

# USO DE REDES NEURAS ARTIFICIAIS (RNA) DO TIPO MULTILAYER PERCEPTRONS (MLP) MODIFICADO COM PROCESSAMENTO ESTATÍSTICO EM PARALELO PARA ESTUDO DO PROBLEMA DE CLASSIFICAÇÃO DA ORIGEM DE VINHO TINTO<sup>1</sup>

Mauro Almeida Lima<sup>1</sup>, Gilmar Gonçalves Ferreira<sup>2</sup>, Luana Loren Corrêa Oliveira<sup>3</sup>, Rochele Ferreira Silva Diniz<sup>4</sup>, Cecília Balduino Ferreira<sup>5</sup>

**RESUMO** - Este estudo foi realizado com a criação de um modelo de Rede Neural Artificial (RNA) do tipo Multilayer Perceptron (MLP), visando a análise de um problema de classificação, aliado à determinação estatística do Coeficiente de Variação dos resultados obtidos pela rede. Para este fim, no programa criado, nomeado como NeuroChem, utilizou-se um banco de dados contendo os resultados da análise físico-química de 178 amostras de vinho tinto, oriundos de três cultivares diferentes da Itália. Ao todo, 13 parâmetros físico-químicos foram normalizados e usados para a alimentação do algoritmo visando a determinação da origem (Classe) de cada uma das amostras através de técnicas de processamento em paralelo (Multithread) para análise estatística da saída de rede, como forma alternativa de minimização do erro de Treinamento/Validação. Ficou evidenciado pelos resultados que, para este banco de dados, a melhor arquitetura de rede MLP deve ser composta por três camadas ocultas e o processamento estatístico, feito com a programação em paralelo, representou uma redução de 73% do tempo de espera dos resultados, quando comparado à mesma análise comumente executada em série.

Palavras-chave: Processamento paralelo, Redes Neurais Artificiais, Vinho.

## ***USE OF ARTIFICIAL NEURAL NETWORK (ANN) - MULTILAYER PERCEPTRON (MLP) - MODIFIED WITH A STATISTICAL PARALLEL PROCESSING FOR THE STUDY OF CLASSIFICATION PROBLEM OF THE RED WINE ORIGIN***

**ABSTRACT** - This study was performed with the creation of an Artificial Neural Network (ANN) model in a Multilayer Perceptron class (MLP) in order to make the analysis of a classification problem, together with the statistical analysis of the of Coefficient of Variation results provided by the neural network. To this goal, the program named as Neurochem used a database containing physical-chemical analysis results of 178 samples of wines, from three different cultivars of Italy. These 13 physical-chemical parameters were standardized and used for feeding the computational algorithm in order to determine the origin (class) of the given samples using techniques of parallel processing (multithread) and the statistical analysis of the network output, as an alternative way to minimize the training/validation error. Evidenced by the results, the best MLP network architecture for this database must be composed of three hidden layers and the statistical processing, done with programming in parallel, confirms a 73% reduction of the waiting time for results when compared to the same analysis commonly performed in the serial way type.

Keywords: Artificial Neural Networks, Parallel Processing, Wine.

---

<sup>1</sup> Aluno de Graduação em Química, Universidade Federal de Viçosa – Campus de Rio Paranaíba (UFV), Rio Paranaíba MG, Email: mauro.lima@ufv.br

<sup>2</sup> Professor Química, UFV, Email: gilmargferreira@ufv.br

<sup>3</sup> Aluna de Graduação em Química, UFV, Email: cecilia.ferreira@ufv.br

<sup>4</sup> Aluna de Graduação em Química, UFV, Email: rochele.diniz@ufv.br

<sup>5</sup> Engenheira Ambiental, IFTM (Instituto Federal de Educação, Ciência e Tecnologia do Triângulo Mineiro), Email: luana@iftm.edu.br



### 1. INTRODUÇÃO

O vinho é o produto da fermentação alcoólica do mosto de uvas maduras. Atualmente, a Itália é um dos maiores exportadores mundiais de vinhos, contando com uma extensa diversidade de rótulos. Dentro deste cenário, há um aumento na exigência do consumidor por vinhos de melhor qualidade e, assim, pesquisas que levam a qualificação vinícola têm a sua importância no meio científico pois aprimoram-se as metodologias de identificação, qualificação e a elucidação das características físico-químicas que identificam as melhores safras produzidas facilitando, inclusive, a fixação dos preços no mercado (CORTEZ et al., 2009).

