

# CLASSIFICAÇÃO FUZZY DE PADRÕES NÃO-MOTORES E INDICAÇÃO DA SEVERIDADE DA DOENÇA DE PARKINSON

Thiago Ribeiro, Maria Borges, Rennan Cardoso, Raquel Coelho, Sílvia Costa, Daniel Leite

*Departamento de Engenharia, Universidade Federal de Lavras (UFLA), Lavras, MG, Brasil*

**E-mails:** thiagoth4@hotmail.com, tailani2009@hotmail.com, rennanalves@yahoo.com.br, raquelrc@engautomacao.ufla.br, silvia.ferreira@deg.ufla.br, daniel.leite@deg.ufla.br

**Abstract**— Parkinson's disease is an age-related neurodegenerative disease. About 1% of individuals over the age of 65 develop the disease. Recent research on incipient Parkinson's disease detection has indicated subtle changes in voice, hyposmia and sleep disorders as the first indicators of the disease. This work considers analyses of speech amplitudes in certain frequencies and computational intelligence algorithms for incipient detection of non-motor patterns of the Parkinson's disease. The large number of data and variables involved and the uncertainty about exact values make expert analyses difficult and imprecise. Clustering algorithms, viz. Fuzzy C-Means and Gustafson-Kessel, were implemented to analyse attributes extracted from a database provided by the University of Oxford. The algorithms have presented results regarding the inference of the severity of the Parkinson's disease for each individual considering the UPDRS (Unified Parkinson's Disease Rating Scale). Particularly, the Gustafson-Kessel algorithm has provided the best results in terms of correct classifications according to severity levels.

**Keywords**— Parkinson's Disease, Computational Intelligence, Fuzzy Clustering, Fuzzy C-Means, Gustafson-Kessel

**Resumo**— A doença de Parkinson é uma doença neurodegenerativa relacionada a idade. Cerca de 1% dos indivíduos com idade superior a 65 anos desenvolvem a doença. Pesquisas recentes em detecção incipiente da doença de Parkinson têm apontado como os primeiros indicadores da doença alterações sutis na voz, hiposmia e distúrbios do sono. Este trabalho leva em consideração análises de amplitudes da fala em certas frequências e algoritmos de inteligência computacional para detecção incipiente de padrões não-motores da doença de Parkinson. O grande número de dados e variáveis envolvidas, e a incerteza sobre valores exatos, causam dificuldades e imprecisão à análise especializada. Algoritmos de agrupamento, viz. Fuzzy C-Means e Gustafson-Kessel, foram implementados para análise das características extraídas de uma base de dados disponibilizada pela Universidade de Oxford. Os algoritmos apresentam resultados a respeito da severidade da doença de Parkinson para cada um dos indivíduos considerando a escala UPDRS (Escala Unificada de Avaliação da Doença de Parkinson). Em particular, o algoritmo Gustafson-Kessel tem apresentado melhores resultados em termos de classificações corretas de acordo com níveis de severidade.

**Palavras-chave**— Doença de Parkinson, Inteligência Computacional, Agrupamento Fuzzy, Fuzzy C-Means, Gustafson-Kessel

## 1 Introdução

A Doença de Parkinson (PD) é uma doença causada pela degeneração e morte de neurônios produtores de dopamina. Neurônios dopaminérgicos regulam a transmissão de comandos conscientes provenientes do córtex para os músculos. A doença envolve ainda à disfunção de outras estruturas produtoras de serotonina, noradrenalina e acetilcolina (Cooper, 2008; Jankovic, 2008). Parkinson é, portanto, uma doença degenerativa do sistema nervoso central. Geralmente, a PD prevalece em pessoas com mais de 60 anos e é mais comum em homens do que em mulheres. Ela acomete cerca de 1% dos indivíduos acima de 65 anos de idade, embora possa se manifestar em idades entre 35 e 85 anos (Rustempasic & Can, 2013).

Os primeiros sinais da doença são pouco perceptíveis. Uma sensação de fraqueza ou rigidez em um membro, ou um fino tremor de uma das mãos quando em repouso pode ser um primeiro sinal da PD (Committee, 2002). Quando a doença se agrava, os indivíduos portadores podem apresentar tremores, rigidez muscular, dificuldade de caminhar e equilibrar, instabilidade postural, entre outras dificuldades motoras. Estas dificuldades motoras são atualmente as principais características levadas em consideração para o

diagnóstico clínico (Cooper et al., 2008; Jankovic, 2008). Os sintomas anteriores, i.e. não-motores, da PD incluem a deterioração da fluência da fala (voz monotônica ou gagueira), insônia, disfunção olfatória, distúrbio comportamental do sono REM (*Rapid Eye Movement*), entre outros (Olanow et al. 2011).

