



CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS  
GERAIS

CURSO DE BACHARELADO EM ENGENHARIA ELÉTRICA

TRABALHO DE CONCLUSÃO DE CURSO

**SISTEMAS FUZZY EVOLUTIVOS APLICADOS AO  
DIAGNÓSTICO DE CÂNCER DE MAMA EM  
PACIENTES**

Danielle Abreu Fortunato

# **SISTEMAS FUZZY EVOLUTIVOS APLICADOS AO DIAGNÓSTICO DE CÂNCER DE MAMA EM PACIENTES**

Trabalho apresentado como requisito parcial para a Conclusão do Curso de Bacharelado em Engenharia Elétrica do Centro Federal de Educação Tecnológica de Minas Gerais

## **COMISSÃO EXAMINADORA**

---

Prof. Márcio Wladimir Santana

---

Prof. Israel Teodoro Mendes

---

Daniel Furtado Leite

Nepomuceno, \_\_\_\_ de \_\_\_\_\_ de 2021

## DEDICATÓRIA

“Dedico este trabalho a Deus, o maior orientador da minha vida. Ele nunca me abandonou nos momentos de necessidade.”

## **AGRADECIMENTOS**

A Deus por ter me fortalecido durante todas as etapas da minha vida acadêmica.

Aos meus pais João e Cássia obrigada pelo apoio que me foi dado durante todos esses anos de estudo, e pela compreensão do quanto isso significa para mim.

Ao meu irmão Igor por todo o carinho e compreensão.

A minha avó Ana da Costa Lima (*in memoriam*), por ter me ensinado o valor de se fazer as coisas da melhor maneira possível e por me olhar de seu lugar.

Ao meu orientador Prof. Márcio Wladimir Santana e coorientador Prof. Israel Teodoro Mendes pela grande sabedoria, paciência e principalmente por acreditar em mim, possibilitando concluir mais uma etapa da minha vida acadêmica.

Aos meus amigos de vida toda e meus presentes dados pelo CEFET-MG, obrigada pelo apoio nos momentos mais difíceis, por compreender meus sumiços ao longo do curso, pelo carinho e amizade a mim retribuídos.

Agradeço a todos que em meio à tantas aflições me confortaram com palavras de afeto e amor.

## EPÍGRAFE

“A ciência humana de maneira alguma nega a existência de Deus. Quando considero quantas e quão maravilhosas coisas o homem compreende, pesquisa e consegue realizar, então reconheço claramente que o espírito humano é obra de Deus, e a mais notável.”

(Galileu Galilei)

## RESUMO

O diagnóstico e tratamento do câncer evoluíram, porém ele ainda é considerado um dos grandes problemas de saúde no mundo, sendo considerado a segunda maior causa de morte. De acordo com o Fundo Mundial para a Pesquisa e Câncer (WCRF) houve um aumento de 20% nos casos de câncer na última década. O câncer de mama, em particular, pode ser detectado em fases iniciais, em grande parte dos casos, aumentando assim a possibilidade de tratamentos menos agressivos e com taxas de sucesso satisfatórias. Este trabalho tem por objetivo auxiliar no diagnóstico de câncer de mama em pacientes distintos através de resultados de mamografia e identificar qual o tipo de alteração se enquadra, seja maligna ou benigna, dessa forma realizando a identificação da possibilidade de ocorrência de uma ou outra classe. Para a identificação do câncer de mama em glândulas mamárias são consideradas as Redes Neurais Granulares Evolutivas e Modelagem evolutiva Baseada em Conjuntos Fuzzy. A detecção de duas classes, sendo elas maligna e benigna as quais podem ocorrer nos exames clínicos de mamografia de pacientes foi alcançada com acurácia média de 93,5% em uma base de dados pública.

**Palavras-chave:** Câncer de Mama; Diagnóstico; Mamografia; Sistema Fuzzy Evolutivo.

## SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>8</b>
1.1 Objetivo.....	9
<b>2 REFERENCIAL TEÓRICO.....</b>	<b>10</b>
2.1 Câncer de Mama.....	10
2.2 Patologia para Detectar Doenças da Mama.....	11
2.3 Reconhecimento de Padrões.....	12
2.4 Redes Neurais Granulares Evolutivas.....	14
2.5 Modelagem evolutiva Baseada em Conjuntos.....	17
2.5.1 Modelagem Fuzzy.....	18
2.5.2 Aprendizagem Recursiva On-line.....	20
2.5.3 Criação de Regras.....	21
2.5.4 Adaptação de Regras.....	22
2.5.5 Ajuste de Granularidade.....	23
2.5.6 Compactação da Estrutura Resultante.....	23
2.5.7 Remoção de Grânulos.....	24
<b>3 METODOLOGIA.....</b>	<b>26</b>
3.1 Descrição do Conjunto de Dados.....	26
3.2 Sistema de Monitoramento Evolutivo proposto.....	27
<b>4 RESULTADOS ESPERADOS.....</b>	<b>29</b>
4.1 Detecção de Anomalias.....	29
4.2 Comparação entre Métodos.....	30
4.3 Matriz de Confusão.....	31
<b>5 CONCLUSÃO.....</b>	<b>36</b>
<b>BIBLIOGRAFIA.....</b>	<b>37</b>

## 1 INTRODUÇÃO

O câncer de mama é o mais frequente em mulheres no Brasil, atrás apenas do câncer de pele não melanoma, com taxa ainda ascendente de mortalidade. O rastreamento por meio da mamografia é a estratégia mais implementada no mundo para a detecção precoce da doença. Essa estratégia consiste na repetição periódica de mamografias de rotina em mulheres sem sinais ou sintomas suspeitos desse câncer (INCA, 2021a).

O diagnóstico e tratamento do câncer evoluíram, porém ele ainda é considerado um dos grandes problemas de saúde ao redor do mundo, sendo considerado a segunda maior causa de morte (INCA, 2021b). De acordo com o Fundo Mundial para a Pesquisa e Câncer (WCRF – *World Cancer Research Fund International*), houve um aumento de 20% nos casos de câncer na última década (WENTZ, 2019). Esses diagnósticos incorretos podem custar a vida dos pacientes, pois a identificação de pacientes com câncer de mama auxilia no tratamento precoce e pode prolongar o tempo de sobrevivência do mesmo. Dessa forma, têm-se como propósito a investigação de métodos com o potencial de diminuir a taxa de erro nos diagnósticos do câncer de mama nas pacientes. Na mortalidade proporcional por câncer em mulheres, em 2019, os óbitos por câncer de mama ocupam o primeiro lugar no país, representando 16,1% do total de óbitos (INCA, 2021b).

A detecção precoce do câncer de mama aumenta consideravelmente as chances de cura da paciente e evita as complicações e sequelas causadas pelos estágios mais avançados da doença, o que diminui a taxa de mortalidade em decorrência desse tipo de câncer. Modelos de detecção e classificação evolutivos, isto é, modelos equipados com algoritmos incrementais *online* de aprendizado de máquina, são capazes de identificar e classificar variações em exames clínicos de mastologia. Tais modelos têm sua base de regras, estrutura conexionista, *clusters*, folhas em árvores, ou grânulos de informação adaptados ao longo do tempo de acordo com o fluxo de dados (SILVA et al., 2018; SOARES et al., 2019; LEITE, 2019; LEITE et al., 2020).

## 1.1 Objetivo

O objetivo desse trabalho é utilizar inteligência computacional para auxiliar o diagnóstico de presença de tumores das regiões mamárias através de atributos extraídos a partir de amostras de imagem de mamografia com presença de uma massa na região da mama. Eles descrevem as características dos núcleos celulares presentes na imagem. O conjunto de dados contendo as características dos exames são disponibilizadas pelo *UCI Machine Learning Repository*. O sistema deverá ser capaz de auxiliar no diagnóstico da presença de uma massa na mama das pacientes, classificando-as com presença de uma massa benigna ou maligna. Para o desenvolvimento deste trabalho serão utilizados sistemas *fuzzy* evolutivos. As abordagens que serão utilizadas são a rede *neuro-fuzzy* evolutiva (eGNN – *Evolving Granular neural networks*) e a modelagem evolutiva baseada em conjuntos *fuzzy* (FBeM - *Fuzzy set Based Evolving Modeling*). Em geral, tais sistemas *fuzzy* têm sido utilizados com sucesso para resolver problemas de classificação, previsão, modelagem e controle em diversas áreas, como sistemas de energia, medicina, agronomia e previsão meteorológica (ANGELOV; FILEV; KASABOV, 2008; BOUCHACHIA, 2010; GABRYS; PETRAKIEVA, 2004; LEITE, 2012). Neste trabalho a classificação das amostras extraídas de exames clínicos ocorreram em modo *online* e em tempo real. Além disso, o trabalho também propõe o compartilhamento de informações sobre o câncer de mama.

## 2 REFERENCIAL TEÓRICO

### 2.1 Câncer de Mama

Câncer é o nome dado às doenças nas quais as células começam a se comportar de maneira atípica, aumentando, crescendo e se dividindo desordenadamente (ARAÚJO, 2009). Os efeitos desse comportamento sobre o hospedeiro são agressivos, visto que o câncer, ou tumor, pode destruir tecidos adjacentes e se espalhar pelo corpo, fenômeno chamado de "metástase" (BRASIL, 1990). Os tumores benignos são pontuais, crescem lentamente, são muito similares aos tecidos saudáveis e podem, na maioria dos casos, ser totalmente removidos por meio de cirurgia. Os tumores malignos multiplicam-se rapidamente e possuem a capacidade de provocar metástases, a cura depende do diagnóstico precoce e do tratamento adotado (PONTES, 2013).

