

Fundamentos da Aprendizagem de Máquina

Lista de Atividades

Aula - Redes Neurais Artificiais

Ricardo Augusto (ricardojunior@inatel.br)

Inatel

Índice

I	Redes Neurais Artificiais	
1	Atividades	7
1.1	Exercícios Conceituais e de Múltipla Escolha	8
1.2	Exercícios Computacionais	15



Redes Neurais Artificiais

Redes Neurais Artificiais

1. Atividades

Esse arquivo consiste em uma lista de atividades a serem realizadas para o módulo sobre redes neurais artificiais (RNA), do curso introdução à ciência de dados e decisões. A lista é composta pelas seguintes atividades:

- **Exercícios Conceituais**
 - São três (3) questões abertas sobre conceitos de redes neurais artificiais.
- **Exercícios de Múltipla Escolha**
 - São dez (7) questões de múltipla escolha sobre os fundamentos de RNA.
- **Exercícios Computacionais**
 - São três (3) exercícios computacionais sobre a aplicação de RNAs como modelos de aprendizado de máquina.

A composição da nota avaliativa desse módulo, denotada como N2 é dada pela combinação linear das atividades citadas, considerando pesos equilibrados, de acordo com

$$\begin{aligned} N2 = & 0.20 \times \text{Exercícios Conceituais} + 0.40 \times \text{Exercícios de Múltipla Escolha} \\ & + 0.40 \times \text{Exercícios Computacionais} \end{aligned} \quad (1.1)$$

1.1 Exercícios Conceituais e de Múltipla Escolha

1. Exercício 1 (Conceitos sobre Redes Neurais Artificiais)

Descreva com suas palavras o funcionamento de uma rede neural artificial.

2. Exercício 2 (Aprendizagem de Parâmetros em Redes Neurais Artificiais)

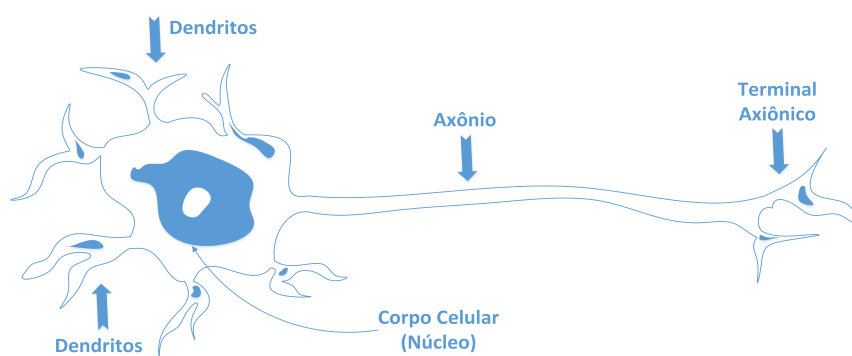
Faça um breve resumo sobre o funcionamento do algoritmo da retropropagação na aprendizagem de parâmetros em redes neurais artificiais.

3. Exercício 3 (Conceitos sobre Redes Neurais Artificiais)

Cite algumas áreas de aplicação das RNAs e os principais benefícios alcançados com essa técnicas de ML.

4. Exercício 4 (Neurônio Biológico)

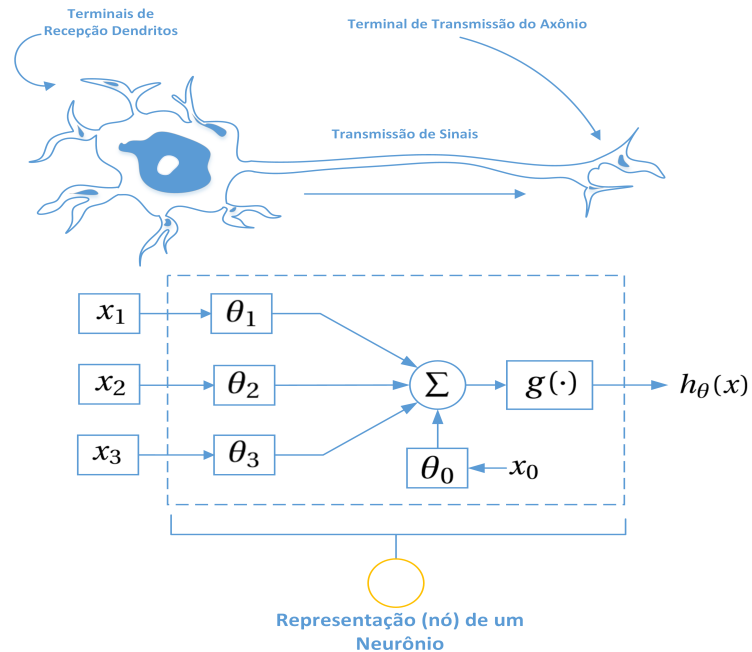
Marque a alternativa correta sobre os principais elementos que compõem um neurônio biológico (apresentado na figura abaixo), que consiste em uma unidade de referência para o avanço das pesquisas em RNAs.



