

## 6ª. LISTA DE EXERCÍCIOS

- (1) **Usando o Weka, crie os classificadores listados abaixo usando a representação *Bag of Words* e mostre uma tabela comparativa considerando as seguintes métricas Precision, Recall, F-measure e Acurácia.**

Classificadores a se avaliar:

- Naïve Bayes
- k-NN (use  $k = 5$ )
- Decision Tree (J48)
- Decision List (PART)
- Support Vector Machine (SMO)
- Logistic Regression

### Considere:

- O dataset: *Polarity Detection*.
- Uso de *Stoplist* e *Stemming* como pré-processamento
- Usar os métodos de ponderação das palavras (cálculo de frequência)
  - *Multinomial*, onde se registra, na matriz de representação dos documentos, a frequência do token (peso = 0 até n), como está nos slides da aula.
  - *Binário*, considera apenas se o termo aparece ou não (peso = 0 ou 1)
- Avalie o seu classificador usando 5-fold Cross Validation ( $N = 5$ )

### Material extra de apoio para os exercícios

Ver os links no arquivo **Links.txt**.