

# Extração de Características e Naive-Bayes

## ENG121 - Reconhecimento de Padrões

Autor: Matheus Castro Silva

### RESUMO

O trabalho a seguir trata da resolução de um problema de reconhecimento de padrões na qual a base de dados utilizada consiste em um grupo de 5 pessoas que contém 10 fotos de cada pessoa, utilizando um classificador e um método para extração de características. O ambiente de programação utilizado para desenvolver o código e gerar os gráficos foi o RStudio. A classificador escolhido foi o de Naive Bayes e as características abordadas na classificação foi a média e o desvio padrão da luminância.

**Palavras Chaves** - algoritmos, Naive Bayes, classificação, reconhecimento de padrões, faces.

### I. INTRODUÇÃO

Nesta atividade, foi utilizado um subconjunto de imagens retiradas da base de dados The Database of Faces, da universidade de Cambridge para realização do trabalho. São selecionadas imagens referentes a 5 pessoas. As imagens estão disponíveis em arquivo .zip no Moodle da disciplina. E foi aplicado o classificador de Naive-Bayes para identificação de cada uma das classes.

### II. DEFINIÇÃO DO PROBLEMA

Dado o conjunto de pessoas a serem identificadas, para cada uma delas estão disponíveis 10 arquivos de imagem, sendo 5 com nome em letras maiúsculas e 5 com nome em minúsculas. Para as 50 imagens, foi pedido que:

- Extrairmos duas características referentes às imagens que possam ser utilizadas no contexto de problemas de classificação (ou seja, características com boa capacidade de discriminação).
- Gerarmos o gráfico mostrando as amostras para os atributos extraídos e o contorno de separação;

### III. EXTRAÇÃO DE CARACTERÍSTICAS

Os dois atributos escolhidos que separam bem os dados são média e desvio padrão das luminâncias.

#### A. Média da Luminância

Consiste em pegar cada ponto da matriz de luminância referente a cada imagem e realizar a média desses pontos. Como a fórmula mostra abaixo :

$$Média = \frac{\sum_{i=1}^N \sum_{j=1}^M M_{ij}}{N * M}$$

#### B. Desvio Padrão da Luminância

Consiste em realizar o desvio padrão de todos os pontos da matriz de luminância referente a cada imagem. Como a fórmula mostra abaixo :

$$Desvio\ Padrão = \sqrt{\frac{\sum_{i=1}^N \sum_{j=1}^M (M_{ij} - M_{médio})^2}{N * M}}$$

### IV. ESTIMAÇÃO DA FUNÇÃO DENSIDADE

A função densidade de cada uma das classes foi gerado utilizando o estimador de densidade de Naive-Bayes mostrado na equação abaixo:

$$p(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right)$$

em que  $\mu$  e  $\Sigma$  são, respectivamente, média e matriz de covariância para a classe em questão.

### V. RESULTADOS

As imagens a seguir conterá os resultados encontrados para os tópicos III e IV.

Ilustração das características extraídas

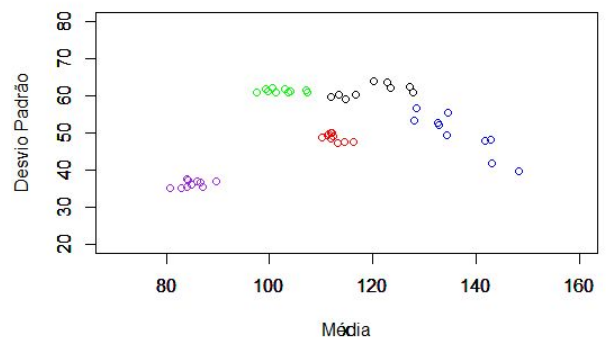


Imagem 1: Ilustração das características extraídas no plano 2D.

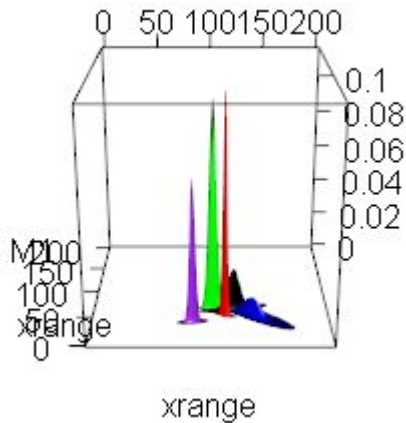


Imagem 2: Ilustração das curvas de densidades referente a cada classe no plano 3D vista frontal.

