

Métodos de Aprendizado de Máquina Aplicados ao Reconhecimento Autoral e Temporal de Bandas de *Rock and Roll*

Oliver C Jorge, Heitor S. Lopes

Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial (CPGEI)

Universidade Tecnológica Federal do Paraná (UTFPR)

Av. 7 de setembro, 3165, 80230-901 Curitiba (PR), Brasil

olivercj@gmail.com, hslopes@utfpr.edu.br

Resumo—Embalando gerações desde o final da década de 40, o *Rock and Roll* é um dos estilos musicais mais conhecidos mundialmente, com diversas melodias de sucesso e mensagens que influenciaram gerações. Músicas podem ser traduzidas como sequências vibratórias numéricas e, portanto, apresentam padrões interpretáveis por métodos computacionais. Este artigo apresenta técnicas de Mineração de Dados utilizadas para identificação autoral e temporal de músicas de cinco bandas famosas de *Rock and Roll*. Foram extraídos 34 metadados de arquivos digitais de toda a discografia das bandas. As informações coletadas foram redimensionadas em um campo bi-dimensional com o algoritmo t-SNE para agrupamento e identificação de coerência de produção artística por banda para, enfim, serem aplicados quatro métodos de classificação de dados: por árvore de decisão, regras de classificação baseado em exceções, regras de classificação de acurácia máxima e redes neurais. Os resultados identificaram que acurácia e compreensibilidade humana são inversamente proporcionais, obtendo 81,1% como a melhor taxa de acerto para regras de classificação, contra 94,7% para redes neurais nos testes de classificação autoral. Nos testes de classificação temporal as taxas foram de 80,0% para regras de classificação contra 90,2% para redes neurais. Ao final deste artigo percebeu-se que os resultados obtidos são diretamente influenciados pela constância na produção das bandas ao longo de sua atividade profissional, havendo melhor taxa de acerto para aquelas bandas cujas músicas produzidas são mais coerentes entre si.

Index Terms—Classificação de Dados, Rock and Roll, Estilos Musicais, Reconhecimento de Padrões.

I. INTRODUÇÃO

A música tem o poder de influenciar diretamente o corpo humano. Segundo Langslet [1], a música pode nos proporcionar bem-estar e estimular o pensamento ou ainda pode nos irritar e fazer mal fisicamente. No viés econômico, o mercado musical movimenta uma receita anual gigantesca. De acordo com dados da federação internacional da indústria fonográfica, IFPI [2] (*International Federation of the Phonographic Industry*), o mercado musical gerou uma receita de US\$ 17,3 bilhões no ano de 2017, e esta receita vem crescendo consecutivamente nos últimos anos.

Além das questões físicas e econômicas, a música é capaz de gerar um fascínio e magia que acompanham o ser humano desde as culturas mais primitivas, sendo uma forma singular de expressão de sentimentos, pensamentos, emoções e histórias.

Se uma música é capaz de gerar tamanha admiração ao ser humano, como representar e ensinar uma inteligência computacional suas muitas variações e nuances?

O som pode ser representado como uma sequência vibratória mensurável de forma numérica que, interpretada da forma correta, gera o resultado esperado. É a partir desta representação numérica que músicas podem ser armazenadas de forma digital e, inversamente à esta lógica, também podem produzir resultados numéricos que as representem. A partir de algumas regras matemáticas aplicadas a arquivos digitais de músicas, é possível extrair características numéricas, ou metadados, que podem servir como uma impressão digital única. Se tais impressões digitais forem extendidas para determinadas bandas ou épocas, tornam-se representações que extrapolam a individualidade de uma única produção, podendo ser, então, comparadas e aprendidas para fins de classificação ou de identificação.

Neste artigo foram aplicados métodos de inteligência computacional a metadados extraídos de arquivos digitais de música, com os seguintes objetivos: (1) reconhecer a banda à qual uma música pertence, demonstrando que uma banda apresenta uma coerência musical que a defina, e (2) identificar a época que a música foi composta, supondo-se que cada época possua uma determinada característica melódica distinta que permita esta classificação (como a *disco music* que representou os anos 70).

A. Trabalhos Correlatos

O interesse em aprendizado de máquina para classificação de arquivos musicais não é recente. Estudos e técnicas foram desenvolvidas e aplicadas de forma a facilitar o reconhecimento e classificação para fins pessoais ou comerciais.

