Proceeding Series of the Brazilian Society of Applied and Computational Mathematics, Vol. 5, N. 1, 2017.

Trabalho apresentado no CNMAC, Gramado - RS, 2016.

**Proceeding Series of the Brazilian Society of Computational and Applied Mathematics**

# Memórias Associativas Recorrentes Exponenciais Fuzzy Generalizadas Aplicadas à Classificação de Padrões

Aline Cristina de Souza[1]
Departamento de Matemática Aplicada, IMECC - UNICAMP, Campinas, SP.
Marcos Eduardo Valle[2]
Departamento de Matemática Aplicada, IMECC - UNICAMP, Campinas, SP.

**Resumo**. Memórias associativas são modelos matemáticos biologicamente inspirados para o armazenamento e recordação de um conjunto finito de itens. As memórias associativas recorrentes exponenciais *fuzzy* generalizadas (GRE-FAMs) podem ser efetivamente utilizadas para o armazenamento e recordação de uma família finita de conjuntos *fuzzy*. Neste trabalho, descrevemos como uma GRE-FAM pode ser aplicada em um problema de classificação de padrões. Experimentos computacionais mostraram que a acurácia obtida pelos classificadores definidos com base nas GRE-FAMs é competitiva comparada a outros classificadores da literatura.

**Palavras-chave**. Memórias Associativas, Medidas de Similaridade, Classificação de Padrões.

## 1 Introdução

Memórias associativas são modelos matemáticos motivados pela capacidade do cérebro humano de armazenar e recordar informações por meio de associações [2,4]. Tais modelos têm como objetivo armazenar um conjunto de associações $\mathcal{X} = \{(\mathbf{x}^1, \mathbf{y}^1), \cdots, (\mathbf{x}^p, \mathbf{y}^p)\}$, chamado conjunto das memórias fundamentais. Além disso, espera-se que uma memória associativa recupere um item armazenado $\mathbf{y}^\xi$ mesmo a partir da apresentação de uma versão corrompida $\widetilde{\mathbf{x}}^\xi$ de $\mathbf{x}^\xi$. No caso em que $\mathbf{x}^\xi = \mathbf{y}^\xi$, para todo $\xi = 1, \cdots, p$, a memória é dita autoassociativa. Memórias associativas projetadas para o armazenamento e recordação de conjuntos *fuzzy* são chamadas memórias associativas *fuzzy* [5,8].

As memórias associativas recorrentes exponenciais *fuzzy* generalizadas (GRE-FAMs, do inglês *generalized recurrent exponential fuzzy associative memories*) podem ser efetivamente utilizadas para implementar uma memória autoassociativa para o armazenamento e recordação de uma família finita de conjuntos *fuzzy* [7].

Um problema de classificação de padrões consiste em associar a cada padrão um rótulo, o qual representa sua classe. Para sintetizar um classificador, geralmente utilizamos um conjunto de padrões já rotulados, denominado conjunto de treinamento. Neste trabalho, apresentaremos uma aplicação das GRE-FAMs a problemas de classificação de padrões.

[1]s.alinedesouza@gmail.com
[2]valle@ime.unicamp.br

Este trabalho está organizado da seguinte maneira: na próxima seção relembramos alguns conceitos importantes e apresentamos as memórias associativas recorrentes exponenciais *fuzzy* generalizadas. Na Seção 3, descrevemos como as GRE-FAMs podem ser aplicadas a problemas de classificação de padrões. Os resultados dos experimentos computacionais realizados para avaliar o desempenho dos classificadores baseados nas GRE-FAMs são apresentados na Seção 4. Finalizamos o trabalho na Seção 5 com as considerações finais.

## 2 Memórias Associativas Recorrentes Exponenciais *Fuzzy* Generalizadas

Iniciaremos esta seção recordando alguns conceitos importantes para o desenvolvimento do trabalho, incluindo as definições de conjuntos *fuzzy* e medidas de similaridade. Posteriormente, apresentaremos a definição das memórias associativas recorrentes exponenciais *fuzzy* generalizadas.