O câncer de mama, ou carcinoma, não possui métodos simples de rastreamento, por isso faz-se necessário a realização de exames clínicos diversos, entre eles o auto-exame, mamografia ou ultrassonografia. Os cânceres de mama benignos, são conhecidos como fibroadenomas, enquanto os malignos possuem duas classificações que dependem do seu local de origem, sendo os carcinomas ductais (se formado nos ductos ou canais lactíferos) ou lobulares (se formado nos lóbulos, respectivamente). Enquanto o câncer não se espalhou pelo tecido ao redor de sua origem ou para outras partes do corpo (conhecido como câncer "invasivo"), ele é conhecido como "*in situ*" (CHAVES, 2019).

O câncer de mama é o mais frequente em mulheres no Brasil, atrás apenas do câncer de pele não melanoma, com taxa ainda ascendente de mortalidade. O rastreamento por meio da mamografia é a estratégia mais implementada no mundo para a detecção precoce da doença. Essa estratégia consiste na repetição periódica de mamografias de rotina em mulheres sem sinais ou sintomas suspeitos desse câncer. No Brasil, o rastreamento mamográfico passou a ser recomendado como política pública em 2004 (INCA, 2021a). A Figura 1 apresenta a ilustração da área da mama com presença de tumor.

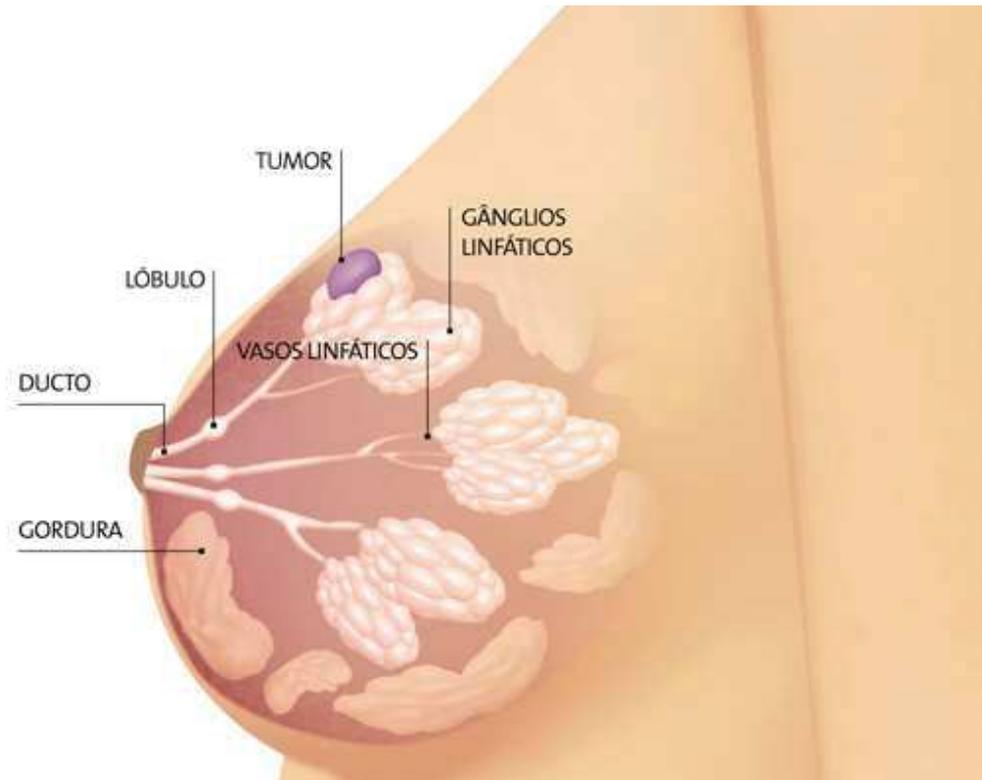


Figura 1 - Ilustração da área da mama com presença de tumor.  
Fonte: (Lima, R. 2017).

## 2.2 Patologia para Detectar Doenças na Mama

As Doenças das Mamas compõem um amplo quadro de acontecimentos, tanto benignos quanto malignos, e que costumam levar as mulheres a procurar atendimento médico – principalmente em devido as queixas de dor mamária, nódulos palpáveis ou saída de secreção do mamilo com apresentações. Atualmente, em decorrência das campanhas anuais, muitas mulheres têm buscado os serviços ambulatoriais para prevenir ou até mesmo com medo do câncer de mama (SBP, 2021). A Figura 2 apresenta o fluxograma de patologia da mama, ou seja, qualquer desvio anatômico e/ou fisiológico, em relação à normalidade, que constitua uma doença ou caracterize determinada doença.

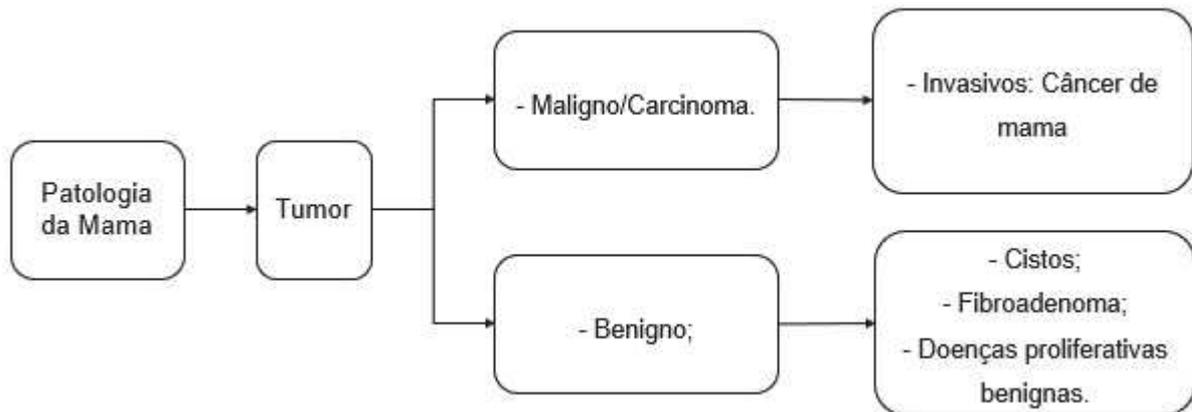


Figura 2 – Patologia da Mama.  
Fonte: (SBP, 2016).

Nota-se pela Figura 2 que patologia consiste nas alterações estruturais, da mama, buscando explicar os processos por meio dos quais surgem sinais e sintomas de doenças. Dessa forma, quando ocorre uma alteração na mama, ela pode ser caracterizada de duas maneiras; como tumor benigno ou maligno."

Existem tumores benignos, que não são câncer, tais como cisto, fibroadenoma e doenças proliferativas que crescem de forma lenta e organizada. Já os tumores malignos (ou cancerígenos) acontecem quando uma célula do organismo passa a se replicar de maneira desordenada e descontrolada. Esses tumores, em geral, crescem rapidamente e podem se espalhar pelo corpo, invadindo tecidos vizinhos, sendo chamado de câncer de mama. Nem todo tumor pode ser um câncer, porém, quanto mais cedo se descobre o tumor maligno, e quanto mais adequado é o tipo de tratamento, maiores são as chances de cura.

### 2.3 Reconhecimento de Padrões

Reconhecimento de padrões é uma área da ciência cujo objetivo é a classificação de objetos em um número de categorias ou classes. Dependendo da aplicação, estes objetos podem ser imagens, formas de onda de sinal ou qualquer tipo de medições que necessitam ser distinguidas (THEODORIDIS; KOUTOUMBAS, 1999; SANTANA, 2017). O reconhecimento de padrões é uma habilidade extremamente desenvolvida nos seres humanos e em alguns animais.

O termo "reconhecimento de padrões" foi introduzido no início da década de 60 e originalmente significava a detecção de formas simples. Há duas grandes

motivações para estudos nesta área: a necessidade das pessoas em se comunicar com máquinas computacionais através de linguagens naturais; e o interesse na ideia de projetar e construir autômatos (máquinas inteligentes) que possam realizar certas tarefas com habilidades comparáveis à performance humana. Estas tarefas envolvem percepção invariante em relação à equivalência de estímulos, posição, deslocamento, rotação, perspectiva, oclusão parcial, entre outros (REIS; ALBURQUERQUE; CASTRO, 2012). Entre as áreas de aplicação do reconhecimento de padrões estão:

- Comunicação do homem com a máquina: reconhecimento automático da fala, reconhecimento da escrita, compreensão da fala, compreensão das imagens, processamento da linguagem natural;
- Aplicações militares: reconhecimento, orientação e controle automáticos de alvos;
- Medicina: diagnose médica, análise de imagens, classificação de doenças;
- Veículos: controladores de automóveis, aviões, trens, barcos;
- Polícia e investigação: detecção criminal a partir da fala, escrita manual, impressões digitais, fotografias;
- Estudo e estimativa de recursos naturais: agricultura, extrativismo, geologia, ambiente;
- Sistemas domésticos: utensílios;
- Computadores: hardware e software difusos.

Entende-se por padrão as propriedades que possibilitam o agrupamento de objetos semelhantes dentro de uma determinada classe ou categoria mediante a interpretação de dados de entrada (TOU; GONZÁLES, 1981). Entende-se por classe um conjunto de atributos comuns do objeto de estudo. A ideia é de que o grau de associação seja maior entre dados da mesma categoria e menor entre dados de estruturas diferentes. Os dados de entrada são, geralmente, provenientes de sensores e selecionados segundo o conteúdo de informações relevantes para a decisão. Em geral, os dados passam por um processo de redução de dimensionalidade, para que possam ser usados por um classificador, que designará a classe que melhor o representa (MARQUES, 1999; MASCARENHAS, 1987; PRADO; ELFES, 1975).

Classificação pressupõe a designação de um objeto a uma determinada

classe. O classificador “aprende” a distinguir, entre as classes, aquela à qual o objeto pertence. Padrões de uma mesma classe aglomeram-se em agrupamentos (MARQUES, 1999).

