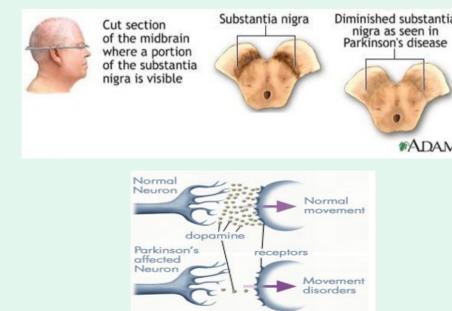


Thiago J. Ribeiro (UFLA – DEG) thiagoth4@hotmail.com; Daniel F. Leite (UFLA – DEG) daniel.leite@deg.ufla.br

A Doença de Parkinson

A doença de Parkinson (PD) é uma doença neurodegenerativa relacionada a idade. Cerca de 1% dos indivíduos com idade superior a 65 anos desenvolvem a doença. Indivíduos portadores da PD 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. Sintomas não-motores da doença incluem deterioração da fluência da fala (voz monotônica ou gagueira), insônia, disfunção olfatória, distúrbio comportamental do sono REM, entre outros. Sintomas não-motores surgem de forma sutil em um estágio incipiente, e.g. anos antes dos sintomas motores. A detecção incipiente e monitoramento da PD poderiam: atrasar seu progresso; sugerir medicação individualizada; estimular o desenvolvimento de novos tratamentos; e melhorar a qualidade de vida dos pacientes.



Classificação Fuzzy de Sinais de Voz

Há um grande número de 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 causam dificuldades à análise especialista. Algoritmos de agrupamento, viz. K-Means e Fuzzy C-Means, foram implementados neste trabalho para reconhecimento de padrões em uma base de dados de sinais de voz disponibilizada pela Universidade de Oxford. A base de dados considera 58 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 experimentos foram realizados em ambiente ideal, em laboratório. Os algoritmos de agrupamento adotados neste trabalho têm apresentado resultados que inferem a severidade da PD para cada um dos indivíduos considerando a escala UPDRS (*Unified Parkinson's Disease Rating Scale*) – uma escala mundialmente aceita.

Proposta

Agrupamento Fuzzy C-means

O algoritmo de agrupamento Fuzzy C-Means é, talvez, o algoritmo mais conhecido e usado na construção dos termos antecedentes de modelos de regras fuzzy. Trata-se de uma extensão fuzzy do também bem conhecido algoritmo K-Means. Ambos os algoritmos consideram espaços Euclidianos, i.e. espaços dotados da métrica L2. Distâncias entre amostras e centros de grupos são calculadas e usadas para atualizar centros e funções de pertinência. Amostras podem pertencer simultaneamente a diferentes grupos. Abaixo se encontra o pseudo-código do algoritmo Fuzzy C-Means adotado neste trabalho.

Procedimento Fuzzy C-means

Faça

1: Calcule os centros:

$$v_i = \frac{\sum_{k=1}^n \mu_{ik}^q x_k}{\sum_{k=1}^n \mu_{ik}^q}$$

2: Calcule as distâncias:

$$d_{ik} = [x_k - v_i]^T [x_k - v_i]$$

3: Atualize a matriz de partição:

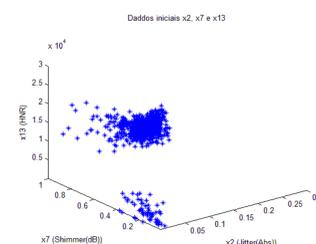
$$\mu_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{1}{q-1}}}$$

Até $\|\Delta U\| < \epsilon$

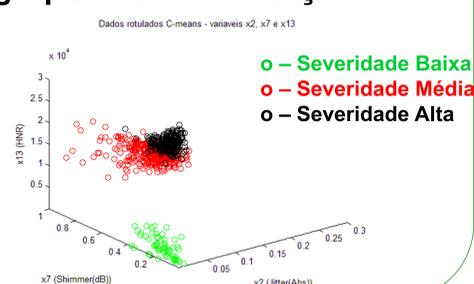
Resultados

Selecionou-se as variáveis mais importantes a partir de uma base de dados original com 132 variáveis. Os algoritmos de agrupamento consideraram as 16 variáveis menos correlacionadas entre si. As figuras abaixo mostram o espaço dos dados antes e após o agrupamento para as 3 variáveis mais importantes no processo discriminativo. São elas: Jitter(Abs), Shimmer(dB) e HNR. As classes foram rotuladas como: 'estágio 1, 2 e 3' de desenvolvimento da PD.

Dados iniciais



Agrupamento e rotulação



Análise Comparativa

Os algoritmos K-Means e Fuzzy C-Means foram processados e comparados em relação ao desempenho em classificações corretas para dados de treinamento / teste, e tempo de processamento, vide tabela abaixo. Observa-se que o classificador C-Means teve um melhor desempenho nas classificações das amostras. Porém, teve um tempo maior de processamento, devido o classificador C-Means permitir que amostras pertençam a grupos diferentes simultaneamente, permitindo diagnosticar a severidade (estágios) da PD através do fator de UPDRS.

Fuzzy C-means x K-means

	C-Means	K-Means
Tempo_treino (seg.)	2.98	2.84
Tempo_teste (seg.)	3.25	2.72
Acerto_treino (%)	82.76	71.43
Acerto_teste (%)	79.73	70.81

Conclusão

O algoritmo Fuzzy C-Means tem apresentado resultados interessantes em termos de classificações corretas de condições de pacientes de acordo com níveis de severidade da PD na escala UPDRS. O classificador Fuzzy C-Means, entretanto, envolve um maior número de parâmetros e é computacionalmente mais lento que o algoritmo K-Means. Trabalhos futuros considerarão a implementação do algoritmo de agrupamento fuzzy de Gustafson e Kessel, uma generalização do algoritmo Fuzzy C-Means onde grupos são desenvolvidos no espaço de Mahalanobis e podem assumir formas elípticas.

Referências Bibliográficas

- [1] Jankovic, J. "Parkinson's disease: clinical features and diagnosis". Journal of Neurology, Neurosurgery, and Psychiatry, vol. 79-4, pp. 368–376, 2008.
- [2] Leite, D.; Costa, P.; Gomide, F. "Evolving granular neural networks from fuzzy data streams." Neural Networks, vol. 38, pp. 1-16, 2013.
- [3] Little, M. A.; et al. "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease". IEEE Trans. on Biomedical Eng., vol. 56, issue 4, pp. 1015-1022, 2009.