



Especificação da Primeira Entrega (E1) do Projeto Prático

Descrição

Ao longo da disciplina os estudantes desenvolverão um **projeto prático**, que conta com três entregas com apresentações (E1, E2 e E3). Este projeto prático visa o estudo de um conjunto de dados escolhido individualmente por cada grupo fazendo uso de técnicas de visualização.

Todos os estudantes devem se dividir em **grupos de até três estudantes** para executar o projeto prático. Os nomes dos membros de cada grupo devem ser informados conforme instruções disponíveis no Moodle da disciplina.

A **primeira entrega (E1) deste projeto prático serão nos dias 29 de abril e 2 de maio de 2022**. Na primeira entrega (E1), os grupos devem apresentar em aula:

- Conjunto de dados que escolheram para fazer seu projeto prático
- Técnicas de pré-processamentos que eventualmente já tenham sido executadas.

Lembre-se também que este será o conjunto de dados usado também na segunda e terceira entregas (E2 e E3), então escolha-o com cuidado. Não serão aceitas entregas atrasadas.

Ordem das Apresentações

A ordem de apresentação dos grupos será definida por sorteio a ser realizado no dia 25 de abril de 2022 durante a atividade síncrona nesse dia.

Especificações Mínimas

Escolha um **conjunto de dados de no mínimo 100 instâncias e 8 atributos**.

Observem que atributos derivados que denotem a mesma informação (por exemplo: data de nascimento de um indivíduo e idade do mesmo indivíduo) não contam duas vezes no número de atributos.

Dicas para sua Apresentação

Durante sua apresentação dos dados, leve os seguintes itens sob consideração:

- Sobre o conjunto de dados escolhido:

- Sobre que tipo de informação é o seu conjunto de dados?
- Descreva como foi coletado ou obtido o conjunto de dados. Cite fontes.
- Em que categoria se enquadra o seu conjunto de dados (multivariado, espacial, topológico ou hierárquico, textual, etc.)? Lembrando que um conjunto de dados pode pertencer a mais de uma categoria.
- Quantas instâncias o conjunto de dados apresenta?
- Quantos e quais atributos estão presentes nos dados? Que tipo de informação cada atributo registra?
- Que tipo de informação/conhecimento você busca explorar ou confirmar por meio das visualizações da sua base de dados?
- Descrever algum pré-processamento que já tenha sido realizado.
- Recomendo o uso de uma apresentação em slides para auxiliar em sua apresentação.

Avaliação

A primeira entrega de cada grupo será avaliada quanto à:

1. Conjunto de Dados (80 pontos):

- a. Identificação do **nome** atribuído e da **fonte** do conjunto de dados. (20 pontos)
- b. Apresentação das **características dos conjuntos de dados**. Exemplo: número de instâncias, número de atributos, categorias onde se enquadra. (30 pontos)
- c. Apresentação dos **atributos** explicando o que representa e o intervalo de valores. (30 pontos)
- d. **Apresentação de demais características do conjunto e/ou de seu pré-processamento**. Exemplos: se valores faltantes estão presentes nos dados, como são representados, quantas instâncias dentro da base possuem valores faltando, e como os valores faltando foram tratados; o conjunto de dados é classificado, quais são as classes e o número (porcentagem) de instâncias em cada classe. (10 pontos)

2. Comparação com os demais grupos observando a qualidade da apresentação. (20 pontos).