## 2.4 Redes Neurais Granulares Evolutivas

O conceito de redes neurais granulares (GNN) foi inicialmente estabelecido por Pedrycz e Vukovich (2001) para designar redes neurais com algum aspecto que envolva incerteza. Por exemplo: (i) os dados processados por uma rede granular não precisam ser necessariamente numéricos, mas podem ser intervalos, números *fuzzy*, intervalos *fuzzy*, distribuições de probabilidade, palavras; (ii) as operações das conexões entre neurônios ou do corpo neural podem não ser pontuais, mas levar em conta incerteza; (iii) as estimações providas pela rede podem ser invólucros em torno do valor-alvo; e/ou (iv) a representação dos dados no modelo neural é feita de forma modular, a partir do desenvolvimento de modelos locais (grânulos de informação encontrados no fluxo de dados *online*). A partir do ambiente GNN original, foram propostas várias outras abordagens neurais granulares. Dentre elas, se destaca a rede *evolving* GNN, eGNN (LEITE; COSTA Jr.; GOMIDE, 2009; LEITE, 2019), pois ela foi equipada com algoritmo de aprendizado de máquina incremental *online* e capacidade de adaptação estrutural do modelo, além da capacidade de atualização paramétrica usual. Além disso, eGNN contempla os quatro aspectos (i)-(iv) mencionados.

O aprendizado em GNN e eGNN segue um princípio comum que, geralmente, envolve dois estágios, conforme mostra a Figura 3. Primeiro, grânulos de informação – intervalos ou conjuntos *fuzzy* – são construídos a partir de uma base de representação numérica. Note que, algumas vezes, os dados são originalmente granulares. Em seguida, a aprendizagem – construção e refinamento – da rede neural é baseada nos grânulos de informação ao invés de ser baseada nos dados originais. Assim, a rede eGNN não é exposta a todos os dados de treinamento, muito mais numerosos que os grânulos formados. Por exemplo, quando amostras de dados não transportam novas informações, os exemplos são incorporados pelos grânulos já conhecidos (LEITE, 2019).

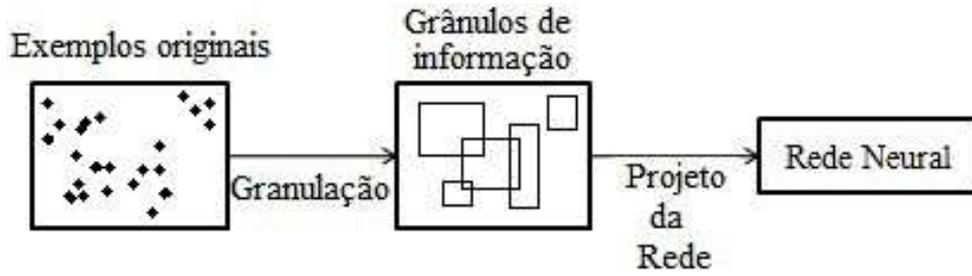


Figura 3 - Projeto de rede neural granular, GNN ou eGNN  
 Fonte: (Leite; Costa Jr.; Gomide, 2009).

Fundamentalmente, modelos eGNN processam dados observando um fluxo somente uma vez. A eGNN começa à aprender a partir de uma base de regras vazia, sem neurônios e sem conhecimento prévio das propriedades estatísticas dos dados e classes (para o caso de problema de classificação). A abordagem consiste em formar limites discriminantes entre classes a partir da granulação do espaço dos dados usando hiper-caixas *fuzzy* (LEITE, 2019). A escolha dos hiper-parâmetros iniciais de algoritmos é crucial para a obtenção de boas soluções. Esta questão é de certa forma facilitada se os algoritmos de aprendizagem contemplam mecanismos que adaptam continuamente os hiper-parâmetros iniciais. O tamanho máximo,  $\rho$ , que um grânulo pode assumir no espaço dos dados define a capacidade não-linear (flexibilidade) do modelo em capturar limites discriminantes entre classes. Portanto,  $\rho$  é o hiper-parâmetro chave do algoritmo.

Um procedimento para adaptar  $\rho$  continuamente é proposto. Primeiro, considere que  $\{\gamma_1, \dots, \gamma^\Theta\}$  sejam os grânulos criados em um certo número de passos de tempo,  $h_g$ . Se a quantidade de grânulos  $\Theta$  cresce a uma taxa maior que um limiar  $\eta$ , então  $\rho$  é aumentado conforme à equação (1):

$$\rho(novo) = \left(1 + \frac{\Theta}{h_g}\right) \rho(velho) \quad (1)$$

Ao contrário, se  $\Theta$  cresce à uma taxa menor que  $\eta$ , então,  $\rho$  é reduzido a partir de

$$\rho(novo) = \left(1 + \frac{\eta - \Theta}{HG}\right) \rho(velho) \quad (2)$$

Redes eGNN aprendem a partir de um fluxo de dados  $x^{[h]}$ ,  $h = 1, 2, \dots$ . Os exemplos de treinamento podem ou não ser acompanhados de um rótulo de classe  $C^{[h]}$ . Cada grânulo de informação  $\gamma^i$  da coleção finita dos grânulos existentes em um dado momento,  $\gamma = \{\gamma^1, \dots, \gamma^c\}$ , definido no espaço de atributos  $X \subseteq R^n$ , é associado a uma classe  $C_k$  da coleção finita de classes  $C = \{C_1, \dots, C_m\}$  no espaço de saída  $Y \subseteq N$ . eGNN associa os espaços de atributos e de saída usando grânulos extraídos do fluxo de dados, e duas camadas de neurônios T-S (T-norma, S-norma, null-norma, unimorma, ou qualquer outra família de normas a escolher). A rede neural tem uma estrutura em 5 camadas, como mostra a Figura 4.

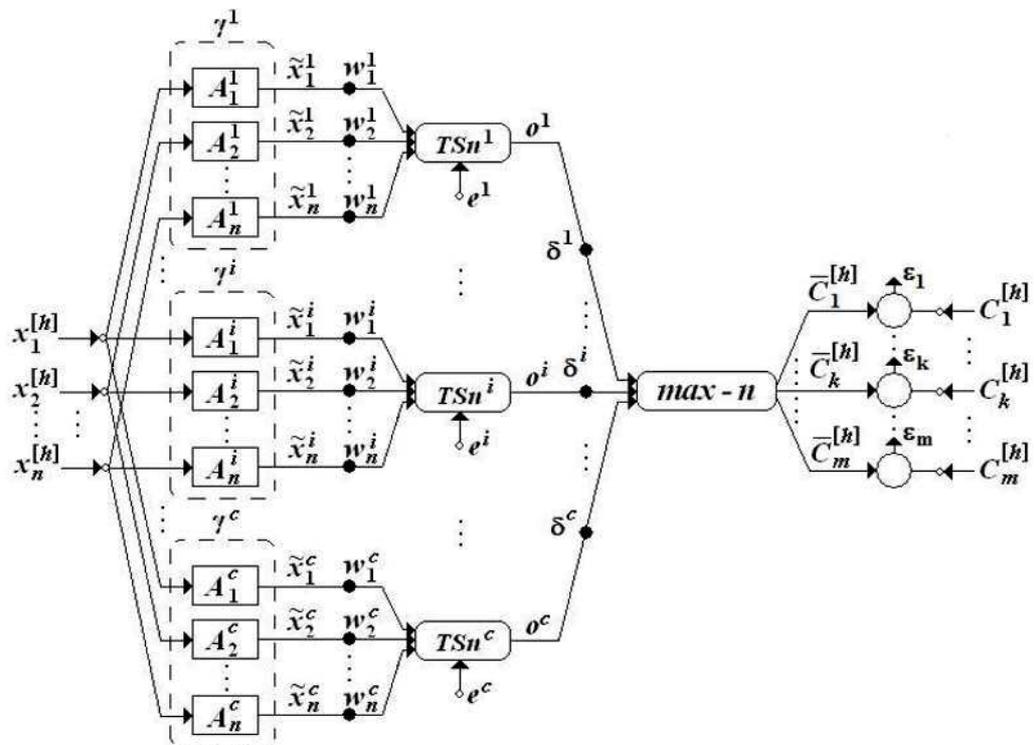


Figura 4 - Rede neuro-fuzzy granular evolutiva (eGNN).  
 Fonte: (Leite; Costa Jr.; Gomide, 2009).

A camada de entrada apresenta vetores característicos  $x^{[h]} = (x_1, \dots, x_j, \dots, x_n)^{[h]}$ ,  $h = 1, \dots$ , à rede neural; a camada granular consiste de um conjunto de grânulos  $\gamma^i \forall i$  formado como um escopo do fluxo de dados. Sobreposição parcial de grânulos (*fuzzy*) são permitidas; a camada de agregação contém neurônios T-S,  $T-Sn^i \forall i$ . Eles agregam graus de pertinência para gerar medidas de compatibilidade  $o^i \forall i$  entre exemplo e grânulos; na camada de decisão, as medidas de compatibilidade são comparadas, e a classe  $C_k$  associada ao grânulo  $\gamma^i$  que apresentou a maior

compatibilidade para um dado exemplo é induzida na saída da rede; a camada de saída compreende indicadores de rótulos de classe. Todas as camadas, exceto a camada de entrada, evoluem dado o fluxo  $x^{[h]}$ ,  $h = 1, 2, \dots$ . A adaptação paramétrica e estrutural do classificador eGNN pode ser realizada de diferentes maneiras dependendo da aplicação, isto é, o número de classes pode ser automaticamente controlado quando tem-se tal informação *a priori*. O número de grânulos na estrutura do modelo pode também ser limitado se memória e tempo de processamento são restrições do problema. Em ambientes desconhecidos, ambos, o número de grânulos de classes, podem ser automaticamente gerenciados pelo algoritmo de aprendizagem (LEITE, COSTA, GOMIDE, 2010; LEITE, 2019).