Embora não exista atualmente cura para a PD, existem tratamentos, como medicação e cirurgia para gerenciar seus sintomas. A detecção incipiente e monitoramento da PD podem: atrasar seu progresso; sugerir medicação individualizada; estimular o desenvolvimento de novos tratamentos; e melhorar a qualidade de vida dos pacientes.

Pesquisas recentes em detecção incipiente da PD têm apontado como os primeiros indicadores da doença alterações sutis na voz, hiposmia e distúrbios do sono (Olanow et al., 2011; Cooper et al., 2008; Jankovic, 2008; Rustempasic & Can, 2013; Tsanas et al., 2010; Tsanas et al., 2012; Leite et al., 2013). Estes surgem anos antes dos sintomas motores. Portanto, um dos métodos para o diagnóstico da doença é detectar e analisar sinais de voz usando ferramentas acústicas que registram as mudanças na pressão nos lábios ou dentro do trato vocal. Há muitos dados e variáveis envolvidas no problema de detecção incipiente da doença de Parkinson baseado em padrões não-motores. Além disso, a incerteza sobre valores exatos de certas medições causa dificuldades à análise especializada.

Max Little, do *Media Labs - Massachusetts Institute of Technology* (MIT), desenvolveu em 2009 um software para detectar diferenças nos padrões de voz afim de classificar indivíduos saudáveis e indivíduos portadores da PD. Em (Little, 2009) obteve-se um desempenho de classificação correto de 91,4% usando o método de análise de regressão LASSO (*Least Absolute Shrinkage and Selection Operator*). Além disso, a proposta de (Gil & Johnson, 2009) forneceu uma precisão de aproximadamente 90% no diagnóstico de PD usando ambos, uma Rede Neural Artificial do tipo feed-forward (RNA) e Máquina de Vector Suporte (SVM). Um classificador Neuro-Fuzzy Adaptativo (ANFIS) com *hedges* linguísticos proposto em (Caglar, 2010) gerou resultados de reconhecimento da PD com acurácia de 94,7% de desempenho para dados de teste.

No presente trabalho, algoritmos de agrupamento fuzzy, viz. Fuzzy C-Means e Gustafson-Kessel, foram implementados para reconhecimento de padrões em uma base de dados de sinais de voz disponibilizada pela *Parkinson's Voice Initiative* em [www.parkinsonsvoice.org](http://www.parkinsonsvoice.org), ver (Tsanas et al., 2010; Tsanas et al., 2012). A base de dados considera 42 indivíduos em estágio inicial de desenvolvimento da doença e aproximadamente 6 mil amostras de dados de fonações sustentadas da vogal 'A'. As variáveis envolvidas dizem respeito ao espectro do sinal no domínio da frequência. Os algoritmos de agrupamento são avaliados neste trabalho com respeito ao tempo de processamento e inferências corretas da severidade da PD na escala UPDRS (Escala Unificada de Avaliação da Doença de Parkinson) – uma escala internacionalmente aceita. Quanto maior o grau UPDRS, mais graves são as deficiências do paciente (Leite et al., 2013; Little et al., 2009).

## 2 Métodos de Agrupamento Fuzzy e Validação

Em campos que lidam com detecção de anomalias e diagnóstico, procura-se encontrar uma estrutura representativa nos dados obtidos a partir de observações de um processo ou fenômeno físico. Ao encontrar a estrutura, os dados são classificados de acordo com similaridades entre os valores dos atributos. Os parâmetros e a estrutura do modelo de classificação são definidos por um algoritmo de agrupamento não-supervisionado.

O processo de agrupamento decompõe uma coleção de amostras em subgrupos ou *clusters* com base em uma medida de similaridade ou distância. O objetivo é dividir o conjunto de dados de tal forma que os elementos pertencentes ao mesmo cluster sejam semelhantes entre si, e os elementos pertencentes a diferentes clusters sejam diferentes. Em agrupamento fuzzy, amostras de dados podem pertencer a mais de um cluster com um grau de pertinência diferente relacionado aos diferentes clusters. Neste trabalho, foi utilizado os algoritmos Fuzzy C-Means e Gustafson-Kessel para realizar o agrupamento dos dados de PD dis-

ponibilizados pela iniciativa da Universidade de Oxford intitulada *Parkinson's Voice Initiative*. Os *clusters* resultantes são rotulados com diferentes níveis de severidade da doença.