Um subconjunto *fuzzy* $X$ de um universo $U$ é definido por meio de sua função de pertinência $X : U \to [0,1]$, de forma que, dado um elemento $u \in U$, $X(u)$ indica o grau de pertinência de $u$ ao subconjunto *fuzzy* $X$. Nos casos em que o universo $U$ é finito, digamos $U = \{u_1, \cdots, u_n\}$, um subconjunto *fuzzy* $X$ de $U$ pode ser identificado com um vetor $X = [X(u_1), \cdots, X(u_n)]^T$ pertencente ao hipercubo unitário n-dimensional $[0,1]^n$. Denotaremos por $\mathcal{F}(U)$ o conjunto dos subconjuntos *fuzzy* de $U$ e por $\mathcal{P}(U)$ o conjunto das partes de $U$.

Dados $A, B \in \mathcal{F}(U)$, dizemos que $A$ é subconjunto de $B$ e escrevemos $A \subset B$, se $A(u) \leq B(u)$, para todo $u \in U$. Além disso, o complemento de um conjunto *fuzzy* $A \in \mathcal{F}(U)$, denotado por $\bar{A}$, é definido por $\bar{A}(u) = 1 - A(u)$, para todo $u \in U$.

Uma medida de similaridade é uma função que associa a cada par de conjuntos *fuzzy* um número pertencente ao intervalo $[0,1]$, o qual indica o quanto esses conjuntos são iguais. Neste trabalho, adotaremos uma versão normalizada da definição proposta por Xuecheng [11]:

**Definição 2.1** (Medidas de Similaridade)**.** *Uma medida de similaridade é uma função* $\mathcal{S} : \mathcal{F}(U) \times \mathcal{F}(U) \to [0,1]$ *que satisfaz, para quaisquer conjuntos fuzzy* $A, B, C, D \in \mathcal{F}(U)$*, as seguintes propriedades:*

1. $\mathcal{S}(A, B) = \mathcal{S}(B, A)$.
2. $\mathcal{S}(A, A) = 1$.
3. *Se* $A \subseteq B \subseteq C \subseteq D$*, então* $\mathcal{S}(A, D) \leq \mathcal{S}(B, C)$.
4. $\mathcal{S}(A, \bar{A}) = 0$*, para todo subconjunto clássico* $A \in \mathcal{P}(U)$.

*Além disso, dizemos que* $\mathcal{S}$ *é uma* ***medida de similaridade forte*** *se* $\mathcal{S}(A, B) = 1$ *implica* $A = B$.

As memórias associativas recorrentes exponenciais *fuzzy* generalizadas (GRE-FAMs, do inglês *generalized recurrent exponential fuzzy associative memories*) são projetadas para o armazenamento e a recordação de uma família finita de conjuntos *fuzzy*. Especificamente, dado um conjunto *fuzzy* inicial $X_0 \in \mathcal{F}(U)$, as GRE-FAMs produzem uma sequência de conjuntos *fuzzy* $X_t \in \mathcal{F}(U)$, os quais são obtidos a partir de uma combinação afim das memórias fundamentais conforme a definição a seguir [7]:

**Definição 2.2** (GRE-FAM)**.** *Considere uma família finita de memórias fundamentais* $\mathcal{A} = \{A^1, A^2, \cdots, A^p\} \subseteq \mathcal{F}(U)$*, um número real* $\alpha > 0$ *e uma medida de similaridade* $\mathcal{S} : \mathcal{F}(U) \times \mathcal{F}(U) \to [0,1]$*. Além disso, considere* $C = (c_{ij})$ *a matriz definida por* $c_{ij} = e^{\alpha \mathcal{S}(A^i, A^j)}$*, para todo* $i, j = 1, \cdots, p$ *e defina* $G = C^{-1}$*. Dado um conjunto fuzzy inicial,* $X_0 \in \mathcal{F}(U)$*, a GRE-FAM produz uma sequência de conjuntos fuzzy* $\{X_t\}_{t \geq 0}$ *definidos, para quaisquer* $u \in U$ *e* $t = 0, 1, \ldots,$ *por:*

$$X_{t+1}(u) = \varphi\left(\sum_{\xi=1}^{p} w_{\xi t} A^{\xi}(u)\right), \tag{1}$$

