

Classificação de Gêneros Musicais Latinos e suas Emoções: Abordagens Bayesiana e Fuzzy[†]

G.M. BRESSAN* e B.C.F. DE AZEVEDO

Recebido em 21 agosto, 2016 / Aceito em 30 agosto, 2017

RESUMO. O objetivo deste trabalho é investigar métodos de classificação inteligente para a tarefa de classificação de gêneros musicais latinos a partir de um conjunto de dados reais. Neste estudo, as características emocionais predominantes de cada gênero também são consideradas, com objetivo de possibilitar ao usuário escolher músicas de acordo com a emoção que deseja estimular. Os métodos propostos para tal tarefa são baseados no método de classificação Bayesiana, o qual utiliza o algoritmo *BayesRule* e é capaz de investigar incertezas probabilísticas nos dados, e no método de classificação *fuzzy*, cujas funções tem seus parâmetros ajustados por um sistema *neuro-fuzzy* e permite avaliar fusões entre os diferentes gêneros musicais. Ambas as metodologias extraem regras de classificação linguísticas, o que possibilita que seja feita uma comparação entre os resultados obtidos, além da classificação inteligente do conjunto de dados considerando incertezas e fusões entre os gêneros musicais.

Palavras-chave: gêneros musicais, classificação Bayesiana, classificação *fuzzy*.

1 INTRODUÇÃO

A crescente expansão dos recursos midiáticos despertou a atenção de inúmeros pesquisadores para a catalogação de mídias digitais [24, 27]. O interesse pela classificação musical destaca-se devido à sua vasta aplicabilidade e descobertas obtidas por meio de estudos. Entretanto, sem o uso de métodos eficientes para organização sistêmica de bibliotecas digitais, seria difícil para os usuários elencarem a melhor opção diante da grande quantidade de músicas disponíveis.

Classificar música de forma automática, utilizando emoção como variável principal, requer uma complexa e simultânea multidisciplinaridade de fatores que dificultam a classificação da emoção [6], tais como: humor particular, personalidade, idade, cultura e informações extra musicais do ouvinte [6, 14, 28]. Essas características contribuem para que haja uma difração da emoção sentida por duas pessoas ao ouvirem a mesma música. Desta forma, é importante o emprego de modelos matemáticos computacionais que consigam avaliar os dados de forma

[†] Apoio financeiro concedido pela Fundação Araucária.

*Autor correspondente: Gláucia Maria Bressan – E-mail: galbressan@gmail.com

Universidade Tecnológica Federal do Paraná – UTFPR, Departamento de Matemática, Av. Alberto Carazzai, 1640, 86300-000 Campus Cornélio Procopio, PR, Brasil. E-mail: beatrizflamiadx@hotmail.com

que a subjetividade humana não interfira nos resultados. Tais modelos podem ser obtidos por meio da representação, extração e análise de atributos dos quais dependem os resultados da classificação [11].

Diante deste cenário, o objetivo deste trabalho é classificar de forma automática o conjunto de músicas provenientes do *Latin Music Database* (LMD) [10], por meio da emoção predominante de cada gênero musical obtido em [23], no qual a base de dados LMD foi analisada por meio da distribuição da emoção nos diferentes gêneros musicais. Para tanto, são empregados dois métodos que tratam incertezas e sobreposições entre as classes: o método de classificação Bayesiana, utilizando o algoritmo *BayesRule* [8], o qual extrai um conjunto de regras linguísticas de uma rede Bayesiana, e o método de classificação *fuzzy*, no qual os parâmetros das funções de pertinência são ajustados por um sistema *neuro-fuzzy*. Estas duas metodologias extraem regras de classificação do tipo “se-então”, o que possibilita que seja feita uma comparação entre os resultados obtidos, além da classificação inteligente do conjunto de dados considerando incertezas e fusões entre os gêneros musicais. Tais fusões são geradas pela sobreposição das fronteiras entre gêneros musicais. Li & Ogihara (2005) [16] afirmam que a medida que a indústria midiática cresce, as fronteiras entre os gêneros tornam-se desfocadas, dificultando o processo de classificação. Ambas as metodologias empregadas neste trabalho se mostram promissoras na tarefa da classificação automática de gêneros musicais, considerando o conjunto de dados utilizado.

Este trabalho está organizado da seguinte forma: a Seção 2 descreve alguns trabalhos que abordam a classificação musical considerando incertezas. A Seção 3 descreve a fundamentação teórica e os métodos de classificação adotados. O procedimento de obtenção do conjunto de dados e os detalhes experimentais são fornecidos na Seção 4. A Seção 5 exhibe os modelos de classificação propostos e os resultados numéricos obtidos. As conclusões são discutidas na Seção 6.

2 REVISÃO BIBLIOGRÁFICA

Tradicionalmente, a associação entre música e emoção musical é feita manualmente de acordo com características comuns. Contudo, esse processo requer tempo e esforço demasiados, pois os gêneros possuem características não exclusivas e tendem a se misturar, criando sub-gêneros e variados estilos, o que gera um processo interminável de classificação. Por esta razão, muitas técnicas tem sido exploradas, visando otimizar o tempo e a confiabilidade da classificação, como *Support Vector Machine* (SVM) [17], Lógica Fuzzy [5, 6], Classificador *Fuzzy k-Nearest Neighbours* (FKNN) [28], Redes Bayesianas [4, 13].

A classificação de emoção musical por redes Bayesianas é pouco explorada na literatura, se comparada a outras metodologias frequentemente empregadas na classificação de atributos musicais [14]. Contudo, o uso de métodos Bayesianos, como as redes, permite a integração de informações, bem como a incorporação de novos conhecimentos ao modelo, sendo essa uma grande vantagem em relação aos demais métodos de classificação aplicados na literatura. Em Donnelly & Sheppard (2014) redes Bayesianas foram aplicadas para identificação do timbre de

instrumentos musicais. Quatro estruturas de redes Bayesianas, incluindo Naïve Bayes, foram examinadas e comparadas com os classificadores SVM e *k-Nearest Neighbours* (KNN). Os experimentos mostraram que as redes Bayesianas são uma abordagem válida para classificação de instrumentos musicais, dada a eficiência dos resultados.

