th NAFIPS international conference (Cat. No. 01TH8569)**. [S.l.], 2001. v. 3, p. 1349–1354.

- PEDRYCZ, W. **Granular computing: analysis and design of intelligent systems**. [S.l.]: CRC press, 2018.
- PEDRYCZ, W.; BARGIELA, A. Granular clustering: a granular signature of data. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, IEEE, v. 32, n. 2, p. 212–224, 2002.
- PEDRYCZ, W. et al. Granular computing-the emerging paradigm. **Journal of uncertain systems**, Citeseer, v. 1, n. 1, p. 38–61, 2007.
- PEDRYCZ, W.; SKOWRON, A.; KREINOVICH, V. **Handbook of granular computing**. [S.l.]: John Wiley & Sons, 2008.
- PRATAMA, M. et al. pclass: an effective classifier for streaming examples. **IEEE Transactions on Fuzzy Systems**, IEEE, v. 23, n. 2, p. 369–386, 2014.
- PRATAMA, M. et al. A novel meta-cognitive-based scaffolding classifier to sequential non-stationary classification problems. In: IEEE. **2014 IEEE international conference on fuzzy systems (FUZZ-IEEE)**. [S.l.], 2014. p. 369–376.
- PRATAMA, M.; LU, J.; ZHANG, G. Evolving type-2 fuzzy classifier. **IEEE Transactions on Fuzzy Systems**, IEEE, v. 24, n. 3, p. 574–589, 2015.
- PRATAMA, M.; PEDRYCZ, W.; LUGHOFER, E. Evolving ensemble fuzzy classifier. **IEEE Transactions on Fuzzy Systems**, IEEE, v. 26, n. 5, p. 2552–2567, 2018.
- RENDÓN, E. et al. Internal versus external cluster validation indexes. **International Journal of computers and communications**, v. 5, n. 1, p. 27–34, 2011.
- RENDÓN, E. et al. A comparison of internal and external cluster validation indexes. In: **Proceedings of the 2011 American Conference, San Francisco, CA, USA**. [S.l.: s.n.], 2011. v. 29, p. 1–10.
- SILVA, A. M. et al. A fast learning algorithm for evolving neo-fuzzy neuron. **Applied Soft Computing**, Elsevier, v. 14, p. 194–209, 2014.
- SILVA, L. E. B. D.; MELTON, N. M.; WUNSCH, D. C. Incremental cluster validity indices for online learning of hard partitions: Extensions and comparative study. **IEEE Access**, IEEE, v. 8, p. 22025–22047, 2020.
- SILVA, L. E. B. da; WUNSCH, D. C. A study on exploiting vat to mitigate ordering effects in fuzzy art. In: IEEE. **2018 International Joint Conference on Neural Networks (IJCNN)**. [S.l.], 2018. p. 1–8.
- SILVA, P. C. et al. Probabilistic forecasting with fuzzy time series. **IEEE Transactions on Fuzzy Systems**, IEEE, 2019.
- ŠKRJANC, I. et al. Evolving fuzzy and neuro-fuzzy approaches in clustering, regression, identification, and classification: A survey. **Info Sci**, Elsevier, v. 490, p. 344–368, 2019.
- SOARES, E. et al. Ensemble of evolving data clouds and fuzzy models for weather time series prediction. **Applied Soft Computing**, Elsevier, v. 64, p. 445–453, 2018.

- SOARES, E. et al. Cloud-based evolving intelligent method for weather time series prediction. In: IEEE. **2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)**. [S.l.], 2017. p. 1–6.
- SOUZA, P. V. de C. et al. Evolving fuzzy neural networks to aid in the construction of systems specialists in cyber attacks. **Journal of Intelligent & Fuzzy Systems**, IOS Press, v. 36, n. 6, p. 6743–6763, 2019.
- TAKAGI, T.; SUGENO, M. Fuzzy identification of systems and its applications to modeling and control. **IEEE transactions on systems, man, and cybernetics**, IEEE, n. 1, p. 116–132, 1985.
- TJOA, E.; GUAN, C. A survey on explainable artificial intelligence (xai): towards medical xai. **arXiv preprint arXiv:1907.07374**, 2019.
- TUNG, S. W.; QUEK, C.; GUAN, C. et2fis: An evolving type-2 neural fuzzy inference system. **Information Sciences**, Elsevier, v. 220, p. 124–148, 2013.
- WOLF, C. T. Explainability scenarios: towards scenario-based xai design. In: **Proceedings of the 24th International Conference on Intelligent User Interfaces**. [S.l.: s.n.], 2019. p. 252–257.
- WU, D.; MENDEL, J. M. A vector similarity measure for linguistic approximation: Interval type-2 and type-1 fuzzy sets. **Information Sciences**, Elsevier, v. 178, n. 2, p. 381–402, 2008.
- XIE, X. L.; BENI, G. A validity measure for fuzzy clustering. **IEEE Transactions on pattern analysis and machine intelligence**, v. 13, n. 8, p. 841–847, 1991.
- XU, R.; XU, J.; WUNSCH, D. C. A comparison study of validity indices on swarm-intelligence-based clustering. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, IEEE, v. 42, n. 4, p. 1243–1256, 2012.
- XU, Y.; WONG, K.; LEUNG, C. Generalized recursive least square to the training of neural network. **IEEE Transactions on Neural Networks**, v. 17, n. 1, p. 19–34, 2006.
- YAGER, R. R.; ZADEH, L. A. **An introduction to fuzzy logic applications in intelligent systems**. [S.l.]: Springer Science & Business Media, 2012. v. 165.
- YANG, J. Dynamic clustering of evolving streams with a single pass. In: IEEE. **Proceedings 19th International Conference on Data Engineering (Cat. No. 03CH37405)**. [S.l.], 2003. p. 695–697.
- YANG, X. et al. A kernel fuzzy c-means clustering-based fuzzy support vector machine algorithm for classification problems with outliers or noises. **IEEE Transactions on Fuzzy Systems**, IEEE, v. 19, n. 1, p. 105–115, 2010.
- YAO, J. T.; VASILAKOS, A. V.; PEDRYCZ, W. Granular computing: perspectives and challenges. **IEEE Transactions on Cybernetics**, IEEE, v. 43, n. 6, p. 1977–1989, 2013.
- ZADEH, L. A. The concept of a linguistic variable and its application to approximate reasoning-iii. **Information sciences**, Elsevier, v. 9, n. 1, p. 43–80, 1975.
- ZADEH, L. A. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. **Fuzzy sets and systems**, Elsevier, v. 90, n. 2, p. 111–127, 1997.

ZHAO, Q.; FRÄNTI, P. Wb-index: A sum-of-squares based index for cluster validity. **Data & Knowledge Engineering**, Elsevier, v. 92, p. 77–89, 2014.