### 2.1 Algoritmo Fuzzy C-Means

O algoritmo de agrupamento Fuzzy C-Means (FCM) é, talvez, o algoritmo mais conhecido e usado na construção dos termos antecedentes de modelos de regras fuzzy (Bezdek, 1981; Rustempasic & Can, 2013). As distâncias entre amostras e centros de grupos são calculadas pelo algoritmo e usadas para atualizar centros e funções de pertinência, podendo uma amostra pertencer simultaneamente a diferentes grupos. Considera-se espaços Euclidianos para o cálculo das distâncias. Logo, os *clusters* resultantes são essencialmente esféricos.

Visto que algoritmo de agrupamento FCM é amplamente conhecido na literatura, abaixo apresentamos seu pseudocódigo de forma breve.

#### Procedimento Fuzzy C-Means

1. Considere um conjunto de  $n$  amostras de dados a serem agrupadas,  $x_i, i = 1, \dots, n$ ;
2. Assuma um número de *clusters* (classes),  $c$ , sendo  $2 \leq c < n$ ;
3. Escolha o fator de fuzzicidade  $m > 1$ . Em geral,  $m = 2$  é o valor padrão;
4. Inicialize a matriz de partição  $U$  de dimensão  $n \times c$  com valores aleatórios tal que  $u_{ik} \in [0, 1]$  e  $\sum u_{ik} = 1$ ;
5. Calcule o centro dos *clusters*  $v_j$  usando

$$v_{ij} = \frac{\sum_{k=1}^n u_{ik}^m x_{ki}}{\sum_{k=1}^n u_{ik}^m} \quad (1)$$

para  $j = 1, \dots, c$ ;

6. Calcule a distância de uma amostra para todos os *clusters* a partir de

$$d_{ik} = d(x_k - v_i) = \sqrt{\sum_{j=1}^m (x_{kj} - v_{ij})^2} \quad (2)$$

$i = 1, \dots, c$  e amostras  $k = 1, \dots, n$ .

7. Atualize a matriz de partição  $U$  de acordo com as distâncias  $d_{ik}$ . Se  $d_{ik} > 0$ , então

$$u_{ik}^{(r+1)} = 1 / \sum_{j=1}^c \frac{d_{ik}^{2(r)}}{d_{jk}^{2(r)}} \quad (3)$$

8. Repetir os passos de 5 a 7 até que um critério de parada de treinamento seja alcançado. Adota-se como critério a variação mínima da matriz de partição, i.e.  $\Delta U < \delta$ , sendo  $\delta$  um limiar de tolerância ou convergência.

## 2.2 Algoritmo Gustafson-Kessel

O algoritmo de agrupamento Gustafson-Kessel (GK) é uma extensão do algoritmo FCM no sentido que ele inclui um mecanismo de adaptação de matriz de covariância capaz de induzir *clusters* elipsoidais, não-necessariamente alinhados aos eixos dos atributos e contendo diferentes distribuições espaciais para diferentes *clusters* (Babuska et al. 2002; Mota et al., 2017).

O aprendizado em GK segue passos semelhantes aos passos do procedimento FCM, exceto pelo fato de que a distância entre a  $k$ -ésima amostra de dados e o  $i$ -ésimo centro de *cluster* é calculada baseada em uma matriz adaptativa independente,  $A_i$ . A métrica de distância em questão é conhecida como distância de Mahalanobis. Ela é dada por

$$D_{kiA_i}^2 = (x_k - v_i)^T A_i (x_k - v_i) \quad (4)$$

Dado um conjunto de dados  $X$ , escolhe-se o número de clusters  $c$ , a constante de fuzziçidade  $m > 1$ , e o critério de parada ou convergência dos centros de *clusters*. Adotou-se novamente o valor padrão para  $m$ , i.e.  $m = 2$ . Para amostras  $l = 1, 2, \dots$ ,

$$v_i^{(l)} = \frac{\sum_{k=1}^N u_{ki}^{(l-1)m} x_k}{\sum_{k=1}^N (u_{ki}^{(l-1)})^m} \quad (5)$$

$i = 1, \dots, c$ .

A matriz de covariância associada a um *cluster* é obtida a partir de

$$A_i^{(l)} = p_i \det(F_i)^{-1} F_i^{-1} \quad (6)$$

onde

$$F_i = \frac{\sum_{k=1}^N u_{ki}^{(l-1)m} (x_k - v_i^{(l)})(x_k - v_i^{(l)})^T}{\sum_{k=1}^N (u_{ki}^{(l-1)})^m} \quad (7)$$

$i = 1, \dots, c$ .

Por fim, atualiza-se a matriz de partição  $U$  de acordo com (3) usando (4) e (6).

