



UNIVERSIDADE FEDERAL DO CEARA  
CAMPUS DE SOBRAL  
ENGENHARIA DA COMPUTAÇÃO E ENGENHARIA ELÉTRICA

## **TÓPICOS ESPECIAIS EM TELECOMUNICAÇÕES I (ECO0080)**

### **TEMA: RECONHECIMENTO DE PADRÕES**

**Prof.: C. Alexandre Rolim Fernandes**

#### **Trabalho 5 – Regressão SVM com PCA**

- Trabalho Individual
- Sugere-se o uso do MATLAB ou Python para a prática de simulação.
- O código deve estar bem organizado e comentado, para que seja possível entendê-lo e corrigi-lo. Códigos que estejam desorganizados ou sem os devidos comentários explicativos terão penalização na nota.
- Fazer todas as questões em **um só arquivo.**
- O seu código deve **gerar automaticamente todos os gráficos e resultados** solicitados.
- Enviar no SIGAA **apenas o arquivo do código (.m, .mlx ou .ipynb).**
- **O nome do arquivo deve ser igual ao seu nome.**
- **Não enviar código em formato PDF ou de texto (DOC, ODT, TXT etc).**
- **Não enviar arquivo compactado** (zip,rar etc).
- Prazo e forma de entrega: dia 30/07/25 às 23h59, no SIGAA.

#### **Prática de Simulação: Previsão do Desempenho dos Estudantes Usando SVM e PCA**

- A base de dados usada neste trabalho está fornecida no arquivo Student\_Performance.csv. Trata-se de um conjunto de dados que tem como objetivo permitir estudar a relação entre o desempenho de estudantes e

algumas informações relativas ao aluno. A base é composta por um total de 1000 amostras.

- Variáveis de entrada:
  - Horas de Estudo: Número total de horas dedicadas ao estudo por cada estudante.
  - Notas Anteriores: As notas obtidas pelos estudantes em testes anteriores.
  - Atividades Extracurriculares: Indica se o estudante participa ou não de atividades extracurriculares (Sim ou Não).
  - Horas de Sono: Número médio de horas de sono que o estudante teve por dia.
  - Provas de Exercício Praticadas: Quantidade de provas de exercícios simuladas que o estudante praticou.
- Variável de Entrada:
  - Índice de Desempenho: Uma medida do desempenho geral de cada estudante. O índice de desempenho representa o rendimento acadêmico do aluno e foi arredondado para o número inteiro mais próximo. O índice varia de 10 a 100, sendo que valores mais altos indicam melhor desempenho.
- O trabalho consiste em implementar o regressor SVR usando esta base de dados. Seu código deve fazer a otimização (*tunning*) dos seguintes hiperparâmetros: tipo de função *kernel*, constante de relaxamento C e do parâmetro *kernel scale*.
- Para a otimização dos hiperparâmetros, você deverá fazer um *grid search*, testando os seguintes valores:
  - Função kernel: linear, RBF e polinomial (grau 2).
  - Constante de relaxamento C: 10<sup>-2</sup>, 10<sup>-1</sup>, 1, 10, 100
  - Kernel scale: 10<sup>-2</sup>, 10<sup>-1</sup>, 1, 10, 100
- Usar validação cruzada K-fold com K=10.
- Além de implementar o SVR “sozinho”, você também deverá implementar o SVR em conjunto com o PCA. Neste caso, você deverá testar o PCA com 1, 2, 3, 4 e 5 componentes. No total, você deverá implementar 6 regressores.
- O PCA deverá implementado apenas com os melhores hiperparâmetros obtidos com o SVR “sozinho”, ou seja, não precisa fazer o ajuste dos hiperparâmetros com o PCA.
- É permitido usar uma função pronta para o classificador SVR e para o K-fold. Não é permitido usar funções prontas para o grid search nem para o PCA.

- O código deve fornecer como saída:
  - O coeficiente de determinação ( $R^2$ ), para cada um dos 6 regressores
  - O erro quadrático médio para cada um dos 6 regressores, para cada um dos 6 regressores
  - A razão entre a variância da entrada e da saída do PCA (ver pag. 81 do slides sobre PCA), , para cada um dos 6 regressores
  - Matriz com os coeficientes de correlação das variáveis de entrada
  - Comentários explicando os resultados obtidos.