

# Identificação Automática de Espécies de Pássaros

---

Felipe Felix ( $f^2$ )

July 24, 2017

IME-USP

Processo de classificação:

1. O canto do pássaro é gravado no campo;
  2. O áudio é pré-processado para melhorar a qualidade do sinal;
  3. Extrai-se as características do sinal de áudio;
  4. Então, utiliza-se as características em algoritmos de aprendizagem de máquina para produzir um procedimento de decisão para novas gravações.
- *Framework* clássico de aprendizagem supervisionada.

# Formulação probabilística do problema

- Dado um sinal de áudio  $\mathcal{S}$  contendo um canto de pássaro, precisamos escolher uma classe  $\hat{b}$  de um conjunto finito de classes  $\mathcal{B}$  que melhor representa a espécie que produz aquele canto.
- Seja  $\bar{X}$  o vetor de características derivado de  $\mathcal{S}$ , queremos determinar a classe  $\hat{b} \in \mathcal{B}$  tal que:

$$\hat{b} = \arg \max_{b \in \mathcal{B}} P(b \mid \bar{X}) \quad (1)$$

# Formulação probabilística do problema

- Pelo Teorema de Bayes:

$$\hat{b} = \arg \max_{b \in \mathcal{B}} \frac{P(\hat{X} | b)P(b)}{P(\hat{X})} \quad (2)$$

- Sabemos que  $\sum_{b \in \mathcal{B}} P(b | \bar{X}) = 1$ . Então  $P(\bar{X}) = \sum_{b \in \mathcal{B}} P(\bar{X} | b)P(b)$ .
- Então a probabilidade desejada é:

$$P(b | \bar{X}) = \frac{P(\bar{X} | b)P(b)}{\sum_{b \in \mathcal{B}} P(\bar{X} | b)P(b)} \quad (3)$$

- Como o denominador da Eq. 3 é o mesmo para todas as classes, a solução é dada pela classe

$$\hat{b} = \arg \max_{b \in \mathcal{B}} P(\hat{X} | b)P(b) \quad (4)$$

- Espécies de pássaros de uma região geográfica comum: Sul da Costa Atlântica Brasileira.
- *Dataset* principal: 1619 gravações de 73 espécies de pássaros.
- Gravações baixadas do Xeno-Canto<sup>1</sup>.
- *Dataset* secundário: derivado do principal. As gravações do *dataset* principal foram divididas em *pulsos*.
- Pulso: Intervalo pequeno de som com altas amplitudes.

---

<sup>1</sup><http://www.xeno-canto.org/>

- *Framework* Marsyas<sup>2</sup>;
- 50+ *features*;

---

<sup>2</sup><http://marsyas.info/doc/sourceDoc/html/index.html>

- *5-fold cross-validation*: os resultados são obtidos de 5 repetições aleatórias;
- Os experimentos variam em três dimensões: sinal completo ou uso dos pulsos; uso de diferentes classificadores; número de classes.

**TABLE I**  
F-MEASURE ON THE FULL AUDIO RECORDS DATASET (%)

Classifier	Number of considered species				
	3	5	8	12	20
Naïve Bayes	61.5	50.7	27.0	25.3	25.4
$kNN$ ( $k = 3$ )	61.4	53.4	41.5	33.1	33.0
J4.8	50.6	41.7	29.4	28.2	26.9
MLP	69.6	69.6	55.0	<b>48.8</b>	<b>47.4</b>
SMO (Polynomial)	<b>73.2</b>	<b>73.2</b>	<b>57.3</b>	47.2	46.4
SMO (Pearson)	67.6	59.5	51.8	42.3	42.7

**TABLE II**  
F-MEASURE ON THE PULSES DATASET (%)

Classifier	Number of considered species				
	3	5	8	12	20
Naïve Bayes	45.9	32.9	27.4	24.8	17.6
$kNN$ ( $k = 3$ )	93.4	88.1	83.8	81.9	77.3
J4.8	87.4	76.9	74.1	67.3	60.2
MLP	94.6	88.4	82.4	76.2	68.3
SMO (Polynomial)	85.5	75.0	72.2	65.8	59.6
SMO (Pearson)	<b>95.1</b>	<b>89.3</b>	<b>85.7</b>	<b>82.9</b>	<b>78.2</b>



TABLE III

F-MEASURE ON THE FULL AUDIO DATASET – RANDOM CLASSES (%)

Classifier	Number of considered species		
	3	5	8
Naïve Bayes	68.9	59.2	39.0
$kNN$ ( $k = 3$ )	80.9	66.2	42.2
J4.8	60.4	51.8	36.9
MLP	<b>88.6</b>	71.4	<b>65.3</b>
SMO (Polynomial)	87.5	<b>75.2</b>	60.9
SMO (Pearson)	60.0	65.0	42.9

TABLE IV

F-MEASURE ON THE PULSES DATASET – RANDOM CLASSES (%)

Classifier	Number of considered species		
	3	5	8
Naïve Bayes	53.1	33.6	33.8
$kNN$ ( $k = 3$ )	95.1	78.3	87.2
J4.8	86.4	69.9	72.8
MLP	<b>96.4</b>	<b>83.1</b>	87.4
SMO (Polynomial)	94.4	69.9	76.3
SMO (Pearson)	95.4	78.5	<b>89.7</b>

[1] Lopes, Marcelo T., et al. "Automatic bird species identification for large number of species." Multimedia (ISM), 2011 IEEE International Symposium on. IEEE, 2011.