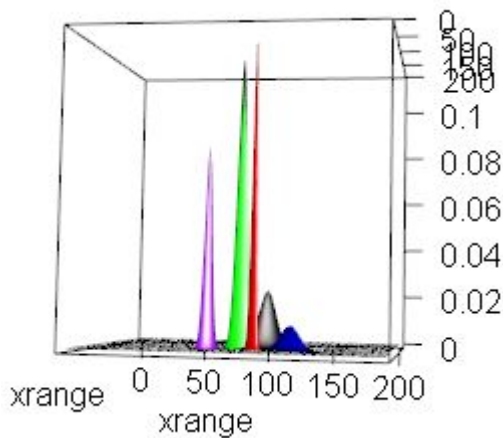


Imagem 3: Ilustração das curvas de densidades referente a cada classe no plano 3D vista lateral.

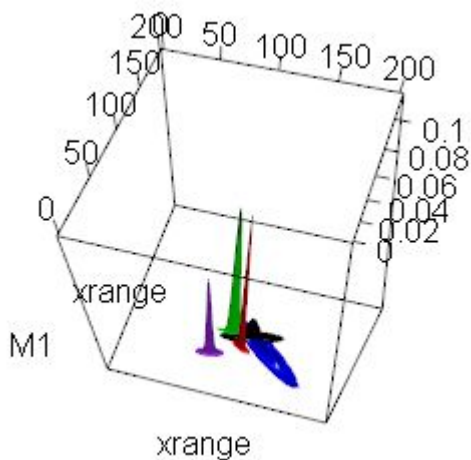


Imagem 4: Ilustração das curvas de densidades referente a cada classe no plano 3D vista superior.

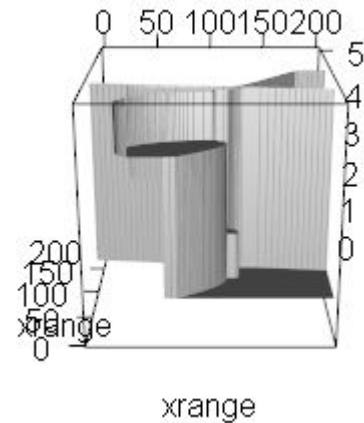


Imagem 5: Ilustração da curva de separação das curvas de densidades referente a cada classe no plano 3D vista frontal.

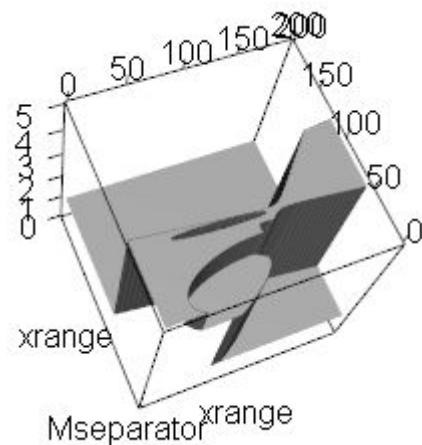


Imagem 6: Ilustração da curva de separação das curvas de densidades referente a cada classe no plano 3D vista superior.

#### Ilustração das curvas de separação

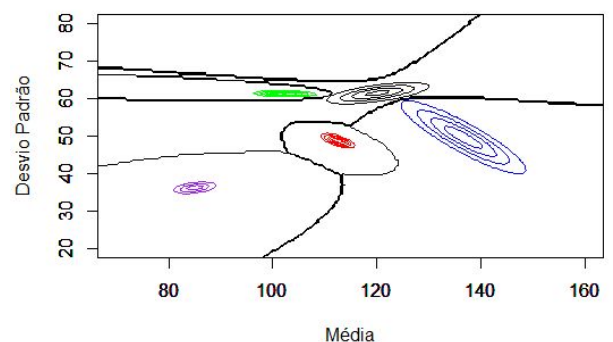


Imagem 7: Ilustração da curva de separação das curvas de densidades referente a cada classe no plano 2D vista superior.

## VI. CONCLUSÕES

Os classificadores Bayes ingênuos são uma família de simples classificadores probabilísticos com base na aplicação do teorema de Bayes com fortes pressupostos de independência entre os recursos.

Apartir das imagens 2 a 8, é de fácil visualização que as classes estão próximas umas das outras, porém não sobrepostas e com isso o classificador conseguiu gerar uma curva de separação bastante interessante para o problema proposto.

Utilizando este classificador e do conjunto de imagens referente a cada classe, foi possível ter ótimos resultados e bastante eficiência perante o classificador de Bayes.

## VII. REFERÊNCIAS

[1] R tutorial, <http://www.r-tutor.com/r-introduction>, acessado 30/08/2017.

[2] The Database of Faces - <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>, acessado 30/08/2017.