Neste cenário, o desenvolvimento e a criação de bancos de dados específicos para produtos agrícolas, contribui cada vez mais para o desenvolvimento de trabalhos científicos nesta área. O repositório on-line para aprendizado de máquina *UCI Machine Learning Repository* (LICHMAN, 2013) possui um acervo de dados experimentais de diversas áreas, inclusive relacionadas à *Origem* e a *Qualidade* de Vinhos produzidos mundialmente. Estes bancos de dados têm sido utilizados nos últimos anos por diversos autores cujos trabalhos têm objetivado principalmente o desenvolvimento de algoritmos para aprendizagem de máquina voltados para problemas numéricos e de classificação, utilizando técnicas diversas como por exemplo Redes Neurais Artificiais, Algoritmos Genéticos e técnicas de Clustering (FORINA et al., 1986; AEBERHARD et al., 1992; AEBERHARD et al., 1994; MUTTER et al., 2005; ZHONG et al., 2007; LIN, 2014; TANG et al., 2015).

Um banco de dados específico do UCI, sobre a *Origem* de vinhos, foi utilizado neste trabalho. Estes dados são o resultado de uma série de análises químicas de vinhos produzidos na mesma região, na Itália, mas derivados de três cultivares diferentes. As análises efetuadas pelos autores do banco de dados determinaram a presença de 13 constituintes diferentes encontrados nos três tipos de vinhos produzidos (FORINA et al., 1986). Mais informações sobre o banco de dados e uma lista completa, com todos os parâmetros analisados, está disponível na referência citada anteriormente.

Uma Rede Neural Artificial (RNA) faz parte de uma grande variedade de ferramentas da área de Inteligência Artificial (AI) e é constituída, em sua definição, por uma série de unidades básicas de processamento

interconectadas, os chamados *Neurônios*, também conhecidos como *Perceptrons*. RNAs do tipo Multilayer Perceptrons (MLP) são sistemas paralelos distribuídos, compostos por camadas de pequenas unidades *Perceptrons*, cada um contendo *funções de ativação* específicas e baseando-se geralmente em uma forma de aprendizado *supervisionado*, no qual o *erro* da análise deve ser minimizado constantemente via algoritmos do tipo *Backpropagation* (FERNANDES et al., 1996; TONSIG, 2000; GONÇALVES et al., 2015). Uma das possíveis configurações para uma rede MLP é mostrada na figura 1 em que é apresentada uma rede composta por quatro camadas, sendo uma camada de entrada, uma camada de saída e duas camadas *ocultas*.

Em uma RNA a camada inicial consiste na camada onde ocorre a *entrada* da informação, ou seja, onde os dados iniciais são inseridos; já a camada final é a camada de saída da *resposta* da rede, ou seja, onde o *erro de saída* é determinado. Em um processo de treinamento de uma RNA, também chamado de *Aprendizado*, as *conexões*, ou *pesos*, entre as unidades neuronais são modificadas afim de que a resposta final calculada pela rede seja semelhante às respostas desejadas. Após um processo de aprendizado, uma RNA pode correlacionar dados complexos de uma equação e prever seu resultado final corretamente pela minimização do erro entre a *saída* e o *valor desejado* no treinamento (SOUZA et al., 2004). A importância do uso de RNAs reside no seu poder de resolução de problemas do tipo *não linear*, como os problemas de classificação, por exemplo. Neste contexto, diversas

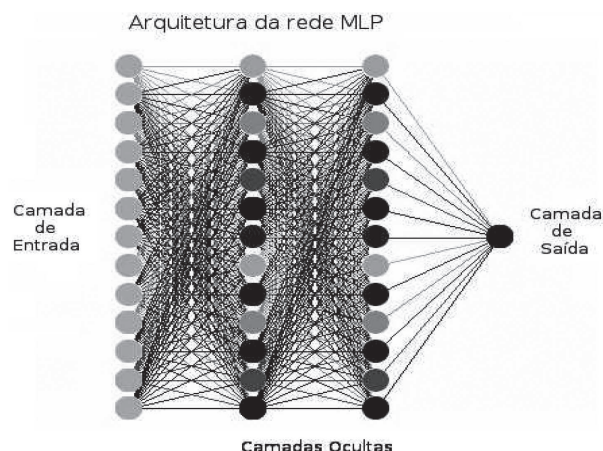


Figura 1 - Arquitetura de uma rede Multi-Layer Perceptrons (MLP) com duas Camadas Ocultas.

pesquisas no meio agrícola e agropecuário têm sido feitas nos últimos anos utilizando RNAs diversas, ou seja, métodos de inteligência artificial (SANGOI, 2014; SOARES et al., 2014; PEREIRA, 2013).