- a) O corpo do neurônio é conhecido como soma e consiste na localização das sinapses.
- b) O terminal axônico de um neurônio se conecta, diretamente, ao axônios de outros neurônios.
- c) O papel do axiônio é transmitir os sinais recebidos via estímulos nervosos que chegam aos dendritos e atravessam o núcleo do neurônio.
- d) Os neurônios biológicos não conseguem receber curtos sinais elétricos (impulsos) em função do tamanho do axiônio.

5. Exercício 5 (Modelo Matemático de Neurônios Artificiais)

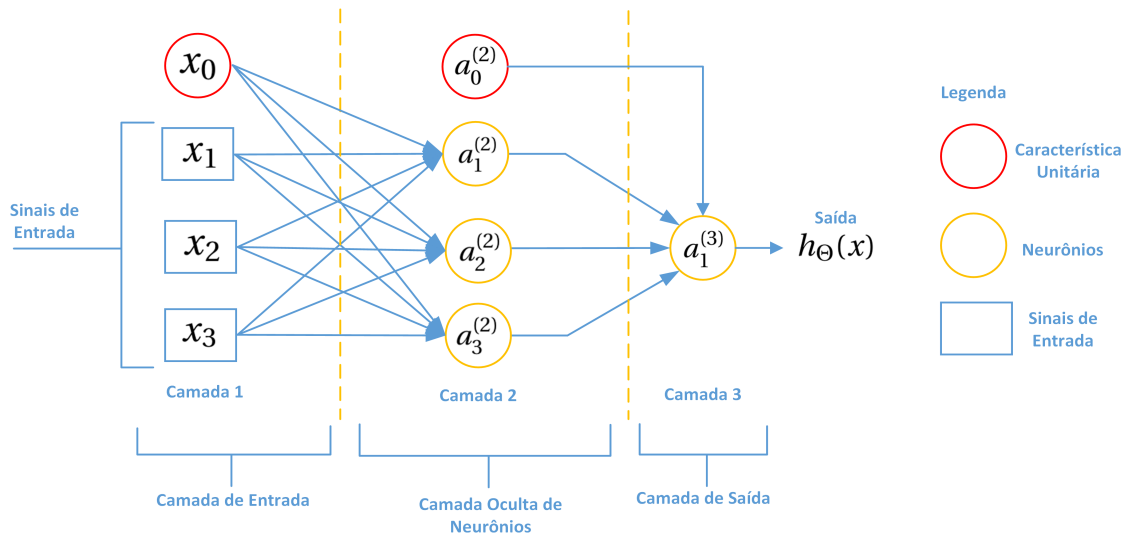
Marque a alternativa correta sobre a modelagem matemática dos neurônios na composição da unidade lógica de processamento de uma RNA.



- a) Os parâmetros $\theta_1, \theta_2, \theta_3$ modelam os efeitos de cada estímulo capturado pelo neurônio artificial.
- b) A função $g(\cdot)$ modela os efeitos dos sinais de entrada no neurônio artificial.
- c) A função $g(\cdot)$ não tem influência sobre o sinal de saída produzido pelo neurônio
- d) Os parâmetros $\theta_1, \theta_2, \theta_3$ consistem nas variáveis de entrada do neurônio conhecidas a partir da função $g(\cdot)$.

