



Inteligência Artificial Computacional

Trabalho Computacional AV2: Redes Neurais Artificiais.

Professor: Prof. Msc. Paulo Cirillo Souza Barbosa

Observações iniciais.

Solicita-se que realize a leitura por completo do presente documento. Na AV3 as bibliotecas permitidas são:

1. numpy;
2. matplotlib;
3. seaborn;

Introdução.

O presente trabalho é composto por duas etapas em que deve-se utilizar os modelos de IA, que são bioinspirados no funcionamento do cérebro humano. Tais modelos são redes neurais artificiais (RNA) lineares e não-lineares. A primeira etapa do trabalho, trata-se da solução de um problema simples, e serve como um validador das implementações dos modelos solicitados. A segunda etapa, trata-se da resolução de um problema de reconhecimento facial de imagens.

Primeira etapa - Regressão e Classificação para problemas bidimensionais. [3,0 pts]

Tarefa de Regressão

Solicita-se que faça o acesso ao conjunto de dados disponibilizado na plataforma AVA, chamado **Spiral3d.csv**. Para o problema de regressão solicita-se que faça o acesso ao conjunto de dados disponibilizado na plataforma AVA, chamado **aerogerador.dat**. A variável independente é uma medida de velocidade do vento, e a variável dependente é uma observação de potência gerada pelo aerogerador. O presente conjunto de dados possui $p = 1$, $N = 2250$. Assim, faça o que se pede:

1. Faça a organização do conjunto de dados para serem apresentados às redes neurais.
2. Faça uma visualização inicial dos dados através do gráfico de espalhamento.
3. Os modelos de RNA a serem implementados nessa etapa serão: **ADALINE** e **Perceptron de Múltiplas Camadas (MLP)**. Para cada modelo, deve-se discutir como os hiperparâmetros foram escolhidos.
4. Para o modelo MLP, pede-se que identifique o *underfitting* e *overfitting*. Neste caso, sua equipe deve explorar diferentes topologias da rede, ao modificar os hiperparâmetros associados a quantidade de neurônios e camadas (realizando um subdimensionamento e superdimensionamento). Faça composição do resultado ao descrever as topologias encontradas e exiba curva de aprendizado para os dois casos (superdimensionado e subdimensionado).
5. Para validar os modelos (com suas topologias definidas) utilizados na tarefa, sua equipe deve projetar a validação utilizando as simulações por Monte Carlo. Nessa etapa, defina a quantidade de rodadas da simulação igual a $R = 250$. Em cada rodada, deve-se realizar o particionamento em 80% dos dados para treinamento e 20% para teste. A medida de desempenho utilizada é a média de desvios quadráticos (MSE). Para produzir um resultado em tabela, calcule a média, desvio-padrão, maior valor e menor valor de MSE para as 250 rodadas.

Modelos	Média	Desvio-Padrão	Maior Valor	Menor Valor
ADALINE				
Perceptron de Múltiplas Camadas				

Tarefa de Classificação

Solicita-se que faça o acesso ao conjunto de dados disponibilizado na plataforma AVA, chamado **Spiral3d.csv** que trata-se de dados sintetizados que possuem duas classes (as quais estão rotuladas no próprio arquivo na quarta coluna). O presente conjunto de dados possui $p = 3$, $N = 2000$ e $C = 2$. Assim, faça o que se pede:

1. Faça a organização do conjunto de dados para serem apresentados às redes neurais.
2. Faça uma visualização inicial dos dados através do gráfico de espalhamento.
3. Os modelos de RNA a serem implementados nessa etapa serão: **Perceptron Simples**, **Perceptron de Múltiplas Camadas (MLP)** e **Rede Função de Base Radial (RBF)**. Para cada modelo, deve-se discutir como os hiperparâmetros foram escolhidos.
4. Para os modelos MLP e RBF, pede-se que identifique o *underfitting* e *overfitting*. Neste caso, sua equipe deve explorar diferentes topologias das duas redes, ao modificar os hiperparâmetros associados a quantidade de neurônios e camadas (realizando um subdimensionamento e superdimensionamento). Para os dois casos, faça a composição do resultado ao descrever as topologias encontradas e calcule os desempenhos utilizando matriz de confusão, acurácia, especificidade, sensibilidade e curva de aprendizado para cada um dos casos (modelos MLP e RBF bem como underfitting e overfitting).
5. Para validar os modelos (com suas topologias definidas) utilizados na tarefa, sua equipe deve projetar a validação utilizando as simulações por Monte Carlo. Nessa etapa, defina a quantidade de rodadas da simulação igual a $R = 250$. Em cada rodada, deve-se realizar o particionamento em 80% dos dados para treinamento e 20% para teste. As medidas de desempenho de cada modelo, devem ser: **acurácia**, **sensibilidade** e **especificidade**.
6. Das 250 rodadas do item anterior, selecione aquelas que proporcionaram o maior e o menor valor de acurácia. Para esses dois casos, construa a matriz de confusão. **Obs:** sua matriz de confusão deve ser implementada pela equipe (ou seja, realizar a soma de seus elementos), no entanto, para realizar seu plot, pode-se utilizar a biblioteca *seaborn*. Exemplo: [link heatmap](#). Além disso, para os dois casos (maior e menor acurácia), pede-se que plote a curva de aprendizado de todos os modelos.
7. Ao final das $R = 250$ rodadas calcule para cada modelo utilizado, média aritmética, desvio-padrão, valor maior, valor menor, para as seguintes métricas de desempenho: **acurácia**, **sensibilidade** e **especificidade**. Coloque os resultados obtidos em uma tabela (UMA PARA CADA MÉTRICA) e **discuta os resultados obtidos**. **Obs:** O resultado não precisa ser limitado a tabela, como pode ser expresso via gráficos (BOXPLOT, VIOLIN PLOT E OUTROS). Como exemplo, segue uma tabela que serve para uma das três métricas solicitadas. Segue um exemplo para apenas uma das tabelas:

Modelos	Média	Desvio-Padrão	Maior Valor	Menor Valor
Perceptron Simples				
Perceptron de Múltiplas Camadas				
Rede RBF				

Classificação Multiclasse para Problema Multidimensional [7,0 pts]

A segunda etapa do trabalho, trata-se da classificação entre condições diferentes da coluna vertebral em pacientes. Para cada um dos pacientes (representando as $N = 310$ amostras) existem os seguintes atributos

(preditores):

$$\mathbf{x} \in \mathbb{R}^{p \times 1} = \begin{bmatrix} x_1 = & \text{Incidência Pélvica} \\ x_2 = & \text{Inclinação Pélvica} \\ x_3 = & \text{Ângulo da Lordose Lombar} \\ x_4 = & \text{Ângulo Sacral} \\ x_5 = & \text{Raio Pélvico} \\ x_6 = & \text{Grau de Espondilolistese} \end{bmatrix}$$

O conjunto de dados está disponibilizado no AVA (arquivo com nome **"coluna vertebral.csv"**) e consiste em 310 amostras com 6 preditores. A tarefa de classificação trata-se de um problema real de auxílio médico para identificação de 3 condições diferentes da coluna vertebral. A primeira condição rotulada nos dados como **"NO"**, trata-se da categoria **normal** (paciente sem alguma das duas patologias). A segunda condição rotulada nos dados como **"DH"**, trata-se da categoria de pacientes que possuem a patologia **Hérnia de Disco**. A última condição que é rotulada nos dados como **"SL"**, trata-se da categoria de pacientes que possuem a patologia **Espondilolistese**. Após esse contexto apresentado, faça o que se pede:

1. De posse do conjunto de dados, organize a massa de dados \mathbf{X} com as dimensões $\mathbb{R}^{p+1 \times N}$.
2. Faça a codificação utilizando a estratégia *one-hot-encoding* da seguinte forma:

$$\text{Normal} = \begin{bmatrix} +1 \\ -1 \\ -1 \end{bmatrix} \quad \text{Hérnia de Disco} = \begin{bmatrix} -1 \\ +1 \\ -1 \end{bmatrix} \quad \text{Espondilolistese} = \begin{bmatrix} -1 \\ -1 \\ +1 \end{bmatrix}$$

Assim, a massa de dados \mathbf{Y} deve possuir as dimensões $\mathbb{R}^{C \times N}$.

