



UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE TECNOLOGIA

Peso: 6

Disciplina: Mineração de Dados (ELC1098)

Professor: Dr. Joaquim Assunção

Trabalho prático 1



Duplas.



Descrição

O **kendo** é uma arte marcial japonesa moderna que significa literalmente “*caminho da espada*”. Deriva das antigas técnicas dos samurais e busca unir treinamento físico, disciplina mental e desenvolvimento do caráter. Os praticantes usam espadas de bambu (**shinai**) e vestem armaduras protetoras (**bogu**) para simular combates de forma segura. Em competições, há somente 4 tipos de golpes que valem pontos: **Men (面)** – golpe na cabeça, **Kote (小手)** – golpe no punho, **Dō (胴)** – golpe no tronco, na lateral do corpo, e **Tsuki (突き)** – estocada na garganta.

Um grupo de praticantes resolveu salvar o resultado das lutas em um arquivo. Após alguns anos, gerou-se o dataset “kendō_matches_work.csv”. O dataset é passível de erros, caso encontre algum erro ou inconsistência, desconsidere a tupla.

Sua tarefa é minerar regras que respondam as seguintes perguntas (Obs. ignore a métrica coverage).

- 1) Qual golpe, em regra, com mais de mais de 90% de confiança, implica o não uso de outro golpe?
- 2) Qual a confiança da regra: um atleta que aplicou um *men* levou um *ippon*?
- 3) Qual(is) golpe(s), em regra, com mais de 85% de confiança, tende a demorar mais de 25 segundos após o primeiro golpe?
- 4) Se o primeiro golpe demorou mais do que 15 segundos, e o segundo golpe demorou mais do que 25 segundos. Qual golpe provavelmente não foi usado com sucesso? Mostre a regra que suporta sua resposta.

Entrega

Um documento contendo um resumo do que foi feito como pré-processamento (2 a 4 páginas), as respostas, e as regras que suportam as respostas. Em anexo deve estar seu script para transformação dos dados e geração das regras (caso haja). O envio deve ser por e-mail, com o título do PDF contendo o nome dos membros no seguinte formato “[DM] alunoA, alunoB”.

Avaliação

A avaliação será dada de acordo com o PDF entregue, contendo a descrição do processo e as regras geradas com suas respectivas métricas. As regras valem 5 pontos, os outros 5 pontos serão referentes ao processo, o script e a qualidade do PDF entregue. ... Caso encontre alguma outra regra interessante, isso contará positivamente.

Caso seja feito por ferramentas externas (e.g., WEKA), o documento deve ser maior, contendo cada passo realizado.

Prazo

O trabalho deve ser entregue por **e-mail** até dia **23/set** às 11:59.