*em que os pesos* $w_{\xi t}$ *são dados pela expressão*

$$w_{\xi t} = \frac{\displaystyle\sum_{\mu=1}^{p} g_{\xi\mu} e^{\alpha \mathcal{S}(A^{\mu}, X_t)}}{\displaystyle\sum_{\eta=1}^{p}\sum_{\mu=1}^{p} g_{\eta\mu} e^{\alpha \mathcal{S}(A^{\mu}, X_t)}} \tag{2}$$

*e a função* $\varphi : \mathbb{R} \to [0,1]$*, definida por*

$$\varphi(x) = \max\{0, \min\{1, x\}\}, \tag{3}$$

*garante que* $X_{t+1}(u) \in [0,1]$*, para quaisquer* $t \geq 0$ *e* $u \in U$*.*

**Observação 2.1.** *Na definição 2.2, consideramos* $G = C^{-1}$*, pois dessa forma, as memórias fundamentais são pontos fixos da GRE-FAM [9]. Caso a matriz* $C$ *não seja invertível, consideramos* $G = C^{\dagger}$*, em que* $C^{\dagger}$ *denota a pseudo-inversa de* $C$ *[3].*

# 3 Classificação de Padrões utilizando GRE-FAMs

Em problemas de classificação, o objetivo é atribuir uma determinada classe a um dado padrão $X$. Tal tarefa é realizada com base em um conjunto de padrões já rotulados, ao qual chamaremos de conjunto de treinamento. Nesta seção, inspirados pelos classificadores de representação esparsa [10], definiremos um classificador com base nas GRE-FAMs. Em outras palavras, associaremos cada subconjunto *fuzzy* $X$ de $U$ a uma classe $\ell \in \mathcal{L}$ utilizando as GRE-FAMs.

Consideremos o conjunto de treinamento

$$\mathcal{A}_{\mathcal{L}} = \{(A^{\xi}, \ell_{\xi}) : \xi = 1, \ldots, p\} \subseteq \mathcal{F}(U) \times \mathcal{L}, \tag{4}$$

 

onde $\mathcal{L}$ é um conjunto de rótulos e $\mathcal{A}^\xi$ são subconjuntos *fuzzy* de $U$ distintos e não vazios. Os modelos de classificadores de representação esparsa [10] se baseiam na hipótese de que uma amostra $X$ de uma determinada classe $i \in \mathcal{L}$ pode ser escrita, aproximadamente, como combinação linear dos padrões de treinamento pertencentes a esta classe, isto é,

$$X(u) = \sum_{\xi;\, \ell_\xi = i} \alpha_\xi A^\xi(u) = \sum_{\xi=1}^{p} \alpha_\xi A^\xi(u), \quad \forall u \in U, \tag{5}$$

em que $\alpha_\xi = 0$, se $\ell_\xi \neq i$.

Agora, se $X$ é uma versão ruidosa de um padrão pertencente à classe $i$, esperamos que o padrão $X_1$, produzido pela GRE-FAM em uma iteração com $X_0 = X$, também pertença a esta classe. Por um lado, das equações (1) e (2) com $t = 0$, temos que:

$$X_1(u) = \varphi\left(\sum_{\xi=1}^{p} w_{\xi 0} A^\xi(u)\right), \quad \forall u \in U. \tag{6}$$

Por outro lado, inspirados pelos modelos de representação esparsa, esperamos que

$$X_1(u) \approx \sum_{\xi=1}^{p} \alpha_\xi A^\xi(u), \quad \forall u \in U, \tag{7}$$

com $\alpha_\xi = 0$, se $\ell_\xi \neq i$.