O algoritmo de aprendizado está fora do escopo deste trabalho. Ele está reportado com detalhes em (LEITE, 2019). Em essência, o algoritmo compreende procedimentos de criação de grânulos, arraste de grânulos, expansão e encolhimento, mesclagem, e deleção de grânulos e neurônios inativos. A rede eGNN é de propósito geral, independente da natureza da aplicação, e também pode ser empregada em problemas de regressão e previsão. Um modelo é gerado autonomamente tal que dados não precisam ser armazenados para posterior análise.

## 2.5 Modelagem Evolutiva Baseada em Conjuntos *Fuzzy*

Modelagem evolutiva baseada em conjuntos *fuzzy* (FBeM, *Fuzzy set Based evolving Modeling*) emprega grânulos de informação tipo *fuzzy* para construir mapas granulares que de maneira geral associam dados granulares de entrada a dados granulares de saída. Grânulos *fuzzy* garantem a generalidade da estrutura dos dados e proveem algoritmos com matemática simples e regras linguísticas descrevendo seu comportamento. Basicamente, um sistema FBeM pode perceber os dados de um fluxo sob diferentes resoluções e decidir entre adotar granularidades mais simples ou mais detalhadas (LEITE; GOMIDE, 2012; SANTANA; LEITE, 2020).

Modelos FBeM são criados e evoluídos quando requisitados pelo fluxo de dados. O algoritmo de aprendizagem de FBeM cria, expande e retrai grânulos recursivamente. Eventualmente, a estrutura granular resultante pode ser melhorada de acordo com relações inter-granulares. FBeM combina sistemas *fuzzy* linguísticos e funcionais para prover aproximações singulares e granulares de funções não estacionárias. Sistemas *fuzzy* funcionais são, geralmente, mais precisos enquanto

que os sistemas *fuzzy* linguísticos são mais interpretáveis. Através da combinação destes, FBeM aproveita as vantagens de ambos os sistemas simultaneamente. Em nível prático, especialistas preferem que sistemas *online* deem resultados aproximados tão bem quanto limites de tolerância nas aproximações (LEITE; GOMIDE, 2012).

### 2.5.1 Modelagem Fuzzy

FBeM é uma abordagem de modelagem incremental a partir de dados detalhados e um algoritmo de aprendizagem. Sua resposta global sucede da agregação de respostas locais mais específicas. O algoritmo incremental de FBeM molda sua estrutura de regras para aceitar novidades, lidar com incerteza, prover aproximações singulares e granulares de funções. FBeM lida com questões como bases de dados ilimitadas e escalabilidade (LEITE; GOMIDE, 2012). Modelos FBeM consistem de regras obtidas a partir de um sistema complexo. A aprendizagem em FBeM não requer pré-concepção de regras. Regras são criadas e adaptadas dinamicamente, consoante com o comportamento da função do processo ou da fronteira de decisão ao longo do tempo. Sempre que amostras de dados são disponibilizadas, um mecanismo de decisão pode acrescentar novas regras à estrutura FBeM ou adaptar parâmetros de regras existentes. Especialistas podem desejar prover uma descrição verbal sobre o processo a partir da intuição e experiência (LEITE; GOMIDE, 2012). Em modelos FBeM, regras  $R^i$  governando grânulos de informação  $\gamma^i$  são do tipo:

$$SE (x_1 \acute{e} A_1^i) E \dots E (x_j \acute{e} A_j^i) E \dots E (x_n \acute{e} A_n^i), ENT\tilde{A}O (y_1 \acute{e} B_1^i) E \acute{y}_1 = p_1^i(x_j \forall j) E \dots$$

$$\dots E (y_k \acute{e} B_k^i) E \acute{y}_k = p_k^i(x_j \forall j) E \dots E (y_m \acute{e} B_m^i) E \acute{y}_m = p_m^i(x_j \forall j)$$

onde  $x_j$  e  $y_k$  são variáveis do fluxo de dados  $(x, y)^{|h|}$ ,  $h = 1, \dots$ ;  $A_j^i \text{ e } B_k^i$  são funções de pertinência construídas a partir dos dados disponibilizados; e  $p_k^i$  são polinômios de aproximação.

As regras  $R^i, i = 1, \dots, c$ , formam a base de regras. Note que uma regra FBeM combina consequentes linguístico e funcional. O consequente linguístico envolve funções e provê interpretabilidade aos resultados. O consequente funcional oferece aproximação singular e precisão. Com esta estrutura, FBeM toma vantagem de

ambos, sistemas linguísticos e funcionais, em uma plataforma de modelagem única. Um aspecto a ser levado em conta em granulação tipo espalhamento refere-se a busca por uma quantidade factível de partições, posições e tamanhos de grânulos. Essencialmente, FBeM emprega funções de pertinência Gaussianas como objetos granulares formais para envolver a incerteza dos dados (LEITE; GOMIDE, 2012). Conforme Leite e Gomide (2012), um conjunto *fuzzy* Gaussiano  $A_j^i = G(\mu_j^i, \sigma_j^i)$  é caracterizado pelo valor modal  $\mu_j^i$  e espalhamento  $\sigma_j^i$ . Características que fazem esta representação apropriada incluem: (i) facilidade de aquisição dos parâmetros. O valor modal e de espalhamento são capturados diretamente a partir do fluxo de dados; (ii) suporte infinito. Visto que o domínio dos dados é desconhecido anteriormente ao aprendizado, o suporte de Gaussianas estende-se ao longo de todo o domínio; (iii) suavidade e superfície continuamente diferenciável. O consequente de regras FBeM inclui funções locais afins do tipo:

$$p_k^i = a_{0k}^i + \sum_{j=1}^n a_{jk}^i x_j \quad (3)$$

Em geral, funções  $p_k^i$  podem ser de tipos diferentes e não requerem linearidade. O algoritmo Mínimos Quadrados Recursivo (RLS) é usado para determinar os coeficientes locais  $a_{jk}^i$ . A representação Gaussiana permite sobreposição de todos os grânulos. Consequentemente, cada regra FBeM contribui à saída do sistema. A saída singular de FBeM é determinada como um valor médio ponderado sobre todas as regras:

$$p_k = \frac{\sum_{i=1}^c \min(A_1^i, \dots, A_n^i) p_k^i}{\sum_{i=1}^c \min(A_1^i, \dots, A_n^i)} \quad (4)$$

isto assegura transição suave entre funções de pertinência.

De forma similar a abordagem para agrupar dados em conjuntos antecedentes  $A_j^i$ , conjuntos consequentes  $B_k^i$  usam granulação tipo espalhamento, por exemplo, funções Gaussianas  $B_k^i = G(\mu_k^i, \sigma_k^i)$ . A saída granular  $B_k^i$  provê informação mais importante que a saída numérica  $p_k$ . Não obstante, em problemas de classificação, o termo consequente de uma regra FBeM é essencialmente uma classe (LEITE; GOMIDE, 2012).

## 2.5.2 Aprendizagem Recursiva On-line

FBeM aprende a partir de um fluxo  $(x, y)^{[h]}$ ,  $h = 1, \dots$ , onde  $y^{[h]}$  é conhecido dado  $x^{[h]}$  ou se tornará conhecido alguns passos adiante. Cada par  $(x, y)$  é uma observação da função  $f$ . Quando  $f$  muda com o tempo, dizemos que a função é não estacionária. O procedimento de aprendizagem para evoluir modelos FBeM é sumarizado, como segue, no Algoritmo 1.

Algoritmo 1: Atualização da estrutura do modelo FBeM
1: <b>Iniciar</b>
2: Definir parâmetros $\rho, h_r, \eta, \psi, c = 0$ ;
3: Ler amostra $(x, y)^{[h]}$ , $h = 1$ ;
4: Criar grânulo $\gamma^{c+1}$ ;
5: <b>Para</b> $h = 2, \dots$ do
6:   Ler amostra $(x, y)^{[h]}$ ;
7:   Fornecer aproximação de valor único $p(x^{[h]})$ ;
8:   Fornecer aproximação granular $B^{i*}$ ;
9:   Calcular erro de saída $\epsilon^{[h]} = mp(y^{[h]}) - p(x^{[h]})$ ;
10: <b>Se</b> $x^{[h]}$ ou $y^{[h]}$ não estão nas regiões expandidas dos grânulos $E^i \forall$
11:     Criar grânulo $\gamma^{c+1}$ ;
12: <b>Senão</b>
13:     Adaptar o grânulo mais ativo $\gamma^i$ , $i = \max_i(S(x, A^1), \dots, S(x, A^c))$ ;
14:     Adaptar os parâmetros da função local $a_j^i$ usando RLS;
15: <b>Se</b> $h = \alpha h_r$ , $\alpha = 1, 2, \dots$
16:     Combine grânulos quando viável;
17:     Atualizar granularidade do modelo $\rho$ ;
18:     Remover grânulos inativos;
19: <b>Fim</b>

Os passos 2 e 18 enfatizam a essência de algoritmos dirigidos a fluxos, isto é, amostras são lidas e descartadas uma por vez. Dados históricos são descartados. Modelos granulares evoluem sempre que novas informações surgem nos dados, passo 6. Quando uma nova amostra não condiz com o conhecimento atual, o procedimento cria um grânulo e uma regra, passo 11. Ao contrário, se uma nova

amostra ajusta-se ao conhecimento atual, o procedimento adapta grânulos e regras existentes, passo 14. Eventualmente, a estrutura resultante pode ser otimizada de acordo com relações intergranulares, passo 17.

