

# Engenharia de Características Baseadas em Cifras para a Classificação de Gêneros na Música Brasileira

Bruna Wundervald e Rogério Hultmann

Orientador: Prof. Dr. Walmes Marques Zeviani

Departamento de Estatística  
Universidade Federal do Paraná

Junho de 2018



# Sumário

[Introdução](#)

[Objetivos](#)

[Materiais e Métodos](#)

[Resultados](#)

[Considerações Finais](#)

[Referências Bibliográficas](#)

1

# Introdução

# Música

## ► Música e Harmonia

- ▶ **Música:** elemento cultural intrínseco em nossa sociedade.
- ▶ Acordes: grupo de 3 ou mais notas com uma estrutura específica.
- ▶ Harmonia: ciência dos sons que são expressos simultaneamente.

## ► *Music Information Retrieval*

- ▶ Desenvolve e aplica ferramentas computacionais combinadas com teoria e conceitos musicais.
- ▶ Pretende ampliar a compreensão e utilidade dos dados de música.

# Dados e Gêneros Musicais

- ▶ Cada formato de dados tem suas respectivas propriedades
  - ▶ MIDI: traz as informações de quais notas são tocadas em cada tempo, principalmente.
  - ▶ Cifras: trazem as informações dos acordes.
  - ▶ Partituras: representação padronizada mundial de música.
- ▶ Gêneros musicas
  - ▶ Classificações atribuídas às músicas de forma geral.
  - ▶ Não existe uma definição exata para cada classe.

2

## Objetivos

# Objetivos

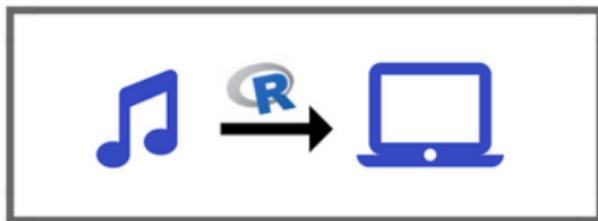
1. Propor um método eficaz de extração de cifras musicais e disponibilizar um pacote em R.
2. Extrair os dados e realizar engenharia de características.
3. Realizar uma breve análise da estrutura harmônica de composições brasileiras.
4. Avaliar, através de modelagem, quais são as características mais relevantes para a classificação das músicas nos gêneros musicais.

3

## Materiais e Métodos

# Web Scraping

- ▶ Técnica de extração de dados da *web*.
- ▶ Captura os dados através das estruturas em *HTML* das páginas.
- ▶ Dados coletados do Cifraclub  
[\(https://www.cifraclub.com.br/\)](https://www.cifraclub.com.br/).
- ▶ Pacote `chorrrds`.



# Base Inicial

Tabela: Extrato dos dados não tratados.

musica	acorde	escala	artista
chico buarque a banda	D6/9	D	chico buarque
chico buarque a banda	A7	D	chico buarque
chico buarque a banda	F#m7	D	chico buarque
chico buarque a banda	B7	D	chico buarque
chico buarque a banda	E7(9)	D	chico buarque
chico buarque a banda	A7	D	chico buarque
chico buarque a banda	D6/9	D	chico buarque

# O círculo das quintas

Indica os acordes mais prováveis nas músicas:

- Se uma música está na escala de Dó (ou C), os acordes mais frequentes devem ser o Dó (C), Sol (G) e Fá (F).

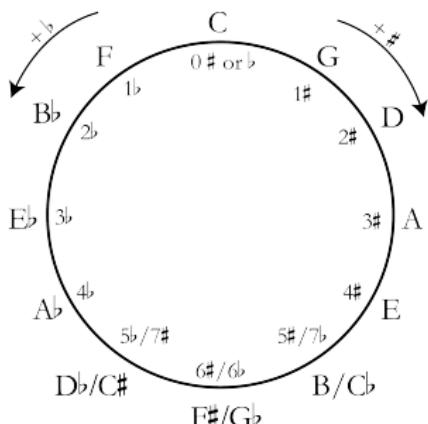


Figura: Círculo das quintas

# Engenharia de Características

- ▶ Transformação dos variáveis em características que melhor representam o problema adjacente.
- ▶ Automatizada ou não; interpretável ou não.
- ▶ Motivação: avaliar quais conjuntos de características são mais importantes na classificação em gêneros.

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix} = f \left( \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \right) \quad (1)$$

# Árvores de Classificação

- ▶ Partições do espaço das covariáveis em regiões retangulares e ajuste de um modelo simples em cada uma delas.
- ▶ O ajuste é prever a classe de cada região como sendo a mais comum nela.
- ▶ Completamente não paramétrico.

Busca minimizar o critério de impureza de Gini:

$$Gini = 1 - \sum_{i=1}^r p_i^2 \quad (2)$$

aonde cada  $p_i$  é a proporção com que a categoria  $i$  é observada na base. É minimizado quando existe uma classe predominante na região.

