INTELIGÊNCIA COMPUTACIONAL II
Prof. Carlos Eduardo Pedreira
PESC/COPPE/UFRJ

Trabalho Prático II

Data de Entrega: 10 de Agosto 2018

O Trabalho Prático II visa implementar os exercícios computacionais descritos no **Homework #2** disponibilizado em:
https://work.caltech.edu/homework/hw2.pdf

Você deverá implementar os exercícios computacionais na linguagem de sua escolha (preferencialmente em Matlab). Um relatório prático sucinto deve ser escrito e entregue com o seguinte conteúdo:
- Você deve informar o seu nome no topo do relatório (primeira informação);
- Uma breve introdução sobre o assunto;
- O código fonte da implementação devidamente comentado;
- Resultados alcançados e sua interpretação dos mesmos.

* O código deverá ser apresentado e executado em tempo oportuno, para fins de avaliação, com possível arguição do mesmo. A ser agendado.

Descrição das tarefas

*Uma tradução livre dos enunciados presentes no **Homework #2** foi gerada. Caso ocorram dúvidas na interpretação recorram ao texto original em inglês.

Regressão Linear

Nestes problemas nós vamos explorar como a Regressão Linear para classificação trabalha. Da mesma maneira com uso do Algoritmo de Aprendizagem Perceptron no **Homework #1**, você vai criar a sua própria função *target* (alvo) f e o conjunto de dados D. Utilize $d = 2$ para que você possa visualizar o problema, e assuma $\mathcal{X}$ = [-1,1] × [-1,1] com probabilidade uniforme selecionando cada $\mathbf{x}$ $\in \mathcal{X}$. Em cada execução, escolha uma reta aleatória no plano como sua função *target* f (faça isso selecionando dois pontos aleatórios, uniformemente distribuídos em [-1,1] × [-1,1] e use a reta que passa entre eles), de forma que a reta mapeie +1 por um lado e -1 pelo outro. Escolha as entradas $\mathbf{x_n}$ do conjunto de dados de pontos aleatórios (uniformemente em $\mathcal{X}$), e avalie a função *target* em cada $\mathbf{x_n}$ para encontrar a saída correspondente y_n.

1) Utilize N = 100. Use Regressão Linear para encontrar g e avaliar E_{in}, a fração de pontos dentro da amostra que foram classificados incorretamente. Repita o experimento 1000 vezes e use o valor médio (guarde as g's que serão usadas novamente no Problema 2). Qual é o valor médio aproximado de E_{in}? (aproximado é a opção que faz a expressão | sua resposta - dada opção| próxima a 0. Use esta definição aqui e sempre).

a) 0 b) 0.001 c) 0.01 d) 0.1 e) 0.5

2) Agora, gere 1000 novos pontos e os use para estimar o erro fora da amostra E_{out} de g que você fez no Problema 1 (número de pontos classificados incorretamente/ número total de pontos fora da amostra). Novamente, execute o experimento 1000 vezes e guarde a média. Qual é o valor médio aproximado de E_{out}?

a) 0 b) 0.001 c) 0.01 d) 0.1 e) 0.5

3) Agora, utilize N = 10. Posteriormente, procurando os pesos usando Regressão Linear, os use como um vetor de pesos inicial para o Algoritmo de Aprendizagem Perceptron. Execute PLA até que convirja para o vetor final de pesos que separe completamente todos os pontos dentro da amostra. Entre as opções abaixo, qual é o valor mais próximo do número médio de iterações (mais de 1000 execuções) que o PLA leva para convergir? (Quando estiver implementando o PLA, escolha um ponto aleatório para o conjunto classificado incorretamente para cada iteração).

a) 1 b) 15 c) 300 d) 5000 e) 10000

Regressão Não-Linear

Nestes problemas, nós vamos novamente aplicar Regressão Linear para classificação. Considere a função *target*:

$f(x_1, x_2) = sign(x_1^2 + x_2^2 - 0.6)$.

Gere um conjunto de treinamento de N = 1000 pontos em $\mathcal{X}$ = [-1,1]×[-1,1] com probabilidade uniforme escolhendo cada $\mathbf{x} \in \mathcal{X}$. Gere um ruído simulado lançando o sinal de saída aleatoriamente selecionando 10% do subconjunto de treinamento gerado.

4) Execute a Regressão Linear sem transformação usando o vetor de atributos:

$(1, x_1, x_2)$,

para encontrar o peso $\mathbf{w}$. Qual é o valor aproximado de classificação do erro E_{in} dentro da amostra? (Execute o experimento 100 vezes e use o valor médio de E_{in} para reduzir a variação nos seus resultados.)

a) 0 b) 0.1 c) 0.3 d) 0.5 e) 0.8

5) Agora, transforme os N = 1000 dados de treinamento seguindo o vetor de atributos não-linear:

$(1, x_1, x_2, x_1x_2, x_1^2, x_2^2)$.

Encontre o vetor $\widetilde{\mathbf{w}}$ que corresponde a solução da regressão linear. Quais das hipóteses a seguir é a mais próxima que você encontrou? Neste caso, próximo significa o valor que mais entra em acordo com sua hipótese (existe uma alta probabilidade de estar acordando com um ponto aleatoriamente selecionado). Em média algumas execuções serão necessárias para assegurar uma resposta estável.

a) $g(x_1, x_2) = sign(-1 - 0.05x_1 + 0.08x_2 + 0.13x_1x_2 + 1.5x_1^2 + 1.5x_2^2)$
b) $g(x_1, x_2) = sign(-1 - 0.05x_1 + 0.08x_2 + 0.13x_1x_2 + 1.5x_1^2 + 15x_2^2)$
c) $g(x_1, x_2) = sign(-1 - 0.05x_1 + 0.08x_2 + 0.13x_1x_2 + 15x_1^2 + 1.5x_2^2)$
d) $g(x_1, x_2) = sign(-1 - 1.5x_1 + 0.08x_2 + 0.13x_1x_2 + 0.05x_1^2 + 0.05x_2^2)$
e) $g(x_1, x_2) = sign(-1 - 0.05x_1 + 0.08x_2 + 1.5x_1x_2 + 0.15x_1^2 + 0.15x_2^2)$

6) Qual o valor mais próximo do erro de classificação fora da amostra E_{out} de sua hipótese no Problema 5? (Estime isso gerando um novo conjunto de 1000 pontos e adicione ruído, como antes. Em média 1000 execuções reduzem a variação em seus resultados).
a) 0 b) 0.1 c) 0.3 d) 0.5 e) 0.8