Choi et al [3] apresentaram os benefícios de utilizar uma rede neural convolucional baseada em *transfer learning* para classificação de estilos musicais e, assim, aumentar o desempenho a partir dos benefícios desta técnica. Jeong et al [4] apresentaram uma nova rede neural baseado em *deep learning* que, a partir de metadados de arquivos musicais, faz uma análise temporal individual da música, ou seja, analisa segundo

a segundo as características obtendo, assim, uma acurácia melhor para a classificação de gêneros.

Bahuleyan [5] utilizou espectrogramas para fins classificatórios. A partir destes, ele comparou o desempenho de uma rede CNN com outros modelos clássicos de aprendizado de máquina, como o *Random Forest*, sendo capaz de apresentar as vantagens e desvantagens de cada modelo.

Recentemente, surgiram trabalhos que objetivaram o gênero musical com base no sentimento que as mesmas despertam no ouvinte. Por exemplo, [6] utilizaram máquinas de vetor de suporte (SVM) e redes neurais para a classificação de gênero. Por outro lado, [7] fizeram a classificação de sentimento despertado no ouvinte (feliz, triste, pacífico, nervoso) utilizando SVM e árvores de decisão.

Este trabalho apresenta uma abordagem mais específica, com técnicas de aprendizado de máquina aplicadas com o fim de reconhecer padrões autorais e temporais em bandas de um mesmo gênero musical, onde as semelhanças sonoras e rítmicas tendem a ser mais evidentes, ao invés de fazer uma classificação mais ampla.

II. MÉTODOS

Esta pesquisa limitou-se ao estilo musical de *Rock and Roll* que, por ser um estilo existente desde o final da década de 40, é muito extenso havendo milhares de produções nacionais e internacionais. Para este estudo foram escolhidas apenas algumas bandas para a composição do *dataset*.

A. Seleção de Bandas

Os critérios de escolha das bandas foram os seguintes: (1) que tivessem produção musical relativamente extensa, (2) que fossem bandas conhecidas com repertório influente, e (3) que tivessem um intervalo temporal de criação musical compatível, de modo a ampliar as possibilidades de comparações temporais. Assim, as bandas selecionadas foram: *AC/DC*, *Iron Maiden*, *Led Zepellin*, *Pink Floyd* e *Metallica*.

Pink Floyd é definida como uma banda de *rock* progressivo ou, em seus primeiros álbuns lançados, *rock* psicodélico. Isto significa que suas músicas possuem uma representação instrumental muito melódica, com experimentações sonoras e ritmos pouco acelerados.

Led Zepellin por sua vez mudou muito seu estilo musical ao longo do tempo, debutando como uma banda de *rock and roll* clássica muito influenciada pelo *blues* e pelo *folk*. Com o passar dos anos a banda mudou o seu estilo para produções mais rápidas com presença marcante de solos de guitarra e bateria mais rápidos, sendo reconhecida como uma das precursoras dos estilos *heavy metal* e *hard rock*, finalizando sua produção com músicas mais dançantes com forte influência do *groove* dos anos 70.

As bandas *Iron Maiden* e *Metallica* são reconhecidas bandas do estilo *heavy metal* com muitas similaridades entre suas produções musicais. A velocidade e energia melódica das guitarras e baterias de ambas as bandas são características marcantes em toda a sua discografia.

Finalmente, *AC/DC* é geralmente descrita como uma banda de *hard rock*. Suas músicas possuem uma identidade sonora facilmente reconhecida, com ritmos marcantes de bateria e solos de guitarra frequentemente velozes além de um vocal agudo único.

A Figura 1 mostra que todas as bandas apresentam um período de produção musical concomitante. Além das características melódicas e temporais, estas bandas apresentam uma boa quantidade de produção musical. A discografia completa destas bandas gerou um total de 638 músicas. A Tabela I mostra os dados quantitativos de produção das bandas.

Tabela I
COMPARAÇÃO ENTRE AS BANDAS E SUA PRODUÇÃO MUSICAL.

Banda	Período de Produção	Num. álbuns	Num. músicas
AC/DC	1976-2000	13	130
Iron Maiden	1980-2015	16	158
Led Zepellin	1969-1982	9	81
Metallica	1983-2008	9	95
Pink Floyd	1967-2014	15	174

B. Extração de Características

Uma etapa fundamental para classificação de dados é a extração de metadados dos arquivos de áudio, de modo a criar um *dataset* numérico. Foi utilizada a biblioteca Python *pyAudioAnalysis* [8] que é capaz de gerar um total de 34 metadados numéricos a partir de um arquivo em extensão WAV. A biblioteca requer que todos os áudios sejam convertidos para canal único (músicas em formato monoaural) e, para normalização dos valores numéricos, todos os arquivos foram parametrizados para uma frequência de 16.000 Hz. A lista dos metadados extraídos e suas respectivas descrições pode ser vista na Tabela II.