### 2.5.3 Criação de Regras

Em FBeM, regras não existem *a priori*. Elas são criadas e evoluem à medida que os dados são disponibilizados. Um novo grânulo  $\gamma^{c+1}$  e a regra  $R^{c+1}$  que o governa são criados quando as regras existentes não são suficientemente ativadas para uma amostra  $x^{[h]}$ . FBeM assume que a amostra traz uma nova informação sobre o processo. Seja,  $\rho \in [0,1]$  um limiar que determina quando criar ou adaptar regras (LEITE; GOMIDE, 2012). Se

$$\min(A_1^i, \dots, A_n^i) \leq \rho \quad \forall i, \quad (5)$$

então a estrutura FBeM é expandida.

Note que se  $\rho$  é 0, então o sistema é estruturalmente estável e incapaz de capturar mudanças de conceito. Ao contrário, se  $\rho$  é 1, FBeM cria uma regra para cada nova amostra, o que não é prático. Adaptabilidade é alcançada em situações intermediárias (compromisso estabilidade/plasticidade) (LEITE; GOMIDE, 2012).

O papel de  $\rho$  é fundamental na determinação da granularidade de modelos FBeM. Escolhas de  $\rho$  impactam na precisão e transparência de modelos, resultando em diferentes visões granuladas e níveis de detalhes do mesmo processo. Um novo grânulo  $\gamma^{c+1}$  é inicialmente representado por funções de pertinência,  $A_j^{c+1}$  e  $B_k^{c+1}$ , com parâmetros:

$$\mu_j^{c+1} = x_j^{[h]}, \quad (6)$$

$$\mu_k^{c+1} = y_k^{[h]}, \quad (7)$$

$$\sigma_j^{c+1} = \sigma_k^{c+1} = 1/2\pi, \quad (8)$$

Esta é a abordagem de Stiegler para funções Gaussianas padrões (STIEGLER, 1982). Os coeficientes de polinômios locais  $p_k^{c+1}$  são:

$$a_{0k}^{c+1} = y_k^{[h]} \quad (9)$$

$$a_{jk}^{c+1} = 0, j \neq 0 \quad (10)$$

Com esta parametrização inicial a preferência é dada ao projeto de grânulos balanceados ao longo de suas dimensões ao invés de grânulos com geometria desbalanceada.

#### 2.5.4 Adaptação de Regras

Conforme Leite e Gomide (2012) a adaptação de regras consiste em (i) expandir ou contrair  $A_j^i$  e  $B_k^i$  para acomodar novos dados; (ii) mover grânulos  $\gamma^i$  na direção de regiões de dados mais densas; e (iii) ajustar os coeficientes de funções locais  $p_k^i$ . Uma regra  $R^i$  é adaptada sempre que é suficientemente ativada por uma amostra  $x^{[h]}$  de acordo com:

$$\min(A_1^i, \dots, A_n^i) > \rho \quad (11)$$

Geometricamente, a amostra pertence a uma região influenciada pelo grânulo  $\gamma^i$ . Para incluir  $x^{[h]}$ , FBeM atualiza o valor modal e o espalhamento das funções de pertinência  $A_j^i$  correspondentes como segue:

$$\mu_j^i(novo) = \frac{(\varpi - 1)\mu_j^i(velho) + x_j}{\varpi^i}, \quad (12)$$

$$\sigma_j^i(novo) = \frac{(\varpi^i - 1)}{\varpi^i} \sigma_j^i(velho) + \frac{1}{(\varpi^i - 1)} (x_j - (\mu_j^i(novo)))^2, \quad (13)$$

onde  $\varpi^i$  refere-se ao número de vezes que o  $i$ -ésimo grânulo foi ativado pelo fluxo de dados.

Nota-se que os valores são calculados recursivamente e, portanto, não demandam acumulação de dados. Apenas a regra mais ativa para  $x^{[h]}$  é escolhida para adaptação. A adaptação de conjuntos *fuzzy* consequentes  $B_k^i$  usa dados de saída  $y_k^{[h]}$ . Coeficientes polinomiais  $a_{jk}^i$  são atualizados usando o algoritmo RLS (LEITE; GOMIDE, 2012).

### 2.5.5 Ajuste de Granularidade

O limiar  $\rho$  assume valores no intervalo unitário de acordo com erros de predição. Níveis de ativação de regras para uma dada entrada  $x^{[h]}$  são comparados com o valor de  $\rho^{[h]}$  e definem mudança paramétrica ou estrutural de modelos FBeM. Valores de  $\rho$  influenciam a granularidade e inteligibilidade de modelos. No caso mais geral, FBeM começa a aprender a partir de uma base de regras vazia, sem qualquer conhecimento sobre as propriedades dos dados. Consequentemente, é justo iniciar  $\rho$  em uma condição intermediária para permitir estabilidade e plasticidade estrutural igualmente. Usa-se  $\rho^{[0]} = 0.5$  como valor padrão.

Seja  $E$  o erro quadrado máximo entre predições  $p_k(x^{[h]})$  e valores reais  $y_k^{[h]}$ , então:

$$e_k = \left( y_k^{[h]} - p_k(x^{[h]}) \right)^2 \quad k = 1, \dots, m, \quad (14)$$

e

$$E = \max(e_1, \dots, e_k, \dots, e_m). \quad (15)$$

Admita  $E_D$  como o erro máximo aceitável. Logo,  $\rho$  aprende valores para si mesmo a partir de:

$$\rho(\text{novo}) = \rho(\text{velho}) + \alpha(E_D - E), \quad (16)$$

onde  $\alpha$  é a taxa de aprendizagem.

Especialistas escolhem o valor de  $E_D$  e podem desejar que ele seja zero. Valores muito pequenos de  $E_D$  conduzem  $\rho$  a zero e levam ao sobre ajuste (*overfitting*) de modelos. A prática sugere abrir mão de certa precisão para alcançar um erro de aproximação aceitável, abstrações granulares úteis e compactação da base de dados em regras interpretáveis. A adaptação recursiva da granularidade alivia escolhas arbitrárias do quão rápido e quão frequente a estrutura dos dados muda (LEITE; GOMIDE, 2012).

### 2.5.6 Compactação da Estrutura Resultante

Segundo Leite e Gomide (2012), relacionamentos entre pares de grânulos podem ser fortes o suficiente para justificar a formação de um grânulo maior e mais

abstrato que herda a essência e natureza de grânulos menores e mais detalhados. A análise de relações intergranulares requer uma métrica para medir a distância entre objetos *fuzzy*. A distância entre dois grânulos,  $\gamma^{i1}$  e  $\gamma^{i2}$ , pode ser calculada da seguinte forma:

$$D(\gamma^{i1}, \gamma^{i2}) = \frac{1}{n} \sum_{j=1}^n \|\mu_j^{i1} - \mu_j^{i2}\|^2 + \sigma_j^{i1} + \sigma_j^{i2} - 2\sqrt{\sigma_j^{i1}\sigma_j^{i2}}. \quad (17)$$

Esta medida considera objetos Gaussianos e a especificidade da informação que, por vez, é inversamente proporcional ao espalhamento. FBeM combina grânulos que apresentam o menor valor de  $D(\cdot)$  para qualquer par de grânulos da coleção atual e um critério de decisão. A decisão pode ser baseada em valor limiar  $\Delta$  ou julgamento especialista a respeito da conveniência da mescla destes grânulos.

Um novo grânulo  $\gamma^i$ , combinação de  $\gamma^{i1}$  e  $\gamma^{i2}$ , é construído por funções de pertinência Gaussianas com valor modal:

$$\mu_j^i = \frac{\frac{\sigma_j^{i1}}{\sigma_j^{i2}} \mu_j^{i1} + \frac{\sigma_j^{i2}}{\sigma_j^{i1}} \mu_j^{i2}}{\frac{\sigma_j^{i1}}{\sigma_j^{i2}} + \frac{\sigma_j^{i2}}{\sigma_j^{i1}}}, j = 1, \dots, n, \quad (18)$$

e espalhamento

$$\sigma_j^i = \sigma_j^{i1} + \sigma_j^{i2}, j = 1, \dots, n. \quad (19)$$

Estas são relações heurísticas que basicamente levam em conta a proporção de incerteza em cada grânulo para determinar a localização e tamanho do novo grânulo. As mesmas relações de combinação valem para variáveis de saída  $k$ . Os coeficientes dos novos polinômios locais são:

$$a_{jk}^i = \frac{1}{2} (a_{jk}^{i1} + a_{jk}^{i2}), j = 0, \dots, n, \quad (20)$$

Naturalmente, combinação de grânulos reduz o número de regras e redundância (LEITE; GOMIDE, 2012).

### 2.5.7 Remoção de Grânulos

Um grânulo deve ser removido da estrutura de FBeM se ele parece ser inconsistente com o conceito atual. Estratégias comuns de remoção consideram (i)

apagar grânulos criados ao longo de muitas iterações; (ii) excluir grânulos baseados em valores de erro local ou (iii) apagar grânulos inativos. Neste trabalho, optamos pela estratégia de remover grânulos inativos. Grânulos antigos ainda podem ser úteis no ambiente atual enquanto que grânulos que geram erros podem ser revistos através do procedimento de adaptação de parâmetros.

Grânulos FBeM são excluídos quando se tornam inativos durante um número de iterações,  $h_r$ . Se a aplicação requer memorização de eventos raros ou se sazonalidades são esperadas, então pode ser o caso de não excluir grânulos. Remover os grânulos inativos periodicamente ajuda a manter a base de regras atualizada (LEITE; GOMIDE, 2012).

### 3 METODOLOGIA

Para este trabalho foram utilizadas a rede *neuro-fuzzy* granular evolutiva e modelagem *fuzzy* evolutiva baseada em conjuntos *fuzzy* para a detecção de câncer de mama em exames clínicos de pacientes femininos. A configuração do computador utilizado: processador Intel® Core™ i3-2328M CPU 2.2GHz, memória RAM 8,00 GB, HDD de 500 GB e sistema operacional de 32 bits.