6. Exercício 6 (Arquitetura de RNA)

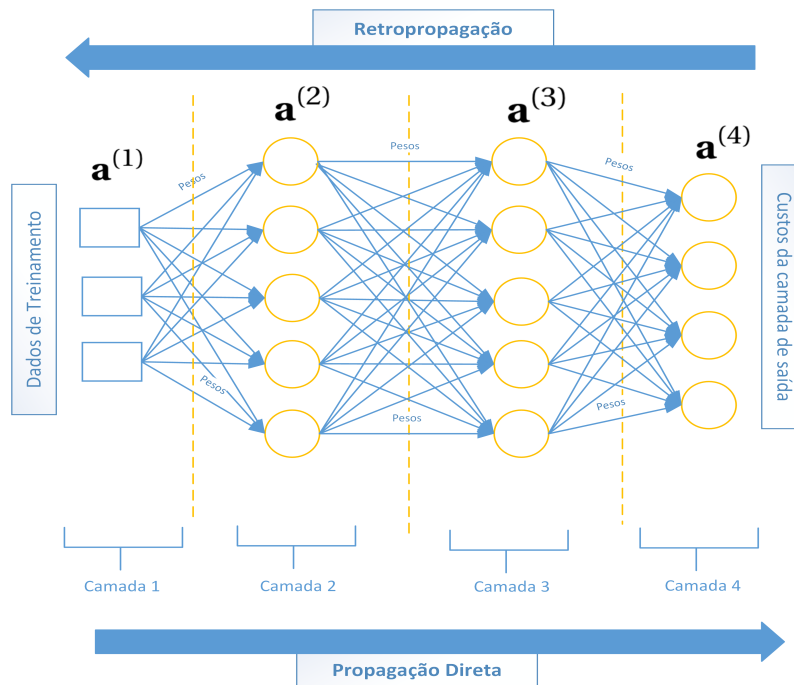
A figura abaixo ilustra a composição da arquitetura direta das redes neurais artificiais, considerando a simbologia para os dados de entrada e os neurônios da rede. Marque a alternativa correta sobre as características e o funcionamento desse tipo de rede neural artificial.



- a) Na arquitetura de RNA direta existe conectividade entre os neurônios da mesma camada, além da conexão entre camadas.
- b) Nas RNAs diretas é importante ressaltar que existe a ponderação de parâmetros nos sinais que entram na camada de entrada.
- c) A arquitetura de uma RNA direta é definida pela disposição dos neurônios na estrutura em camadas da rede, bem como o tipo de processamento que ocorre em cada neurônio.
- d) Não existem camadas ocultas em RNAs diretas, isto é, só existem as camadas de entrada e saída da rede.

7. Exercício 7 (Aprendizagem de Parâmetros em RNA)

Marque a alternativa correta sobre os conceitos do processo de retropropagação na aprendizagem de parâmetros em RNAs, conforme ilustrado na figura abaixo.



- a) O processo de retropropagação em uma RNA não envolve a propagação direta de sinais na rede.
- b) A retropropagação é baseada na obtenção de gradientes da camada de entrada e saída, que são posteriormente usados por um otimizador como o gradiente descendente.
- c) A obtenção de gradientes a partir dos termos de erro obtidos com a retropropagação consiste em uma parte do processo de aprendizagem, que ocorre após a propagação direta de sinais na rede.
- d) A retropropagação é o processo que maximiza a função custo da RNA.

8. Exercício 8 (Função Custo em RNAs)

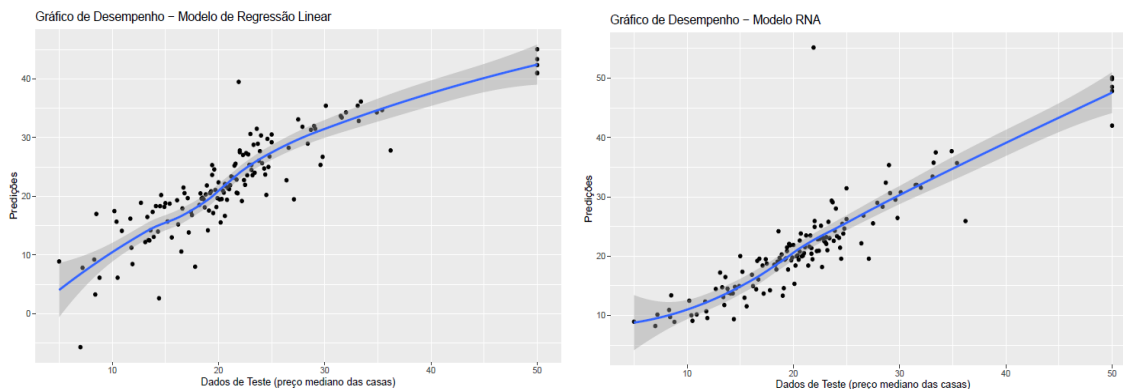
Marque a alternativa correta sobre as características da função custo, apresentada abaixo,