3. Os modelos de RNA a serem implementados nessa etapa serão: **ADaptive LINear Element, Perceptron de Múltiplas Camadas (MLP)** e **Rede RBF**. Para cada modelo, deve-se discutir como os hiperparâmetros foram escolhidos. Neste caso, utiliza a experiência sua adquirida na primeira etapa do trabalho.
4. Para validar os modelos utilizados na tarefa, sua equipe deve projetar a validação utilizando as simulações por Monte Carlo. Nessa etapa, defina a quantidade de rodadas da simulação igual a $R = 100$. Em cada rodada, deve-se realizar o particionamento em 80% dos dados para treinamento e 20% para teste. As medidas de desempenho de cada modelo, devem ser: **acurácia, sensibilidade e especificidade**.
5. Das 100 rodadas do item anterior, selecione aquelas que proporcionaram o maior e o menor valor de acurácia. Para esses dois casos, construa a matriz de confusão. **Obs:** sua matriz de confusão deve ser implementada pela equipe, no entanto, para realizar seu plot, pode-se utilizar a biblioteca *seaborn*. Exemplo: [link heatmap](#). Com essas matrizes de confusão criadas, tente entender para qual categoria os modelos fazem previsões piores e melhores. Tente verificar se há um padrão percebido nos modelos em questão (sempre erram ou sempre acertam as mesmas categorias). Além disso, para os dois casos (maior e menor acurácia), pede-se que plote a curva de aprendizado de todos os modelos.
6. Ao final das $R = 100$ rodadas, calcule para cada modelo utilizado, média aritmética, desvio-padrão, valor maior, valor menor, para as seguintes métricas de desempenho: **acurácia**. Coloque os resultados obtidos em uma tabela **ediscuta os resultados obtidos**. **Obs:** O resultado não precisa ser limitado a tabela, como pode ser expresso via gráficos (BOXPLOT, VIOLIN PLOT E OUTROS). Como exemplo, segue uma tabela que serve para uma das três métricas solicitadas.

Modelos	Média	Desvio-Padrão	Maior Valor	Menor Valor
ADaptive LINear Element.				
Perceptron de Múltiplas Camadas				
Rede RBF				

Informações Adicionais aos modelos de RNA

De maneira geral, para as duas etapas do trabalho, sua equipe deve considerar que tais modelos possuem hiperparâmetros e procedimentos específicos. Para tais modelos, é necessário que os dados sejam normalizados (caso eles não já estejam). Além disso, todos os modelos precisam de um critério de convergência baseado no número máximo de épocas. Essa última informação é importante, pois, o modelo Perceptron Simples tradicional não possui esse critério. Assim, tal modificação deve ser feita pela sua equipe.

Considere que para os modelos Perceptron Simples e ADALINE, a função de ativação é linear (degrau ou degrau bipolar). A função de ativação utilizada na rede MLP deve ser escolhida entre a tangente hiperbólica e a sigmoide logística. Para simplicidade do projeto, faça com que todos os neurônios possuam a mesma função.

Para o caso das redes MLP e RBF, considere escolha dos hiperparâmetros, conforme as discussões realizadas em sala (ou através do que está descrito nos slides). Os hiperparâmetros são :

1. Quantidade de camadas escondidas (apenas para MLP).
2. Quantidade de neurônios nas camadas escondidas (na rede RBF existe apenas uma camada oculta).
3. Quantidade de neurônios na camada de saída.
4. Valor da precisão (ou critério de parada da convergência).
5. Número máximo de épocas.
6. Taxa de aprendizagem.

5) Relatório.

Além das implementações, o presente trabalho deve ser entregue em modelo de relatório. Este deve possuir as características descritas nos slides de apresentação do curso. Desta maneira, deve possuir:

1. Título (2,5%).
2. Resumo (2,5%).
3. Metodologia (42,5%).
4. Resultados (42,5%).
5. Conclusões (10%).

O modelo para trabalho pode ser encontrado neste [LINK](#)

6) Observações.

- **Obs1:** O envio das implementações é **obrigatório**. Caso a equipe não realize esta entrega, será atribuído nota **zero** para os respectivos alunos.
- **Obs2:** A data estipulada para entrega do trabalho, também é um critério avaliativo. Assim, caso haja atraso na entrega do trabalho, será aplicada: **de 00:15h até 24h: penalidade de 20% ; 24:15h até 48h: penalidade de 40% ; acima de 48h: penalização máxima (100%)**.
- **Obs 3:** As implementações devem ser realizadas utilizando apenas as bibliotecas descritas no início do presente documento. Caso alguma biblioteca diferente seja utilizada, será atribuída a nota **zero** para todos os membros da equipe.
- **Obs 4:** QUAISQUER alterações dos modelos, que não foram discutidas em sala de aula, ou que não estão presentes no conjunto de slides, devem ser explicadas pela equipe. Caso a equipe não saiba o motivo pelo qual as adequações foram realizadas, será atribuída a nota **zero** aos respectivos alunos.
- Caso não sejam utilizados os modelos descritos no presente documento, será considerado que sua equipe não atendeu aos requisitos, e assim, a nota para os membros da equipe será **zero**.

- **Obs 5:** A apresentação do trabalho no formato de arguição, é obrigatória. A data da apresentação está definida no AVA. Se a equipe não participar deste momento, a nota do trabalho será ZERO.
- **Obs 6:** Os trabalhos e implementações serão enviadas a um software anti-plágio. Qualquer caracterização de plágio ocasionará em nota zero para ambas equipes.