Yang & Chen (2011) avaliam a emoção musical por meio de um modelo probabilístico, o qual incorpora a subjetividade da percepção da emoção musical do ouvinte. É proposta a distribuição probabilística da emoção musical pelo teorema de Bayes, sendo a preferência musical do ouvinte a informação *a priori* [27].

Outras abordagens Bayesianas envolvendo análise musical podem ser visto em [24], o qual apresenta diversos estudos e aplicações de modelos probabilísticos e métodos Bayesianos na recuperação de dados musicais.

Sistema Baseado em Regras *Fuzzy* (SBRF) tem sido muito utilizado na área de controle e apoio à tomada de decisões, nos casos em que existe subjetividade associada aos dados. Fernandéz et al. (2011) apresentam uma tentativa preliminar de utilização de um SBRF, juntamente com algoritmos evolucionários para classificar gêneros musicais. De acordo com os autores, essa foi a primeira vez que utilizou-se aproximação *fuzzy* para explorar a imprecisão desse tipo de problema [5].

Yang et al. (2006) exploraram dois modelos de classificação: FKNN e *Fuzzy Nearest Mean* (FNM). As emoções musicais foram agrupadas em 4 classes, seguindo o modelo de descrição de emoção musical de Thayer [25]. Em comparação com a classificação que inclui a amostra em apenas uma classe, os classificadores *fuzzy* indicam fusões entre as classes, já que emoção musical é subjetiva [28].

De acordo com o exposto na literatura, os métodos aplicados geralmente classificam apenas dois ou poucos gêneros musicais, conforme afirma [1]. Além disso, utilizam apenas frequência ou timbre como entradas do sistema de classificação e o modelo de Thayer [25] ou de Russel [22] para classificar a emoção. Em vista disso, a principal contribuição deste trabalho é proporcionar a classificação automática de 10 gêneros musicais latinos, ainda não encontrados na literatura, considerando múltiplos grupos como variáveis de entrada (batidas sonoras, timbre e frequência). Juslin & Laukka (2004) sugerem que a música possui “poderes emocionais”, pois é capaz de induzir emoções [12]. Portanto, a motivação de se conhecer a emoção musical é poder identificar a provável emoção que cada gênero musical pode despertar no ouvinte e permitir a este escolher a música de acordo com a emoção requerida.

3 FUNDAMENTAÇÃO TEÓRICA DOS MÉTODOS DE CLASSIFICAÇÃO

Nesta seção é descrita a fundamentação teórica dos dois métodos de classificação nos quais a proposta deste trabalho se baseia: o método de classificação *Bayesiano*, que aplica o método *BayesRule*, e o método de classificação *fuzzy*, que utiliza os SBRFs.

3.1 Classificação Bayesiana

O modelo de inferência Bayesiana consiste em analisar de maneira probabilística variáveis condicionadas por meio do *Teorema de Bayes*, descrito conforme Box & Tiao [2]: Suponha que $y' = (y_1, \dots, y_n)$ seja um vetor de n observações no qual a distribuição de probabilidade $p(y | \theta)$ dependa do valor de k parâmetros $\theta' = (\theta_1, \dots, \theta_k)$. Considere que θ tenha distribuição de probabilidade $p(\theta)$. Assim, dada a observação y , a probabilidade condicional de θ é representada pela equação (3.1), conhecida como *Teorema de Bayes*

$$p(\theta | y) = \frac{p(y | \theta) \times p(\theta)}{p(y)}. \quad (3.1)$$

Em (3.1), $p(\theta)$, chamado de distribuição *a priori*, fornece o conhecimento sobre θ sem o conhecimento dos dados e $p(\theta | y)$, conhecido como distribuição *a posteriori* de θ dado y , fornece o conhecimento sobre θ dado o conhecimento dos dados. Considerando que a informação sobre y é conhecida, $p(y | \theta)$ pode ser considerado uma função de θ , conhecida como *verossimilhança* de θ dado y . O teorema de Bayes então estabelece que a distribuição *a posteriori* de θ dado y é proporcional ao produto da distribuição *a priori* de θ dado o conhecimento dos dados pela verossimilhança de θ dado y . Ou seja: *posteriori* \propto *verossimilhança* \times *priori*.

Na maioria dos sistemas reais, a variável de interesse é dependente ou está condicionada a várias outras. Mesmo podendo ser expandido para mais de uma variável, o teorema de Bayes unicamente não é recomendado em casos envolvendo múltiplas variáveis, devido à complexidade dos cálculos [19]. Diante dessa dificuldade Pearl (1988) desenvolveu redes que avaliam as interligações das variáveis por meio de suas estruturas, conhecidas como *Redes Bayesianas*.

Redes Bayesianas podem ser definidas como um modelo gráfico probabilístico usado para representar o conhecimento sobre o domínio dos dados. Estas redes aprendem relações de causa e consequência e podem combinar conhecimento *a priori* com padrões aprendidos a partir dos dados. Além disso, permitem que o usuário ou especialista interfira nos nós da rede e insira um conhecimento que se propaga nos outros nós. São compostas então por uma estrutura de rede, que consiste em um grafo acíclico direcionado, e um conjunto de tabelas de probabilidade. Os nós da estrutura da rede representam as variáveis e os arcos entre os nós representam relações de dependência entre as variáveis correspondentes. Um arco começando em um nó A e terminando em um nó B estabelece A como pai de B e B como filho de A . Uma rede Bayesiana pode ser utilizada como um classificador, calculando a probabilidade condicional de um nó, chamado nó classe, dados os valores das probabilidades dos outros nós.