# Árvores de Classificação

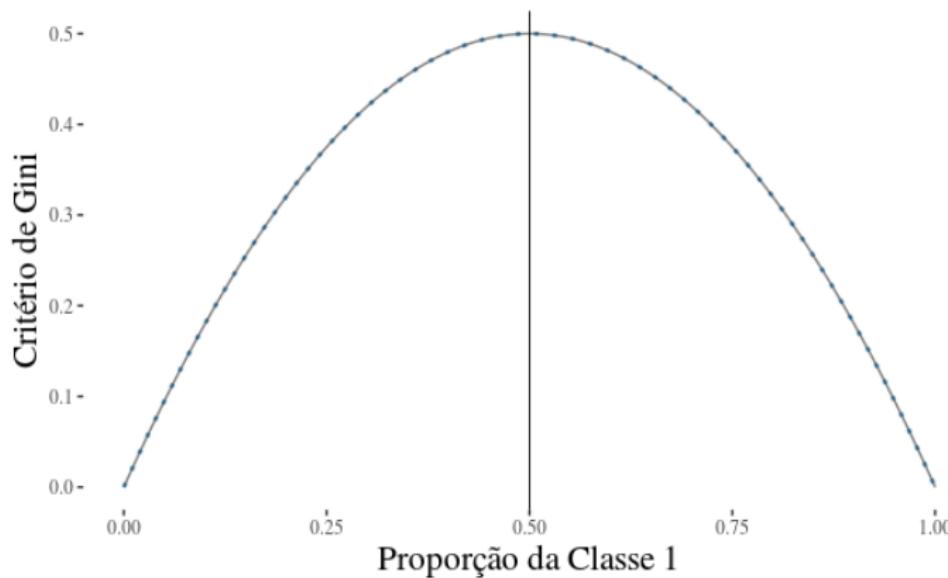


Figura: Comportamento do critério de Gini.

# Árvores de Classificação

Algoritmo:

1. Em uma preditora é feita uma partição, escolhendo-se um ponto de quebra.
2. Para cada partição, encontra-se a classe mais comum naquela região, que é a previsão.
3. Busca-se o ponto de quebra que cause a maior diminuição no critério de Gini.
4. As partições são sub-partidas, de forma recursiva.

Muitos pontos de quebra são avaliados, mas os cálculos em cada partição são simples → permite que o modelo seja ajustado sem maiores esforços computacionais.

# Árvores de Classificação

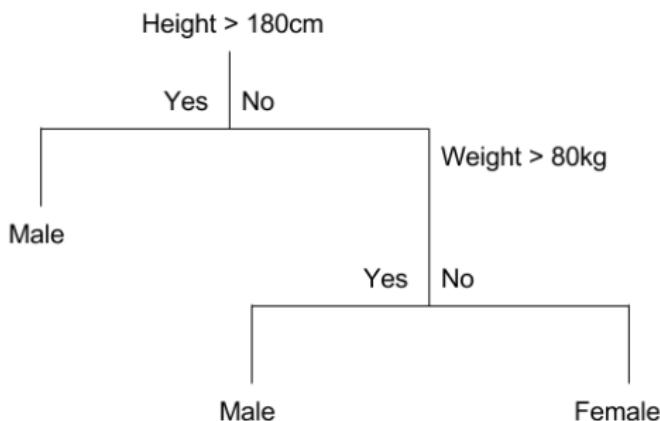


Figura: Exemplo de árvore de classificação binária.

# Árvores de Classificação

Vantagens:

- ▶ Fácil interpretação.
- ▶ Forma um conjunto de regras próximo da decisão humana.
- ▶ Captura bem regiões de previsão não lineares.

Qualidade de predição: avaliada através da acurácia.

$$\text{Acurácia} = \frac{1}{n} \sum_{i=1}^n I(y_i \neq \hat{y}_i) \quad (3)$$

aonde  $I$  indica se a predição feita pelo modelo,  $\hat{y}_i$ , é compatível com o que foi observado,  $y_i$ , e  $n$  é o tamanho amostral.

## Florestas Aleatórias

São obtidos  $B$  conjuntos de treinamento, provenientes da amostra original, através de *bootstrap*. Esta técnica é chamada de *bagging*:

- ▶ Para cada conjunto, obtém-se um modelo.
- ▶ Os modelos são combinados, gerando o ajuste final.
- ▶ A previsão é aquela que estiver em concordância com a maioria dos  $B$  modelos.
- ▶ Reduz a variância das árvores para  $\sigma^2/n$ .

$$\hat{f}_{\text{bagging}}(x) = \frac{1}{B} \sum_{b=1}^B \hat{f}^{*b}(x) \quad (4)$$

## Florestas Aleatórias

Florestas Aleatórias: *bagging* com descorrelacionamento das árvores.

- ▶ A cada divisão, apenas  $\sqrt{p} \approx m$  variáveis são consideradas,  $p$  sendo o total de variáveis.
- ▶ As  $m$  candidatas são selecionadas aleatoriamente e sem reposição.

Evita que as variáveis muito fortes/discriminantes apareçam na primeira divisão da maioria das árvores, pois isso torna seus resultados correlacionados.

# Importância de Variáveis

- ▶ A cada nova divisão, a diminuição no critério de Gini é contabilizada para as variáveis.
- ▶ Ao fim do algoritmo, a média desta diminuição, para cada variável, é acessada.
- ▶ Esta média informa o quanto cada variável é importante para a classificação.