O critério de parada para o GK é a variação mínima da matriz de partição, ou seja,  $\Delta U < \delta$ , sendo  $\delta$  um limiar de tolerância ou convergência.

## 2.3 Validação dos Modelos Classificadores

Medidas de validação são geralmente usadas para determinar a quantidade ideal de *clusters* existentes em uma base de dados (Babuska et al., 2002; Pal & Bezdek, 1995; Mota et al., 2017). Neste trabalho adotou-se duas medidas de validação usuais, conforme descrito a seguir.

1. Critério de validação de Xie-Beni: o número ideal de *clusters* é determinado segundo o valor mínimo para o critério de Xie e Beni (XB) (Xie & Beni, 1991):

$$XB = \frac{\sum_{i=1}^c \sum_{k=1}^N u_{ki}^m \|x_k - v_i\|}{N \min_{v_{k,i}} \|x_k - v_i\|} \quad (8)$$

O numerador da Eq. (8) refere-se à minimização da distância Euclidiana entre as amostras e os centros de *clusters*,  $\|\cdot\|$  é a norma 2. O denominador refere-se à maximização da distância entre centros de *clusters*. Pequenos valores de  $XB$  são esperados para *clusters* compactos e bem separados. Os valores de  $XB$  dependem da constante de fuzziçidade;  $m = 2$  é o padrão adotado.

2. Coeficiente de Partição: o critério Coeficiente de Partição (PC) é definido segundo

$$PC = \frac{1}{N} \sum_{i=1}^c \sum_{k=1}^N u_{ki}^2 \quad (9)$$

O valor de PC mede a quantidade de sobreposição entre *clusters* (Bezdek, 1981; Mota et al., 2017). Os valores variam em  $[1/c, 1]$ . Quanto mais próximo do valor 1 for o coeficiente PC, mais “nítido” será o *cluster*. Um valor próximo de  $1/c$  indica que não há tendência da existência de qualquer agrupamento na base de dados ou que o algoritmo em questão não conseguiu revelá-los. O número ideal de *clusters* gera o maior valor do coeficiente PC.

## 3 Metodologia

### 3.1 Sobre a Base de Dados

O conjunto de dados considerado foi criado pela *Parkinson's Voice Initiative* em colaboração com 10 centros médicos nos Estados Unidos e a Intel Corporation, que desenvolveram um dispositivo de tele monitoramento para registrar sinais de fala de indivíduos em estágio incipiente de PD, vide (Tsanas et al., 2010; Tsanas et al., 2012) para detalhes. Os atributos das amostras de dados são listados na Tabela 1.

Os dados foram produzidos a partir de 5875 gravações de voz, fonações sustentadas da vogal ‘A’, durante um período de 3 segundos. Quarenta e duas pessoas portadoras da PD em estágio inicial participaram das gravações. Existem aproximadamente duzentas gravações por paciente. O teste foi realizado durante o período de 6 meses. Foi usado um dispositivo de tele monitoramento para gravações remotas da progressão dos sintomas. As gravações foram obtidas em ambiente doméstico, não ideal, para evitar idas contínuas à centros médicos. Um total de 16 atributos de entrada de modelos classificadores foi considerado, conforme mostrado na Tabela 1. A saída dos modelos classificadores é o UPDRS total, sendo o UPDRS motor desconsiderado neste trabalho. Além disso, os 4 primeiros atributos da Tabela 1, ‘*subject*’, ‘*gender*’, ‘*age*’ e ‘*test\_time*’ foram desconsiderados para fim de

adaptação de parâmetros dos modelos classificadores das gravações de voz.

Tabela 1: Descrição dos Atributos da Base de Dados

Nome de Atributo	Descrição
Subject	Valor inteiro que identifica cada indivíduo
Age	Idade do indivíduo
Gender	Gênero – ‘0’ - Masculino, ‘1’ - Feminino
Test_time	Tempo desde o recrutamento para o julgamento
Motor_UPDRS	Pontuação da UPDRS motora linearmente interpolada
Total_UPDRS	Pontuação total da UPDRS linearmente interpolada
Jitter(%), Jitter(Abs), Jitter:RAP, Jitter:PPQ5, Jitter:DDP	São 5 características de medidas de variação da frequência fundamental
Shimmer, Shimmer(dB), Shimmer:APQ3, Shimmer:APQ5, Shimmer:APQ11, Shimmer:DDA	São 6 características de medidas de variação da amplitude da fonação sustentada em certas frequências
NHR, HNR	São duas medidas de proporção de ruído para componentes tonais da voz
RPDE	Medida da complexidade dinâmica não linear
DFA	Expoente de escala fractal do sinal de fala
EPI	Medida não linear da variação da frequência fundamental

Segundo (Tsanas et al., 2010; Little et al., 2009; Tsanas et al., 2012), entre os atributos apresentados na Tabela 1, uma pessoa portadora de PD apresenta valores baixos para HNR. Já para os demais atributos, i.e. Jitter(%), Jitter(Abs), Jitter:RAP, Jitter:PPQ5, Jitter:DDP, Shimmer, Shimmer(dB), Shimmer:APQ3, Shimmer:APQ5, Shimmer:APQ11, Shimmer:DDA, NH, RPDE, DFA, EPI, valores mais altos representam um estágio mais avançado da PD.