Em Cheng et al. (2002) uma rede Bayesiana é representada por $BN = \langle N, A, \Theta \rangle$, em que $\langle N, A \rangle$ é um grafo acíclico direcionado em que cada nó $x_i \in N$ representa uma variável e cada arco $a \in A$ entre nós representa uma dependência probabilística entre os nós associados. Uma distribuição de probabilidade condicional θ_i é associada com cada nó $x_i \in N$, $i = 1, \dots, n$, coletivamente representado por $\Theta = \theta_i$, que quantifica o quanto um nó depende de seus pais [3].

Conforme descrito em Pearl (1988), a independência condicional, que é a condição de Markov, permite o cálculo da distribuição conjunta de todas as variáveis, dado pela equação (3.2)

$$P(x_1, x_2, \dots, x_n \mid BK) = \prod_{i=1}^n P(x_i \mid \pi_{x_i}, BK), \quad (3.2)$$

em que BK representa o conhecimento antecedente, x_i é a i -ésima variável ou nó, e π_{x_i} é o conjunto de pais de x_i . Portanto, uma rede Bayesiana pode ser usada como uma representação do conhecimento que permite inferências. Os nós que fazem parte do Markov Blanket do nó M são os que exercem influência na distribuição de probabilidade condicional de M . São estes: os nós pais de M , os nós filhos de M e os pais dos filhos de M [20].

Redes Bayesianas consistem em uma forma eficiente de trabalhar com informações complexas; contudo, o conhecimento representado não é tão compreensível como outras formas de representação que expressam conhecimento semelhante ao raciocínio humano, como regras de classificação. Pensando nisso, um método chamado *BayesRule* desenvolvido por Hruschka Jr et al. (2007), usa o conceito de probabilidade máxima *a posteriori* para extrair um conjunto de regras probabilísticas do tipo “se-então” que descrevem a classificação.

O método *BayesRule* faz uso do algoritmo K2, o qual emprega a busca heurística para induzir uma rede Bayesiana usando como entrada uma base de dados e uma lista ordenada de variáveis. O algoritmo utiliza um método *greedy* para procurar a melhor estrutura de rede que represente os dados [8]. Além de utilizar a probabilidade máxima *a posteriori* o *BayesRule* emprega o conceito do Markov Blanket para reduzir o número e a complexidade de regras de classificação geradas pelo processo de extração. Isto permite a redução no tempo requerido para construir modelos quando mais de uma variável é designada como classe. Seguindo esta aproximação, um classificador Bayesiano evidencia que o algoritmo de propagação deve ser usado para propagar os valores das variáveis e inferir o valor da variável classe. A confiança da regra pode ser definida utilizando resultados inferenciais. Com isso, a probabilidade dada à classe inferida pode ser usada como um valor de confiança e é embutido no algoritmo de inferência. O algoritmo *BayesRule* pode ser visto em [8].

3.2 Classificação Fuzzy

A classificação por SBRFs é basicamente constituída pela definição de parâmetros de funções de pertinência, variáveis de entrada e saída, e base de regra “se-então”. Primeiramente, é fundamental definir um *conjunto fuzzy* [29].

Definition 3.1 (Conjunto Fuzzy). *Um conjunto fuzzy é caracterizado por uma função de pertinência que leva os elementos de um domínio, espaço ou universo de discurso \mathbf{X} ao intervalo unitário $[0,1]$, ou seja, $A: \mathbf{X} \rightarrow [0,1]$.*

Portanto, um conjunto *fuzzy* A de \mathbf{X} pode ser representado como um conjunto de pares ordenados de um elemento genérico $x \in \mathbf{X}$ e seu grau de pertinência: $A = \{(A(x)/x) \mid x \in \mathbf{X}\}$.

Função de pertinência, por sua vez, é definida como uma curva que define como cada ponto na variável de entrada esta representada por um valor de pertinência (ou grau de pertinência) entre 0 e 1, no qual esses valores expressam o quanto cada objeto é compatível com as propriedades ou características do conjunto [21].

A princípio qualquer função na forma $A: \mathbf{X} \rightarrow [0,1]$ descreve uma função de pertinência associada ao conjunto *fuzzy* A , que depende não só do conceito representado, mas também do contexto em que é usado [21]. Sendo m um valor modal, a e b os limites inferior e superior, respectivamente, para valores não nulos de $A(x)$, define-se os principais tipos de funções de pertinência de acordo com as equações (3.3), (3.4) e (3.5) [21].

– Funções Triangulares

$$A(x) = \begin{cases} 0 & \text{se } x \leq a \\ \frac{x-a}{m-a} & \text{se } x \in [a, m] \\ \frac{b-x}{b-m} & \text{se } x \in [m, b] \\ 0 & \text{se } x \geq b \end{cases} \quad (3.3)$$

– Funções Trapezoidais

$$A(x) = \begin{cases} 0 & \text{se } x < a \\ \frac{x-a}{m-a} & \text{se } x \in [a, m] \\ 1 & \text{se } x \in [m, n] \\ \frac{b-x}{b-n} & \text{se } x \in [n, b] \\ 0 & \text{se } x > b \end{cases} \quad (3.4)$$

– Funções Gaussianas

$$A(x) = e^{-k(x-m)^2}; k > 0 \quad (3.5)$$

Os gráficos destas funções assumem formatos, respectivamente, triangular (constituída de três pontos formando um triângulo), trapezoidal (cujo topo é planificado) ou Gaussiano (construído com base na distribuição da curva Gaussiana).

A base de regras consiste de expressões linguísticas do tipo “se-então” que descrevem a relação entre as variáveis de entrada e as de saída. As entradas são combinadas logicamente utilizando operadores para produzir valores de saída para todas as entradas esperadas. As conclusões são então combinadas em uma soma lógica para cada função de pertinência [21]. Estas somas lógicas

devem ser combinadas em um processo de defuzzificação para produzir a saída numérica. O processo de inferência *fuzzy*, ilustrado na Figura 1, é o processo de mapear, a partir da entrada, uma saída do conjunto *fuzzy*. Mais detalhes podem ser vistos em [21].

