

Descrição do Trabalho 3 – Aplicação de algoritmos de aprendizagem

Objetivo: Aplicar e comparar algoritmos de aprendizagem

Instruções:

- 1- O grupo deverá selecionar um “dataset” (em sites de IA ou através da busca de artigos nacionais ou não), onde se indica ou disponibiliza datasets e se descreve a aplicação dos algoritmos de aprendizagem Árvore de Decisão e/ou Redes Neurais.
- 2- Independente do uso de artigos, o grupo deverá realizar experimentação, atribuindo diferentes valores aos parâmetros necessários à execução de cada algoritmo. Mas, se algum artigo foi consultado, pode-se extrair destes os parâmetros necessários para aplicação dos algoritmos. **Vale ressaltar** que a experimentação é necessária, independente de ter os parâmetros fornecidos por artigos.
- 3- A experimentação deve ser concluída, quando obtiver acurácia significativa, levando em consideração o tamanho e o número de atributos do dataset.
- 4- Comparar as experimentações (pelo menos duas) realizadas com cada algoritmo, utilizando medidas estatísticas como a acurácia e a matriz de confusão. Comparar os melhores resultados dos dois algoritmos, em termos destas mesmas medidas. As redes neurais resultantes das experimentações devem ser apresentadas, graficamente.
- 4- Preparar apresentação (em formato para apresentação) contendo descrição do dataset, os parâmetros utilizados em cada experimentação de cada algoritmo, as comparações do item anterior, bem como as medidas estatísticas mencionadas no item anterior, finalizando com conclusão relativa à comparação dos resultados e aos algoritmos utilizados. Caso tenha sido feito uso de artigos, apresentar e comparar o resultado do trabalho com os resultados do artigo, independente de serem diferentes os algoritmos utilizados no artigo.

Restrições:

- 1- Cada grupo deve apresentar dataset diferente dos demais grupos.
- 2- **Selecione artigos onde foram utilizados árvore de decisão ou rede neural.** Preferencialmente, redes neurais pois necessitam de vários parâmetros para execução.
- 3- **Selecione datasets que sejam públicos, mas confiáveis, que os dados dos atributos e classes estejam completos. Além disso, o número de atributos e amostras/exemplos/entradas deveriam ser interessantes. Sugestão: Pelo menos dezenas de atributos e milhares de amostras.**
- 4- **Datasets muito pequenos ou os mnist disponíveis no Python devem ser evitados.**
- 5- **Os algoritmos de aprendizagem não devem ser implementados. Ferramentas para execução dos algoritmos podem ser utilizados. Sugestão: Weka (link disponibilizado em aula) ou Python e suas bibliotecas.**

Observações:

- 1- **Dois algoritmos foram recomendados para o trabalho: Árvore de Decisão e Redes Neurais. Mas, caso consigam utilizar algum outro algoritmo que se mostre mais adequado ao tipo de dados, isto pode ser feito.**
- 2- **Em aula, exemplos de parâmetros do Weka para Redes Neurais e para algoritmos de aprendizagem em geral foram apresentados. Estejam atentos a estes parâmetros, com por exemplo, Validação cruzada.**