#### 3.1 Descrição do Conjunto de Dados

Como forma de auxiliar no diagnóstico precoce de problemas de câncer de mama em pacientes feminino, foi necessário ter conhecimento de uma base de dados. Neste caso, utilizou-se dados do Repositório de Aprendizado de Máquina (UCI - *Machine Learning Repository*). Foi considerado o banco de dados, que tem por título “*Wisconsin Breast Cancer Database*” ou Base de dados de câncer de mama Wisconsin, desenvolvido pelo médico William H. Wolberg, em 1991, e disponibilizadas, em 1992, para o repositório. Tem por base de dados o estudo com 699 casos, cada um com 9 atributos, conforme mostrado na Tabela 1.

Tabela 1 – Descrição dos atributos da base de dados de câncer de mama.

<b>Atributo</b>	<b>Descrição</b>
1	Espessura do glóbulo
2	Uniformidade do tamanho da célula
3	Uniformidade da forma celular
4	Adesão marginal
5	Tamanho de célula epitelial única
6	Núcleos nus
7	Cromatina Suave
8	Nucléolos normais
9	Mitoses

Fonte: *UCI Machine Learning Repository* (1991).

Os atributos foram usados para representar amostras. Cada amostra tem uma de duas classes possíveis: benigna ou maligna. Das 699 amostras de dados considerados, a distribuição de classes é dada por:

- 454 são benignos, representando 64,95% das amostras;
- 245 são malignos, representando 35,05% das amostras.

As amostras foram obtidas periodicamente enquanto o médico William H. Wolberg relatava seus casos clínicos. O banco de dados, portanto, reflete esse agrupamento cronológico dos dados. Essas informações de agrupamento aparecem imediatamente abaixo, tendo sido removidas a partir dos próprios dados:

- Grupo 1: 367 amostras (janeiro de 1989);
- Grupo 2: 70 amostras (outubro de 1989);
- Grupo 3: 31 amostras (fevereiro de 1990);
- Grupo 4: 17 amostras (abril de 1990);
- Grupo 5: 48 amostras (agosto de 1990);
- Grupo 6: 49 amostras (atualizado em janeiro de 1991);
- Grupo 7: 31 amostras (junho de 1991);
- Grupo 8: 86 amostras (novembro de 1991).

### **3.2 Sistema de Monitoramento Evolutivo Proposto**

O sistema de monitoramento evolutivo proposto para o procedimento referente ao rastreamento e a investigação diagnóstica do câncer de mama é mostrado no diagrama da Figura 5. Os atributos de exames clínicos a serem processados pelo classificador é o vetor “dados de entrada”, onde na prática são extraídos de exames clínicos de pacientes femininos. Para o trabalho em questão, foram utilizados os dados do repositório de aprendizado de máquina. A construção do modelo detector de anomalias nos exames é realizada em modo *online* a partir das amostras características de cada exame.

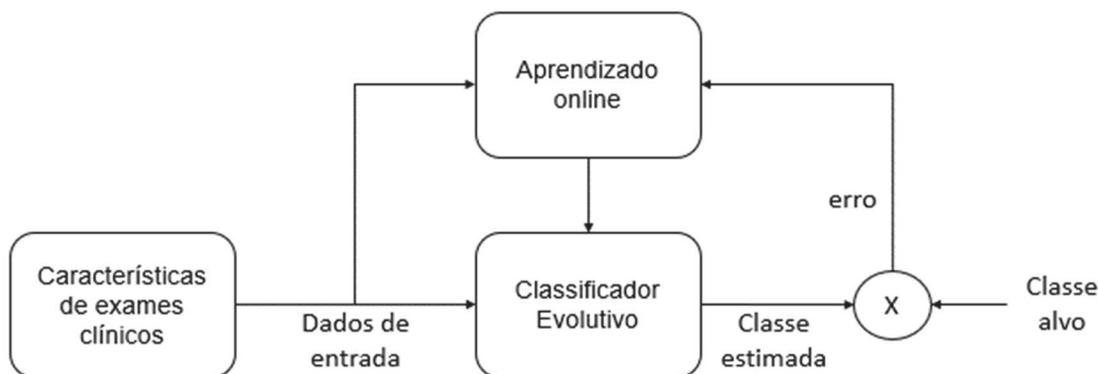


Figura 5 - Sistema de monitoramento evolutivo proposto.

Fonte: (Do autor, 2021).

Através da Figura 5, nota-se que o fluxo de dados denominado “dados de entrada” é apresentado ao classificador evolutivo, que por sua vez, realiza o aprendizado de maneira *online* e estipula a classe estimada de saída. Posteriormente, é avaliado o erro cometido pelo classificador e realizado o aprendizado *online* e adaptativo, evoluindo de acordo com a variabilidade do fluxo de dados.

Os parâmetros iniciais para o modelo evolutivo eGNN são  $\rho = 0,82$ ;  $h_r = 2000$ ;  $\eta = 8$  e  $c = 0$ . Já os parâmetros iniciais para o modelo evolutivo FBeM são  $\rho = 0,12$ ;  $h_r = 2000$ ;  $\eta = 5$ ; e  $c = 0$ . O número total de amostras para o desenvolvimento de ambos os modelos são  $h = 699$ .

## 4 RESULTADOS E DISCUSSÕES

Este capítulo apresenta o desempenho de detecção de anomalias utilizando os modelos evolutivos eGNN e o FBeM. Os modelos foram comparados com MLR (*Multiple Linear Regression*), a rede neural MLP (*Multi-Layer Perceptron*) e o SVM (*Support Vector Machine*). Foi utilizada a mesma base de dados que foi considerada para os modelos evolutivos.

### 4.1 Detecção de Anomalias

Neste item, é analisado o desempenho de detecção utilizando os sistemas evolutivos, em particular. A detecção de anomalias nos gânglios linfáticos compreende basicamente a aplicação de um algoritmo que realize a detecção baseada nos parâmetros extraídos dos exames clínicos.

Considere a acurácia como índice de desempenho,

$$A_{cc} = \frac{VP + VN}{VP + VN + FP + FN} \times 100\% \quad (21)$$

em que  $A_{cc}$  é a acurácia;  $VP$  e  $VN$  são os verdadeiros positivos e verdadeiros negativos, respectivamente;  $FP$  e  $FN$  são os falsos positivos e falsos negativos, respectivamente.

As classes positivas são todas aquelas relativas ao câncer de mama. A classe negativa é a operação normal dos exames. Os modelos eGNN e FBeM realizam a detecção dos atributos de maneira *online*.

A Tabela 2 resume o desempenho de detecção de eGNN e FBeM de câncer de mama em função dos atributos da Tabela 1.

Tabela 2 – Desempenho de classificação do modelo eGNN e FBeM considerando 699 amostras.

Classes	Classificador	A <sub>cc</sub> (%)	Tempo <sub>Total</sub> (s)	Tempo por Amostra (ms)
As duas classes envolvidas	eGNN	92,00	1,15	1,65
	FBeM	95,00	3,41	4,89

Fonte: Do autor (2021).

Nota-se pela Tabela 2 que os resultados para os classificadores eGNN e FBeM são encorajadores. A acurácia para o classificador eGNN foi de 92% com tempo computacional requerido de 1,65 ms por amostra. Enquanto o classificador FBeM teve uma acurácia de 95% e tempo computacional requerido de 4,89 ms por amostra. Isto é considerado muito rápido, pois 699 amostras foram processadas.

## 4.2 Comparação entre Métodos

Nesta seção os modelos evolutivos eGNN e FBeM são comparados com os modelos estatísticos MLR (*Multiple Linear Regression*) e SVM (*Support Vector Machine*) e a rede neural MLP (*Multi-Layer Perceptron*). É utilizada a mesma base de dados que foi considerada para a eGNN e FBeM.

A Regressão Linear Múltipla é uma coleção de técnicas estatísticas para construir modelos que descrevem de maneira razoável relações entre várias variáveis explicativas de um determinado processo (HOFFMANN, 2016). O modelo utiliza a comparação e o ajuste de mínimos quadrados nos dados.

A Máquina de Vetores de Suporte é uma técnica de classificação baseada em aprendizado de máquina na qual é utilizado o aprendizado supervisionado. Isso quer dizer que seu treinamento deve se possuir uma base na qual as instâncias de cada classe estejam previamente classificadas (CÂMARA, 2015). O parâmetro desse modelo utilizado na comparação é baseado na função de ativação polinomial de grau 1, algoritmo de otimização gradiente conjugado escalado.

As Redes Neurais Artificiais (RNA) são modelos inspirados na forma de processamento de informação dos neurônios biológicos. Foi escolhida a RNA *feed-forward* convencional, chamada MLP (Silva, 2016). O arquitetura desse modelo utilizado na comparação possui 5 neurônios na camada oculta, são realizadas 100 iterações, e são usadas função de ativação linear.

A Tabela 3 apresenta o resultado comparativo do desempenho ( $A_{CC}$ ) entre os modelos considerados em relação as amostras obtidas pelo Repositório de Aprendizado de Máquina (UCI - *Machine Learning Repository*) considerando eGNN, FBeM, MLR, SVM e MLP, respectivamente.

Tabela 3 – Resultado Comparativo entre os classificadores.

Classes	Classificador	Acc (%)	Tempo <sub>Total</sub> (s)	Tempo por Amostra (ms)
As duas classes envolvidas	eGNN	92,00	1,15	1,65
	FBeM	95,00	3,41	4,89
	MLR	95,00	1,35	1,93
	SVM	96,00	3,41	4,87
	MLP	96,00	1,43	2,04

Fonte: Do autor (2021).

Percebe-se pela Tabela 3 que todos os classificadores apresentaram resultados bastante satisfatórios tanto em relação ao desempenho quanto para o tempo de processamento. Os classificadores evolutivos eGNN e FBeM obtiveram uma acurácia de 92% e 95% respectivamente. Ambos são capazes de realizar a detecção *online* ou em tempo real, não necessitando de treinamento *offline* das amostras para apresentar a classe estimada final, ou seja, realizam a leitura das amostras uma única vez. Isso significa que os seus desempenhos elevados ocorrem mesmo quando são submetidos à um fluxo de dados não-estacionário, sendo pertinente na aplicação real.