Com o avanço tecnológico, a busca por procedimentos computacionais que buscam o melhor desempenho e que, na medida do possível, apresentam custo relativamente baixo, levam ao emprego de métodos específicos para a obtenção de uma maior capacidade de processamento. Neste cenário, procedimentos que fornecem recursos de programação *paralela* se desenvolveram amplamente e ganharam espaço com o desenvolvimento das chamadas arquiteturas *Multicore* e dentro desta nova área, pesquisas referentes a processamento em *Paralelo* e processamento *Multithread* ganharam amplo desenvolvimento ao empregar técnicas que permite a execução simultânea de vários *threads*, ou *processos*, minimizando assim o tempo de espera dos resultados (DA COSTA, 2002; NAVAUX et al., 2011).

Este estudo foi criado um modelo de Rede Neural Artificial (RNA) do tipo Multilayer Perceptron (MLP), visando a análise de um problema de classificação, aliado à determinação estatística do Coeficiente de Variação dos resultados obtidos pela rede.

## 2. MATERIALE MÉTODOS

Neste trabalho, utilizou-se o banco de dados de Lichman (2013) para problemas de classificação, formado por 13 parâmetros químicos analisados em 178 amostras de vinho tinto, oriundas de três cultivares diferentes de coleta na Itália, para alimentação de uma rede neural artificial visando a determinação estatística do *Coeficiente de Variação* dos testes efetuados em três arquiteturas de rede distintas, durante o processo de *classificação* da região de origem das amostras de vinho.

Os testes de classificação da origem das amostras de vinhos, foram realizadas em um computador desktop com processador *Intel i7* de 3.6 GHz, dotado de 16 gigabytes de memória RAM, e sistema operacional *Linux Debian 8*. A programação foi efetuada em ambiente *Scilab* versão 5.5.1 e sua toolbox de Redes Neurais Artificiais, versão 0.4.2.5. O programa escrito neste ambiente foi intitulado *NeuroChem* e é voltado principalmente para as análises estatísticas dos resultados obtidos pela sua rede neural. A arquitetura

escolhida para a RNA durante os testes foi a MultiLayer perceptrons (MLP).

Os parâmetros de análise química, descritos na Tabela 1, foram utilizados para alimentar a camada de entrada de uma RNA, cujo processamento estatístico foi efetuado em *Série* e em *Paralelo* usando o banco de dados descrito anteriormente. As etapas de *Treinamento* e *Validação* foram efetuadas com 70% e 23% do banco de dados. Os 7% restantes foram utilizados para a fase final de *Teste*.

Os valores dos dados da camada de entrada, utilizados para a alimentação da rede, foram inicialmente normalizados para o intervalo [0,1] para as 178 amostras, segundo a Equação 1, descrita abaixo:

$$\text{Normalização} = \frac{(L_{\max} - L_{\min})(V - V_{\min})}{(V_{\max} - V_{\min})} + L_{\min} \quad (1)$$

Nesta equação  $L_{\max}$  e  $L_{\min}$  se referem, respectivamente, ao limite *máximo* e *mínimo* da normalização, ou seja, o intervalo fechado [0,1]. Já os valores  $V_{\max}$  e  $V_{\min}$  se referem aos valores *máximo* e *mínimo* do parâmetro de entrada da rede. Esta etapa de normalização é muito importante para se garantir a estabilidade numérica da rede neural em função dos grandes intervalos variáveis dos dados de entrada (BRANCO & SAMPAIO, 2008).

A *função de ativação* escolhida para todos os neurônios da rede foi a *função sigmoide* e o algoritmo de treinamento escolhido foi o de “retro propagação de erros” ou *Backpropagation*. Inicialmente, foram

Tabela 1 - Parâmetros químicos usados no problema de classificação da origem de Vinhos

1) Álcool
2) Ácido maleico
3) Cinza
4) Alcalinidade de cinzas
5) Magnésio
6) Fenóis totais
7) Flavonoides
8) Fenóis não flavonoides
9) Proantocianidinas
10) Cor
11) Matiz
12) OD280/OD315 de vinhos diluídos
13) Prolinas

Fontes: Forina, M. et al., (1986); Lichman, 2013.

definidas três arquiteturas de redes MLP específicas para os testes: A primeira, foi composta por uma camada de *Entrada*, composta de 13 neurônios, uma camada *oculta*, também com 13 neurônios, e uma camada de *saída*, composta por apenas 1 neurônio. Esta configuração pode ser representada pela sequência [13 13 1]. A sequência [13 13 13 1] e [13 13 13 13 1] se referem, respectivamente, as arquiteturas de 2 e 3 camadas ocultas.

