

Atividade 4 – Experimentador do WEKA.

1 Descrição

Executar uma série de experimentos de validação cruzada (VC), utilizando o WEKA Experimenter. Para cada combinação (algoritmo \times conjunto de dados) um total de 5 (cinco) experimentos de VC devem ser realizados. Cada experimento de VC deve usar 10 *folds*.

Você deve utilizar 10 classificadores diferentes. Dentre os classificadores utilizados, você deve incluir, pelo menos, um classificador do tipo **Support Vector Machine (SMO)**, pelo menos um do tipo **Random Forest**, pelo menos um do tipo **JRip**, pelo menos um do tipo **J48**, e pelo menos um do tipo **Multilayer Perceptron** (disponível na pasta **functions** dos classificadores do WEKA).

Perceba que você pode ter dois ou mais classificadores de um mesmo tipo (por exemplo, do tipo J48), desde que as configurações de hiperparâmetros sejam distintas. Por exemplo, podemos ter um classificador J48 com `reducedErrorPruning=False` e outro com `reducedErrorPruning=True`.

Os conjuntos de dados usados nos experimentos devem ser os seguintes:

- **breast-cancer**: disponível na pasta **data** do WEKA;
- **credit-g**: disponível na pasta **data** do WEKA;
- **Ecoli**: disponível no UCI Machine Learning repository;
- **ILPD (Indian Liver Patient Dataset)**: disponível no UCI Machine Learning repository.

2 Entrega

- Pontuação total pela atividade: 1,5 pontos (um ponto e meio).
- Entregáveis:
 - Arquivo em formato CSV, gerado automaticamente pelo WEKA Experimenter, contendo todas as informações dos experimentos de validação cruzada realizados. O arquivo **não deve ser do tipo ARFF** (que é a opção padrão do WEKA Experimenter), mas **sim do tipo CSV**.
 - Relatório em formato PDF contendo uma tabela com a acurácia (`Percent_correct`) média de cada um dos algoritmos em cada um dos conjuntos de dados. O relatório deve incluir as **configurações de hiperparâmetros** usadas para cada um dos 10 algoritmos (no caso de valores *default*, é suficiente especificar “*default*”).
- Data de entrega: **4 de junho** de 2019, 23:59:59, via SIGAA. **Não** via e-mail.
- Data de entrega com adiamento automático: **5 de junho** de 2019, 23:59:59, via SIGAA. **Não** via e-mail.

3 Trilha Sonora Sugerida

<https://www.youtube.com/watch?v=pl3vxEudif8>

4 Mais detalhes

- A atividade pode ser feita individualmente ou em dupla.