

Douglas Siqueira Gonçalves
Orientadora: Deborah Maria Vieira Magalhães

Seleção de Características para Predição da Emoção na Música

Picos - PI
30 de Maio de 2019

Douglas Siqueira Gonçalves
Orientadora: Deborah Maria Vieira Magalhães

Seleção de Características para Predição da Emoção na Música

Trabalho de Conclusão de Curso em Bacharelado em Sistemas de Informação pela Universidade Federal do Piauí. Nesta monografia é abordada a seleção de um conjunto de características musicais relevantes na predição de emoções.

Universidade Federal do Piauí
Campus Senador Heuvídio Nunes de Barros
Bacharelado em Sistemas de Informação

Picos - PI
30 de Maio de 2019

SELEÇÃO DE CARACTERÍSTICAS PARA PREDIÇÃO DA EMOÇÃO NA MÚSICA

DOUGLAS SIQUEIRA GONÇALVES


Monografia Aprovada como exigência parcial para obtenção do grau de Bacharel em Sistemas de Informação.

Data de Aprovação:

Picos – PI, 12 de junho de 2019



Profa. Deborah Maria Vieira Magalhães



Prof. Romuere Rodrigues Veloso e Silva



Prof. Antônio Oseas de Carvalho Filho

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca José Albano de Macêdo

S237a Gonçalves, Douglas Siqueira.
Seleção de características para predição da emoção na música. / Douglas Siqueira Gonçalves. -- Picos, PI, 2019.
36 f.
CD-ROM: 4 ¾ pol.

Trabalho de Conclusão de Curso (Bacharelado em Sistemas de Informação) – Universidade Federal do Piauí, Picos, 2019.
Orientador(A): Prof^ª. Deborah Maria Vieira Magalhães.

1. Áudio - Processamento. 2. Regressores (Algoritmos). 3. Caracterização do Humor - Música. I. Título.

CDD 005.1

Agradecimentos

Primeiramente agradeço a Deus, pois sem a ajuda dele não iria chegar a este momento. Agradeço a minha esposa Naara que me apoiou em todas as fases deste trabalho, como também no meu dia-à-dia.

Agradeço a meus pais, que deram todo o suporte para que eu conseguisse.

Agradeço a minha orientadora Deborah, que me direcionou neste trabalho de TCC.

Agradeço aos meus amigos Matheus Lima, Davi Luis, Tomaz Bisneto, Estevão Viana, Samuel Campos, Brena Maia, Railane Roque, Mateus Garcia, Milton Júnior, João Paulo, Raylla Vitoria, Wellington João, Tiago Mendes, Paulo Henrique, que me ajudaram nessa trajetória.

Também agradeço aos meus amigos de trabalho Beatriz, Viviane, Rafael, Alexandre, Nicolas, Lucas, Elizabeth e Bruno, que sempre me ajudaram, flexibilizando tempo para que pudesse desenvolver este trabalho.

A Bíblia (2 Timóteo 4:7)

Combati o bom combate, acabei a carreira, guardei a fé.

Resumo

A música é uma das artes existente que conseguem expressar as emoções, seja ela positiva ou negativa. As emoções são estudadas nas mais diversas áreas, dentre elas, a computação. Na computação, as emoções despertadas por meio da música, são utilizadas para entender o comportamento humano. Com este fim, foi desenvolvida uma área da computação para a extração e identificação de emoções musicais. Esta área de pesquisa utiliza métodos ou algoritmos existentes na Inteligência Artificial. Um dos algoritmos usados são os regressores. Estes recebem informações através das bases de dados e realizam a predição da emoção. Para obter melhores resultados, em alguns casos é necessário realizar uma etapa anterior chamada de seleção de características. Com a seleção, é possível reduzir a quantidade de características afim de obter um resultado tão bom quanto o da base completa. Além disso, pode obter uma redução da dimensão inicial da base de dados, o que na área de reconhecimento das emoções na música pode proporcionar em uma quantidade de característica mais efetivas em determinadas emoções. No entanto, algumas bases de dados públicas da área de reconhecimento de emoções na música não realizam este processo. Por este motivo, este trabalho realizar a seleção de características na base de dados pública. A base escolhida foi a PMemo, esta contém 794 músicas e determinou as emoções a partir do modelo de Theyers, como também, não realiza a seleção de características. O presente trabalho aplica a etapa de seleção de características com os seletores *Classifier Attribute Eval*, *Correlaion Attribute Eval*, *Relief Attribute Eval* e *CFS Subset Eval*, por conta dos dados na base serem continuou foi realizada a regressão com os métodos: *XGBoost*, *SVR* e o *MLR*. As métricas utilizadas para obter os resultados foram o *RMSE* e o *Correlation Coeficient*. Para definir as emoções foi utilizado o modelo de Theyers, que divide em *Arousal* e *Valence*. Por fim, este trabalho conseguiu um melhor resultado na predição comparado a base PMemo e determinou que as características de *audspec*, *MFCC* e *PCM* são mais impactante para *Arousal* e *Valence*.