C. Softwares Utilizados

Para a aplicação dos métodos de aprendizado de máquina e mineração de dados, foram utilizados os *softwares* Orange [9] e Weka [10]. Ambos são de domínio público e rodam em diversos sistemas operacionais.

O Orange, desenvolvido em Python, se destaca pela sua interface gráfica, produzindo bons resultados visuais que facilitam a compreensão dos resultados. O Weka, desenvolvido em Java, possui uma extensa lista de bibliotecas e funções gratuitas provenientes de diversos desenvolvedores, tornando-o um *software* com ampla variedade de métodos aplicáveis.

D. Metodologia aplicada

A elaboração deste trabalho consistiu de três fases:

- Análise preliminar individual das bandas de forma a identificar coerência entre as produções ao longo de sua discografia. Esta análise serve de apoio para os resultados das classificações onde espera-se que uma banda coerente seja facilmente classificada.
- Aplicação de métodos de classificação para identificação autoral de bandas.

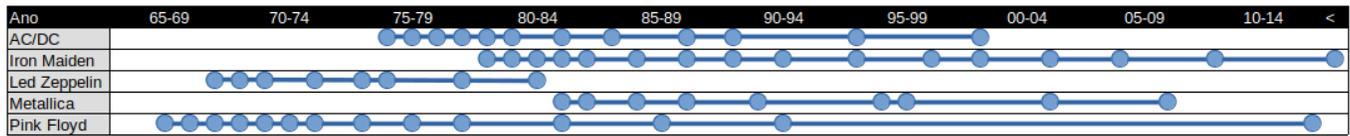


Figura 1. Linha do tempo para produção musical das 4 bandas incluídas neste estudo. A linha contínua identifica período ativo de produção musical enquanto que os círculos representam lançamentos de álbuns.

Tabela II
METADADOS EXTRAÍDOS A PARTIR DA BIBLIOTECA *pyAudioAnalysis*.

Característica	Descrição
Taxa de Cruzamentos por Zero	Pode ser interpretada como a medição de “barulho” de uma música. Tende a apresentar valores altos para sinais barulhentos.
Energia	Identifica a potência de um sinal de áudio. Batidas, vocais e guitarras fortes elevam o valor da Energia.
Entropia de Energia	É a medida da variação abrupta da Energia. Quanto menor o valor da Entropia, mais agressiva é a variação da Energia.
Centróide Espectral	Mede o centróide do espectrograma, onde está a maior frequência de Energia do espectrograma. Quanto maior o valor do centróide, mais Energia está concentrada nas frequências mais altas, ou seja, nos sons mais agudos.
Propagação Espectral	Mede o formato do espectrograma. Quanto menor o valor, menor é a variação do espectrograma em comparação com o centróide.
Entropia Espectral	Identifica a imprevisibilidade e desorganização de um espectrograma. Quanto mais variado é um espectrograma, maior a sua Entropia.
Fluxo Espectral	Mede a taxa de variação de um som ao longo de seu espectrograma.
Saída Espectral	A frequência abaixo da qual 90% da distribuição do espectrograma está concentrada, ou seja, a frequência onde está a menor concentração de som.
MFCCs	Um conjunto de 13 coeficientes que juntos identificam o espectrograma de um som de acordo com a “escala de mel” (uma representação em Hertz de melodias sonoras, o que aproxima os resultados sonoros ao conhecido pelo ser humano).
Vetor Cromático	Uma representação de 12 elementos da Energia de um espectrograma, em que cada uma das 12 entradas do vetor representa cada um dos semitons das classes de notas temperadas iguais aos da música do tipo ocidental (espaçamento de semitons).
Varição Cromática	O desvio padrão dos 12 coeficientes cromáticos.

- Aplicação de métodos de classificação para identificação temporal.

As análises individuais foram feitas com o *software* Orange aplicando o algoritmo t-SNE (*T-distributed Stochastic Neighbor Embedding*). Este algoritmo facilita a visualização de dados multi-dimensionais reduzindo-os a um espaço bi-dimensional, sendo ideal para este problema, visto que cada música possui 34 características numéricas que, ao final, podem ser visualizadas em um plano Cartesiano bi-dimensional. Os dados similares são representados mais próximos uns aos outros, enquanto que dados diferentes dos demais são representados mais distantes. Bandas com produção musical coerente devem apresentar dados próximos.