4

## Resultados

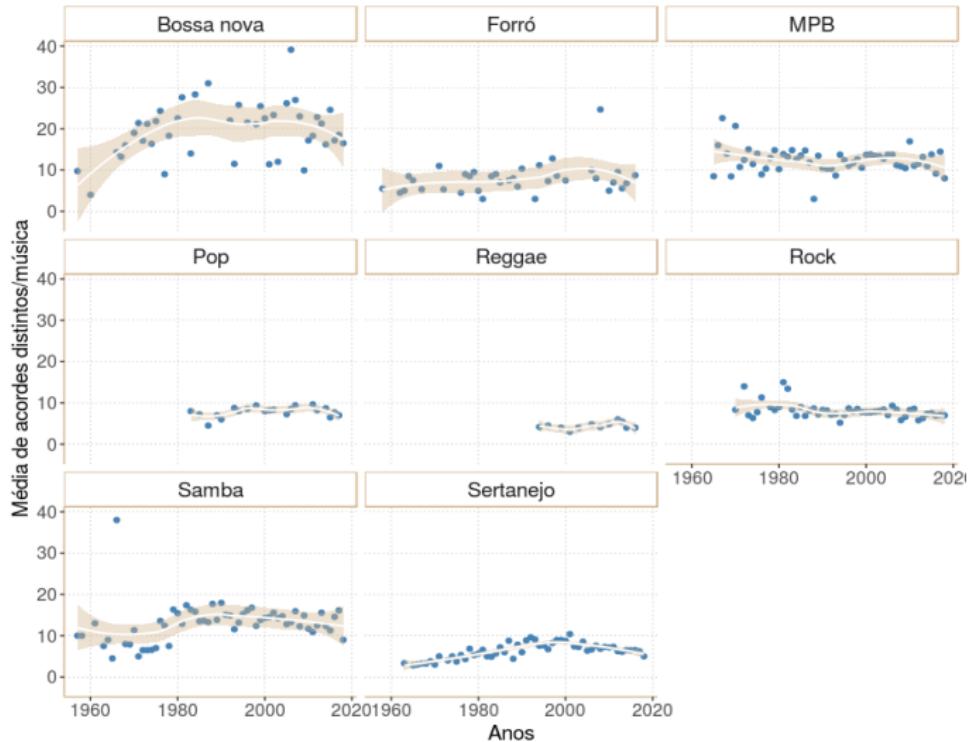
# Recursos Computacionais



R, versão 3.4.4

- ▶ **chorrrds**: pacote desenvolvido para a extração de cifras no R (Wundervald (2018)). Disponível no *CRAN*.
- ▶ **r-music**: organização que tem o objetivo de concentrar estudos sobre música e estatística com o R (Wundervald e Trecenti (2018)).

# Análise Exploratória



**Figura:** Quantidade média de acordes distintos em cada ano, com separação por gêneros musicais.

# Análise Exploratória

- ▶ Samba, MPB e bossa nova têm uma quantidade média de acordes distintos por música maior do que os demais.
- ▶ Os outros gêneros são mais uniformes harmonicamente.
- ▶ Gêneros mais simples apareceram em períodos posteriores (CALDAS, 2010).

# Análise Exploratória

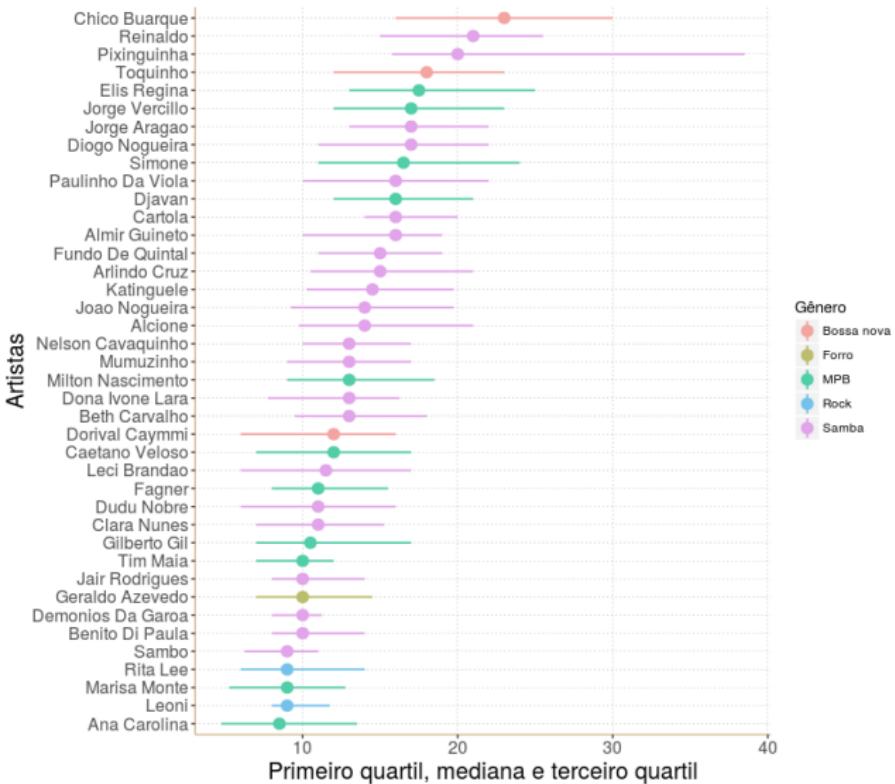


Figura: Mediana de acordes distintos por artista, com intervalo quantílico

## Análise Exploratória

- ▶ A primeira posição é de um dos principais musicistas brasileiros, referência internacional em refinamento musical: Chico Buarque.
- ▶ Em seguida, temos o Reinaldo, um dos maiores sambistas que o Brasil já conheceu, e o Pixinguinha, conhecido como gênio do choro.
- ▶ Artistas de outros gêneros, como o rock e sertanejo, não estão no topo da lista.