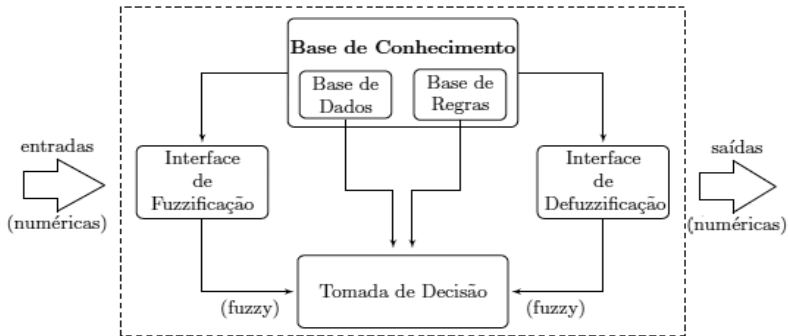


Figura 1: Sistema Baseado em Regras Fuzzy.

O ajuste dos parâmetros das funções de pertinência dos SBRFs pode ser realizado com auxílio de sistemas híbridos como o *neuro-fuzzy*, a partir do aprendizado de dados. Com apoio computacional, o Sistema de Inferência *Neuro-fuzzy* Adaptativo (ANFIS) do MATLAB¹ efetua o treinamento do sistema *neuro-fuzzy* com um método de inferência do tipo Sugeno, no qual a saída é linear ou constante. O método de treinamento da rede neural contida no sistema é um algoritmo baseado em uma abordagem híbrida, o qual combina o método *backpropagation* com o método dos mínimos quadrados.

A estimação desses parâmetros é facilitada pelo vetor gradiente, que fornece a medida de quão bem o SBRF modela as entradas/saídas dos dados com os parâmetros obtidos. Uma vez obtido o vetor gradiente qualquer uma das várias rotinas de otimização podem ser aplicadas para ajustar os parâmetros, assim como reduzir alguns erros de medidas. Uma estrutura *neuro-fuzzy* pode ser vista na Figura 2, a qual mostra as camadas da Rede *neuro-fuzzy*, sendo: 1 – Entradas numéricas; 2 – Processo de Fuzzificação; 3 – Base de Regras linguística, 4 – Consequentes das regras, 5 – Defuzzificação.

Os dados de treinamento de cada linha são os pares de entrada/saída do sistema que será modelado. O número de linhas é igual ao número de pares de dados de treinamento, e o número de colunas é igual ao número de entradas mais um. O processo de treinamento termina quando for atingido o número de épocas ou quando o critério de erro é alcançado.

4 CONJUNTO DE DADOS E DETALHES EXPERIMENTAIS

O *Latin Music Database* (LMD) consiste em um conjunto de dados numéricos referentes a atributos musicais, representados de forma matricial, no qual as linhas correspondem às gravações musicais e as colunas correspondem aos atributos dessas gravações, extraídos a partir do *software*

¹Desenvolvido pela MathWorks (<http://www.mathworks.com/products/matlab/>)

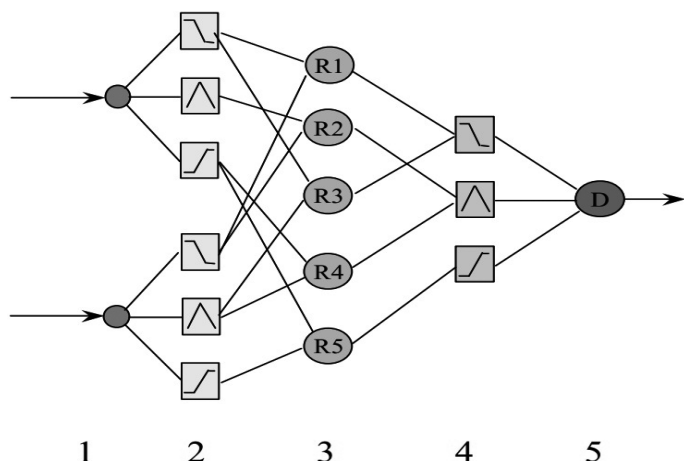


Figura 2: Estrutura do Sistema *Neuro-fuzzy*.

MARSYAS [26]. Cada linha da matriz é chamada de instância, a qual contém as informações numéricas sobre cada atributo e o gênero correspondente como saída. No total, o LMD é composto de 3000 instâncias, divididas igualmente em 10 gêneros musicais latinos: tango, bachata, bolero, merengue, salsa, forró, pagode, sertanejo, gaúcha, e axé. São selecionados 10% dos dados de cada gênero para validação dos resultados, ou seja, para constituir um conjunto de teste. Os demais são usados para treinamento e geração dos resultados.

Neste trabalho, portanto, são considerados 30 atributos numéricos que descrevem os gêneros musicais latinos, selecionados por Silla Jr et al. (2008). Esses atributos são divididos, como pode ser visto em Silla Jr et al. (2009), em três grupos: batidas sonoras (atributos 1-6), timbre (atributos 7-25) e frequência (atributos 26-30). De acordo com Silla Jr et al. (2009), a determinação de tais atributos é feita mediante a aplicação de algoritmos genéticos [18] e o cálculo para sua obtenção engloba conceitos como transformada de Fourier, coeficiente cepstral Mel-frequência (MFCC), picos e batidas, histogramas dos ritmos e descritores estatísticos.

Para elaboração dos dois modelos de classificação propostos neste trabalho, foram selecionados os 3 atributos mais representativos de cada um dos 3 grupos de atributos de entrada, do trabalho de Silla Jr et al. (2009). A seleção foi feita de forma a viabilizar a aplicação do método *BayesRule*, por meio do método *Wrapper* [9], ou seja, constrói-se diferentes redes com vários subconjuntos de atributos e são realizados testes entre tais atributos até encontrar a rede que apresente a melhor relação entre os nós. Os valores numéricos dos atributos de entrada, quando treinados, proporcionaram os melhores relacionamentos entre os nós da Rede Bayesiana e o menor erro no processo de classificação. Logo, os atributos selecionados por meio do método *Wrapper* de cada um dos 3 grupos de atributos são os seguintes:

- *Grupo 1* (referente ao atributo *batidas sonoras*): Amplitude relativa do segundo pico do histograma (Atributo 2); razão entre as amplitudes do segundo pico e do primeiro pico (Atributo 3); soma global dos histogramas (Atributo 6).
- *Grupo 2* (referente ao atributo *timbre*): Média do centroide espectral (Atributo 7); desvio padrão para a atenuação espectral (Atributo 12); desvio padrão para o segundo coeficiente cepstral Mel-frequência (MFCC) (Atributo 22).
- *Grupo 3* (referente ao atributo *frequência*): Soma global do histograma (Atributo 26); amplitude do pico máximo do histograma (Atributo 28); período do pico máximo do histograma (Atributo 29).