A fase de classificação de dados do trabalho utilizou quatro métodos fornecidos pelo *software* Weka:

- **C45**: que gera uma de árvore de decisão.
- **RIDOR**: gera regras de classificação baseado em exceções, começando por uma regra de classificação válida e expandindo a mesma listando suas exceções.
- **RIPPER**: gera regras de classificação baseado em acurácia máxima, começando com uma regra de classificação válida de acurácia mínima pré-definida e expandindo a mesma com refinamentos sucessivos até atingir a melhor acurácia possível.
- **Multi Layer Perceptron (MLP)**: para classificação a partir de uma rede neural com diversas camadas e o algoritmo de *backpropagation*.

Para todos os métodos citados, foram utilizados os seus parâmetros padrão. Para a rede neural MLP, foi utilizada uma camada oculta 34 neurônios e, em uma segunda execução, com 68 neurônios. Todos os métodos de classificação aplicados no Weka foram validados a partir de validação cruzada com dez *folds*, ou seja, o *dataset* foi automaticamente dividido em dez partes de igual tamanho onde nove partes foram usadas para treinamento e uma parte para validar o método encontrado. Isto é repetido 10 vezes comutando-se as partes [11].

A partir do *dataset* original foi criado uma partição de treino com 512 registros e uma arquivo de teste, com 126 registros (uma proporção de 80% - 20% respectivamente do total original de dados). Todos os registros foram selecionados de forma aleatória e respeitando a mesma proporção por banda.

Para a classificação por banda, o atributo-alvo é o nome da banda. Já para a classificação temporal, foram feitos dois testes sendo um com intervalo temporal de 5 anos (levando a 10 grupos) e outro de 10 anos (levando a 5 grupos). Para conseguir este intervalo definido, foi utilizado o método para discretização de classes do Weka *Discretize*.

A Figura 1 e a Tabela I mostram que as proporções de músicas por banda e por período temporal não são uniformes e, por isto, houve a necessidade de se balancear o *dataset* utilizando-se o método *classBalancer*. O método foi escolhido por efetuar um balanceamento aplicando pesos aos grupos alvo para classificação de forma que não há perda nem replicação de material, mantendo o *dataset* com seus dados originais.

III. RESULTADOS

Os primeiros resultados analisados foram os testes preliminares aplicados sobre as bandas de forma individual, a fim de identificar similaridades ao longo de sua produção musical. As bandas *AC/DC*, *Iron Maiden*, *Metallica* e *Pink Floyd* retornaram uma linha relativamente próxima em todo o seu eixo

diagonal, mostrando que deve existir uma correlação numérica entre as duas dimensionalidades geradas pelo método t-SNE. Percebe-se que nestas bandas o valor do eixo X reflete no valor do eixo Y de forma direta, e que esta correlação também deve existir conforme o ano de lançamento da música. Para as bandas *Iron Maiden*, *AC/DC* e *Metallica*, quanto menor o valor do eixo Y, mais recente é o ano de lançamento da música. A banda *Pink Floyd* apresentou comportamento inverso, havendo uma concentração das músicas mais recentes quanto maior é o valor do eixo Y. A Figura 2 representa o resultado do método aplicado sobre a banda *Iron Maiden*, onde é possível perceber facilmente a linha diagonal resultante, bem como a concentração de músicas mais recentes na base da diagonal, representadas pela cor amarelada, e a concentração de músicas mais antigas no topo da diagonal, representadas pela cor azulada. A Figura 3 representa o resultado da aplicação do método t-SNE sobre a produção musical da banda *Led Zeppelin*. Nota-se que o resultado gerado é mais difuso, com seus dados dispersos ao longo dos eixos mostrando não haver correlação clara entre estes. Também não há uma divisão temporal visível, havendo músicas de todas as datas de lançamento ao longo do gráfico.