### 3.2. Considerações Experimentais

Os algoritmos de agrupamento fuzzy, FCM e GK, foram usados para encontrar similaridades entre as amostras de dados. O número de *clusters* no conjunto de dados é um parâmetro crucial. No caso dos algoritmos em questão, o número de *clusters* deve ser fornecido como parâmetro inicial. Diferentes quantidades de clusters foram avaliadas e as medidas de validação XB e PC apontaram a quantidade de *clusters* existentes na base de dados segundo o formato geométrico permitido por suas medidas de distância. Em outras palavras, FCM produz esferas no espaço de 16 dimensões dos atributos (distância Euclidiana) enquanto GK produz elipsoides (distância de Mahalanobis). As melhores partições são indicadas pelo menor valor de XB e pelo maior valor de PC.

Uma parcela de 60% das amostras da base de dados foi utilizada para realizar o treinamento (ajuste de parâmetros) dos modelos classificadores. Os 40%

dos dados restantes foram úteis para fim de teste e validação dos modelos. Os dados obtidos foram normalizados em  $[0, 1]$  para considerar os atributos na mesma proporção. A normalização do  $i$ -ésimo elemento do  $k$ -ésimo vetor de dados é dada por:

$$x_{i(k)}^r = \frac{x_{i(k)} - \min_j x_{i(j)}}{\max_j x_{i(j)} - \min_j x_{i(j)}} \quad (10)$$

onde  $j = 1, \dots, k - 1$ , e  $x_{i(k)}^r \in [0, 1]$ .

Os dados de saída, relativos ao UPDRS total dos indivíduos, foram divididos em níveis de severidade da PD, linguísticos, de acordo com a quantidade de clusters do modelo classificador em questão.

Os resultados obtidos no agrupamento foram comparados com os valores de UPDRS originais para encontrar a acurácia dos algoritmos. O tempo de processamento para as diferentes configurações de classificadores também foi computado.

A acurácia dos classificadores é calculada como segue. Inicialmente, erros,  $E$ , e acertos,  $A$ , das classes estimadas são definidos nulos. Se a classe estimada para a amostra  $x_k$  é errada, i.e., se  $C_k - \bar{C}_k \neq 0$ , então  $E = E + 1$ . Ao contrário,  $A = A + 1$ . Logo,

$$Acc (\%) = \frac{A}{N} \cdot 100\% \quad (11)$$

onde  $N = A + E$  é o número total de amostras em avaliação. Para acurácia do treinamento dos modelos,  $N = 3525$  (60%); e para acurácia de teste  $N = 2350$  (40%).

## 4 Resultados

Primeiro, os índices de validação XB e PC foram utilizados para avaliar a quantidade ideal de *clusters* a serem utilizados no processo de agrupamento para cada algoritmo. Variou-se o número de *clusters* entre 2 e 10. Valores mínimos de XB e máximo de PC sugerem a melhor partição dos dados. As Tabelas 2 e 3 apresentam os valores obtidos pelos algoritmos FCM e GK, respectivamente.

Tabela 2: Validação do classificador FCM segundo os índices Xie-Beni e Coeficiente de Partição

Clusters	XB	PC	Tempo(s)
2	67.1796	<b>0.8212</b>	0.8625
3	26.8884	0.7658	<b>0.4193</b>
4	32.5769	0.7210	0.5581
5	54.5680	0.7014	0.7501
6	<b>16.5568</b>	0.6851	0.6531
7	21.4429	0.61224	0.6124
8	25.8061	0.6363	0.9509
9	19.3115	0.6240	0.9911
10	19.3679	0.6088	1.0337

Tabela 3: Validação do classificador GK segundo os índices Xie-Beni e Coeficiente de Partição

Clusters	XB	PC	Tempo(s)
2	58.3132	<b>0.8882</b>	1.1490
3	35.4573	07319	0.7835
4	33.4759	0.6691	<b>0.7460</b>
5	39.6960	0.6281	0.9473
6	35.5639	0.6154	1.1802
7	21.6335	0.5703	2.4046
8	20.6347	0.5748	1.6224
9	<b>15.2495</b>	0.5277	1.4212
10	37.1441	0.5151	1.9378