Palavras-chaves: Seleção de Características, Regressão, Processamento de Áudio, Caraterização do Humor da Música

Abstract

Music is one of the existing arts that can express emotions, be it positive or negative. Emotions are studied in the most diverse areas, among them, computation. In computing, the emotions awakened through music are used to understand human behavior. To this end, an area of computation was developed for the extraction and identification of musical emotions. This research area uses methods or algorithms in Artificial Intelligence. One of the algorithms used is the regressors. They receive information through the databases and perform the prediction of emotion. For best results, in some cases it is necessary to perform an earlier step called a feature selection. With the selection, it is possible to reduce the amount of characteristics in order to obtain a result as good as the complete base. In addition, it can obtain a reduction of the initial dimension of the database, which in the area of recognition of the emotions in the music can provide in a quantity of characteristic more effective in certain emotions. However, some public databases of the area of emotion recognition in music do not perform this process. For this reason, this work perform the selection of features in the public database. The base chosen was the PMEmo, which contains 794 songs and determined the emotions from the Theyers model, as well as does not perform the selection of characteristics. The present work applies the feature selection stage with the Classifier Attribute Eval, Correlaion Attribute Eval, Relief Attribute Eval, and CFS Subset Eval selectors, because the data in the database are continued, and the regression was performed with the methods: XGBoost, SVR and MLR. The metrics used to obtain the results were RMSE and Correlation Coefficient. To define the emotions was used the model of Theyers, which divides into Arousal and Valence. Finally, this work achieved a better result in prediction compared to PMEmo base and determined that the characteristics of audspec, MFCC and PCM are more impacting for Arousal and Valence.

Keywords: Character Selection, Regression, Audio Processing, Characterization of Music Mood.

Lista de ilustrações

| | |
|--|----|
| Figura 1 – Etapas utilizadas no trabalho | 14 |
| Figura 2 – Níveis de separação das características musicais. | 18 |
| Figura 3 – Modelo dimensional | 19 |