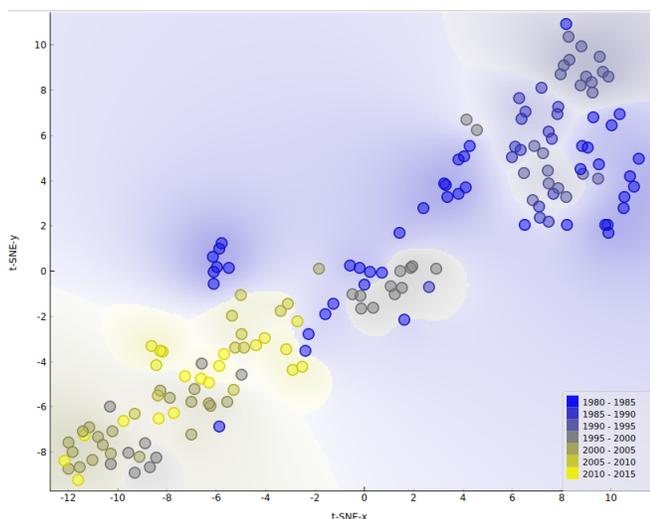


Figura 2. Aplicação do método t-SNE sobre a produção musical da banda *Iron Maiden*.

A. Classificação por Banda

Os métodos de classificação foram executados na sequência apresentada na Seção *Metodologia Aplicada*. Inicialmente, foram aplicados os métodos de classificação com resultados mais compreensíveis ao ser humano, sejam por árvore de decisão (C45) ou por regras de classificação (RIDOR e RIPPER), mas os resultados não foram totalmente satisfatórios. A melhor taxa de classificação encontrada dentre estas técnicas foi obtida com o método RIDOR, com uma acurácia média de 81,1% para a banda *Metallica*, e de apenas 58,0% para a banda *Led Zeppelin*.

Em seguida foi aplicado o método *Multi Layer Perceptron*. A rede com 34 neurônios na camada oculta obteve resultados

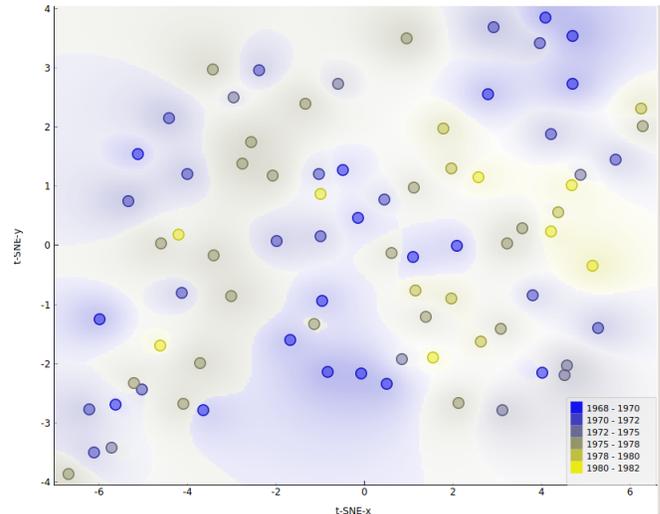


Figura 3. Aplicação do método t-SNE sobre a produção musical da banda *Led Zeppelin*.

bem interessantes onde a melhor e pior classificação foram novamente obtidas com as bandas *Metallica*, com 94,7% de acurácia, e *Led Zeppelin*, com 67,9% de acurácia.

Inicialmente o método *Multi Layer Perceptron* com 34 neurônios na camada escondida, com aplicação de peso de aprendizado 0,3 e momento aplicado aos pesos de 0,2 foi executado usando o arquivo para treino. Este modelo foi salvo para em seguida ser testado usando o arquivo de teste. Os resultados são mostrados na Tabela III, com representação gráfica dos mesmos na Figura 4, para melhor visualização.

Tabela III
ACURÁCIA DA CLASSIFICAÇÃO COM MLP PARA TREINO (ESQUERDA) E TESTE (DIREITA). TP SIGNIFICA *True Positive* E FP SIGNIFICA *False Positive*.

Banda	Treino TP	Treino FP	Teste TP	Teste FP
a = AC/DC	0,913	0,040	0,885	0,007
b = Iron Maiden	0,929	0,039	0,871	0,015
c = Led Zeppelin	0,692	0,036	0,750	0,053
d = Metallica	0,921	0,006	1,000	0,019
e = Pink Floyd	0,836	0,056	0,765	0,088
Média	0,858	0,039	0,854	0,036

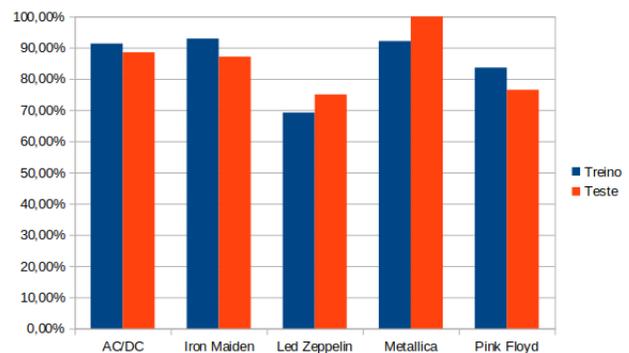


Figura 4. Acurácia obtida para treino e teste utilizando MLP para classificação autoral de bandas.