$$J(\Theta) = -\frac{1}{m} \sum_{i=1}^m \sum_{k=1}^K \left[y_k^{(i)} \log(h_{\Theta}(x^{(i)}))_k + (1 - y_k^{(i)}) \log(1 - h_{\Theta}(x^{(i)}))_k \right] + \frac{\lambda}{2m} \sum_{l=1}^{L-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (\Theta_{ji}^{(l)})^2 \quad (1.2)$$

- a) O resultado da RNA é expresso pelo termo $h_{\Theta}(x^{(i)})_k$ e consiste em um único valor escalar obtido pelo processamento dos neurônios da RNA, considerando $K > 1$.
- b) A função custo apresentada pode ser aplicada somente em casos de RNAs que visam o problema de classificação binária.
- c) O termo de regularização é modificado a fim de selecionar os parâmetros da matrizes $\Theta^{(l)}$, com $l = 1, \dots, L$, incluindo os pesos que representam o viés ou característica unitária dos neurônios
- d) O resultado da RNA é expresso pelo termo $h_{\Theta}(x^{(i)})_k$ e consiste em um vetor cuja composição envolve os valores de cada k -ésimo neurônio da camada de saída, considerando $K > 1$.

9. Exercício 9 (Interpretação dos Resultados de Predição)

A figura abaixo traz os resultados relacionados ao Exercício Computacional 2, apresentando as previsões de preços de casas com os modelos de regressão linear e redes neurais artificiais profundas. Marque a alternativa correta sobre a interpretação dos resultados obtidos.



- a) Os pontos sólidos (em preto) consistem na posições dos dados de treinamento dos modelos de regressão linear e RNA.
- b) O gráfico de desempenho relaciona os dados de teste com as previsões geradas pelos modelos de regressão e RNA. Dessa forma, se os pontos estiverem distantes da curva em azul, temos uma bom desempenho de generalização do modelo de DL.
- c) Quanto mais oscilante forem as curvas em azul, maior a variabilidade dos modelos de DL e, conseqüentemente, melhor o desempenho. Nos gráficos, é possível verificar a superioridade das RNAs em relação à regressão linear múltipla.
- d) O gráfico de desempenho relaciona os dados de teste com as previsões geradas pelos modelos de regressão e RNA. Cada ponto sólido (em preto) consiste em um par entre o respectivo dado de teste e sua previsão. Dessa forma, se os pontos recaírem sobre a curva em azul linearmente, temos uma bom desempenho de generalização do modelo de DL.

10. Exercício 10 (Frameworks de Redes Neurais Artificiais)

A Figura abaixo mostra os passos de iteração executados na fase de treinamento de uma rede neural profunda em conjunto com uma instrução de treino do modelo DL no framework Keras. Nota-se dois parâmetros na instrução: **batch_size** e **epochs**. Marque a alternativa que indica corretamente o conceito e as diferenças entre esses parâmetros da instrução.