O ambiente Linux Debian 8 ([www.debian.org](http://www.debian.org)) oferece uma particularidade muito importante neste contexto, o que não ocorre com facilidade em outros sistemas operacionais: O acesso do algoritmo *NeuroChem*, escrito em linguagem de programação de alto nível *Scilab*, é facilitado a todos os recursos (núcleos) de processamento da máquina. Isto é feito através de uma rotina específica com o comando interno *paralell\_run*. Assim, uma sequência de processos, ou *threads*, executados apenas em um núcleo de processamento em outros sistemas operacionais, podem ser distribuídos pelos núcleos virtuais do sistema, otimizando todo o processamento das informações, ou seja, através do chamado processamento *Multithread*.

O Erro Quadrático Médio (EQM) (Equação 2) foi medido a cada *época* de treinamento verificando a diferença entre as respostas encontradas pela RNA e os dados originais de saída. Sobre cada arquitetura testada foi feita uma análise do EQM, que é utilizado principalmente devido a sua boa resposta à pequenas variações entre os dados de entrada e saída de uma rede neural (MOREIRA et al., 2006). De forma semelhante, trabalhou-se ainda com o chamado *Erro Percentual Médio Absoluto* (EPMA) (Equação 3) que foi utilizado para medir a diferença entre as previsões feitas pela RNA e os dados originais, uma vez que o EPMA é mais sensível que o EQM, e seu resultado final é dado em percentagem (DE OLIVEIRA et al., 2010). O Coeficiente de Variação (Equação 4) é a razão entre o desvio padrão estimado e a média, avaliando a precisão da RNA em função do banco de dados; quanto menor for o seu valor menor é o erro associado ao conjunto calculado (GOMES, 1991).

$$EQM = \frac{1}{N} \sum_{i=1}^N (y_i - y_{pi})^2 \quad (2)$$

$$EQMA = \frac{1}{N} \sum_{i=1}^N \left| \frac{(y_i - y_{pi})}{y_i} \right| \quad (3)$$

$$EQMA = \frac{\sigma}{\bar{x}} \quad (4)$$

Nestas equações,  $y_i$  consiste no alvo do Treinamento/Validação,  $y_{pi}$  é o valor previsto pela rede neural,  $N$  é o número de previsões,  $\sigma$  é o desvio padrão previsto e  $\bar{x}$  é a média obtida nos testes.

### 3. RESULTADOS E DISCUSSÃO

A eficiência do treinamento de uma Rede MLP geralmente é calculada pela medida do EQM durante as etapas de Treinamento e Validação. Um fator importante neste contexto se refere ao parâmetro *Taxa de Aprendizado* que define quão rápido vai ser o processo de ajuste dos pesos sinápticos entre os neurônios de todas as camadas da rede neural, durante o aprendizado via *backpropagation*.

Pela análise da Figura 2 pode ser visto que, utilizando uma taxa de aprendizado igual a 0,01, a rede apresentou uma significativa perda de *generalização*, por volta de 1500 épocas (ciclos) de treinamento. Isto significa que a rede não foi capaz de inferir com precisão adequada a resposta dada pela sua camada de saída, elevando o EQM em dado ponto do aprendizado. Isto é conhecido como *overtraining* e indica que a rede foi treinada em excesso e/ou perdeu a capacidade de generalização durante a atualização dos pesos sinápticos, para o banco de dados usado.

Aumentando-se a taxa de aprendizado para 0,1, observa-se pela Figura 3, que a rede consegue, até o limite estimado de 5.000 épocas, manter a sua capacidade de generalização, o que reflete em um EQM cada vez menor para as etapas de Treinamento e Validação.