Nota-se pelos resultados apresentados na Tabela III que as bandas *AC/DC*, *Iron Maiden* e *Metallica* alcançaram acurácia sempre acima de 91,0% na fase de treinamento, e sempre acima de 87,0% na fase de testes, sendo que a banda *Metallica* obteve acurácia de 100,0% nesta última fase. Estas são as bandas que se mantiveram constantes durante seu período de atividade. Tal como dito na Seção *Seleção de bandas*, *Iron Maiden* e *Metallica* sempre foram bandas de *heavy metal*, enquanto que *AC/DC* sempre foi uma banda de *hard rock*. Esta constância de estilo deve ter resultado em maior coerência de produção musical tornando-as bandas fáceis de serem reconhecidas e classificadas.

Por sua vez, a banda *Led Zeppelin*, que possui um histórico de produção musical muito distinto entre si variando do *blues* introspectivo em suas primeiras produções a um som mais dançante em seus últimos álbuns, obteve uma acurácia inferior às outras bandas na classificação, alcançando valores abaixo de 70% na fase de treinamento. Visto que a comparação entre as produções musicais é feita totalmente de forma numérica, estes resultados muito provavelmente estão relacionados com os apresentados nas análises individuais feitas com as reduções numéricas de suas características com o t-SNE, como foi observado anteriormente nas Figuras 2 e 3.

B. Classificação por Época

Para classificar as músicas temporalmente, o *dataset* foi discretizado em 10 grupos, com períodos próximos de 5 anos. Os mesmos métodos usados anteriormente (por árvore de decisão e regras de classificação) foram aplicados sobre esta nova base. Destes, o método RIDOR obteve o melhor resultado, alcançando uma acurácia máxima de 80,0% para o período de 1998 a 2004 e uma acurácia mínima de 23,1% para o período pré-1970, ao custo de 714 regras de classificação.

A quantidade de regras encontradas mostra um grande desbalanceamento entre acurácia e compreensibilidade. Apesar do valor da acurácia ser razoavelmente bom, a capacidade de compreensão de 714 regras é praticamente nula. Para tentar diminuir este *overfitting*, os parâmetros de poda aplicáveis aos métodos foram ajustados e os testes refeitos, resultando em um conjunto de 97 regras de classificação. Após a poda, a acurácia foi bastante afetada conseguindo, em média, 37,1% de acerto. Isto mostra claramente a dificuldade de classificação temporal das músicas.

Por outro lado, foi aplicado o método *Multi Layer Perceptron* (MLP) com os mesmos parâmetros aplicados anteriormente para classificação de bandas, com 34 neurônios em uma camada escondida. O melhor resultado apresentou uma acurácia máxima de 90,2% para o período de 1988 a 1991 e uma acurácia mínima de 30,6% para o período de 1970 a 1974. Apesar de ser um resultado melhor do que o encontrado nos métodos anteriores, ainda se manteve com uma média muito baixa como mostrado na Tabela IV.

Objetivanto melhorar os resultados para a predição temporal, uma nova análise foi feita. Neste caso, o *dataset* foi discretizado em 5 grupos, em períodos próximos de 10 anos. A mesma metodologia anterior foi seguida e, após ajustes de

Tabela IV
ACURÁCIA PARA CLASSIFICAÇÃO TEMPORAL UTILIZANDO MLP COM O *dataset* DISCRETIZADO EM 10 PERÍODOS DE APROXIMADAMENTE 5 ANOS.

Período	Verdadeiro Positivo	Falso Positivo
Pré - 1970	0,400	0,069
1970 - 1974	0,306	0,061
1975 - 1978	0,522	0,045
1979 - 1980	0,548	0,058
1981 - 1983	0,625	0,055
1984 - 1987	0,731	0,033
1988 - 1991	0,92	0,033
1992 - 1996	0,597	0,031
1997 - 2004	0,846	0,014
Pós - 2004	0,629	0,035
Média	0,611	0,043

parâmetros de poda para encontrar um equilíbrio entre acurácia e compreensibilidade, o método que apresentou melhor resultado foi o C45 com uma árvore composta de 173 nós e 87 folhas. Nesta árvore a acurácia máxima foi 78,7% para as músicas criadas pós-1997 e a acurácia mínima foi 48,8% para as músicas criadas pré-1974. Alterações no tamanho da árvore resultaram em diminuição brusca da acurácia dos resultados ao serem aplicadas podas mais agressivas, ou melhora mínima obtida ao custo de árvores muito maiores.

