

Disciplina: Aprendizagem de Máquina
Período: 2025.1
Professor: César Lincoln Cavalcante Mattos

Lista 4 - Redes Neurais Artificiais

Instruções

- Com exceção dos casos explicitamente indicados, os algoritmos e modelos devem ser implementados do início em qualquer linguagem de programação (Python, R, Octave...).
- Pacotes auxiliares (sklearn, matplotlib, etc) podem ser usados somente para facilitar a manipulação dos dados e criar gráficos.
- A entrega da solução pode ser feita via pdf ou Jupyter notebook pelo SIGAA.

Observações

- **Graduação:** Você pode usar uma das implementações de redes neurais a seguir: Tensorflow/Keras, PyTorch/Lightning ou Jax/Flax.
- **Pós-graduação:** Você deve implementar seus próprios algoritmos em Numpy.

Questão 1

Considere o conjunto de dados disponível em **concrete.csv**, organizado em 9 colunas, sendo as 8 primeiras colunas os atributos e a última coluna a saída. Os 8 atributos referem-se à caracterização de diferentes tipos de concreto para construção civil. A saída é a resistência à compressão do concreto (em megapascals, MPa). Maiores detalhes sobre os dados podem ser conferidos em <https://www.openml.org/d/4353>.

- Considere um modelo de regressão não linear baseado em redes neurais artificiais. Separe os dados aleatoriamente em treino, validação e teste (por exemplo, 60%, 20% e 20%). Nesse cenário, treine e avalie o modelo abaixo:
 - **MLP (multilayer perceptron):** 1 camada oculta e treinamento em *minibatch* via gradiente descendente estocástico com termo de *momentum*. Utilize o conjunto de validação para ajustar os hiperparâmetros.
- Apresente as curvas da função custo nos conjuntos de treinamento e validação ao longo das épocas. Reporte também para os conjuntos de treino, validação e teste as métricas abaixo:

- **RMSE (root mean squared error)**: $\sqrt{\frac{1}{N} \sum_i (y_i - \hat{y}_i)^2}$;
- **MAE (mean absolute error)**: $\frac{1}{N} \sum_i |y_i - \hat{y}_i|$;
- **MRE (mean relative error)**: $\frac{1}{N} \sum_i \left| \frac{y_i - \hat{y}_i}{y_i} \right|$.

Questão 2

Considere o conjunto de dados disponível em **vowel.csv**, organizado em 11 colunas, sendo as 10 primeiras colunas os atributos e a última coluna a saída. Os 10 atributos referem-se à caracterização de amostras da fala de britânicos. A saída é o fonema de vogal correspondente, dentre as 11 possibilidades. Maiores detalhes sobre os dados podem ser conferidos em <https://www.openml.org/d/307>.

- Considere um modelo de classificação não linear baseado em redes neurais artificiais. Separe os dados aleatoriamente em treino, validação e teste (por exemplo, 60%, 20% e 20%). Nesse cenário, treine e avalie o modelo abaixo:
 - **MLP (multilayer perceptron)**: 1 camada oculta e treinamento em *minibatch* via gradiente descendente estocástico com termo de *momentum*. Utilize o conjunto de validação para ajustar os hiperparâmetros.
- Apresente as curvas da função custo nos conjuntos de treinamento e validação ao longo das épocas. Reporte também a acurácia obtida para os conjuntos de treino, validação e teste. Por último, mostre a matriz de confusão obtida no conjunto de teste.