Em geral, verifica-se que conforme o número de *clusters* aumenta, os valores de XB e PC tendem a reduzir. Porém, não se trata de uma redução monótona; percebe-se mínimos locais de XB. Os valores em negrito nas Tabela 2 e 3 indicam a melhor partição FCM e GK, respectivamente, segundo um índice de validação. De acordo com o índice XB, o classificador FCM usando 6 *clusters* é mais compacto e seus *clusters* mais separados entre si. Já o classificador GK usando 9 *clusters* é indicado como o mais interessante segundo XB. De acordo com o índice PC, o número ideal de *clusters* para os dois modelos classificadores é 2. Conclui-se a partir da análise de PC que não há uma tendência clara de formação de *clusters* já que os dados se encontram bastante sobrepostos em uma região do espaço. Dessa forma, baseado apenas na análise de PC, a abordagem fuzzy para o problema parece interessante no sentido da indicação da pertinência de uma amostra em diferentes *clusters* e da adoção de rótulos linguísticos, mais abstratos que os dados. A partição fuzzy evita fronteiras abruptas entre classes como àquelas geradas por abordagens de agrupamento *crisp*, e.g. algoritmo K-Means. De uma maneira geral, as indicações de modelos classificadores fornecidas pelo índice XB, isto é, FCM com 6 *clusters* e GK com 9 *clusters*, foram admitidas. A velocidade de processamento foi alta em todos os cenários. Os *clusters* convergiram em menos de 3 segundos para todas as configurações, ou seja a matriz de partição  $U$  convergiu.

As Figuras 1 e 2 exemplificam as fronteiras de decisão suaves fornecidas a partir do uso de *clusters* fuzzy com diferentes geometrias. O processo de agrupamento FCM e GK foi realizado no espaço original dos dados, i.e., no espaço contendo 16 dimensões/atributos. Logo, as figuras mostram apenas os atributos mais discriminativos de classes de severidade da PD, i.e.  $\bar{x}_{(1)}$ , Jitter(Abs); e  $\bar{x}_{(2)}$ , DFA. Estes atributos foram considerados os mais discriminativos no sentido de que são, em média, menos correlacionados com os demais atributos de entrada, e, ao mesmo tempo, mais correlacionados com o grau UPDRS segundo a correlação de Pearson. Esta análise é útil somente para fins de visualização gráfica dos *clusters* FCM e GK resultantes.

Ademais, com relação as Figuras 1 e 2, sabe-se que quanto mais próximo do ponto (0,0) está um *cluster* ou uma amostra, menor é a severidade da PD.

Ao contrário, quanto mais próximo do ponto (1,1), mais avançado é o estágio da PD de um indivíduo. Nota-se nas figuras que a maior parte dos dados se encontra mais próxima da posição (0,0). Portanto, constata-se, de fato, que a maior parte dos indivíduos analisados se encontram em estágio inicial de desenvolvimento da doença. Nota-se também que os *clusters* GK na Figura 2 são mais alongados para capturar a orientação dos dados quando comparados com os *clusters* FCM vistos na Figura 1.

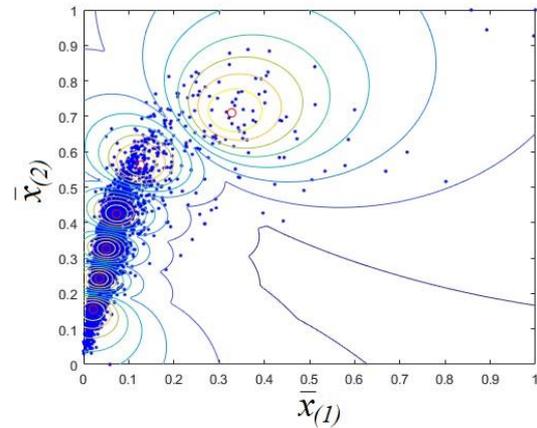


Figura 1: Fronteira de decisão fuzzy para 6 clusters FCM. Os pontos vermelhos são centros de clusters, e os azuis são amostras