Os classificadores MLR e SVM que são ferramentas estatísticas obtiveram acurácia de 95% e 96%, respectivamente. Enquanto a rede neural MLP obteve acurácia de 96%. O desempenho dos classificadores MLR, SVM e MLP são consideráveis, mas por sua vez, precisam realizar o treinamento das amostras para apresentar o seu desempenho final. Dessa forma, não possuem a capacidade de trabalhar com a variação de fluxo de entrada, porque precisam treinar as amostras para posteriormente validarem as mesmas.

### 4.3 Matriz de Confusão

A matriz de confusão é uma forma intuitiva de saber como o classificador está se comportando. Ela oferece uma medida efetiva ao mostrar o número de classificações corretas entre detecção ou não do distúrbio em um sinal. O nome deriva do fato de que ela torna mais fácil ver se o sistema está confundindo as classes. As entradas da matriz são representadas por  $M(C_i, C_j)$ , indicando o número de exemplos

de  $N$  que são da classe  $C_i$  mas que foram classificados pela hipótese  $h$  como sendo da classe  $C_j$  como mostra a equação (22):

$$M(C_i, C_j) = \sum_{\{x \in N: y=C_i\}} \mathbb{1}_{h(x)=C_j} \quad (22)$$

onde o número de acertos, para cada classe, localiza-se na diagonal principal  $M(C_i, C_i)$  da matriz; os demais elementos  $M(C_i, C_j)$ , para  $i \neq j$ , representam erros na classificação. A matriz de confusão possibilita analisar a porcentagem de assiduidade das diversas amostras associadas a uma classe, assim como permite o entendimento e a quantificação das amostras classificadas de forma errada (TAN, et al., 2006). A matriz de confusão é utilizada, para avaliar o modelo de detecção de cada classificador em relação as classes envolvidas. As classes das matrizes apresentadas a seguir são distribuídas em duas classes distintas, sendo (classe 0) amostras benignas e (classe 1) amostras malignas. Para cada análise os modelos utilizados são eGNN, FBeM, SVM, MLR e MLP.

A Figura 6 apresenta a matriz de confusão que é para avaliar o índice de acertos para detecção do câncer de mama em 699 amostras para a rede *neuro-fuzzy* granular evolutiva.

**Matriz de Confusão**

<b>Classe de Saída</b>	0	412	14	96,70%
	1	42	231	84,60%
		90,70%	94,30%	92%
		0	1	
		<b>Classe Alvo</b>		

Figura 6 – Matriz de confusão para a rede neuro-fuzzy granular evolutiva (eGNN).  
Fonte: (Do autor, 2021).

Nota-se através da Figura 6 que os acertos estão na diagonal principal da matriz, em cor verde. Os erros estão identificados pela cor branca, onde, das 245 amostras que eram malignas, 14 amostras foram classificadas erradas, como sendo benignas; e 231 foram classificadas corretamente como sendo malignas, enquanto das 454 amostras benignas, 42 amostras foram classificadas erradas, como sendo malignas; e 412 amostras classificadas de maneira correta como benignos. A classe 0 teve o pior desempenho com 90,70% de acurácia, com 42 amostras classificadas erradas. Dessa forma, a rede *neuro-fuzzy* granular evolutiva teve um desempenho de 92% na detecção de anormalidades nas amostras de exames clínicos. O tempo de processamento dos dados pela eGNN foi de 1,1586 s para todos os eventos.

A Figura 7 apresenta a matriz de confusão que é para avaliar o índice de acertos para detecção do câncer de mama em 699 amostras para a Modelagem Evolutiva Baseada em Conjuntos *Fuzzy*.

**Matriz de Confusão**

<b>Classe de Saída</b>	0	442	22	95,30%
	1	12	223	94,90%
		97,40%	91,00%	95%
		0	1	
		<b>Classe Alvo</b>		

Figura 7 – Matriz de confusão para a Modelagem Evolutiva Baseada em Conjuntos *Fuzzy* (FBeM).  
Fonte: (Do autor, 2021).

Percebe-se pela Figura 7 que a classe 0 teve melhor desempenho em relação a classe 1 com 97,40% de acurácia, onde das 445 amostras que eram benignas, 442 foram classificadas de maneira correta e 12 amostras foram classificadas de maneira errada, como sendo maligna. Enquanto das 245 amostras malignas, 223 foram classificadas corretamente e 22 amostras foram classificadas de maneira errada,

como benigna. Dessa forma, a modelagem evolutiva baseada em conjuntos *fuzzy* teve um desempenho de 95% na detecção das amostras. O tempo de processamento dos dados pela FBeM foi de 3,4194 s para todos os eventos.

A Figura 8 apresenta a matriz de confusão que é utilizada para avaliar o índice de certos para detecção do câncer de mama em 699 amostras para a) MLR, b) SVM e c) MLP.

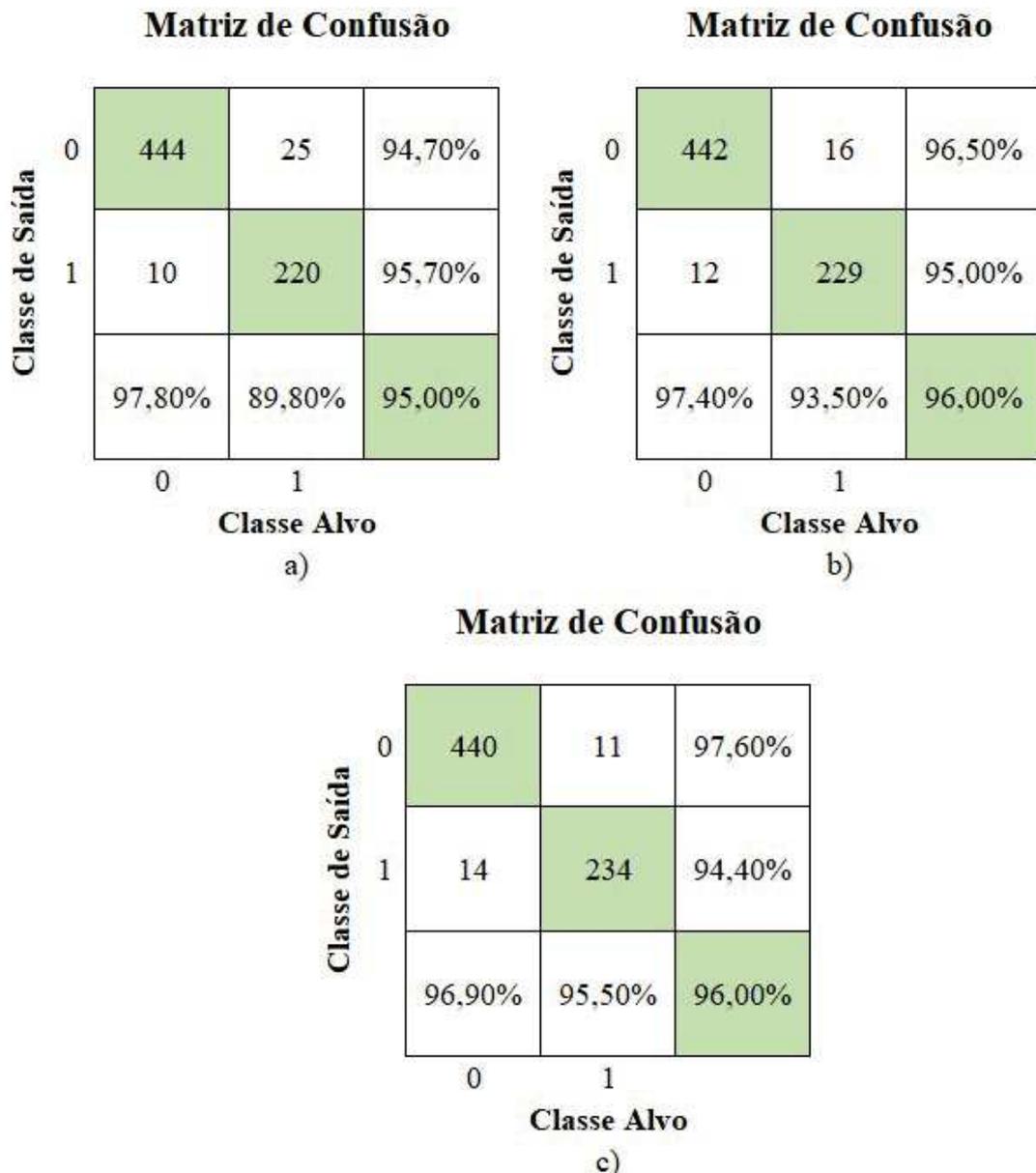


Figura 7 – Matriz de confusão para a) MLR b) SVM e c) MLP.  
 Fonte: (Do autor, 2021).

Através da Figura 8 pode-se observar que os acertos estão na diagonal principal das matrizes, em cor verde. Para a MLR nota-se que a classe 0 teve melhor desempenho em relação a classe 1, com 97,80% de acurácia. Das 454 amostras que eram benignas, 10 amostras foram classificadas como malignas, enquanto das 245 amostras que eram malignas, 20 amostras foram classificadas erradas, como sendo benignas. Assim apenas 220 amostras foram classificadas certas, como sendo malignas, obtendo um desempenho de 95% na detecção de anormalidades nos exames para o modelo MLR com tempo de processamento dos dados de 0,45 s para todos os eventos.