Motivados pelas equações (6) e (7), definimos os coeficientes $\alpha_\xi$, $\xi = 1, \cdots, p$ em (7) da seguinte forma:

$$\alpha_\xi = w_{\xi 0} \chi_i(\ell_\xi), \tag{8}$$

em que $\chi_i : \mathcal{L} \to \{0, 1\}$ é definida por $\chi_i(x) = 1$, se $x = i$ e $\chi_i(x) = 0$ caso contrário. Assim, temos que $\alpha_\xi = w_{\xi 0}$, se $\ell_\xi = i$ e $\alpha_\xi = 0$, se $\ell_\xi \neq i$.

Dessa forma, se o padrão $X$ pertence a classe $i$, teremos

$$X_1(u) \approx \sum_{\xi=1}^{p} w_{\xi 0} \chi_i(\ell_\xi) A^\xi(u), \quad \forall u \in U. \tag{9}$$

Na prática, porém, não sabemos a que classe um dado padrão $X$ pertence. Então, atribuímos a $X$ a classe $\ell$ que minimiza a distância entre $X_1$ e as combinações lineares $\sum_{\xi=1}^{p} w_{\xi 0} \chi_\ell(\ell_\xi) A^\xi$. Formalmente, atribuímos a $X$ a classe $\ell$, em que $\ell$ satisfaz:

$$d_2\left(X_1, \sum_{\xi=1}^{p} w_{\xi 0} \chi_\ell(\ell_\xi) A^\xi\right) \leq d_2\left(X_1, \sum_{\xi=1}^{p} w_{\xi 0} \chi_i(\ell_\xi) A^\xi\right), \quad \forall i \in \mathcal{L}, \tag{10}$$

onde $d_2$ denota a distância $L_2$.

 

# 4 Experimentos Computacionais

Experimentos computacionais foram realizados para avaliar a acurácia do classificador baseado na GRE-FAM. Consideramos o parâmetro $\alpha = 15$ e a matriz $G = C^{-1}$ conforme a Definição 2.2. Utilizamos a medida de similaridade de Gregson [12] definida por:

$$\mathcal{S}_G(A,B) = \frac{\displaystyle\sum_{j=1}^{n} \min\{a_j, b_j\}}{\displaystyle\sum_{j=1}^{n} \max\{a_j, b_j\}}. \tag{11}$$

Realizamos os experimentos com seis conjuntos de dados disponíveis em [6], a saber: *Haberman survival*, *Liver disorders*, *Pima Indians diabetes*, *Breast Cancer Wisconsin (Original)*, *Heart disease* e *Hepatitis disease.*

Primeiramente, cada conjunto de dados foi dividido em duas partes, sendo uma delas, formada por 70% dos padrões, utilizada como conjunto de treinamento e os 30% restantes utilizados como conjunto de teste. Depois, aplicamos uma transformação nos dados a fim de representá-los por conjuntos *fuzzy*. Especificamente, dado um par $(\mathbf{v}, \ell)$ pertencente a um conjunto de dados, associamos $\mathbf{v} = [v_1, \cdots, v_n]^T$ ao conjunto *fuzzy* $X = [x_1, \cdots, x_n]^T$ definido por:

$$x_i = \frac{1}{1 + e^{-(v_i - \mu_i)/\sigma_i}} \in [0,1], \quad \forall i = 1, \cdots, n, \tag{12}$$

onde $\mu_i$ e $\sigma_i$ representam, respectivamente, a média e o desvio padrão da $i$-ésima componente de todos os padrões pertencentes ao conjunto de treinamento. Dessa forma, obtivemos como conjunto de treinamento uma família de conjuntos *fuzzy* correspondentes a 70% dos dados.