Novamente o método *Multi Layer Perceptron* apresentou resultados significativamente melhores. Ao final do treinamento, o método obteve uma acurácia máxima de 79,5% para o período pós-1997 e uma acurácia mínima de 58,1% para o período de 1975 a 1980, como mostrado na Tabela V.

Tabela V
ACURÁCIA PARA CLASSIFICAÇÃO TEMPORAL UTILIZANDO MLP COM O *dataset* DISCRETIZADO EM 5 PERÍODOS DE APROXIMADAMENTE 10 ANOS.

Período	Verdadeiro Positivo	Falso Positivo
Pré - 1974	0,630	0,133
1975 - 1980	0,581	0,122
1981 - 1987	0,674	0,055
1988 - 1996	0,732	0,056
Pós - 1997	0,795	0,031
Média	0,682	0,040

Nas duas análises observa-se que o período anterior a 1980 é de difícil classificação. Confrontando esta análise com a demonstração de produção musical por período apresentada na Figura 1, percebe-se que a produção deste período é quase que exclusivamente das bandas *Led Zeppelin* e *Pink Floyd*. A banda *Led Zeppelin* apresenta baixa coerência musical neste período, como mostrado na Figura 3, por causa de sua grande variação de estilos musicais ao longo de sua produção que também são refletidas na dificuldade de classificação temporal. Portanto, não há um estilo predominantemente marcante para esta classificação. Já as produções da banda *Pink Floyd* deste período foram, em sua maioria, da época psicodélica, ou seja, com excesso de experimentação instrumental e variações sonoras típicas deste estilo artístico.

O período que, em média, apresentou a melhor acurácia foram as décadas de 1980 e 1990, período de grande quantidade de exemplos disponíveis pelas produções musicais das bandas *AC/DC*, *Iron Maiden* e *Metallica*. Aqui um misto

de grande quantidade de exemplos provenientes de bandas coerentes pode ter elevado a acurácia média de classificação no período.

Dos anos 2000 em diante, os exemplos musicais do *dataset* foram quase que exclusivamente das bandas *Iron Maiden* e *Metallica*, com apenas uma exceção, sendo o último álbum lançado da banda *Pink Floyd* após um hiato de mais de vinte anos. Aqui, a acurácia média da classificação temporal pode ter sido elevada pela exclusividade de exemplos disponíveis, podendo ter saturado o treinamento com pouca variedade.

IV. CONCLUSÕES

A partir da biblioteca *pyAudioAnalysis* foi possível extrair características e metadados numéricos importantes de arquivos musicais. Com estes metadados foi possível analisar produções inteiras de bandas distintas, classificá-las e achar um padrão recorrente que pode ser interpretado como uma marca única que as identificam.

A análise feita com o algoritmo t-SNE identificou que as bandas que não tiveram variações em seus estilos musicais ao longo de seu período ativo de produção são mais coerentes numericamente, com músicas com padrões matemáticos mais evidentes.

Também ficou claro que a compreensibilidade das regras de classificação a partir dos metadados extraídos é inversamente proporcional à sua acurácia. Quanto mais compreensível ao ser humano, pior é a capacidade classificatória encontrada. Assim, os melhores resultados foram obtidos a partir dos testes feitos com redes neurais MLP.

As bandas de *heavy metal* *Metallica* e *Iron Maiden* e de *hard rock* *AC/DC* apresentaram uma grande coerência musical tornando sua identificação muito eficiente. Para estas bandas, os testes executados neste trabalho alcançaram taxas de acerto de 100,0%, 87,1% e 88,5%, respectivamente. Isto prova que mesmo havendo alterações notáveis na formação de seus integrantes, como a troca de vocalista nas bandas *AC/DC* e *Iron Maiden*, ou em trocas de guitarristas, baixistas e bateristas nas três bandas, alterações e adaptações comuns em bandas de longo tempo de existência, as mesmas possuem uma identidade musical que é única e que as representam independentemente de suas formações.

