

## Tarefa: aprender a classificar textos pela abordagem NB

A partir do conjunto de dados fornecidos, construir um classificador de textos segundo a abordagem Naïve Bayes (NB).

Para tanto, organize o conjunto de dados para permitir a validação cruzada de 10 vezes (10-fold-cv).

Construa 10 classificadores NB para cada subconjunto de treinamento e teste o seu desempenho preditivo sobre o respectivo subconjunto de teste.

Apresente a matriz de confusão média correspondente ao teste, juntamente com os respectivos desvios padrões.

A partir da matriz de confusão média, calcule as seguintes métricas:

Precisão, Taxa de Verdadeiros Positivos (TVP), Taxa de Falsos Positivos (TFP), Medida-F

A partir dos arquivos disponibilizados, cada aluno deverá usar os arquivos listados abaixo para formar o seu conjunto de dados para esta tarefa.

A tarefa deve ser realizada usando qualquer linguagem de programação, mas **não** utilizando funções específicas para aprendizado bayesiano, como as disponíveis na ferramenta Weka, por exemplo.

Entregar relatório contendo detalhes da implementação e os resultados obtidos, conforme previsto na tarefa, bem como os códigos fonte e executável. **Prazo de entrega: 18/05/2015.**

Nome	Arquivos positivos 24000	Arquivos negativos 24000
André Luis	0 – 23999	0 – 23999
Bernardo	1 – 24000	1 – 24000
Cássio Alan Garcia	2 – 24001	2 – 24001
Cássio Pazinato	3 – 24002	3 – 24002
Cesar	4 – 24003	4 – 24003
Cristiano	5 – 24004	5 – 24004
Éderson Frozi	6 – 24005	6 – 24005
Eduardo	7 – 24006	7 – 24006
Guilherme	8 – 24007	8 – 24007
Jackson	9 – 24008	9 – 24008
Jhonny	10 – 24009	10 – 24009
José Luis	11 – 24010	11 – 24010
Leonardo	12 – 24011	12 – 24011
Márcio	13 – 24012	13 – 24012
Matheus	14 – 24013	14 – 24013
Renata	15 – 24014	15 – 24014
Rodrigo	16 – 24015	16 – 24015
Vinicius	17 – 24016	17 – 24016