Lista de tabelas

| | | |
|-----------|--|----|
| Tabela 1 | – Listagem dos Trabalhos Relacionados. | 17 |
| Tabela 2 | – Combinações possíveis a partir da quantidade de características selecionadas, método de regressão e método de seleção. | 24 |
| Tabela 3 | – Resultado do <i>RMSE</i> na combinação do <i>Classifier Attribute Eval MLR</i> | 25 |
| Tabela 4 | – Resultado do <i>Correlation Coeficient</i> na combinação do <i>Classifier Attribute Eval</i> mais <i>MLR</i> | 26 |
| Tabela 5 | – Resultado do <i>RMSE</i> na combinação do <i>Classifier Attribute Eval</i> mais <i>SVR</i> | 26 |
| Tabela 6 | – Resultado do <i>Classifier Coeficient</i> na combinação com o <i>Correlation Attribute Eval</i> mais <i>SVR</i> | 26 |
| Tabela 7 | – Resultado do <i>RMSE</i> na combinação com o <i>Classifier Attribute Eval</i> mais <i>XGBoost</i> | 26 |
| Tabela 8 | – Resultado do com a combinação do <i>Classifier Attribute Eval</i> mais o <i>XGBoost</i> | 26 |
| Tabela 9 | – Resultado do <i>RMSE</i> na combinação do <i>Correlation Attribute Eval</i> mais <i>MLR</i> | 27 |
| Tabela 10 | – Resultado do <i>Correlation Coeficient</i> na combinação do <i>Correlation Attribute Eval</i> mais <i>MLR</i> | 27 |
| Tabela 11 | – Resultado do <i>RMSE</i> na combinação com <i>Correlation Attribute Eval</i> mais <i>SVR</i> | 27 |
| Tabela 12 | – Resultado <i>Correlation Coeficient</i> na combinação com o <i>Correlation Attribute Eval</i> mais <i>SVR</i> | 27 |
| Tabela 13 | – Resultado do <i>RMSE</i> com o <i>Correlation Attribute Eval</i> mais o <i>XGBoost</i> | 28 |
| Tabela 14 | – Resultado do <i>Correlation Coeficient</i> na combinação com <i>Correlation Attribute Eval</i> mais o <i>XGBoost</i> | 28 |
| Tabela 15 | – Resultado do <i>RMSE</i> na combinação do <i>Relief Attribute Eval</i> mais <i>MLR</i> | 28 |
| Tabela 16 | – Resultado do <i>Correlation Coeficient</i> na combinação do <i>Relief Attribute Eval</i> mais <i>MLR</i> | 28 |
| Tabela 17 | – Resultado do <i>RMSE</i> na combinação com o <i>Relief Attribute Eval</i> mais <i>SVR</i> | 29 |
| Tabela 18 | – Resultado do <i>Correlation Coeficient</i> na combinação com o <i>Relief Attribute Eval</i> mais <i>SVR</i> | 29 |
| Tabela 19 | – Resultado do <i>Relief Coeficient</i> na combinação com <i>Correlation Attribute Eval</i> mais o <i>XGBoost</i> | 29 |

| | | |
|-----------|--|----|
| Tabela 20 | –Resultado do <i>Relief Coeficient</i> na combinação com <i>Correlation Attribute Eval</i> mais o <i>XGBoost</i> | 29 |
| Tabela 21 | –Resultado do <i>RMSE</i> na combinação do <i>CFS Subset Eval</i> mais <i>MLR</i> | 30 |
| Tabela 22 | –Resultado do <i>Classifier Coeficient</i> na combinação com o <i>CFS Subset Eval</i> mais <i>MLR</i> | 30 |
| Tabela 23 | –Resultado do <i>RMSE</i> na combinação <i>CFS Subset Eval</i> mais <i>SVR</i> | 30 |
| Tabela 24 | –Resultado do <i>Correlation Coeficient</i> na combinação <i>CFS Subset Eval</i> mais <i>SVR</i> | 30 |
| Tabela 25 | –Resultado do <i>RMSE</i> na combinação com <i>CFS Subset Eval</i> mais o <i>XGBoost</i> | 30 |
| Tabela 26 | –Resultado do <i>Correlation Coeficient Eval</i> na combinação com <i>CFS Subset Eval</i> mais o <i>XGBoost</i> | 30 |
| Tabela 27 | –Resultado das melhores características para <i>Arousal</i> | 31 |
| Tabela 28 | –Resultado das melhores características para <i>Valence</i> | 31 |

Lista de abreviaturas e siglas

| | |
|-------|--|
| REM | Reconhecimento das Emoções na Música |
| MER | <i>Music Emotion Recognition</i> |
| RIM | Recuperação das Informações na Música |
| SVR | <i>Support Vector Regression</i> |
| MLR | <i>Multiple Linnear Regression</i> |
| MIDI | <i>Musical Instrument Digital Interface</i> |
| PCA | <i>Principal Component Analysis</i> |
| GLM | <i>General Linear Model</i> |
| RMSE | <i>Root-mean-square deviation</i> |
| LWR | <i>Locally Weighted Regression</i> |
| MP3 | <i>MPEG Layer 3</i> |
| MFCC | <i>Mel Frequency Cepstral Coefficient</i> |
| PCM | <i>Pulse-code modulation</i> |
| LR | <i>Logistic Regression</i> |
| LR-FS | <i>Logistic Regression Feature Selection</i> |
| L1-LR | <i>Level 1 Linear Regression</i> |
| RFR | <i>Random Forest Regression)</i> |
| RNN | <i>Recurrent Neural Networks</i> |