As bandas *Led Zeppelin* e *Pink Floyd* apresentaram uma coerência menor de produção musical, mesmo não havendo nenhuma alteração em sua formação original no caso da primeira, ou de apenas um guitarrista no caso da segunda, resultado de suas grandes variedades de estilos musicais em seus períodos de atividade. Esta variedade pode ter tornado as bandas expoentes no meio musical, sendo precursoras de novos estilos ou inovando artisticamente com músicas que marcaram gerações. No entanto, esta mesma variedade afetou a capacidade de classificação pelos métodos aplicados, mas ainda atingindo valores satisfatórios para este tipo de problema (75,0% e 76,5%, respectivamente).

Conclui-se que cada banda possui uma identidade musical, seja na forma rítmica, na potência de seus instrumentos, no virtuosismo de seus músicos, na capacidade vocal de seu

vocalista ou, mais possivelmente, em todas estas características juntas. Independentemente do motivo, é possível converter este conjunto de características em uma representação numérica, sendo possível a identificação e classificação de sua produção musical autoral.

Os resultados da identificação temporal não foram totalmente conclusivos. O *dataset* utilizado possui toda a discografia das bandas selecionadas, o que garantiu grande quantidade de exemplos disponíveis para a classificação proposta anteriormente, mas não foi suficiente para a identificação temporal pois a produção das bandas não foi constante no período analisado. Não se pode afirmar que não exista uma característica que identifique uma produção por período de tempo sem um *dataset* preparado para esta análise. Para obter uma hipótese definitiva, seria necessário uma coleta maior e mais ampla de músicas, desta vez focado exclusivamente na tentativa de explicar mudanças temporais independentemente da banda que a produziu.

Como recomendações de trabalhos futuros, sugere-se fazer uma nova amostragem de músicas focado exclusivamente na produção temporal para que possa ser feita uma análise mais robusta de características ou, ainda, focar em reconhecimento de características de estilos musicais, abrangendo assim várias facetas do *Rock and Roll* ou qualquer outro estilo artístico.

AGRADECIMENTOS

H.S. Lopes agradece ao CNPq pelas bolsas de pesquisa 311778/2016-0 e 423872/2016-8, bem como à Fundação Araucária pelo suporte financeiro através do PRONEX042/2018.

REFERÊNCIAS

- [1] L. Langslet, "Discurso de abertura," *Simpósio Música e Saúde*. In E. Rudd (Org.). *Música e saúde*, 1991.
- [2] IFPI. (2019, fev) Ifpi global music report 2018. [Online]. Available: <https://www.ifpi.org/news/IFPI-GLOBAL-MUSIC-REPORT-2018>
- [3] K. Choi, G. Fazekas, M. Sandler, and K. Cho, "Transfer learning for music classification and regression tasks," *arXiv preprint arXiv:1703.09179*, 2017.
- [4] I.-Y. Jeong and K. Lee, "Learning temporal features using a deep neural network and its application to music genre classification." in *Ismir*, 2016, pp. 434–440.
- [5] H. Bahuleyan, "Music genre classification using machine learning techniques," *arXiv preprint arXiv:1804.01149*, 2018.
- [6] A. Shakyia, B. Gurung, M. S. Thapa, M. Rai, and B. Joshi, "Music classification based on genre and mood," in *Proc. Int. Conf. on Computational Intelligence, Communications, and Business Analytics*. Springer, 2017.
- [7] A. Kartikay, H. Ganesan, and V. M. Ladwani, "Classification of music into moods using musical features," in *Proc. Int. Conf. on Inventive Computation Technologies*. IEEE, 2016.
- [8] T. Giannakopoulos, "pyaudioanalysis: An open-source python library for audio signal analysis." *PLoS ONE*, vol. 10, no. 12, 2015.
- [9] J. Demšar, T. Curk, A. Erjavec, Črt Gorup, T. Hočevar, M. Milutinovič, M. Možina, M. Polajnar, M. Toplak, A. Starič, M. Štajdohar, L. Umek, L. Žagar, J. Žbontar, M. Žitnik, and B. Zupan, "Orange: Data mining toolbox in python," *Journal of Machine Learning Research*, vol. 14, pp. 2349–2353, 2013.
- [10] E. Frank, M. A. Hall, and I. H. Witten, *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"*, 4th ed., M. Kaufmann, Ed., 2016.
- [11] W. Weinert and H. Lopes, "Data mining with a parallel rule induction system based on gene expression programming," *International Journal of Innovative Computing and Applications*, vol. 3, no. 3, pp. 136–143, 2011.