

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS

Instituto de Ciências Exatas e Informática

Curso de Ciência da Computação

Disciplina: Inteligência Artificial

Profa.: Cristiane Neri Nobre — nobre@pucminas.br

Descrição do trabalho Prático

Valor: 20 pontos

Cada grupo deverá escolher um tema a ser resolvido usando técnicas de **Machine Learning**.

Para isso, vocês precisam encontrar uma base de dados na PNS

Veja em: <https://www.pns.iciet.fiocruz.br/bases-de-dados/>

Leia mais sobre a PNS em: <https://www.pns.iciet.fiocruz.br/wp-content/uploads/2021/12/liv101846.pdf>

Caso não encontre nada do seu interesse, fique à vontade para escolher qualquer outro tema.

Serão formados os seguintes grupos:

- 6 grupos de 5 alunos e,
- 2 grupos de 4 alunos

Não é permitida a criação de grupos fora desta configuração.

Assim, favor arrastar o seu nome para o grupo que está no CANVAS, ok?

A partir da base selecionada, cada grupo deverá fazer:

1. Fazer um pré-processamento criterioso na base de dados (selecionar atributos, codificar, balancear, eliminar outlier, etc).
2. Utilizar alguns algoritmos vistos em sala: árvore de decisão, Random Forest, K-means, Redes Neurais, Apriori, etc para extrair conhecimento da base de dados. O grupo poderá investigar outros métodos também. Isso inclusive é muito bem vindo! 😊
3. Gerar o relatório, com os resultados da avaliação

Todas as entregas deverão estar no **formato da ACM**, disponível no CANVAS. Favor conferir o template.

O **relatório final** deverá possuir as seguintes seções:

1) Uma pequena contextualização do problema abordado

2) Explicação da metodologia adotada para resolver o problema

Descrição da base de dados, apresentando todos os atributos da base.

- a. São atributos binários?
- b. São atributos contínuos?
- c. São atributos nominais?
- d. Quais os possíveis valores para cada atributo?
- e. Quais os valores máximo e mínimo de cada atributo?
- f. A base é supervisionada? Ou seja, tem alguma classificação? Exemplo: base onde já tem uma classificação de pessoas gripadas ou não gripadas, com base nos atributos da base.
- g. Que método está sendo utilizado para resolução do problema?
- h. Que etapas de pré-processamento foram realizadas?

3) Resultados e discussões

Explicações dos resultados obtidos.

Apresentar os resultados das métricas de avaliação para o problema abordado.

- a. Se for uma base supervisionada, indicar os valores de precisão, sensibilidade, etc..
- b. Se for uma base não supervisionada, indicar os valores de métricas, como silhouette.

Os resultados, de acordo com os parâmetros alterados

- c. Rede neural necessita de alguns parâmetros e árvore de decisão de outros, por exemplo...

Discutir os resultados obtidos

4) Referências

O **relatório final** deverá ser entregue e apresentado, pelo CANVAS, até o dia **24 de maio**.

Por favor, para se inspirarem, vejam os artigos: **healthinf20241.pdf**, **healthinf20242.pdf** e **healthinf20243.pdf**, **ICMLA2019.pdf**, **SMC2019B.pdf** e **SMC2019A.pdf**, **health** que estão no **CANVAS**.

IMPORTANTE: Favor ler o documento

O trabalho será dividido nos **dois pontos de controle**:

Assunto	Valor	Data
<p>Implementação de um algoritmo de Aprendizado na base escolhida</p> <p>Apresentar todas as etapas de pré-processamento realizadas:</p> <p>Exemplos: Codificação dos atributos, identificação de outlier, inconsistência, visualização dos dados, balanceamento, normalização, redução de dimensionalidade, discretização, etc</p> <p>Implementar algum algoritmo visto em sala na base de dados escolhida e verificar os resultados a partir da base de dados selecionada.</p> <p>Avaliar o desempenho e fazer um pequeno relatório discutindo os resultados e discussões. O relatório deve conter todas as etapas realizadas.</p> <p>Para esta etapa, observe as recomendações contidas neste documento. Além disso, veja o documento abaixo para maiores informações:</p> <p>Template_ACM_Word_EXEMPLO.doc</p>	6 pontos	Até 31/03
<p>Relatório final</p> <p>O relatório final deverá conter:</p> <ol style="list-style-type: none">1) Uma contextualização do problema abordado2) Uma explicação da metodologia utilizada no trabalho. O que foi feito? Como? Que etapas de pré-processamento foram realizadas? Explicar, com detalhes, cada escolha realizada.3) Apresentar os resultados e discussões obtidos.4) O trabalho final precisa utilizar pelo menos 3 algoritmos de aprendizado de máquina diferentes.	14 pontos	Até 24/05

Qualquer dúvida, entre em contato com nobre@pucminas.br