# Análise Exploratória

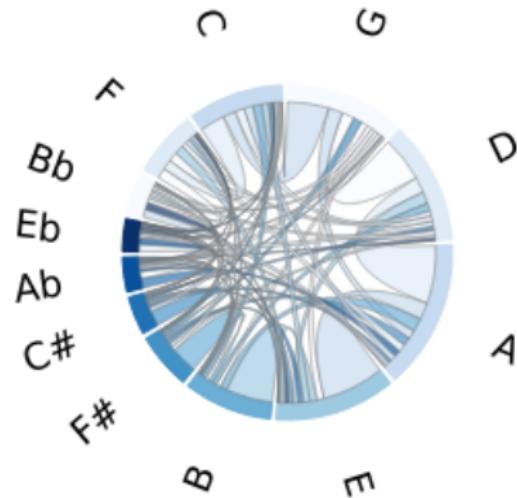


Figura: Diagrama de Cordas - Bossa Nova

# Análise Exploratória

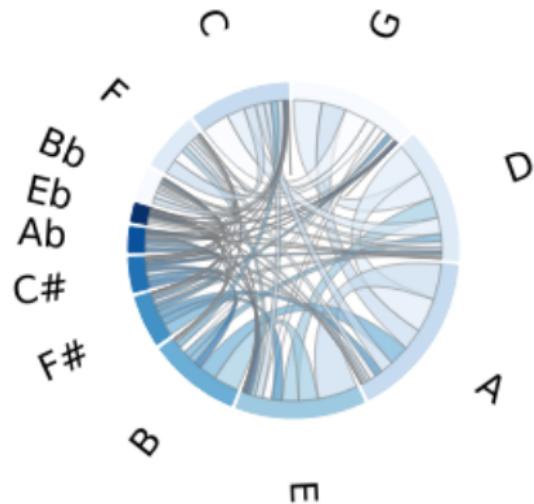


Figura: Diagrama de Cordas - Forró

# Análise Exploratória

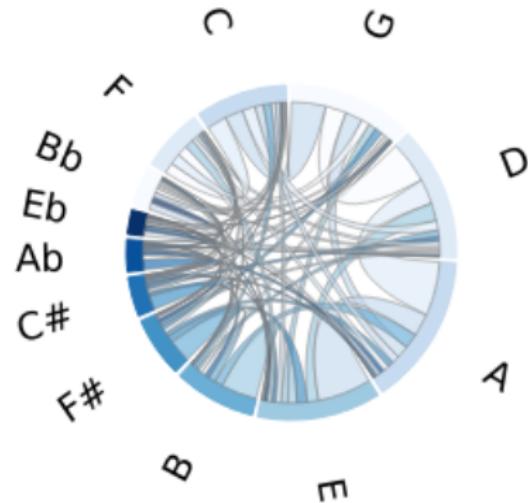


Figura: Diagrama de Cordas - MPB

# Análise Exploratória

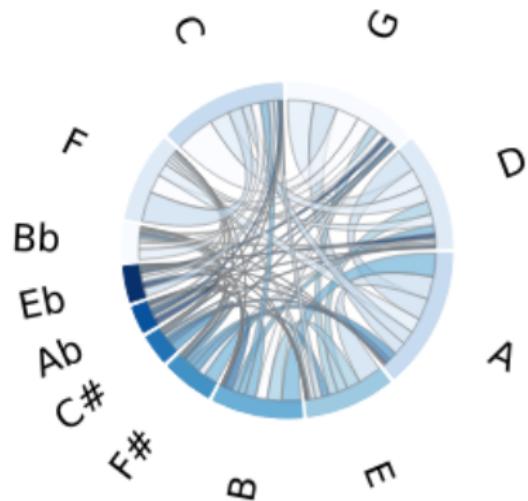


Figura: Diagrama de Cordas - Pop

# Análise Exploratória

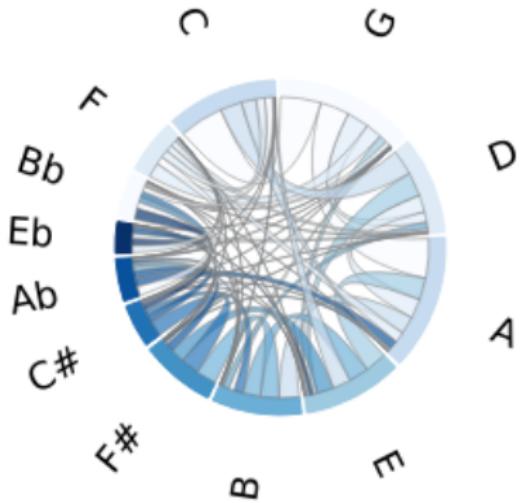


Figura: Diagrama de Cordas - Reggae

# Análise Exploratória

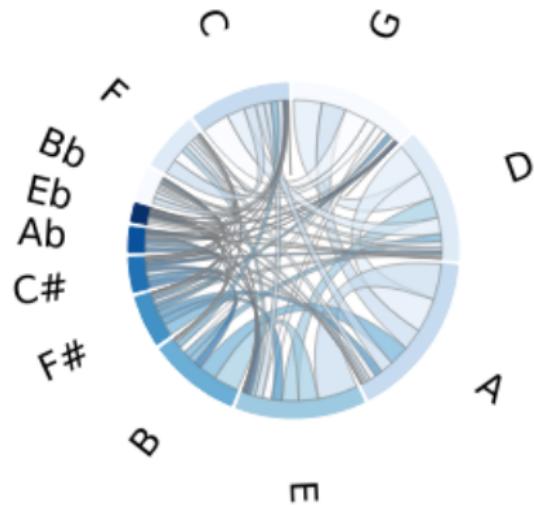


Figura: Diagrama de Cordas - Rock

# Análise Exploratória

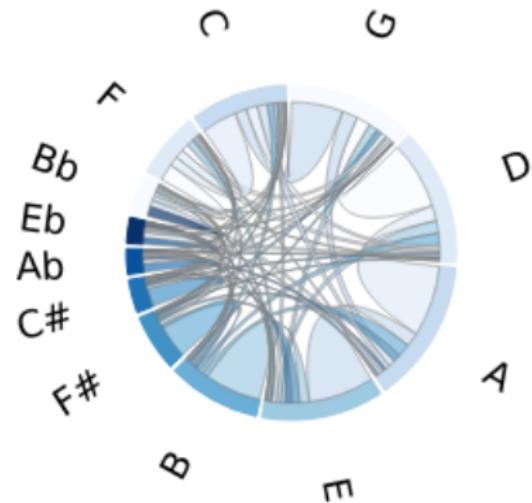


Figura: Diagrama de Cordas - Samba

# Análise Exploratória

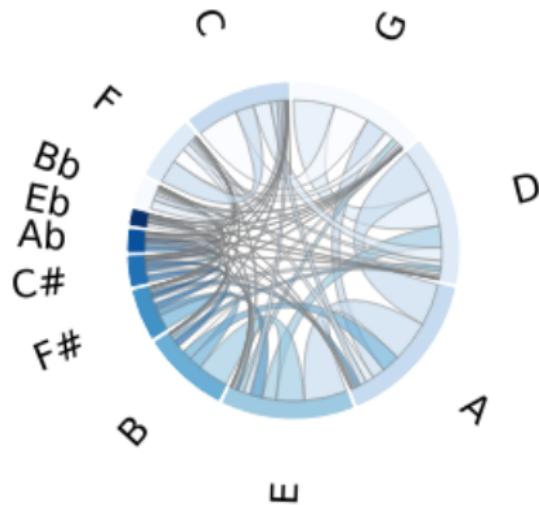


Figura: Diagrama de Cordas - Sertanejo

# Análise Exploratória

- ▶ Os diagramas representam as frequências nas quais as transições dos acordes aconteceram por gênero, ordenadas pelo círculo das quintas.
- ▶ As transições mais frequentes são as das tônicas para suas vizinhas.
- ▶ Os diagramas são diferentes entre os gêneros, de forma geral.