Os mesmos atributos de entrada foram utilizados para modelar os dois sistemas de classificação, o que possibilitou, além a classificação inteligente do conjunto de dados, a comparação entre os resultados obtidos pelas diferentes metodologias.

Após a seleção dos atributos de entrada, uma revisão bibliográfica foi realizada para auxiliar na associação entre os gêneros musicais latinos e as emoções. Especificamente no trabalho [23], as músicas do LMD foram identificadas por suas emoções predominantes com o apoio de especialistas e do conhecimento de músicos brasileiros que auxiliaram na tarefa de criar o *Latin Music Mood Database* (Santos e Silla Jr (2015)). Por meio destes, pode-se associar os gêneros às emoções conforme descrito na Tabela 1.

Tabela 1: Associação entre Gênero e Emoção.

Gênero	Emoção	Gênero	Emoção
Tango	Decepção	Forró	Acolhedor
Bachata	Amor	Pagode	Felicidade
Bolero	Romântico	Sertanejo	Tristeza
Merengue	Paixão	Gaúcha	Alegria
Salsa	Sedução	Axé	Entusiasmo

5 MODELOS PROPOSTOS DE CLASSIFICAÇÃO

Nesta seção são descritas as propostas de classificação que são baseadas, respectivamente, na abordagem Bayesiana e nos SBRFs.

5.1 Classificação por Redes Bayesianas

Sabendo que a categorização musical envolve múltiplas variáveis, propõe-se tratar a estrutura da emoção musical de forma probabilística, baseada na relação dos atributos físicos da música. Assim, à medida que novos conhecimentos venham a surgir, estes poderão ser incorporados facilmente ao modelo, contribuindo para melhora da confiabilidade dos resultados. Partindo do

princípio que existe uma emoção predominante para cada gênero musical [23], e que o gênero pode ser determinado por meio da relação entre atributos, tem-se a Equação (5.1).

$$\text{Emoção} \propto P(\text{Gênero} \mid 1^\circ \text{ Atributo}, 2^\circ \text{ Atributo}, 3^\circ \text{ Atributo}) \quad (5.1)$$

Para a elaboração das redes Bayesianas, foram feitos vários testes com os atributos do LMD, visando encontrar os 3 atributos de cada grupo (descritos na Seção 4) que constituam uma rede Bayesiana que apresente relações significativas e consiga expressar os melhores resultados sobre o parâmetro de inferência. São selecionados 3 atributos para que seja possível aplicar o Método *BayesRule* para extração de regras linguísticas de uma rede Bayesiana. Para isso, os atributos numéricos selecionados de cada categoria passaram um por um por processo de discretização, ou seja, foram transformados em variáveis linguísticas particionadas em três classes, as quais são definidas como: baixo, médio e alto. Os intervalos de cada classe são obtidos com auxílio do *software* GENIE². Também com apoio desta ferramenta, foram geradas as estruturas das redes Bayesianas, conforme Figura 3, a partir da aplicação dos algoritmos *K2* e *Greedy Thick Thinning* [15]. Em seguida, foi aplicado o algoritmo *BayesRule* para obtenção das regras “se-então”, como por exemplo, no Grupo 2, uma regra é: *SE atributo 7 é alto E atributo 12 é baixo E atributo 22 é baixo ENTÃO a saída é Merengue* com 38,70% de probabilidade. Por meio da Tabela 1 associa-se o gênero a sua emoção predominante, que neste caso é *paixão*.

Para validação dos resultados, as instâncias de teste foram discretizadas nos mesmos intervalos das instâncias de treinamento, e os acertos foram contabilizados de acordo com as regras “se-então” geradas pelo algoritmo *BayesRule*. Um pré-processamento dos dados foi aplicado devido a presença de ruídos (improváveis saídas para o precedente da regra) e redundâncias (presença de instâncias repetidas), baseando-se em técnicas de mineração de dados [7]. De acordo com essas técnicas, ruídos, redundâncias e inconsistências comprometem a acurácia dos dados. Desta forma, primeiramente, algumas discrepâncias devem ser eliminadas, como erros humanos e formatos diferentes de entrada dos dados. As redundâncias ocorrem quando um atributo pode ser derivado de outro(s), o que pode ser detectado pelo coeficiente de correlação [7]. Assim, o pré-processamento, também chamado de *procedimento de limpeza*, proporciona um conjunto de dados mais confiável e consistente, com melhor acurácia dos resultados.

Com o modelo proposto de classificação Bayesiana, são obtidas as redes descritas na Figura 3, com as seguintes acurácias: 98,33%, 97,87% e 98,04%, respectivamente, para os sistemas de classificação dos grupos [a]-Batidas Sonoras, [b]-Timbre e [c]-Frequência.