Após a definição da melhor taxa de aprendizado a ser usada, foram efetuados testes específicos para a análise estatística em *Série* e em *Paralelo*. Na Tabela 2, são mostrados os valores para o EPMA e EQM a etapa final de testes com a rede treinada. Foram efetuados 20 testes para cada amostra de vinho desconhecida (13 amostras) com a análise estatística dos seus resultados após 5.000 épocas de treinamento. As redes montadas com três camadas ocultas, ou seja arquitetura [13 13 13 13 1] foram as que obtiveram os menores valores de EPMA e EQM para a classificação da origem das amostras aleatórias de vinho, na fase final de Teste. Pode-se perceber ainda, pelos resultados mostrados na Tabela 2, a grande eficiência do



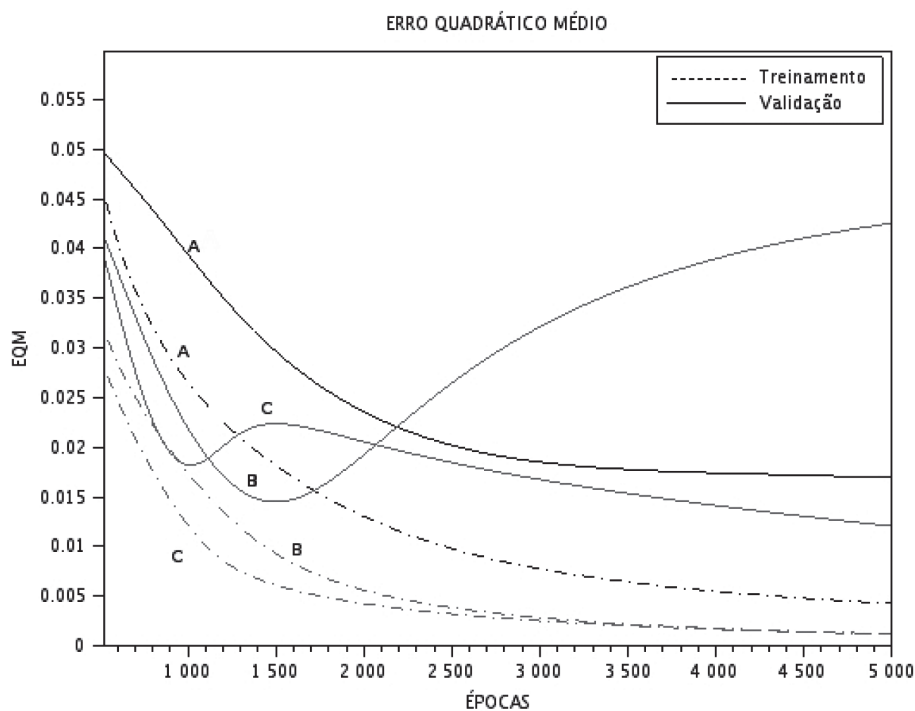


Figura 2 - Perda de Generalização por volta de 1500 épocas durante Treinamento e Validação, com taxa de aprendizado de 0.01. Arquiteturas com uma (A), duas (B) e Três (C) Camadas ocultas.

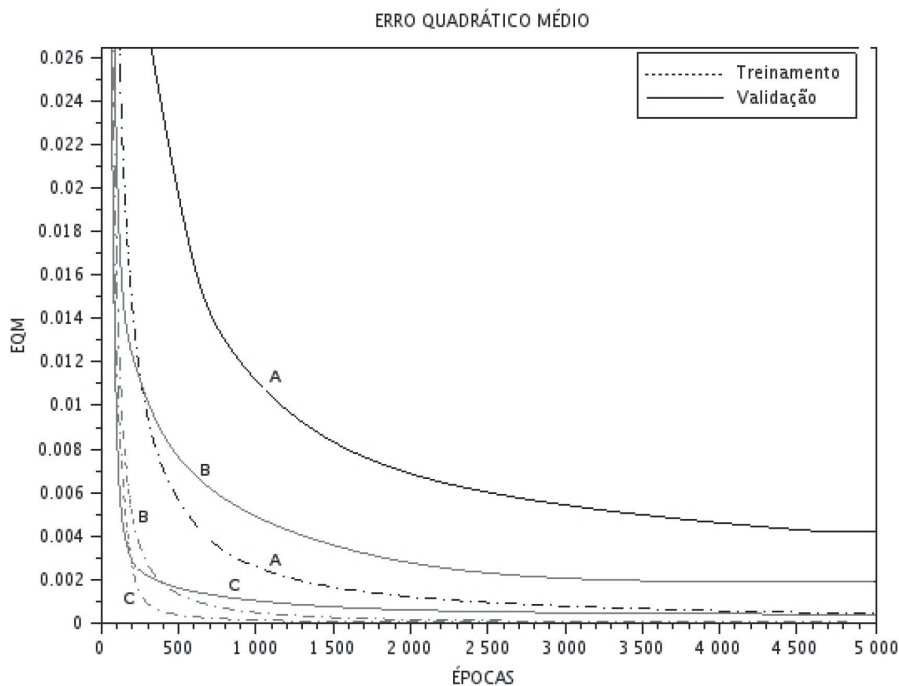


Figura 3 - EQM para a fase de Treinamento e Validação usando-se arquiteturas de rede MLP com uma (A), duas (B) e Três (C) Camadas ocultas. A taxa de aprendizado usada foi de 0,1.

tratamento estatístico em Paralelo com relação ao tempo de processamento computacional: a etapa de teste em paralelo gastou, em média, cerca de 24% do tempo de processamento total gasto pela análise estatística em Série, o que representa um ganho importante quando se trata de análises em bancos de dados muito volumosos.