A Tabela 1 mostra a acurácia do classificador baseado na GRE-FAM, para o conjunto de teste, em cada um dos seis problemas citados acima, bem como a acurácia de outros 21 classificadores da literatura considerados anteriormente por Aldape-Pérez et al. em [1]. A última coluna da tabela é formada pelas médias das acurácias obtidas por cada classificador nos seis problemas. Os valores da acurácia obtidos pelos 21 primeiros classificadores, bem como maiores detalhes sobre os conjuntos de dados e sobre os outros classificadores considerados na Tabela 1 podem ser encontrados em [1]. Os maiores valores de acurácia para cada problema estão destacados em negrito.

Os resultados da Tabela 1 mostram que o classificador baseado na GRE-FAM apresentou bom desempenho. De fato, o classificador baseado na GRE-FAM obteve a maior acurácia em três dos seis problemas de classificação, bem como a maior média considerando as acurácias obtidas em todos os problemas. Foram realizados também experimentos considerando outros problemas de classificação, os resultados podem ser encontrados em [9].

# 5 Considerações Finais

Inicialmente, apresentamos as memórias associativas recorrentes exponenciais *fuzzy* generalizadas [7]. Resumidamente, dado um conjunto *fuzzy* inicial, as GRE-FAMs produzem

 

Tabela 1: Acurácia de classificação obtida para o conjunto de teste utilizando 70% dos dados como conjunto de treinamento. A acurácia dos 21 primeiros algoritmos foram extraídos de [1].

| | Algorithm | Haberman | Liver | Diabetes | Breast | Heart | Hepatitis | Average |
|---|---|---|---|---|---|---|---|---|
| 1. | AMBC | **77.30** | 59.59 | 70.18 | 97.64 | 83.33 | **84.86** | 78.82 |
| 2. | AdaBoostM1 | 73.52 | 68.98 | 74.08 | 95.16 | 82.96 | 62.58 | 76.21 |
| 3. | Bagging | 73.20 | 72.17 | 75.13 | 95.75 | 81.11 | 64.51 | 76.98 |
| 4. | BayesNet | 72.22 | 58.55 | 74.73 | 97.36 | 82.96 | 69.03 | 75.81 |
| 5. | Dagging | 73.85 | 58.26 | 73.95 | 96.63 | 83.33 | 65.16 | 75.20 |
| 6. | DecisionTable | 72.22 | 58.55 | 74.86 | 93.99 | 80.37 | 69.67 | 74.94 |
| 7. | DTNB | 72.22 | 58.55 | 75.52 | 97.07 | 82.59 | 65.80 | 75.29 |
| 8. | FT | 73.20 | 69.85 | 77.21 | 96.92 | 82.22 | 57.41 | 76.14 |
| 9. | LMT | 73.85 | 71.59 | 76.69 | 96.63 | **84.44** | 63.87 | 77.85 |
| 10. | Logistic | 74.18 | 66.37 | 76.69 | 96.48 | 83.33 | 61.29 | 76.39 |
| 11. | MultiClassClassifier | 74.18 | 66.37 | 76.69 | 96.48 | 83.33 | 61.29 | 76.39 |
| 12. | NaiveBayes | 75.49 | 56.23 | 75.78 | 96.33 | 83.33 | 72.25 | 76.57 |
| 13. | NaiveBayesSimple | 75.49 | 55.07 | 75.78 | 96.19 | 84.07 | 70.32 | 76.15 |
| 14. | NveBayesUpdateable | 75.49 | 56.23 | 75.78 | 96.33 | 83.33 | 72.25 | 76.57 |
| 15. | RandomCommittee | 64.70 | 68.40 | 73.30 | 95.60 | 80.37 | 62.58 | 74.16 |
| 16. | RandomForest | 67.64 | 66.37 | 73.56 | 96.33 | 78.88 | 60.64 | 73.90 |
| 17. | RandomSubSpace | 74.50 | 67.53 | 73.43 | 95.90 | 78.14 | 61.93 | 75.24 |
| 18. | RBFNetwork | 73.52 | 61.73 | 74.34 | 96.33 | 82.59 | 71.61 | 76.69 |
| 19. | RotationForest | 74.18 | 70.72 | 75.39 | 97.36 | 83.70 | 63.22 | 77.43 |
| 20. | SimpleLogistic | 73.85 | 67.82 | 76.69 | 96.63 | **84.44** | 65.16 | 77.43 |
| 21. | SMO | 73.52 | 57.68 | 77.47 | 96.48 | 84.07 | 64.51 | 75.62 |
| 22. | GRE-FAM | 70.65 | **75.96** | **79.65** | **98.10** | 78.89 | 80.85 | **80.68** |

