



Departamento de
Sistemas e
Computação

Universidade Federal de Campina Grande
Unidade Acadêmica de Sistemas e Computação
Programa de Pós-Graduação em Ciência da Computação
Disciplina: *Inteligência Artificial*
Profa.: *Joseana Macêdo Fachine Régis de Araújo*

Projeto 03 (Aprendizagem – Redes Neurais)

Equipe (até dois integrantes):

1. **Objetivo Principal do Projeto:** Analisar o resultado obtido com o uso de redes neurais na solução de um problema de classificação.
2. **O Problema: *Letter Image Recognition Data***
“The objective is to identify each of a large number of black-and-white rectangular pixel displays as one of the 26 capital letters in the English alphabet. The character images were based on 20 different fonts and each letter within these 20 fonts was randomly distorted to produce a file of 20,000 unique stimuli. Each stimulus was converted into 16 primitive numerical attributes (statistical moments and edge counts) which were then scaled to fit into a range of integer values from 0 through 15. We typically train on the first 16000 items and then use the resulting model to predict the letter category for the remaining 4000.”
3. **Base de Dados:** [clique aqui](#).
4. **Software: WEKA** (*Waikaito Environment for Knowledge Analysis*). Download: [clique aqui](#).

Formatação dos dados: antes de executar o WEKA com a base de dados, é necessário colocar os padrões no formato ARFF (formato do WEKA).

Obs.: O treinamento da rede neural, como também os resultados experimentais das simulações utilizando o WEKA serão anotados, interpretados e apresentados em forma de relatório (conforme descrição abaixo).

5. Descrição Geral do Experimento

- i) Serão feitos treinamentos de redes MLP usando Backpropagation padrão (com o termo *momentum* fixo de 0.8), variando os seguintes parâmetros:
 - o Quantidade máxima de iterações (ou ciclos)
 - o Quantidade de neurônios intermediários (ou escondidos) da rede, pois serão usadas redes com apenas uma camada intermediária.
 - o Taxa de aprendizado.

- ii) Para cada um destes 3 parâmetros, devem ser usados 3 valores. Sendo assim, a equipe deve escolher 3 quantidades máximas de iterações (épocas), 3 quantidades de neurônios escondidos e 3 taxas de aprendizado.
 - o Exemplo: poderiam ser usados 100, 1.000 e 10.000 iterações; 4, 8 e 12 neurônios escondidos; e taxas de aprendizado de 0.1, 0.01 e 0.001.
 - o Obs.: isto é apenas um exemplo; a própria equipe é quem vai escolher os valores a serem usados. Vale lembrar, que não existem valores ideais que podem ser usados para qualquer problema, de modo que uma taxa de aprendizado de, por exemplo, 0.01 pode ser pequena demais para um dado problema, sendo grande demais para outro problema. O mesmo vale para a quantidade de iterações e de nodos escondidos.
- iii) Deve ser feito 1 treinamento para cada combinação destes 3 parâmetros, sendo anotados os respectivos resultados. Sendo assim, devem ser feitos $3 \times 3 \times 3 = 27$ treinamentos preliminares.
- iv) Deve ser escolhida a melhor rede obtida (ou seja, a melhor combinação destes 3 parâmetros), sendo que a “melhor rede” é a rede que proporcionou os melhores resultados. Uma boa escolha pode ser baseada no erro do conjunto de treinamento e no tamanho da rede.
 - o Exemplo: se o menor erro de treinamento foi obtido usando 1.000 iterações, 8 nodos escondidos e taxa de aprendizado 0.001, então este pode ser considerado o melhor conjunto de parâmetros (ou seja, a melhor rede).
 - o Obs.: para um bom projeto de redes neurais, não é suficiente fazer apenas uma execução para cada conjunto de parâmetros, pois sabemos que, variando a inicialização de pesos, podemos obter resultados finais diferentes. Dessa forma, seriam necessárias várias inicializações de pesos para cada conjunto de parâmetros para escolhermos a melhor rede. Entretanto, o objetivo nesta prática é apenas dar uma noção de como se fazem experimentos com redes neurais. É por isso que só é exigido 1 treinamento para cada um dos 27 conjuntos de parâmetros.
- v) Para este conjunto de parâmetros escolhido (melhor rede), deve ser aplicado o *10-fold cross validation* com 10 execuções com diferentes inicializações de pesos (ou seja, fazer 10 treinamentos, cada um partindo de uma inicialização de pesos distinta). É conveniente organizar estes resultados em uma tabela e verificar os valores de média e desvio-padrão para os resultados.

6. Modelo de Relatório para o Projeto

- 1) Introdução
- 2) Descrição do problema e Base de Dados
- 3) Experimentos (Explicar o comportamento dos parâmetros estudados, dificuldades encontradas, limitações do modelo).
- 4) Análise e Considerações Finais
- 5) Referências bibliográficas

Referências Bibliográficas

- ANNE CANUTO & MARCÍLIO SOUTO, **Disciplina: Inteligência Artificial**. Universidade Federal de Rio Grande do Norte, Departamento de Informática e Matemática Aplicada. Disponível em www.dimap.ufrn.br/~marcilio/IA/IA2004.../Projeto-IA-2004.2.doc. Último acesso em: 10/11/15.
- **WEKA Software**. Disponível em <http://www.cs.waikato.ac.nz/~ml/index.html>. Último acesso em: 10/11/15.
- IAN H. WITTEN and EIBE FRANK. **Data mining: practical machine learning tools and techniques with Java implementations**. Morgan Kaufmann, San Francisco, 2000.