

Fazer as tarefas solicitadas abaixo, usando os algoritmos de aprendizado em fluxo de dados e os geradores de streams disponíveis na ferramenta MOA (<https://moa.cms.waikato.ac.nz/>) e de acordo com as recomendações a seguir:

- 1) Compare e discuta a acurácia da Árvore de Hoeffding para um fluxo gerado por RandomTreeGenerator com 1.000.000 instâncias e frequência de amostra de 10.000, usando três avaliações diferentes:
 - a) Periodic Held Out com 1.000 instâncias para teste
 - b) Interleaved Test Then Train
 - c) Prequential com janela de 800 instâncias.
- 2) Compare e discuta a acurácia da Árvore de Hoeffding para o mesmo fluxo do exercício anterior usando duas avaliações diferentes:
 - a) Prequential com fading factor de 0,99.
 - b) Prequential com fading factor de 0,999.
- 3) Compare e discuta a acurácia da Árvore de Hoeffding para um fluxo gerado por RandomRBFGenerator com 1.000.000 instâncias usando Interleaved Test Then Train e dois algoritmos diferentes:
 - a) Hoeffding Tree com classe majoritária nas folhas
 - b) Hoeffding Adaptive Tree
- 4) Execute uma tarefa de agrupamento com o gerador RandomRBFGeneratorEvents selecionado o algoritmo CluStream como Algoritmo 1 e algoritmo WithKMeans como Algoritmo 2. Utilize os parâmetros default para todas as definições, selecione algumas métricas de avaliação, selecione a aba “Visualize”, execute os algoritmos e responda: Como são representados os pontos, os microgrupos, os grupos encontrados e os grupos de referência (são conhecidos para o gerador utilizado)?
- 5) Execute uma tarefa de agrupamento com as mesmas definições do exercício anterior, apenas incluindo as alterações: velocidade = 1000, speedRange = 10 e nível de ruído = 0.333. Selecione a métrica de avaliação SSQ e observe o que acontece com essa métrica quando aumenta o número de microgrupos (usar 150 e 200) do algoritmo Clustream(maxNumKernels), selecionado como Algoritmo 1.
- 6) Execute uma tarefa de agrupamento com as mesmas definições do exercício anterior, definindo CluStream como Algoritmo 1 com as definições (número de micro cluster) que resultaram no melhor valor de SSQ. Selecione Clustree como Algoritmo 2. Qual deles apresenta melhor valor de SSQ?

Entregar um relatório contendo:

- Para cada exercício de classificação (1 a 3), incluir:
 - o print da janela inferior do MOA, que contém os gráficos e as medidas de avaliação;
 - Uma descrição breve da conclusão observada sobre o aspecto questionado no exercício.
- Para cada exercício de agrupamento (4 a 6), incluir:
 - O print da aba “Visualize” do MOA, que contém as figuras dos clusters e as medidas de avaliação;
 - Uma descrição breve da conclusão observada sobre o aspecto questionado no exercício.

Observações:

- O trabalho pode ser feito em duplas.
- Site do MOA: <https://moa.cms.waikato.ac.nz/>
- Entregar (no moodle): Relatório, com o conteúdo detalhado anteriormente.
- DATA DE ENTREGA: 22/11/2017