A Figura 4 retrata o cenário observado durante as simulações com relação ao uso dos núcleos de processamento de máquina no sistema Linux Debian. Observa-se pela figura que, para os testes em Série, apenas 100% de um núcleo de processamento foi utilizado por vez, mas durante a fase de testes em Paralelo, ocorreu uma distribuição de carga (de processos) entre todos

os núcleos de processamento. Este processo resultou em um pequeno aquecimento do sistema, o que já era esperado.

O Coeficiente de Variação (CV) é uma importante ferramenta estatística e relaciona o *Valor Médio* da saída dos 20 testes efetuados para cada amostra de vinho com o seu respectivo *Desvio Padrão*. Durante a etapa de testes para 13 amostras desconhecidas de vinho pôde ser observado, pela análise da Figura 5, que a arquitetura MLP com três camadas ocultas foi a que apresentou o menor CV indicando assim que, para este tipo de problema de classificação, esta é a melhor configuração a ser usada para este banco de dados específico.

Tabela 2 - Medida de EPMA, EQM e Tempo de processamento para análise estatística em Série e Paralela para 20 testes/amostra, em função do número de camadas ocultas da rede

Tipo de Análise	Nº de Camadas Ocultas	EPMA	EQM	Tempo (Min)
Série	1	0,0267860	0,0102473	58,28
	2	0,0299030	0,0047721	66,63
	3	0,0096288	0,0022111	74,57
Paralela	1	0,0279311	0,0101712	14,04
	2	0,0180413	0,0036372	16,08
	3	0,0082031	0,0029007	17,85

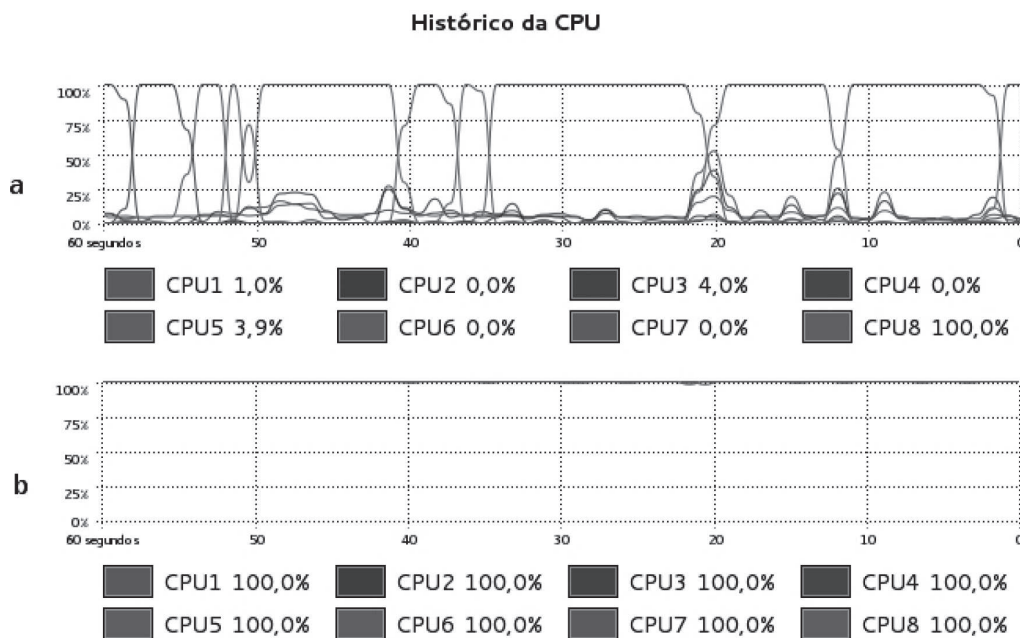


Figura 4 - Registro do monitor do sistema Linux Debian para análise estatística com processamento em Série (a) e em Paralelo (b) para o banco de dados de origem de vinhos.