5.2 Sistema de Classificação Fuzzy

São elaborados três SBRFs, sendo um para cada grupo de atributos de entrada. Para gerar as regras linguísticas de cada sistema, os dados numéricos de cada atributo de entrada foram discretizados também com auxílio do *software* GENIE e então particionados em três classes linguísticas: baixo, médio e alto. Cada um dos 3 sistemas de classificação possuem, conforme descrito an-

²Desenvolvido pela Universidade de Pittsburgh (<https://dslpitt.org/genie/>)

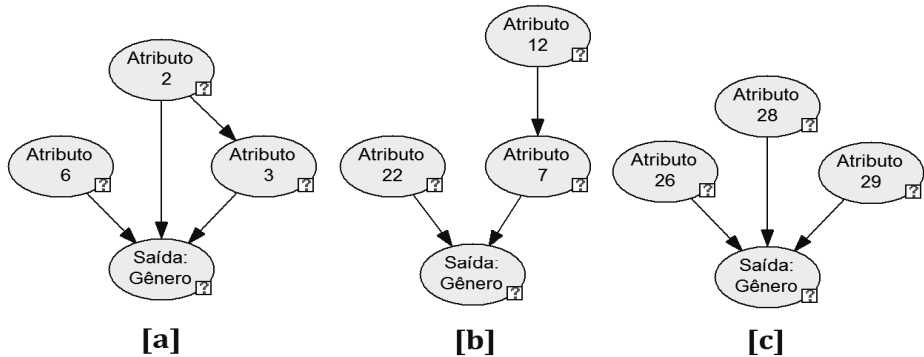


Figura 3: Redes Bayesianas Geradas.

teriormente (Seção 4), 3 atributos de entrada por sua vez particionados em 3 classes. Desta forma, para cada sistema são geradas 27 regras linguísticas do tipo “se-então”. Em cada sistema, cada uma dessas combinações foram contabilizadas no conjunto de treinamento, verificando-se a quantidade de ocorrência de cada combinação em cada gênero musical. Considerando as combinações como os antecedentes das regras, é possível estabelecer se há um gênero predominante para cada uma das 27 combinações, sendo este considerado então o conseqüente daquela regra. A Tabela 2 exibe um exemplo para a combinação do antecedente “SE atributo 26 é baixo E atributo 28 é baixo E atributo 29 é alto”, avaliada no sistema de classificação do Grupo 3 (frequência).

Tabela 2: Número de ocorrências do antecedente “SE atributo 26 é baixo E atributo 28 é baixo E atributo 29 é alto” no Grupo 3.

Gênero	Ocorrência	Gênero	Ocorrência
Tango	147	Forró	141
Bachata	158	Pagode	203
Bolero	152	Sertanejo	131
Merengue	143	Gaúcha	178
Salsa	101	Axé	135

Na Tabela 2, para essa combinação, o gênero *Pagode* apresenta maior número de ocorrências. Conclui-se então que este gênero é o conseqüente desta regra linguística, o que implica na maior probabilidade de acertos do modelo. Neste caso, a regra linguística é descrita da seguinte forma: *SE atributo 26 é baixo E atributo 28 é baixo E o atributo 29 é alto, ENTÃO o gênero é Pagode e*, de acordo com a Tabela 1, associa-se a este gênero a emoção *felicidade*. Desta forma, por meio da extração do conhecimento do conjunto de dados, a base de regras do tipo “se-então” é obtida para cada um dos três SBRFs propostos.

O ajuste dos parâmetros das funções de pertinência das entradas dos SBRFs são obtidos por meio do treinamento dos dados no sistema *neuro-fuzzy* conforme descrito na Seção 3.2. As funções do tipo Gaussianas foram selecionadas por apresentarem maior suavidade entre as fronteiras das classes e menor erro, quando comparadas com testes executados para os tipos trapezoidais e triangulares. Os parâmetros das funções de pertinência obtidos são inseridos nos sistemas de classificação *fuzzy*, que operam com sistema de inferência do tipo *Mamdani* [21], cujas saídas são conjuntos *fuzzy*. A Tabela 3 exibe os parâmetros das funções de pertinência de entrada obtidas para cada um dos 3 grupos de entrada dos sistemas de classificação, enquanto a Figura 4 apresenta as funções de pertinência geradas por estes parâmetros.

Tabela 3: Parâmetros de Entrada das Funções de Pertinência.

GRUPO 1			
Atributos	Parâmetros		
	Baixo	Médio	Alto
Atributo 2	0,2127 ; 0,0003	0,2124 ; 0,5003	0,2073 ; 1,0020
Atributo 3	0,2093 ; 0,0128	0,2096 ; 0,5064	0,2097 ; 0,9999
Atributo 6	0,2130 ; 0,0006	0,2123 ; 0,4999	0,2046 ; 1,0030

GRUPO 2			
Atributos	Parâmetros		
	Baixo	Médio	Alto
Atributo 7	0,2123 ; 0,0000	0,2123 ; 0,5000	0,2123 ; 1,0000
Atributo 12	0,2123 ; 0,0000	0,2123 ; 0,5000	0,2123 ; 1,0000
Atributo 22	0,2123 ; 0,0000	0,2123 ; 0,5000	0,2123 ; 1,0000

GRUPO 3			
Atributos	Parâmetros		
	Baixo	Médio	Alto
Atributo 26	0,1788 ; 0,0035	0,1788 ; 0,4246	0,1788 ; 0,8457
Atributo 28	0,1293 ; 0,0011	0,1293 ; 0,3056	0,1293 ; 0,6101
Atributo 29	0,2123 ; 0,0000	0,2123 ; 0,5000	0,2123 ; 1,0000

As funções de pertinência da saída de cada um dos 3 sistemas de classificação podem ser vistas na Figura 5 e são construídas a partir dos gêneros musicais que ocorrem nas saídas das regras linguísticas. As fronteiras para as funções gaussianas são obtidas dividindo-se o intervalo dos parâmetros proporcionalmente pelo número de ocorrências dos gêneros na saída das regras. A Figura 5 mostra, respectivamente, as representações das saídas dos 3 sistemas de classificação referentes aos grupos: Batidas Sonoras, Timbre e Frequência.