# Características Extraídas

## Conjunto 1 - Triádes

- ▶ % suspensos (e.g. Gsus),
- ▶ % diminutos (e.g. B°),
- ▶ % aumentados (e.g. Baug),
- ▶ % menores (e.g. C#m),
- ▶ % com sétima (e.g. Em7),
- ▶ % com sétima maior (e.g. F7+),
- ▶ % com sétima e menores (e.g. C#m7).

## Conjunto 2 - Tetrádes

- ▶ % com quarta nota (e.g. D4),
- ▶ % com quinta nota aumentada (e.g. C5+),
- ▶ % com quinta nota diminuta (e.g. C5-),
- ▶ % com sexta nota (e.g. E6),
- ▶ % com nona nota (e.g. G9).

# Características Extraídas

## Conjunto 3 - Transições

- ▶ % Primeira transição mais comum,
- ▶ % Segunda transição mais comum,
- ▶ % Terceira transição mais comum.

## Conjunto 4 - Variáveis Extra

- ▶ Total de acordes,
- ▶ Total do acorde mais comum,

- ▶ Popularidade (Spotify)
- ▶ Ano de lançamento (Spotify)
- ▶ Média de distâncias dos acordes ao dó, pelo círculo das quintas,
- ▶ Média de distâncias dos acordes ao dó, por semitons,
- ▶ Indicador(escala = acorde mais comum).

# Características Extraídas

Tabela: Resumo das variáveis extraídas contínuas (1).

Variável	Média	DP	Min	Max
Sus	0.00097	0.013	0	0.46
Sétima	0.3	0.31	0	1
Sétima e Menor	0.0028	0.022	0	1
Sétima Maior	0.038	0.088	0	1
Menores	0.27	0.2	0	1
Diminuto	0.01	0.033	0	0.58
Aumentado	0.022	0.068	0	1
Quarta	0.016	0.016	0	0.68
Sexta	0.025	0.072	0	0.85
Nona	0.059	0.13	0	1
Quinta Dim.	0.0061	0.024	0	0.27
Quinta Aum.	0.0033	0.02	0	0.39

# Características Extraídas

Tabela: Resumo das variáveis extraídas contínuas (2).

Variável	Média	DP	Min	Max
Sus	0.00097	0.013	0	0.46
Transição (1 <sup>a</sup> )	0.19	0.081	0.0078	0.8
Transição (2 <sup>a</sup> )	0.17	0.075	0.016	0.9
Transição (3 <sup>a</sup> )	0.15	0.073	0.01	0.88
Baixo	0.081	0.15	0	1
Total de Acordes	57.12	31.97	1	370
Qtde. acorde mais comum	18.72	11.54	1	202
Popularidade	28.54	14.6	1	83
Dist. ao Dó	2.67	1.18	0	6
Dist. em Semitons	5.48	1.27	0	1

# Características Extraídas

Tabela: Resumo das variáveis extraídas discretas.

Variável	Classes	Mais comum
Ano	1957 a 2017	2015 (7%)
I(escala = acorde mais comum)	Sim ou Não	Não (67%)

Todas as variáveis são calculadas **por música**.

## Variáveis do Spotify:

Extraídas através da API do *software* Spotify.

# Características Extraídas

## Variáveis de porcentagem:

$$\text{Porcentagem} = \frac{\sum_{i=1}^{n_j} I(\text{carac.} \in x_i)}{n_j} \quad (5)$$

onde  $I(\text{carac.} \in x_i)$  indica se a característica está presente na observação  $i$ ,  $i = 1, \dots, n_j$ ,  $j$ -ésima música, que tem tamanho  $n_j$ . A identificação se dá por meio de expressões regulares.

## Variáveis de distâncias:

$$Dist = \frac{\sum_{i=1}^{n_j} d_i}{n_j} \quad (6)$$

aonde os  $d_i$  são as distâncias para os  $i$ -ésimos acordes,  $i = 1, \dots, n_j$ , da  $j$ -ésima música, que tem tamanho  $n_j$ .

# Características Extraídas

Redução de dimensionalidade por estatísticas suficientes:

- ▶ As variáveis indicadoras de presença de atributos seguem uma distribuição de probabilidades  $Bernoulli(\theta_j)$ .
- ▶ O somatório destas observações,  $T = \sum_{i=1}^{n_j} x_i$ , para cada música, segue uma  $Binomial(n_j, \theta_j)$ , e é uma estatística suficiente.
- ▶ Condensamento das informações relevantes da amostra.

# Florestas Aleatórias

Os modelos construídos vão **crescendo em quantidade de variáveis**, seguindo uma estrutura lógica:

1. Conjunto 1 (tríades): variáveis relacionadas a acordes menores, com sétima, aumentados ou diminuto ( as tríades e acorde com sétima).
2. Conjunto 2 (tétrades): acréscimo de variáveis sobre as quartas, sextas, sétimas maiores e nonas notas.
3. Conjunto 3 (transições): inclusão das frações das transições mais comuns.
4. Conjunto 4 (informações extras): ano e popularidade, quantidade de acordes e médias de distâncias.

