

Trabalho Computacional III

Prof. Cristiano Leite de Castro

21 de novembro de 2018

1 TAREFAS PRELIMINARES

1. Ler o Capítulo 12 “Adaptive Neuro-Fuzzy Inference System” do livro texto: Jyh-Shing Roger Jang and Chuen-Tsai Sun. 1996. Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence.
2. Resolver os problemas a seguir.

2 PROBLEMA: APROXIMAÇÃO DE UMA FUNÇÃO

Projete um sistema nebuloso adaptativo (ANFIS) para realizar a aproximação da função $\text{seno}(x)$, para x definido no intervalo de $[0, 2\pi]$.

O sistema nebuloso adaptativo deve ser do tipo *Takagi-Sugeno* de ordem 1 e deve usar, a princípio, o mesmo conjunto de regras definidas no Trabalho II. Use funções de pertinência do tipo Gaussiana para os conjuntos nebulosos do antecedente das regras.

Mostre o gráfico da aproximação em relação à função original e calcule o MSE (*Mean Squared Error*).

3 PROBLEMA: CLASSIFICAÇÃO DE PADRÕES

Projete um sistema nebuloso adaptativo (ANFIS) para a resolução de 3 problemas de classificação de padrões:

1. Dados sintéticos 2D: mesmo conjunto de dados usado no TP anterior. Serve como referência para verificar se a estratégia proposta está funcionando de forma correta. Para este problema, vocês podem plotar a superfície de decisão no espaço dos dados.
2. Breast Cancer Wisconsin (Diagnostic) Data Set, disponível no link:
<https://www.kaggle.com/uciml/breast-cancer-wisconsin-data>.
Observação: a coluna *ID Number* deve ser removida durante a análise pois não traz informação útil para o problema.
3. Iris Species Data Set, disponível no link:
<https://www.kaggle.com/uciml/iris>
Observação: este problema tem originalmente 3 classes. Para fins de simplificação, considere que o objetivo aqui é distinguir a espécie “Iris-setosa” das outras duas espécies existentes. Novamente, a coluna *ID* deve ser removida durante a análise pois não traz informação útil para o problema.

Para cada problema, divida o conjunto de dados em subconjuntos de treinamento (70%) e teste (30%). Utilize o subconjunto de treinamento para treinamento e seleção de modelos. A seleção de modelos consiste em escolher a combinação mais apropriada de hiperparâmetros a partir de um método de busca em grade via Validação Cruzada com k folds. O espaço de busca de hiperparâmetros compreende o número de regras do ANFIS e o tipo de função pertinência a ser usada no antecedente de cada regra. O sistema projetado deverá ser do tipo Takagi-Sugeno de ordem 1. Vocês podem utilizar algum mecanismo para inicialização dos parâmetros das funções de pertinência do antecedente como, por exemplo, o algoritmo de agrupamento *Fuzzy C-means*. Como já visto, ele é útil para encontrar os centros e dispersões iniciais para funções de pertinência do tipo Gaussiana. Após a seleção de modelos, recomenda-se retreinar o ANFIS com a melhor combinação de hiperparâmetros usando todo o conjunto de treinamento. Utilize este modelo treinado para calcular a Acurácia de classificação para o subconjunto de teste.

O modelo fuzzy Takagi-Sugeno de ordem 1 foi inicialmente projetado para resolver problemas de Regressão. Entretanto, como esta parte do trabalho consiste em Classificação, utilize alguma abordagem para adaptar o sistema ANFIS para esta tarefa. Uma abordagem simples é utilizar um limiar na saída do modelo (neste exemplo, o limiar seria de 0.5):

$$\hat{y} = \begin{cases} 1 & \text{se } \hat{y} \geq 0.5 \\ 0 & \text{caso contrario} \end{cases}$$

Descreva no seu relatório as decisões tomadas e discuta os resultados obtidos. Faça um gráfico com o erro de treinamento e de teste em cada experimento realizado e discuta o comportamento do erro de treinamento e teste em função do número de regras do modelo ANFIS.