Sumário

| | | |
|----------|--|-----------|
| 1 | Introdução | 13 |
| 1.1 | Contextualização | 13 |
| 1.2 | Objetivos | 14 |
| 1.3 | Método | 14 |
| 1.4 | Contribuições | 14 |
| 1.5 | Organização do Trabalho | 15 |
| 2 | Trabalhos Relacionados | 16 |
| 3 | Referencial Teórico | 18 |
| 3.1 | Caraterísticas da Música por Níveis | 18 |
| 3.2 | Modelo Emocional | 19 |
| 3.3 | Reconhecimento de Emoção na Música | 20 |
| 4 | Material e Métodos | 21 |
| 4.1 | Base de Dados de REM | 21 |
| 4.2 | Métodos de Seleção | 21 |
| 4.3 | Métodos de Regressão | 23 |
| 4.4 | Métricas de Desempenho | 23 |
| 5 | Resultados | 25 |
| 5.1 | Resultado das Combinações | 25 |
| 5.2 | Impacto no <i>Arousal</i> e <i>Valence</i> | 31 |
| 6 | Conclusão | 33 |
| | Referências | 34 |

1 Introdução

Esta monografia realiza uma seleção de características em uma base de dados pública de reconhecimento de emoções na música. Este capítulo está dividido em nas seções: contextualização, objetivos, metodologia, contribuições e organização do presente documento.

1.1 Contextualização

Com os recentes avanços no campo da recuperação da informação musical (RIM), há um interesse emergente em analisar e compreender o conteúdo emocional da música (HAN *et al.*, 2009). A partir do interesse na compreensão das emoções surgiu a área de REM (Reconhecimento das Emoções na Música).

O reconhecimento de Emoções na Musica é um campo que visa estudar a relação entre música e emoção e é útil na compreensão musical e na informação musical (SEO; HUH, 2019). Com a compreensão da emoção na música é possível desenvolver sistemas, como os de recomendação musical Lopes *et al.* (2018), Abdul *et al.* (2018).

Para identificar as emoções nas músicas é necessário desenvolver ou utilizar uma base de dados. Na montagem da base contém o processo de definição do modelo emocional. Na literatura o mais utilizado é o bidimensional de *Theyers*. O modelo de *Theyers* separa as músicas em duas dimensões, chamadas de *Arousal* e *Valence*. O *Valence* representa o modo como se julga uma situação, de desagradável (emoção negativa) a agradável (emoção positiva) (CHANEL; ANSARI-ASL; PUN, 2007). O *Arousal* expressa o grau de entusiasmo que vai de calmo, sendo essa a tristeza a emocionante, que pode ser a felicidade (CHANEL; ANSARI-ASL; PUN, 2007).

Há alguns desafios em rotular uma base de dados de REM. Um dos desafios é a dependência da participação de pessoas, seja voluntariamente ou com pagamentos. Outro desafio, é a realização de pagamento como nos trabalhos Aljanaki, Yang e Soleymani (2017), Aljanaki, Wiering e Veltkamp (2016), para incentivar as pessoas a darem uma resposta mais precisa. Por fim, problemas com os direitos autorais, o que limita o tamanho da base de dados. Por conta destes problemas, a montagem de uma base de dados se torna um desafio dentro da área de REM. Esses desafios restringe a quantidade de autores em tornar suas bases de dados públicas.

Os desafios na montagem de uma base de dados de REM torna o desenvolvimento da área mais complicado. Pois limita a quantidade de pesquisas, assim, diminui a evolução e futuras contribuições. Mesmo com as dificuldades há trabalhos que publicam bases de dados, mas deixam de realizar a etapa de seleção de características (ALJANAKI; YANG; SOLEYMANI, 2017; CHEN *et al.*, 2016; CHEN *et al.*, 2015). Dessa forma, deixam de trazer contribuições e ter melhores resultados dentro da área de REM.

A partir da brecha de pesquisa deixada por trabalhos que desenvolvem bases de dados públicas de REM é que este trabalho visa realizar a seleção de características. Com isso, é possível verificar o impacto nos resultados obtidos e determinar quais características podem impactar no *Arousal* e no *Valence*.