## Florestas Aleatórias

A base foi particionada aleatoriamente em dados de treino (70%) e teste (30%), com balanceamento por gênero. Um total de 8261 músicas foram capturadas.

Tabela: Quantidade de músicas para cada gênero entre as partições.

Gênero	Treino	Teste	Representatividade
Bossa Nova	305 (68%)	133 (32%)	438 (5.3%)
Forró	115 (73%)	48 (27%)	163 (2%)
MPB	1196 (67.8%)	476 (32.2%)	1679 (20.3%)
Pop	104 (66.4%)	39 (33.6%)	143 (1.7%)
Reggae	46 (68.1%)	24 (31.9%)	70 (0.8%)
Rock	1127 (69.8%)	552 (30.2%)	1679 (20.4%)
Samba	877 (70.8%)	378 (29.2%)	1255 (15.1%)
Sertanejo	1992 (68.2%)	849 (31.8%)	2841 (34.4%)

# Florestas Aleatórias

Tabela: Comparativo de medidas de qualidade dos quatro modelos.

Modelo	Acur. Geral	L.I.	L.S.	Kappa	Valor P
Modelo 1	0.53	0.51	0.55	0.37	< 0.0
Modelo 2	0.57	0.54	0.59	0.42	< 0.0
Modelo 3	0.59	0.56	0.60	0.44	< 0.0
Modelo 4	0.62	0.60	0.64	0.49	< 0.0

- ▶ Estatística *Kappa*: compara a acurácia observada com a acurácia esperada.
- ▶ Acurácia esperada: modelo trivial, que acerta aproximadamente 34% (porcentagem da classe mais comum na base).

# Florestas Aleatórias

- ▶ O aumento da acurácia se dá de forma uniforme nos modelos.
- ▶ A acurácia não se comporta como o  $R^2$ , podendo ser diminuída com a inserção de ruído.
- ▶ O modelo 4, que contém todas as características, possui acurácia de 62%, sendo quase o dobro do modelo basal.

# Florestas Aleatórias

**Tabela:** Matriz de confusão para o modelo com o primeiro conjunto de variáveis.

	Bossa	Forró	MPB	Pop	Reggae	Rock	Samba	Sertanejo
Bossa Nova	<b>0.14</b>	0.00	0.33	0.00	0.00	0.05	0.33	0.15
Forró	0.00	<b>0.00</b>	0.10	0.00	0.00	0.15	0.12	0.62
MPB	0.03	0.00	<b>0.41</b>	0.00	0.00	0.14	0.23	0.20
Pop	0.00	0.00	0.15	<b>0.00</b>	0.00	0.26	0.23	0.36
Reggae	0.00	0.00	0.25	0.00	<b>0.00</b>	0.50	0.04	0.21
Rock	0.01	0.00	0.11	0.00	0.00	<b>0.34</b>	0.07	0.47
Samba	0.02	0.00	0.26	0.00	0.00	0.05	<b>0.57</b>	0.11
Sertanejo	0.00	0.00	0.02	0.00	0.00	0.12	0.02	<b>0.84</b>

# Florestas Aleatórias

**Tabela:** Matriz de confusão para o modelo com o primeiro e segundo conjuntos de variáveis.

	Bossa	Forró	MPB	Pop	Reggae	Rock	Samba	Sertanejo
Bossa Nova	<b>0.29</b>	0.00	0.35	0.00	0.00	0.05	0.19	0.14
Forró	0.00	<b>0.00</b>	0.10	0.00	0.00	0.15	0.12	0.62
MPB	0.03	0.00	<b>0.49</b>	0.00	0.00	0.13	0.17	0.18
Pop	0.00	0.00	0.15	<b>0.00</b>	0.00	0.31	0.18	0.36
Reggae	0.00	0.00	0.17	0.00	<b>0.00</b>	0.50	0.12	0.21
Rock	0.00	0.00	0.13	0.00	0.00	<b>0.36</b>	0.06	0.44
Samba	0.02	0.00	0.20	0.00	0.00	0.04	<b>0.63</b>	0.10
Sertanejo	0.00	0.00	0.02	0.00	0.00	0.12	0.02	<b>0.84</b>

# Florestas Aleatórias

**Tabela:** Matriz de confusão para o modelo com o primeiro, segundo e terceiro conjuntos de variáveis.

	Bossa	Forró	MPB	Pop	Reggae	Rock	Samba	Sertanejo
Bossa Nova	<b>0.29</b>	0.00	0.35	0.00	0.00	0.05	0.17	0.13
Forró	0.00	<b>0.00</b>	0.06	0.00	0.00	0.21	0.08	0.65
MPB	0.03	0.00	<b>0.55</b>	0.00	0.00	0.12	0.15	0.15
Pop	0.00	0.00	0.23	<b>0.00</b>	0.00	0.13	0.21	0.44
Reggae	0.00	0.00	0.38	0.00	<b>0.04</b>	0.46	0.04	0.08
Rock	0.00	0.00	0.14	0.00	0.00	<b>0.35</b>	0.06	0.45
Samba	0.02	0.00	0.21	0.00	0.00	0.03	<b>0.66</b>	0.08
Sertanejo	0.00	0.00	0.02	0.00	0.00	0.09	0.02	<b>0.86</b>

# Florestas Aleatórias

Tabela: Matriz de confusão para o modelo todas as variáveis.