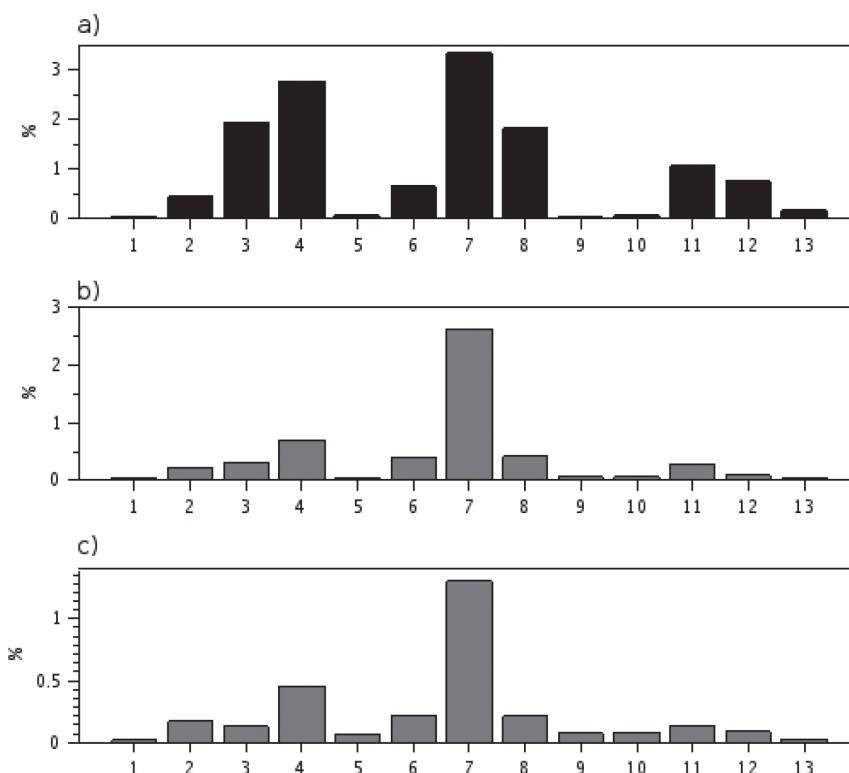


FIGURA 5 - CV da fase de Teste em paralelo, obtido em função das arquiteturas MLP com (a) 1 camada oculta, (b) 2 camadas ocultas e (c) 3 Camadas ocultas para classificação da origem de 13 amostras aleatórias de vinhos.

#### 4. CONCLUSÕES

Neste trabalho, foi proposto um algoritmo computacional chamado *NeuroChem*, para análise estatística em paralelo através do uso de uma rede neural artificial do tipo MLP alimentada com um banco de dados referente para determinação da origem de vinhos. As etapas de treinamento e validação da rede MLP se mostraram eficientes para a taxa de aprendizado igual a 0.1. A arquitetura da rede com três camadas ocultas [13 13 13 13 1] apresentou o menor EQM e o menor coeficiente de variação para as amostras de vinhos, durante a fase final de teste. Ficou evidenciado pelos resultados que o processamento estatístico feito com a programação paralela descrita representou uma redução de 73% do tempo de espera, quando comparado à mesma análise em série. Trabalhos futuros visarão a determinação ótima do número de neurônios por camada na arquitetura da rede MLP e também a otimização do algoritmo *Neurochem* para englobar novas metodologias e protocolos de paralelismo de máquina, como a *Message Passing Interface* (MPI),

para trabalho com bancos de dados extremamente volumosos.

#### 5. LITERATURA CITADA

AEBERHARD, Stefan; COOMANS, Danny; DE VEL, Olivier. The classification performance of RDA. Dept. of Computer Science and Dept. of Mathematics and Statistics, James Cook University of North Queensland, Tech. Rep, p. 92-01, 1992.

AEBERHARD, Stefan; COOMANS, Danny; DE VEL, Olivier. The performance of statistical pattern recognition methods in high dimensional settings. In: Proc. IEEE Signal Process. Workshop Higher Order Statist. 1994. p. 14-16.

BRANCO; Tadeu Sacha; SAMPAIO; Raimundo José Borges de Sampaio; Aplicação de redes neurais artificiais em modelos de previsão de demanda para equipamentos de infraestrutura de telecomunicações. in XXVIII Encontro Nacional de Engenharia de Produção; Rio de Janeiro, RJ, Brasil, 13 a 16 de outubro de 2008.