Para validação dos resultados dos SBRFs, foram contabilizados o número de acertos da classificação, aplicando-se os dados de teste nos sistemas elaborados. Das 3000 instâncias que compõem a base de dados LMD, 2700 instâncias foram utilizadas para o treinamento dos sistemas e, as demais, utilizadas para o teste (validação) da classificação. Com um algoritmo desenvolvido no

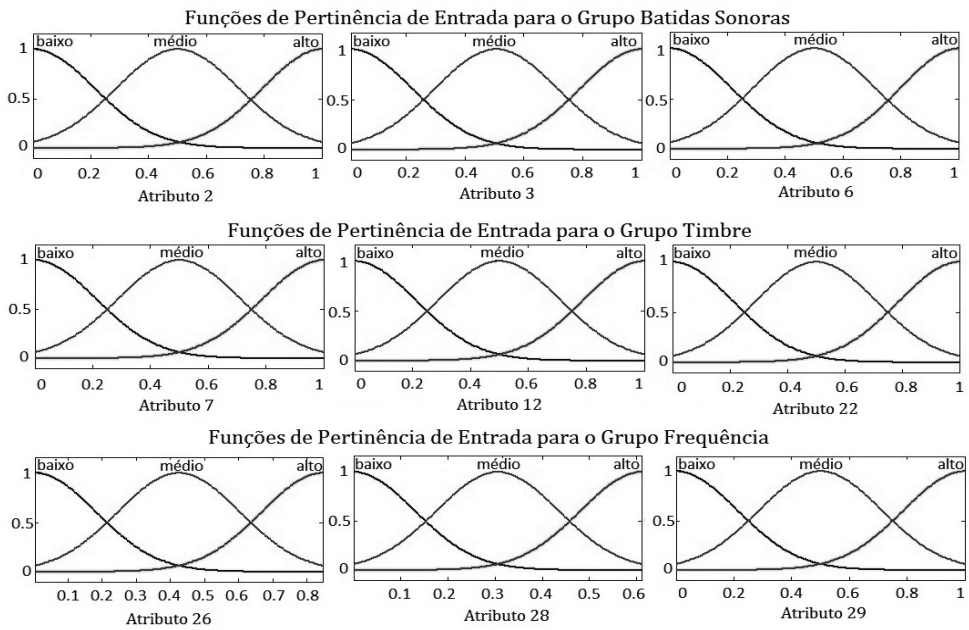


Figura 4: Funções de Pertinência das Entradas dos SBRFs.

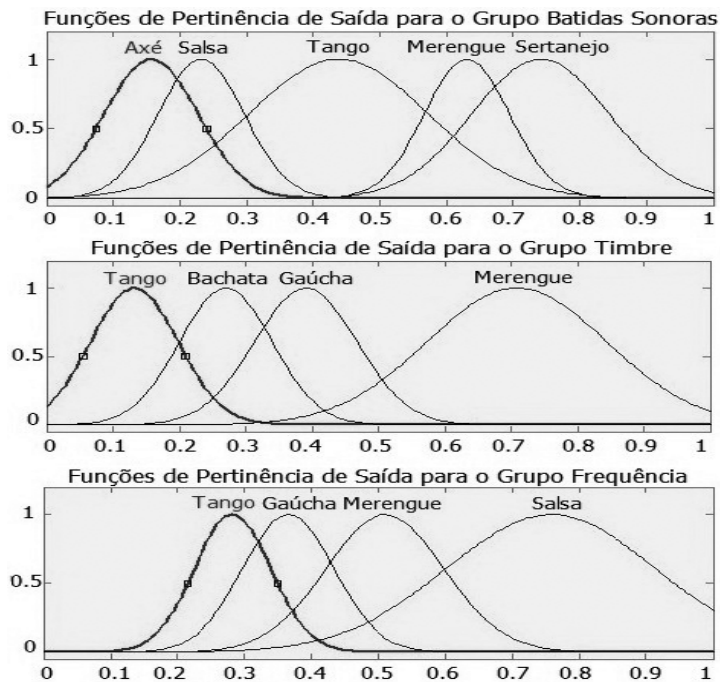


Figura 5: Funções de Pertinência das Saídas dos Sistemas Fuzzy.

software MATLAB, os dados de teste foram inseridos nos respectivos SBRFs propostos para obter um índice de acerto da classificação. Se o parâmetro de saída da instância de teste está contido entre as fronteiras do gênero correspondente, é considerado acerto; caso contrário, é considerado erro. O resultado geral dos testes apontam uma acurácia de 75,33%, 71,67% e 70,83%, respectivamente, para os sistemas de classificação dos grupos Batidas Sonoras, Timbre e Frequência.

6 CONCLUSÃO

Neste trabalho, foram apresentadas duas propostas de modelagem matemática para a classificação automática de gêneros musicais latinos considerando suas emoções. Para tanto, foram empregados o método de classificação Bayesiana, o qual avalia a dependência das variáveis de forma probabilística pela aplicação do teorema de Bayes, juntamente com o algoritmo *BayesRule*, e o método de classificação *fuzzy*, o qual considera fusões entre as fronteiras dos gêneros. Estas abordagens possibilitaram, além da classificação automática dos gêneros musicais, a comparação dos resultados.

Por meio da aplicação dos modelos de classificação propostos, obtém-se um índice de acerto geral, considerando-se a média aritmética dos índices de acerto (acurácias) dos 3 grupos de entrada, de 98,08% com o método de classificação Bayesiana e de 72,61% com o SBRF, o que evidencia que o método de classificação Bayesiana foi melhor aplicado a este problema de classificação, devido ao tratamento probabilístico das variáveis. Vale ressaltar que, apesar dos conjuntos de regras linguísticas terem sido gerados de formas distintas para cada metodologia, ambos apresentaram muitas similaridades entre as regras, demonstrando coerência entre os modelos. Por exemplo, para o grupo de entradas “batida sonora”, a regra linguística *SE atributo 2 é baixo E atributo 3 é baixo E atributo 6 é baixo, ENTÃO o gênero é Axé* é identicamente encontrada em ambos os métodos de classificação. Similarmente no grupo “timbre”, a regra linguística *SE atributo 7 é baixo E atributo 12 é baixo E atributo 22 é baixo, ENTÃO o gênero é Tango* também se encontra em ambos os modelos e, finalmente, para o grupo “frequência”, a regra *SE atributo 26 é baixo E atributo 28 é baixo E atributo 29 é médio, ENTÃO o gênero é Salsa* é uma das quais se repete em ambos os métodos aplicados.