Para o modelo SVM a classe 0 também teve melhor desempenho em relação a classe 1 com 97,40% de acurácia, onde, das amostras benignas, 12 amostra foram classificadas malignas de um total de 454. Enquanto das 245 amostras malignas, 16 foram classificadas erradas, como sendo benignas, obtendo um desempenho de 96% na detecção de anormalidades nos exames para o modelo com tempo de processamento dos dados de 1,20 s para todos os eventos.

Para o modelo MLP, das 454 amostras benignas, 14 foram classificadas de maneira errada, como sendo malignas; enquanto das 245 amostras malignas, 11 foram classificadas de maneira errada, como sendo benignas. Percebe-se que a MLP teve maior confusão de amostras da classe 1 em relação a classe 0, obtendo um desempenho de 96% na detecção de anormalidades nos exames para o modelo MLP com tempo de processamento dos dados de 0,50 s para todos os eventos.

## 5 CONCLUSÃO

Os sistemas evolutivos apresentaram alto desempenho na detecção de amostras benignas e malignas relacionados aos dados apresentados de exames clínicos de mama nos pacientes, possibilitando a análise eficiente de resultados clínicos e baixo custo computacional, ou seja, pouco tempo de processamento, atendendo com excelência aos resultados esperados. A rede *neuro-fuzzy* granular e evolutiva e modelagem evolutiva baseada em conjuntos *fuzzy* foram capazes de trabalhar de maneira *online*, em tempo real, e foram extremamente eficazes em problemas reais e pertinentes proposta neste trabalho.

Em adição ao objetivo proposto em utilizar a inteligência computacional para auxiliar o diagnóstico de presença de tumores das regiões mamárias, o trabalho permitiu o compartilhamento de informações sobre o câncer de mama, o que pode promover a conscientização sobre a doença, proporcionando maior acesso aos serviços de diagnóstico e de tratamento e, dessa forma, contribuir para a redução da mortalidade. Porquanto, quando o câncer de mama é descoberto no início, a doença tem 95% de chance de cura. Vários estudos têm confirmado a importância da mamografia na redução da mortalidade pelo câncer de mama. Por outro lado, existe uma preocupação com os efeitos que a radiação ionizante pode trazer para o organismo. Dessa forma, uma ferramenta que trabalha em tempo real e *online*, permitindo uma maior precisão na detecção de anomalias nos exames, impede que as mulheres sejam expostas inúmeras vezes à radiação desnecessária.

Para trabalhos futuros pretende-se estender os métodos propostos a análise e processamento de imagens de mamografias de pacientes clínicos. A intenção é ressaltar o benefício da aplicação de sistemas evolutivos na assistência médica, a fim de colaborar com o diagnóstico precoce de câncer de mama e demais doenças.

## BIBLIOGRAFIA

ANGELOV, P.; FILEV, D.; KASABOV, N. **Evolving fuzzy systems – preface to the special section**. IEEE Transactions on Fuzzy Systems, New York, v. 6, n. 6, p. 1390-1392, 2008.

ARAÚJO, M. C. de. **Utilização de câmera por infravermelho para avaliação de diferentes patologias em clima tropical e uso conjunto de sistemas de banco de dados para detecção de câncer de mama**. Universidade Federal de Pernambuco, 2009.

BRASIL MINISTÉRIO DA SAÚDE. **Controle do câncer: Uma Proposta de Integração ensino serviço**. Rio de Janeiro, Pro-Onco/NUTES, 1990.

BOUCHACHIA, A. **An evolving classification cascade with self-learning**. Evolving Systems, New York, v. 1, n. 3, p. 143-160, Oct. 2010.

CÂMARA, R. **Entendendo Sua Matemática – Parte 1 – A Margem**. Laboratório Mobilis Computação Móvel. Universidade federal de Ouro Preto. 2015.

CHAVES, E. L. **Detecção de câncer de mama por meio de imagens infravermelhas utilizando Redes Neurais Convolucionais**. Trabalho de Conclusão de Curso de Bacharelado em Ciência da Computação. Universidade Federal de Uberlândia. Uberlândia. 2019

DIAS, M. **O Enfrentamento da Problemática do Câncer de Mama na Estratégia de Saúde da Família**. Trabalho de Conclusão de Curso de Especialização em Atenção Básica em Saúde da Família. Conselheiro Lafaiete. 2009.

GABRYS, B.; PETRAKIEVA, L. **Combining labelled and unlabelled data in the design of pattern classification systems**. International Journal of Approximate Reasoning, Oxford, v. 35, n. 3, p. 251-273, Mar. 2004.

HOFFMANN, R. **Análise Estatística de Relações lineares e não lineares**. 1ª Edição Digital. ISBN: 978-85-921057-1-6. 259 p – Universidade de São Paulo. 2016.

INSTITUTO NACIONAL DE CÂNCER – INCA. **Parâmetros Técnicos Para o Rastreamento do Câncer de Mama**. Ministério da Saúde. Rio de Janeiro. 2021a.

INSTITUTO NACIONAL DE CÂNCER – INCA. **Conceito e Magnitude do Câncer de Mama**. Ministério da Saúde. Rio de Janeiro. 2021b.

LEITE, D; COSTA, P.; GOMIDE, F.; BALLINI, R. **Evolving Fuzzy Granular Modeling from Nonstationary Fuzzy data Streams**. EVOLVING SYSTEMS, v. 3, p. 65-79. 2012.

- LEITE, D. **Comparison of Genetic and Incremental Learning Methods for Neural Network-based Electrical Machine Fault Detection**. In: Lughofer, E.; Sayed30 Mouchaweh, M. (Eds.) Predictive Maintenance in Dynamic Systems, p. 231-268. Springer: Cham. 2019.
- LEITE, D.; DECKER, L.; Santana, M.; Souza, P. EGFC: **Evolving Gaussian Fuzzy Classifier from Never-Ending Semi-Supervised Data Streams – With Application to Power Quality Disturbance Detection and Classification**. IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'20), 9p. arXiv:2004.09986. 2020.
- MASCARENHAS, N. A. **Breve introdução ao reconhecimento estatístico de padrões**. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA PARA O PROGRESSO DA CIÊNCIA, 39., 1987, Brasília. Anais... Brasília: SBPC, 1987.
- PONTES, D. L. de B. **Câncer Benigno e Maligno**. 2013. Disponível em: <https://www.einstein.br/noticias/noticia/cancer-benigno-maligno>. Acessado dia 24/03/2021.
- PRADO, A. J.; ELFES, A. **Um projeto em reconhecimento de padrões de forma**. 1975. 125 p. Monografia (Graduação em Engenharia Eletrônica) - Instituto Tecnológico de Aeronáutica, São José dos Campos, 1975.
- REIS, C. F.; ALBURQUERQUE, M. P.; CASTRO, S. B. **Introdução ao Reconhecimento de Padrões utilizando Redes Neurais**. Centro Brasileiro de Pesquisas Físicas – CBPF. Rio de Janeiro/RJ. 2012.
- SANTANA, M. W.; LEITE, D. F. **Aprendizado de Máquina Fuzzy Incremental para Classificação de Falhas em Sistemas de Potência**. Abakós, v. 8, n. 2, p. 03-28, 24 nov. 2020.
- SANTANA, M. **Agrupamento Incremental de Fluxo de Dados para Análise e Monitoramento da Qualidade de Energia**. 98p. Dissertação (Mestrado em Engenharia de Sistemas e Automação) – U. Federal de Lavras. 2017.
- SILVA, I. N. et al. **Redes Neurais Artificiais para Engenharia e Ciências Aplicadas: Fundamentos Teóricos e Aspectos Práticos**. ISBN 9788588098879. 431p. 2016.
- SILVA, S. et. al. **Evolving Neuro-Fuzzy Network for Real-Time High Impedance Fault Detection and Classification**. Neural Computing and Applications. p. 14p. DOI: doi.org/10.1007/s00521-018-3789-2. 2018.
- SOARES, E. et. al. **Evolving Fuzzy Set-based and Cloud-based Unsupervised Classifiers for Spam Detection**. IEEE Latin America Transactions, v. 17, n. 9, p. 1449-1457. 2019.
- SOCIEDADE BRASILEIRA DE PATOLOGIA. **Científico: Câncer de Mama**. São Paulo. 2016.

STIEGLER, S. M. **A modest proposal: a new standard for the normal.** The American Statistician, Washington, v. 36, n. 2, p. 137-138, Mar. 1982.

TAN, P.-N. et al., **Introduction to data mining.** Noida: Pearson Education India, 2006. 165 p.

THEODORIDIS, S.; KOUTOUMBAS, K. **Pattern recognition.** San Diego: Academic Press, 1999. 689 p.

TOU, J. T.; GONZALEZ, R. C. **Pattern recognition principles.** Massachusetts: Addison-Wesley Publishing Company, 1981. 399 p.

UCI Machine Learning Repository. **Center for Machine Learning and Intelligent Systems.** Disponível em: <https://archive.ics.uci.edu/>, acessado dia 22/03/2021.

WENTZ, V. **Redes Neurais Densamente Conectadas para Detecção de Câncer de Mama em Imagens Histopatológicas.** Trabalho de Conclusão de Curso de Bacharelado em Computação. Universidade Tecnológica Federal do Paraná. Departamento Acadêmico de Computação. Medianeira, 2019.



**RELATÓRIO FINAL DE CURSO Nº 14/2025 - CEENP (11.51.21)**

**(Nº do Protocolo: NÃO PROTOCOLADO)**

**(Assinado digitalmente em 09/04/2025 18:14 )**

MARCIO WLADIMIR SANTANA  
PROFESSOR ENS BASICO TECN TECNOLOGICO  
CTETTNP (11.50.36)  
Matricula: ###520#9

Visualize o documento original em <https://sig.cefetmg.br/documentos/> informando seu número: 14, ano: 2025, tipo:  
**RELATÓRIO FINAL DE CURSO**, data de emissão: 09/04/2025 e o código de verificação: **a9924a4f0a**