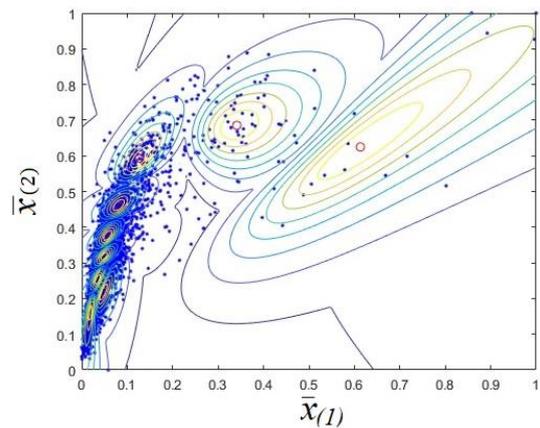


Figura 2: Fronteira de decisão fuzzy para 9 clusters GK. Os pontos vermelhos são centros de clusters, e os azuis são amostras

Ainda sobre as Figuras 1 e 2, nota-se uma ordenação dos centros de clusters aproximadamente ao longo de uma curva espiral ou quadrática partindo da origem em direção a valores mais próximos à coordenada (1,1). Esta ordenação sugere uma rotulação para os clusters e, por conseguinte, a mesma rotulação para as amostras com maior pertinência em um cluster. Para o modelo GK, os 9 clusters foram rotulados segundo a severidade da PD por: 'Mínima', 'Muito baixa', 'Baixa', 'Abaixo da média', 'Média', 'Acima da média', 'Alta', 'Muito alta', e 'Máxima'. Já para o modelo FCM, os 6 clusters receberam as mesmas ro-

tulações, porém, exclusas as classes ‘Mínima’, ‘Média’ e ‘Máxima’. A Tabela 4 sumariza a rotulação dos *clusters* FCM e GK adotadas neste trabalho.

Tabela 4: Níveis de severidade classificados pelos algoritmos FCM e GK

	Valor linguístico	Modelo FCM	Modelo GK
Seve- ridade	1 Mínima	-	x
	2 Muito baixa	x	x
	3 Baixa	x	x
	4 Abaixo da média	x	x
	5 Média	-	x
	6 Acima da média	x	x
	7 Alta	x	x
	8 Muito alta	x	x
	9 Máxima	-	x
UPDRS		6 rótulos	9 rótulos

Finalmente, as acurácias dos modelos foram avaliadas para amostras de treino e teste segundo a Eq. (11). Em treinamento, o modelo FCM obteve 87,33% em classificações corretas, enquanto GK obteve 89,41%. Em teste, os modelos FCM e GK obtiveram 83,89% e 85,44% de classificações pertinentes. Isto demonstra que ambos os algoritmos puderam capturar informações a partir dos dados e generalizar a aproximação do grau UPDRS em Parkinson para amostras de dados nunca vistas antes. Uma ligeira vantagem foi apontada ao modelo GK visto a maior flexibilidade geométrica de seus *clusters*.

Ressalta-se que os desempenhos dos métodos FCM e GK obtidos a partir da configuração experimental empregada neste trabalho são próximos aos desempenhos obtidos em (Little, 2009), de 91,4% usando LASSO; (Gil & Johnson, 2009), de 90% usando RNA e SVM; e (Caglar, 2010), de 94,7% a partir de uma rede ANFIS com *hedges* linguísticos. Não obstante, é importante notar que as condições das análises experimentais e formas de cálculo da acurácia em todos estes trabalhos não são exatamente as mesmas. Ademais, muito pode ser trabalhado em termos de pré-processamento de dados, filtragem, regulação e projeção dos dados originais em diferentes espaços para possíveis melhorias dos métodos FCM e GK estudados. Como vantagem dos métodos de agrupamento fuzzy tem-se a atribuição de rótulos linguisticamente interpretáveis às amostras e a indicação da pertinência da amostra à diferentes classes.

## 5 Conclusão e Trabalhos Futuros

O presente trabalho teve como objetivo a classificação de padrões incipientes, não-motores, da doença de Parkinson e a indicação de sua severidade

na escala UPDRS. Utilizou-se 16 atributos relacionados a amplitudes e frequências de sinais de voz, em particular de fonações sustentadas da vogal ‘A’. Foram desenvolvidos modelos classificadores a partir dos algoritmos de agrupamento Fuzzy C-Means e Gustafson-Kessel. Os índices de validação Xie-Beni e Coeficiente de Partição foram empregados para validação dos modelos. A base de dados foi disponibilizada pela *Parkinson’s Voice Initiative*.

O modelo GK apresentou melhores resultados em termos de classificações corretas do grau de severidade da PD para dados de teste (85,4%). Já o modelo FCM também apresentou um desempenho aceitável, porém inferior (83,9%). As acurácias de classificação foram obtidas a partir do uso de 9 e 6 *clusters* para GK e FCM, respectivamente, levando em consideração a validação de Xie e Beni. Os *clusters* foram rotulados linguisticamente de acordo com sua ordenação no espaço dos atributos mais discriminativos de classes. A abordagem fuzzy permitiu a verificação do grau de pertinência de cada amostra em cada um dos *clusters*, sendo esta a principal vantagem da abordagem frente a abordagens *crisp*, que provêm fronteiras de separação abruptas, e abordagens não-modulares, como modelos de redes neurais *feed-forward*.