Como perspectivas de continuidade deste trabalho, propõe-se o desenvolvimento de modelos estatísticos de predição. Algoritmos de árvores de decisão com critérios de poda serão desenvolvidos e aplicados, comparando-se os resultados obtidos por meio da aplicação de árvores de decisão *fuzzy*.

ABSTRACT. The goal of this paper is to investigate intelligent classification methods for the Latin musical genres classification task from a real data set. In this study, the predominant emotional characteristics of each genre are also considered, with the aim of enabling the user to choose music according to the required emotion. The proposed methods are based on the Bayesian classification method, which uses the algorithm *BayesRule*, able to investigate data probabilistic uncertain, and on the *fuzzy* classification method, whose parameters of the functions are fitted by a *neuro-fuzzy* system, which allows to evaluate mergers between different

musical genres. These methodologies extract linguistic classification rules, which allow comparing results and performing an intelligent classification of the data by taking into account uncertainties and fusion of musical genres.

Keywords: musical genres, Bayesian classification, *fuzzy* classification.

REFERÊNCIAS

- [1] J.J. Aucouturier & F. Pachet. Representing musical genre: A state of the art. *Journal of New Music Research*, **32** (2003), 83–93.
- [2] G.E.P. Box & G.C. Tiao. *Bayesian inference in statistical analysis*. John Wiley and Sons, Canadá, (1992).
- [3] J. Cheng, R. Greiner, J. Kelly, D. Bell & W. Liu. Learning bayesian networks from data: An information-theory based approach. *Artificial Intelligence*, **137** (2002), 43–90.
- [4] P.J. Donnelly & J.W. Sheppard. Classification of musical timbre using bayesian networks. *Computer Music Journal*, **37**(4) (2014), 70–86.
- [5] F. Fernández, F. Chávez, R. Alcalá & F. Herrera. Musical genre classification by means of fuzzy rule-based systems: A preliminary approach. *IEEE Congress on Evolutionary Computation, IEEE CEC*, **13**(2) (2011), 303–319.
- [6] S. Goyal & E. Kim. Application of fuzzy relational interval computing for emotional classification of music. *IEEE Conference on Norbert Wiener in the 21st Century (21CW)*, (1943), 1–8.
- [7] J. Han, M. Kamber & J. Pei. *Data mining: Concepts and techniques*. Morgan Kaufmann, USA, (2012).
- [8] E.R. Hruschka, M.C. Nicoletti, V.A. Oliveira & G.M. Bressan. Markov-blanket based strategy for translating a bayesian classifier into a reduced set of classification rules. *7th International Conference on Hybrid Intelligent Systems*, (2007), 192–197.
- [9] G. James, D. Witten, T. Hastie & R. Tibshirani. *An introduction to statistical learning*. Springer, USA, (2013).
- [10] C.N. Silla Jr, A.L. Koerich & C.A.A. Kaestner. The latin music database. *Proceedings of 9th International Conference on Music Information Retrieval*, (2008), 451–456.
- [11] C.N. Silla Jr, A.L. Koerich & C.A.A. Kaestner. A feature selection approach for automatic music genre classification. *International Journal of Semantic Computing*, **3**(2) (2009), 183–208.
- [12] P.N. Juslin & P. Laukka. Expression, perception, and induction of music emotion: A review and a questionnaire study of everyday listening. *Journal of New Music Research*, **33**(3) (2004), 217–238.
- [13] K. Kashino, K. Nakadai, T. Konoshita & H. Tanaka. Application of bayesian probability network to music scene analysis. *Computational Auditory Scene Analysis*, **1** (1996), 115–137.
- [14] Y. Kim, E.M. Schmidt, R. Migneco, B.G. Morton, P. Richardson, J. Scott, J.A. Speck & D. Turnbull. Music emotion recognition: A state of the art review. *ISMIR*, (2010), 255–266.
- [15] B. Lerner & R. Malka. Investigation of the k2 algorithm in learning bayesian network classifiers. *Applied Artificial Intelligence*, **25** (2011), 74–96.

- [16] T. Li & M. Ogihara. Music genre classification with taxonomy. *IEEE*, **5** (2005), 197–200.
- [17] N.C. Maddage, C. Xu & Y. Wang. An svm-based classification approach to musical audio. *Proceedings of the 4th International Conference on Music Information Retrieval*, (2003).
- [18] M. Mitchell. *An introduction to genetic algorithms*. Massachusetts Institute of Technology Press, USA, (1996).
- [19] R.E. Neapolitan. *Learning bayesian networks*. Prentice Hall Series in Artificial Intelligence. Pearson Prentice Hall, USA, (2003).
- [20] J. Pearl. *Probabilistic reasoning in intelligent Systems: networks of plausible inference*. Morgan Kaufmann, USA, (1988).
- [21] W. Pedrycz & F. Gomide. *An introduction to fuzzy sets*. Massachusetts Institute of Technology Press, USA, (1998).
- [22] J.A. Russel. A circumplex model of affect. *Journal of Personality and Social Psychology*, **39**(6) (1980), 1161–1178.
- [23] C.L. Santos & C.N. Silla Jr. The latin music mood database. *Journal on Audio, Speech and Music Processing*, (2015).
- [24] D. Temperley. *Music and probability*. Massachusetts Institute of Technology Press, USA, (2007).
- [25] R.E. Thayer. *The biopsychology of mood and arousal*. Oxford University Press, USA, (1989).
- [26] G. Tzanetakis & P. Cook. Marsyas: A framework for audio analysis. *Journal Organized Sound*, **4** (1999), 169–175.
- [27] Y. Yang & H.H. Chen. *Music emotion recognition*. CRC Press, USA, (2011).
- [28] Y. Yang, C. Liu & H.H. Chen. Music emotion classification: A fuzzy approach. *Proceedings of the 14th ACM international conference on Multimedia*, (2006), 81–84.
- [29] L.A. Zadeh. Fuzzy sets. *Information and Control*, (1965), 338–353.