



## **COC786: Inteligência Computacional**

### **Relatórios do Projeto**

#### **1. Entrega 1: 07/05/2019: Relatório parcial + discussão em aula**

- Descrição do Problema
- Pesquisa Bibliográfica
- Descrição dos Dados
- Apresentação da tecnologia
- Avaliação preliminar dos dados (caracterização e visualização)

#### **2. Entrega 2: 04/06/2019: Relatório Final + Seminário**

- Descrição da solução do problema
- Descrição dos Dados
- Apresentação da tecnologia
- Resultados de Modelos Lineares
- Resultados de Modelos Não-Lineares
- Análise Comparativa
- Conclusões

## Caracterização e Visualização de Dados

- ✓ Preparar o Relatório contendo os resultados da análise do aluno:
  - Apresentação do software selecionado pelo aluno com informações básicas sobre a instalação, formato de arquivos e utilização.
  - Apresentação dos datasets escolhidos: apresentar as características principais do problema:
    - i. Descrição sucinta da aplicação;
    - ii. Tipo de problema (classificação ou regressão);
    - iii. Número de variáveis e de registros;
    - iv. Existência de valores ausentes;
  - Realizar a análise exploratória dos dados, dentro das possibilidades oferecidas pelo software escolhido e discutir:
    - i. Estatísticas básicas: mínimos, máximos, médias, desvios, etc... de cada variável.
    - ii. Avaliar os histogramas e as características das distribuições das variáveis.
    - iii. Verificar a existência de *outliers*.
    - iv. Verificar correlações.
  - **Apresentar as conclusões da análise realizada.**

## Modelos Preditivos Lineares

### Regressão

1. O objetivo deste exercício é realizar a Análise de Regressão Linear com o dataset escolhido pelo aluno utilizando as funcionalidades oferecidas pelo software.
2. Cada aluno deve explorar as opções disponíveis no software e realizar a análise comparativa dos resultados em validação cruzada de 10 ciclos.
3. Preparar o Relatório contendo os resultados da Análise:
  - Descrever sucintamente a formulação matemática do modelo linear generalizado e o algoritmo de ajuste de parâmetros.
  - Descrever o problema de regularização e os possíveis efeitos do mal condicionamento numérico nos resultados
  - Explorar diferentes alternativas de estrutura e ajuste de parâmetros para os modelos de regressão linear. Avaliar, se possível, os efeitos de “outliers” e sobre-ajuste do modelo (overfitting) nos resultados.
  - Avaliar criticamente os resultados em validação cruzada de 10 ciclos e concluir sobre o melhor modelo, apresentando as justificativas.
  - Discutir sobre os resultados obtidos utilizando elementos da análise estatística (Relatório 1)
  - **Apresentar suas conclusões a partir dos resultados obtidos.**

### Classificação

1. O objetivo deste exercício é realizar a Classificação Supervisionada com o dataset escolhido pelo aluno utilizando as funcionalidades oferecidas pelo software.
2. Cada aluno deve explorar as opções disponíveis no software e realizar a análise comparativa dos resultados em validação cruzada de 10 ciclos.
3. Preparar o Relatório contendo os resultados da Análise:
  - Descrever sucintamente a formulação matemática do modelo utilizado e o respectivo algoritmo de ajuste de parâmetros.
  - Descrever o problema de desbalanceamento e os seus efeitos nos resultados (se for o caso)
  - Explorar diferentes alternativas de estrutura e ajuste de parâmetros para os modelos de classificação. Avaliar, se possível, os efeitos de “outliers” e sobre-ajuste do modelo (overfitting) nos resultados.
  - Avaliar criticamente os resultados em validação cruzada de 10 ciclos e concluir sobre o melhor modelo, apresentando as justificativas.
  - Discutir sobre os resultados obtidos utilizando elementos da análise estatística (Relatório 1)
  - **Apresentar suas conclusões a partir dos resultados obtidos.**

**É proibida a cópia de texto, figuras, tabelas ou equações das notas de aula**

## Classificação e Regressão com Modelos Não Lineares

1. O objetivo deste exercício é realizar a Classificação e Regressão com os datasets escolhidos pelo aluno utilizando algoritmos de redes neurais ou SVM.
2. Para cada dataset (classificação e regressão), o aluno deve explorar diversas opções de topologias de redes neurais ou opções de Kernel e parâmetros do SVM e realizar a análise comparativa dos resultados em validação cruzada de 10 ciclos.
3. Os resultados obtidos devem ser comparados com os resultados de outras técnicas (classificação Bayesiana, modelos lineares e árvore de decisão), apresentadas nos Relatórios anteriores. Caso o aluno utilize um dataset diferente do que foi utilizado anteriormente, os resultados dos métodos anteriores devem ser refeitos para o novo conjunto de dados.
4. Preparar o Relatório contendo os resultados da Análise:
  - a) Descrever sucintamente a formulação matemática do modelo de redes neurais utilizado e o respectivo algoritmo de ajuste de parâmetros.
  - b) Apresentar sucintamente a caracterização estatística dos dados
  - c) Avaliar criticamente os resultados em validação cruzada de 10 ciclos e concluir sobre o melhor modelo, apresentando as justificativas.
  - d) Comparar os resultados obtidos com redes neurais com os resultados de outros modelos (classificação bayesiana, modelos lineares, árvore de decisão). **A comparação de resultados deve ser feita com base em métricas de avaliação definidas.**
  - e) Discutir sobre os resultados obtidos utilizando elementos da análise estatística (Relatório 1)
  - f) **Apresentar suas conclusões a partir dos resultados obtidos.**

**É proibida a cópia de texto, figuras, tabelas ou equações das notas de aula**