- CORTEZ, Paulo; CERDEIRA, Antônio; ALMEIDA, Fernando et al. Modeling wine preferences by data mining from physicochemical properties. *Decision Support Systems*, v. 47, n. 4, p. 547-553, 2009.
- DA COSTA, Celso Maciel; STRINGHINI, Denise; CAVALHEIRO, Gerson Geraldo Homrich. Programa Concorrente: Threads, MPI e PVM. *Escola Regional de Alto Desempenho, II ERAD*, v. 2, 2002.
- DE OLIVEIRA, Anderson C.S.; DE SOUZA, Ademária A.; LACERDA, Wilian S. et al. Aplicação de redes neurais artificiais na previsão da produção de álcool. *Ciênc. Agrotec.*, Lavras, v. 34, n. 2, p. 279-284, mar./abr., 2010
- GOMES, Frederico Pimentel. O índice de variação, um substituto vantajoso do coeficiente de variação. Piracicaba: Ipef, 1991, 4p. (Circular Técnica, 178)
- FERNANDES, Luiz Gustavo L.; PORTUGAL, Marcelo S.; NAVAUX, Philippe O.A. Previsão de séries de tempo: redes neurais artificiais e modelos estruturais. *Pesquisa e Planejamento Econômico*, v. 26, n. 2, p. 253-276, 1996.
- FORINA, M., ARMANINO, C., CASTINO, M., UBIGLI, M. Multivariate data analysis as a discriminating method of the origin of wines. *Vitis*, 25(3), 189-201, 1986.
- GONÇALVES, Luciene R. Modelagem de séries representativas do setor energético brasileiro. Dissertação (Mestrado em Estatística e Experimentação Agropecuária). Lavras, MG: UFLA, 2007. 106 p
- LICHMAN, M. (2013). UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science. Acesso em 12/02/2015 às 15:00h.
- LIN, Kuo-Ping. A novel evolutionary kernel intuitionistic fuzzy-means clustering algorithm. *IEEE Transactions on Fuzzy Systems Fuzzy Systems*, v. 22, n. 5, p.1074-1087, 2014.
- MOREIRA, Michel C.; CECÍLIO, Roberto A.; PINTO, Francisco A.C. et al. Desenvolvimento e análise de uma rede neural artificial para estimativa da erosividade da chuva para o Estado de São Paulo. *Revista Brasileira de Ciência do Solo*, v. 30, n. 6, p.1069-1076, 2006.
- MUTTER, Stefan; HALL, Mark; FRANK, Eibe. Using classification to evaluate the output of confidence-based association rule mining. In: *AI 2004: Advances in Artificial Intelligence*. Springer Berlin Heidelberg, 2005. p. 538-549.
- NAVAUX, Philippe O. A.; DE ROSE, César A.F.; PILLA, Laércio L. Fundamentos das Arquiteturas para Processamento Paralelo e Distribuído. *XI Escola Regional de Alto Desempenho do Estado do Rio Grande do Sul-2011-Porto Alegre, RS*, p. 22-59, 2011.
- PEREIRA, Lilianny Aparecida dos Anjos. Classificação automática de áreas cafeeiras em imagens de satélite, utilizando Redes Neurais Artificiais (Monografia de Graduação). Universidade Federal de Lavras. 2013
- SANGOI, Luiz Fernando. Análises do sistema produtivo suinícola da região central do rio grande do sul (Tese de Doutorado). Universidade Federal do Rio Grande do sul. 2014.
- SOARES, Fátima C.; ROBAINA, Adroaldo D.; PEITER, Márcia X. et al. Redes neurais artificiais na estimativa da retenção de água do solo. *Ciência Rural*, v. 44, n. 2, p. 293-300, 2014.
- SOUZA, M. B., TOMIKAWA, V. O., OLIVEIRA, B. F., POLATI, M. Uso da rede neural artificial no planejamento cirúrgico da correção do estrabismo. *Arq Bras Oftalmol*, 67(5), 459-62, 2004.
- TANG, B., HE, H., DING, Q.; KAY, S. A parametric classification rule based on the exponentially embedded family. *Neural Networks and Learning Systems, IEEE Transactions on*, 26(2), 367-377, 2015.
- TONSIG, Sérgio Luiz. *Redes Neurais Artificiais Multicamadas e o Algoritmo Backpropagation*. 2000.
- ZHONG, Ping; FUKUSHIMA, Masao. Regularized nonsmooth Newton method for multi-class support vector machines. *Optimisation Methods and Software*, v. 22, n. 1, p. 225-236, 2007.

Recebido para publicação em 30/04/2015 e aprovado em 23/06/2016.