	Bossa	Forró	MPB	Pop	Reggae	Rock	Samba	Sertanejo
Bossa Nova	<b>0.28</b>	0.00	0.40	0.00	0.00	0.05	0.16	0.12
Forró	0.00	<b>0.00</b>	0.12	0.00	0.00	0.12	0.10	0.65
MPB	0.01	0.00	<b>0.59</b>	0.00	0.00	0.11	0.13	0.15
Pop	0.00	0.00	0.13	<b>0.00</b>	0.00	0.28	0.15	0.44
Reggae	0.00	0.00	0.25	0.00	<b>0.08</b>	0.46	0.08	0.12
Rock	0.00	0.00	0.16	0.00	0.00	<b>0.43</b>	0.05	0.35
Samba	0.01	0.00	0.20	0.00	0.00	0.03	<b>0.66</b>	0.10
Sertanejo	0.00	0.00	0.02	0.00	0.00	0.07	0.02	<b>0.89</b>

## Florestas Aleatórias

- ▶ Da primeira matriz para a segunda, houve um acréscimo considerável na classificação correta para a Bossa Nova, seguida da MPB e Samba.
- ▶ Na terceira matriz, o aumento também ocorre para a MPB e Samba, mas é mais chamativo para o Reggae, que antes era completamente mal classificado.
- ▶ Na última matriz, o acréscimo ocorre especialmente para o Rock e o Sertanejo.

# Florestas Aleatórias



**Figura:** Importância das variáveis na classificação dos gêneros, para o primeiro modelo.

# Florestas Aleatórias

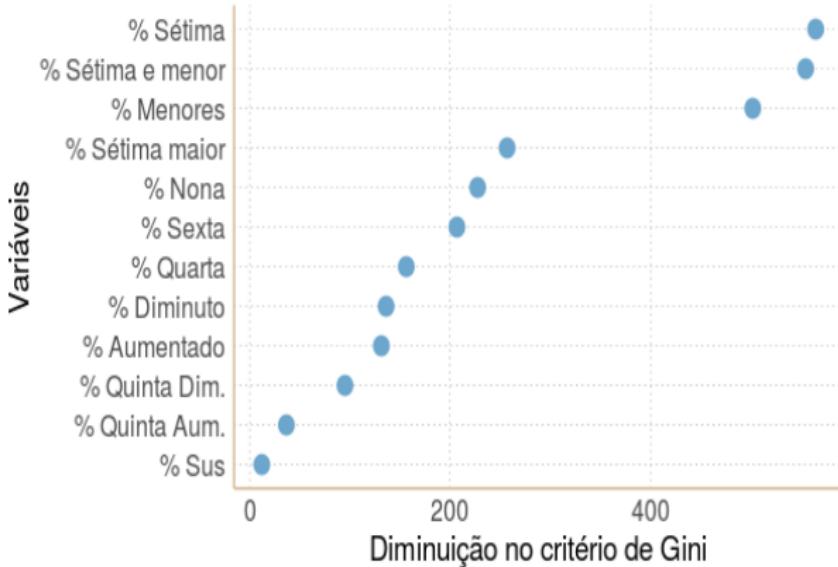


Figura: Importância das variáveis na classificação dos gêneros, para o segundo modelo.

# Florestas Aleatórias

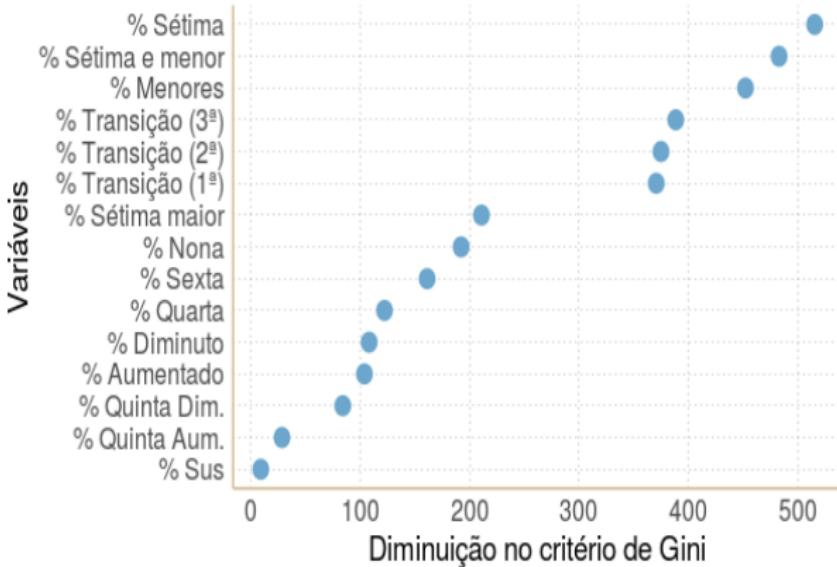


Figura: Importância das variáveis na classificação dos gêneros, para o terceiro modelo.

# Florestas Aleatórias

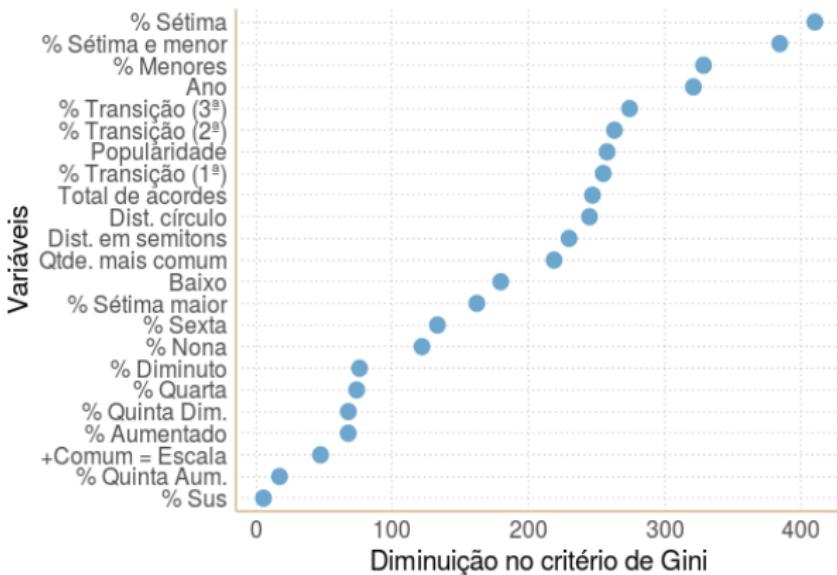


Figura: Importância das variáveis na classificação dos gêneros, para o quarto modelo.