Em trabalhos futuros serão analisados um número maior de atributos; atributos não somente extraídos de fonações sustentadas, mas também da análise de imagens do cérebro. Potenciais correlações entre o espectro de frequências da fala e imagens cerebrais serão levantadas como tentativa de melhorar o desempenho de modelos classificadores. Além disso, métodos de pré-processamento de dados, como métodos de projeção em eixos não-ortogonais serão estudados com o propósito de aumentar a discrepância espacial entre amostras de classes diferentes. Visto o aumento significativo da quantidade de dados, serão investigadas e desenvolvidas versões incrementais de algoritmos de agrupamento fuzzy, próprias para lidar com fluxos de dados online.

## Agradecimentos

Os autores agradecem a Universidade Federal de Lavras pela infraestrutura para a realização do trabalho e à FAPEMIG pelo apoio financeiro.

## Referências Bibliográficas

- Babuka, R.; Van der Veen, P. J.; Kaymak, U. (2002). Improved covariance estimation for Gustafson-Kessel clustering. *IEEE Int Conf on Fuzzy Systems*, vol. 2, pp. 1081-1085.
- Bezdek, J. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum Press.
- Caglar, M. F.; Cetisli, B.; Toprak, I. B. (2010). “Automatic recognition of Parkinson’s

- disease from sustained phonation tests using ANN and adaptive neuro-fuzzy classifier". *Journal of Eng Sci and Design*, vol. 1, no. 2, pp. 59-64.
- Committee, The Global Parkinson's Disease Survey Steering (2002). "Factors impacting on quality of life in Parkinson's disease: results from an international survey". *Mov Disord.*, vol. 17, issue 1, 60-67.
- Cooper, G.; Eichhorn G.; Rodnitzky, R. L. (2008). "Parkinson's Disease". In: *Neuroscience in Medicine*. Totowa, NJ: Humana Press. pp. 508-512.
- Gil, D.; Johnson, M. (2009). "Diagnosing Parkinson by using artificial neural networks and support vector machines". *Global Journal of Computer Science and Technology*, vol. 9, issue 4, pp. 63-71.
- Jankovic, J. (2008). "Parkinson's disease: clinical features and diagnosis". *Journal of Neurology, Neurosurgery, and Psychiatry*, vol. 79, issue 4, pp. 368-376.
- Leite, D.; Costa, P.; Gomide, F. (2013). "Evolving granular neural networks from fuzzy data streams". *Neural Networks*, vol. 38, pp. 1-16.
- Little, M. A.; McSharry, P. E.; Hunter, E. J.; Spielman, J.; Ramig, L. (2009). "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease". *IEEE Trans Biomed Eng*, vol. 56, issue 4, pp. 1015-1022.
- Mota, V. C.; Damasceno, F. A.; Soares, E. A.; Leite, D. F. (2017). "Fuzzy clustering methods applied to the evaluation of compost bedded pack barns". *IEEE Int Conf on Fuzzy Systems*, pp. 6p.
- Olanow, C. W.; Stocchi, F.; Lang, A. E. (Eds.) (2011) *Parkinson's Disease: Non-Motor and Non-Dopaminergic Features*. Wiley-Blackwell.
- Pal, N. R.; Bezdek, J. C. (1995). "On cluster validity for the Fuzzy C-Means model". *IEEE Trans Fuzzy Syst*, vol. 3, issue 3, pp. 370-379.
- Rustempasic, I.; Can, M. (2013). "Diagnosis of Parkinson's disease using fuzzy C-means clustering and pattern recognition". *SouthEast Europe Journal of Soft Computing*, Online available, pp. 42-49.
- Tsanas, A.; Little, M. A.; McSharry, P. E.; Ramig, L. O. (2010). "Accurate telemonitoring of Parkinson's disease progression by non-invasive speech tests". *IEEE Trans Biomed Eng*, vol. 57, no. 4, pp. 884-893.
- Tsanas, A.; Little, M. A.; McSharry, P. E.; Spielman, J.; Ramig, L. O. (2012). "Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease". *IEEE Trans Biomed Eng*, vol. 59, no. 5, pp. 1264-1271.
- Xie, X.; Beni, G. (1991). "A Validity measure for fuzzy clustering." *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 8, pp. 841-847.