

IA013 (1s2014) – ECC 1

Peso 4 – (Individual ou em Grupo de 2)

Data de entrega: 07/04/2014

Questão 1) (4,0 pontos)

Aplique uma **rede neural MLP** a um **problema de classificação de padrões**. Considere problemas com mais de duas classes e uma saída da rede neural por classe. Defina o número n de neurônios na camada intermediária (Sugestão: Faça uma **busca** no intervalo [5, 20]). **Separe o conjunto de dados em Treinamento e Validação** e estabeleça o critério de parada com base no erro de validação. Use o erro de classificação e não o erro quadrático. Compare o desempenho com um classificador linear e com uma *extreme learning machine* (ELM). No caso da ELM, não é necessário fazer uma busca pelo melhor parâmetro de regularização, bastando testar alguns valores.

Questão 2) (4,0 pontos)

Aplique uma **rede neural MLP** a um **problema de regressão** com uma única saída. Defina o número n de neurônios na camada intermediária (Sugestão: escolha um valor entre 5 e 20). **Divida o conjunto de dados em 10 pastas e considere 9 delas para treinamento e a que sobrou para validação**. Faça isso 10 vezes, de modo que todas as pastas sejam empregadas ao menos uma vez para validação. Apresente o desempenho médio obtido pelas 10 MLPs treinadas. Em seguida, componha as 10 MLPs num ensemble, usando a média para combinar as saídas das 10 redes neurais. Nota: Nesta combinação das saídas poderia (mas não é solicitado aqui) também ser empregada a técnica LASSO (<http://www.di.ens.fr/~mschmidt/Software/lasso.html>).

Sugestão: Escolha conjuntos de dados dentre os disponíveis no *UCI Machine Learning Repository*: <http://archive.ics.uci.edu/ml/>

Observação 1: Em treinamento de RNAs a partir de dados amostrados, pode ser necessário **realizar pré-processamento junto aos dados**. Exemplo 1: Normalizar cada variável de entrada. Exemplo 2: Definir o número de atrasos no vetor de entradas da rede neural quando realizar predição de séries temporais.

Observação 2: Os programas para treinamento de redes MLP são fornecidos juntamente com este roteiro. Eles requerem o ambiente Matlab para a sua execução.

Questão 3) (2,0 pontos)

Dentre os algoritmos de auto-organização mais empregados para análise de dados está o *Growing Neural Gas* (GNG), que se assemelha ao mapa de Kohonen, mas tem uma topologia dinâmica. Dependendo do tipo de distribuição de dados e da aplicação, os resultados junto a cada proposta de algoritmo podem ser bem distintos. O DemoGNG v1.5 está disponível no endereço:

<http://www.sund.de/netze/applets/gng/full/tex/DemoGNG/DemoGNG.html>

e permite comparar os resultados de diferentes algoritmos de auto-organização. Execute o DEMO para verificar como o GNG e outros algoritmos de auto-organização operam na prática, junto a várias distribuições de dados. Faça algumas comparações. Em seguida, explique o princípio de funcionamento do GNG (por exemplo, na forma de pseudo-código), incluindo a forma como arestas entre neurônios são incluídas e excluídas.