recursivamente uma sequência de conjuntos *fuzzy* obtidos a partir de uma combinação afim das memórias fundamentais.

Neste trabalho, motivados pelos modelos de representação esparsa, descrevemos como uma GRE-FAM pode ser aplicada a problemas de classificação de padrões. Apresentamos também os resultados dos experimentos computacionais obtidos em seis problemas de classificação disponíveis em [6]. Tais resultados mostraram que a acurácia obtida pelo classificador baseado nas GRE-FAMs é competitiva se comparada a de outros classificadores da literatura [1]. De fato, o classificador baseado nas GRE-FAMs obteve a maior acurácia em três dos seis problemas considerados, bem como a maior acurácia média.

## Agradecimentos

Agradecemos o apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), processo 2015/00745-1, e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), processo 305486/2014-4.

 

## Referências

[1] M. Aldape-Pérez, C. Yáñez-Márquez, O. Camacho-Nieto, and A. J. Argüelles-Cruz. An associative memory approach to medical decision support systems. *Computer methods and programs in biomedicine*, 106(3):287–307, 2012. DOI: 10.1016/j.cmpb.2011.05.002.

[2] J. Austin. Associative Memory. In E. Fiesler and R. Beale, editors, *Handbook of Neural Computation*, pages F1.4:1–F1.4:7. Oxford University Press, 1997.

[3] G. Golub and C. van Loan. *Matrix Computations*. John Hopkins University Press, Baltimore, MD, 3th edition, 1996.

[4] M. H. Hassoun and P. B. Watta. Associative Memory Networks. In E. Fiesler and R. Beale, editors, *Handbook of Neural Computation*, pages C1.3:1–C1.3:14. Oxford University Press, 1997.

[5] B. Kosko. *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Prentice Hall, Englewood Cliffs, NJ, 1992.

[6] M. Lichman. UCI machine learning repository, 2013.

[7] A. C. Souza, M. E. Valle, and P. Sussner. Generalized Recurrent Exponential Fuzzy Associative Memories Based on Similarity Measures. In *Proceedings of the 16th World Congress of the International Fuzzy Systems Association (IFSA) and the 9th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT)*, volume 1, pages 455–462. Atlantis Press, 2015. DOI: 10.2991/ifsa-eusflat-15.2015.66.

[8] P. Sussner and M. E. Valle. Fuzzy Associative Memories and Their Relationship to Mathematical Morphology. In W. Pedrycz, A. Skowron, and V. Kreinovich, editors, *Handbook of Granular Computing*, chapter 33, pages 733–754. John Wiley and Sons, Inc., New York, 2008.

[9] M. E. Valle and A. C. Souza. Pattern classification using generalized recurrent exponential fuzzy associative memories. A ser publicado.

[10] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(2):210–227, Feb 2009. DOI: 10.1109/TPAMI.2008.79.

[11] L. Xuecheng. Entropy, distance measure and similarity measure of fuzzy sets and their relations. *Fuzzy Sets and Systems*, 52(3):305–318, 1992. DOI: 10.1016/0165-0114(92)90239-Z.

[12] R. Zwick, E. Caristein, and D. V. Budescu. Measures of Similarity Among Fuzzy Concepts: A Comparative Analysis. *International Journal of Approximate Reasoning*, 1(2):221–242, apr 1987. DOI: 10.1016/0888-613X(87)90015-6.

 