Treinamento da Rede Neural Profunda (Keras)

```
modelo_RNA_Keras.fit(X_train, y_train, epochs = 100, batch_size = 10)
```

```
Epoch 1/100  
700/700 [=====] - 0s 204us/step - loss: 0.6421 - accuracy: 0.6943  
Epoch 2/100  
700/700 [=====] - 0s 115us/step - loss: 0.6124 - accuracy: 0.7000  
Epoch 3/100  
700/700 [=====] - 0s 117us/step - loss: 0.6114 - accuracy: 0.7000  
Epoch 4/100  
700/700 [=====] - 0s 121us/step - loss: 0.6113 - accuracy: 0.7000  
Epoch 5/100  
700/700 [=====] - 0s 110us/step - loss: 0.6116 - accuracy: 0.7000
```

- a) Os parâmetros **batch_size** e **epochs** possuem o mesmo significado que consiste no número de iterações do modelo de DL.
- b) O parâmetro **batch_size** consiste no número de variáveis exploratórias de entrada e, com isso, é fixo, enquanto o parâmetro **epochs** indica a quantidade de variáveis de saída na RNA.
- c) O parâmetro **batch_size** consiste no tamanho do subconjunto de dados de treinamento usados para as aprendizagens da RNA. Uma iteração sobre todo o conjunto de dados de treinamento, por meio da aprendizagem de vários **batch_sizes** consiste em um **epoch**.
- d) O parâmetro **epoch** consiste no tamanho do subconjunto de dados de treinamento usados para as aprendizagens da RNA. Uma iteração sobre todo o conjunto de dados de treinamento, por meio da aprendizagem de vários **epoch** consiste em um **batch_size**.

1.2 Exercícios Computacionais

1. Exercício 1 (Modelos de Redes Neurais Artificiais)

Contexto do problema: Temos o objetivo de construir um modelo preditivo de ML baseado em Redes Neurais Artificiais (RNAs) que seja capaz de realizar predições dos valores medianos de preços das casas em uma região suburbana de Boston, EUA. Tal como no caso da regressão linear, a variável a ser predita consiste em um valor numérico que representa o preço mediano das casas. Para cada uma das observações, temos diversas variáveis de entrada. Logo, podemos buscar a solução para esse problema usando diferentes técnicas de *Machine Learning*: nesse caso iremos construir, treinar e testar uma rede neural artificial para fazer as predições.

O pacote **neuralnet** pode ser usado para construirmos o modelo de RNA. Na linguagem R podemos usar:
`install.packages("neuralnet")`
`library(neuralnet)`

Dataset: Os dados estão fornecidos na forma de tabela (.xlsx e .csv) retratado pelo dataset Boston, presente na biblioteca **MASS** do software R que apresenta os valores das casas (*Median value of owner-occupied homes*) em 506 vizinhanças da cidade de Boston. Os dados que acompanham o valor mediano dos preços das casas consistem em indicadores de condições socioeconômicas, ambientais, instalações educacionais e alguns outros fatores semelhantes. No ambiente **R**, o comando `?Boston` fornece informações sobre cada uma das *features*.

Ao todo, são 13 *features* e uma variável resposta, denotada como **medv** (preço mediano da casa), baseada em \$1,000 dólares. De forma específica, no conjunto de variáveis explanatórias (i.e., características), temos doze (12) variáveis numéricas e uma (1) variável categórica, que no caso pode assumir 0 (zero) ou 1 (um). Com isso, a planilha de dados apresenta 506 linhas (exemplos de treinamento) e 14 colunas (*features*). Abaixo, estão colocadas cada uma das variáveis características do dataset e seu respectivo significado:

- CRIM: per capita crime rate by town
- ZN: proportion of residential land zoned for lots over 25,000 sq.ft.
- INDUS: proportion of non-retail business acres per town
- CHAS: Charles River dummy variable (= 1 if tract bounds river; 0 otherwise)
- NOX: nitric oxides concentration (parts per 10 million)
- RM: average number of rooms per dwelling
- AGE: proportion of owner-occupied units built prior to 1940
- DIS: weighted distances to five Boston employment centres
- RAD: index of accessibility to radial highways
- TAX: full-value property-tax rate per 10,000
- PTRATIO: pupil-teacher ratio by town
- B: $1000(B_k - 0.63)^2$ where B_k is the proportion of blacks by town
- LSTAT: % lower status of the population
- TARGET: Median value of owner-occupied homes in \$1000's

Abaixo, seguem os itens que devemos solucionar, visando alcançar o objetivo deste exercício:

- **Questões Avaliativas**
 - 1) Faça a exploração dos dados bem como síntese sobre suas principais variáveis explanatórias.
 - 2) Realize o procedimento de normalização mínimo-máximo sobre as características do dataframe
 - 3) Para reprodução dos resultados use o `set.seed(12)`.
 - 4) Realize a divisão do conjunto de treino e teste em 70/30.
 - 5) Construa e treine o modelo preditivo de ML baseado em redes neurais artificiais.
 - 6) Visualize a arquitetura da RNA treinada
 - 7) Aplique o modelo sobre os dados de teste para geração de predições
 - 8) Forneça a estimativa de erro quadrático médio (MSE)
 - 9) Compare os resultados de predição com os dados de teste (graficamente)
- **Dicas para o Exercício**
 - Escolha o ambiente de desenvolvimento e a linguagem que for mais confortável para você (e.g., R/RStudio, Python/Jupyter, MATLAB, Java, entre outras).

2. Exercício 2 (Modelos de Redes Neurais Artificiais)

Considere o desenvolvimento da rede neural artificial do Exercício Computacional 1 - obtenha um comparativo de desempenho desse modelo com a regressão linear múltipla conduzida sobre os mesmos dados de treinamento. O ponto chave aqui é realizar a **interpretação** dos resultados obtidos.

3. Exercício 3 (Modelos de Redes Neurais Artificiais (RNA) com Framework Keras)

Uma instituição financeira nos forneceu um conjunto de dados relacionados à créditos financeiros presentes no banco de dados da instituição. A instituição está trabalhando em um projeto de ciência de dados para previsão de risco de crédito. Nós iremos participar de uma fase específica desse projeto com o objetivo de construir um **classificador**, que possa auxiliar na análise de risco de crédito de diversos clientes da instituição.

Na fase de modelagem do projeto de ciência de dados, temos o objetivo de construir uma rede neural artificial profunda para realização de tarefas de classificação necessárias. Para isso, iremos usar a biblioteca de *Deep Learning* **Keras**, amplamente conhecida no campo de machine learning, e que permite a construção de RNAs profundas sobre os *frameworks* TensorFlow e Theano.

Sobre o problema, a RNA deve prever se um determinado cliente deve ou não receber créditos de produtos financeiros ofertados pela instituição. Isso significa que teremos acesso a um conjunto de dados com informações diversas sobre inúmeros clientes da instituição.

O conjunto de dados consiste em vinte (20) variáveis explanatórias que consistem em informações diversas sobre os clientes incluindo: duração e tamanho do crédito, indicadores de saldo e operações financeiras, além de dados dos clientes como idade, dependentes, emprego e até contatos como telefone. Tais informações são apresentadas com codificação que são processadas pela instituição para posteriores interpretações. Com isso, nossa tarefa consiste em lidar com os dados codificados e a variável de saída **credit.rating**, que indica o estado de **aprovação** (1) ou **desaprovação** (0) de crédito para cada dados de treinamento (registro) do dataset.

• Questões Avaliativas

- 1) Use as instruções `!pip install` para preparar o ambiente Python com a biblioteca Keras e os frameworks necessários.
- 2) Faça a importação dos pacotes da biblioteca Keras.
 - * `keras.models` - modelos
 - * `keras.layers` - estrutura de RNA profundas
 - * `keras.optimizers` - otimizadores
- 3) Realize a importação do dataset .csv de crédito para o ambiente de desenvolvimento.
- 4) Faça a análise exploratória do dataset e verifique a necessidade de normalização dos dados.
- 5) Considerando o item 4) aplique a normalização padrão nas variáveis explanatórias numéricas do dataset.
- 6) Realize a divisão do conjunto de treino e teste em 70/30.
- 7) Para reprodução dos resultados use o `set.seed(101)`.
- 6) Construa e treine um modelo de Deep Learning com o Keras, por meio de três fases:
 - * Construção do modelo (`keras.models` e `keras.layers`)
 - * Compilação do modelo (`keras.optimizers`)
 - * Treinamento do modelo (`keras.fit`)
- 7) Use a biblioteca `ann_visualizer` para visualizar a arquitetura da RNA construída
 - * `from ann_visualizer.visualize import ann_viz;`
- 8) Interprete os resultados do modelo de DL.
- 9) Avalie a acurácia do modelo de DL baseado na RNA construída e treinada sobre os dados de teste.