## Florestas Aleatórias

- ▶ Na adição do primeiro conjunto de variáveis ao segundo, as três variáveis mais significativas se mantiveram.
- ▶ Do segundo ajuste para o terceiro, as variáveis, do terceiro conjunto tomam o lugar das que haviam entrado no segundo passo.
- ▶ Com a inclusão do último conjunto de variáveis, a mais relevante entre estas é o ano de lançamento do álbum das músicas, seguida pela popularidade e total de acordes.

5

## Considerações Finais

## Algumas conclusões

- ▶ É possível caracterizar os gêneros da música brasileira pela estrutura harmônica de suas músicas.
- ▶ As cinco variáveis mais importantes são as porcentagens de acordes com sétima, com sétima e menores, acordes menores, o ano das músicas e a popularidade.
- ▶ Prevalecem as características que podem ser extraídas apenas com os acordes.
- ▶ O levantamento das *features* é feito usando conhecimento estatístico acompanhado de teoria musical.

# Considerações Finais

- ▶ A aplicação dos métodos motivou um profundo aprendizado sobre estes.
- ▶ A realização do trabalho permite que outros pesquisadores façam mais estudos a partir dele.

## Referências Bibliográficas

# Referências Bibliográficas I

-  AGARWAL, B.; MITTAL, N. *Prominent Feature Extraction for Sentiment Analysis*. [S.I.]: Springer International Publishing, 2015. (Socio-Affective Computing). ISBN 9783319253435.
-  ALMADA, C. *Harmonia Funcional*. [S.I.: s.n.], 2012. v. 1.
-  BASILI, R.; SERAFINI, A.; STELLATO, A. Classification of musical genre: a machine learning approach. In: *ISMIR*. [S.I.: s.n.], 2004.
-  BEN-GAL, I. et al. *Bayesian networks, encyclopedia of statistics in quality and reliability*. [S.I.]: Wiley & Sons, 2007.
-  CALDAS, W. *Iniciação à Música Popular Brasileira*. [S.I.: s.n.], 2010. v. 1.

## Referências Bibliográficas II

-  CHIASSON, T.; GREGORY, D. *Data + Design: A Simple Introduction to Preparing and Visualizing Information.* [S.I.]: Infoactive, 2014. ISBN 9781787286474.
-  CIFRA Club - seu site de cifras e tablaturas. <<https://www.cifraclub.com.br/>>.
-  CORREA, D. C.; COSTA, L. d. F.; SAITO, J. H. *Tracking the Beat: Classification of Music Genres and Synthesis of Rhythms.* [S.I.]: IWSSIP, 2010.
-  Ese Xxxx, T. N. et al. Probabilistic Models for Music JEAN-FRANÇ OIS PAIEMENT. 2008.
-  FARAWAY, J. J. *Extending the Linear Model with R (Texts in Statistical Science).* [S.I.]: Chapman & Hall/CRC, 2005. ISBN 158488424X.

## Referências Bibliográficas III

-  FARID, D. M.; HARBI, N.; RAHMAN, M. Z. Combining naive bayes and decision tree for adaptive intrusion detection. *arXiv preprint arXiv:1005.4496*, 2010.
-  GUYON, I.; ELISSEEFF, A. An introduction to feature extraction. In: *Feature extraction*. [S.I.]: Springer, 2006. p. 1–25.
-  KUHN, M.; JOHNSON, K. *Feature Engineering and Selection: A Practical Approach for Predictive Models*. [S.I.: s.n.], 2018.
-  LEÓN, P. J. P. D.; IÑESTA, J. M.; RIZO, D. Mining digital music score collections: melody extraction and genre recognition. In: *Pattern Recognition Techniques, Technology and Applications*. [S.I.]: Intech, 2008.

## Referências Bibliográficas IV

-  MERKELBACH, K. Feature extraction for musical genre classification mus-15. Citeseer, 2015.
-  PAIEMENT, J.-F.; ECK, D.; BENGIO, S. A probabilistic model for chord progressions. In: *Proceedings of the Sixth International Conference on Music Information Retrieval (ISMIR)*. [S.l.: s.n.], 2005.
-  ROLLAND, P.-Y. Discovering patterns in musical sequences. *Journal of New Music Research*, Taylor & Francis, v. 28, n. 4, p. 334–350, 1999.
-  SCARINGELLA, N.; ZOIA, G.; MLYNEK, D. Automatic genre classification of music content: a survey. *IEEE Signal Processing Magazine*, IEEE, v. 23, n. 2, p. 133–141, 2006.

## Referências Bibliográficas V

-  SEAGER, A. Etnografia da música. *Cadernos de Campo (São Paulo, 1991)*, v. 17, n. 17, p. 237–260, 2008.
-  SPOTIFY AB, Regeringsgatan 19, SE-111 53 Stockholm, Sweden. [S.I.: s.n.].
-  TEAM, R. R Development Core Team. *R: A Language and Environment for Statistical Computing*, v. 55, p. 275–286, 2013. ISSN 3-900051-07-0.
-  VIRTANEN, T. Unsupervised learning methods for source separation in monaural music signals. In: *Signal Processing Methods for Music Transcription*. [S.I.]: Springer, 2006. p. 267–296.
-  WILKINSON, L. The grammar of graphics. In: *Handbook of Computational Statistics*. [S.I.]: Springer, 2012. p. 375–414.

## Referências Bibliográficas VI



WUNDERVALD, B. *GitHub - Bruna Wundervald*. 2018.

Disponível em: <<https://github.com/brunaw/chorrrds>>.



WUNDERVALD, B.; TRECENI, J. *R-Music*. 2018.

Disponível em: <<https://github.com/r-music>>.

Agradecemos a atenção!