1.2 Objetivos

O objetivo deste trabalho é investigar e avaliar o impacto da seleção de características na etapa de regressão no contexto da caracterização da música. Os objetivos específicos são:

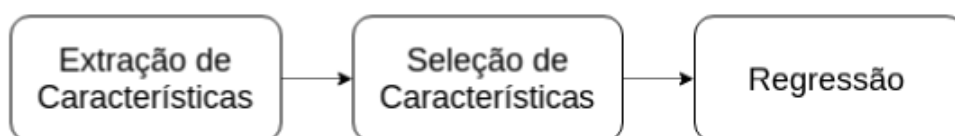
1. Verificar diferentes métodos de seleção de características na predição da emoção na música.
2. Aplicar diferentes regressores para estimar a emoção da música.
3. Estabelecer uma relação das características selecionadas e o eixo *Arousal* e *Valence*.

1.3 Método

Na Figura 1 é apresentada a divisão das etapas para o desenvolvimento de parte dos trabalhos de REM. A definição do fluxograma foi determinada com a leitura de diferentes trabalhos dentro da área de REM.

Na extração é comum realizar a verificação na literatura de quais características podem impactar as emoções. Na seleção de características os objetivos são diminuir o vetor inicial, verificar se com menos características é possível obter melhores resultados. Na regressão, deve-se determinar o resultado obtido a partir de uma das métricas, que podem ser utilizadas.

Figura 1: Etapas utilizadas no trabalho



1.4 Contribuições

Este trabalho tem como contribuições:

1. Identificar características que são relevantes para os eixos *Arousal* e *Valence*;
2. Apresentar métodos de seleção que se destacam na seleção de características em bases de dados de REM;
3. Apresentar regressores que se destacam na predição de emoções em bases de dados de REM;

1.5 Organização do Trabalho

Este trabalho está organizado em capítulos. No Capítulo 2 são apresentados trabalhos que estão relacionados a este. No Capítulo 3, foram descritos alguns conceitos pertinentes a este trabalho. No Capítulo 4, é mostrado o desenvolvimento do presente trabalho. No Capítulo 5 são apresentados os resultados obtidos neste trabalho em comparação com a PMEmo e também apresentados as características mais impactantes para o *Arousal* e o *Valence*. Por fim, são apresentados os principais achados do trabalho e perspectivas futuras no Capítulo 6.

2 Trabalhos Relacionados

Neste capítulo é discutido brevemente alguns trabalhos correlatos. Esta seleção de artigos se deu com base numa pesquisa da literatura que considera os seguintes critérios: o trabalho estar situado dentro do contexto de REM, fazer uso de método de seleção de características, realizar regressão. Tais informações são sumarizadas na Tabela 1.

O trabalho [Huq, Bello e Rowe \(2010\)](#) realizou um estudo sobre as melhores características para serem escolhidas nas bases de dados de músicas para definir o humor da música. O modelo emocional utilizado foi definido por *Thayers*, que divide a música em *Arousal* e *Valence*. Tal trabalho desenvolveu sua própria base de dados, com a participação de 150 pessoas, 50 músicas e 160 características. Com as características definidas houve a seleção a partir dos métodos: *PCA (Principal Component Analysis)* e *PCA-W (Principal Component Analysis Weight)*. Por fim, a regressão com: *Linear Regression*, *Regression Tree*, *LWR (Locally Weighted Regression)*, *Model Tree*, *SVR (Support Vector Regression)*, *SVR-RBF*. Os resultados foram apresentados com o *MAE (Mean Absolute Error)*.

O trabalho ([GREKOW, 2017](#)) tem como objetivo definir a melhor combinação entre as características de baixo nível, ritmo e tonalidade. O trabalho citado determinou que a partir desse modelo é possível indicar melhores formas de determinar o *Valence* e *Arousal*. Para realização do trabalho houve a extração de 654 características nas bases de dados *Essentia*¹ e *Marsyas*². Após isso, foram selecionadas as características extraídas com o método *Wrapper Subset Eval*, que utiliza método de busca *Best First*. Por fim, as características selecionadas foram submetidas a regressão. Os resultados foram avaliados através das métricas de *Correlation Coefficient* e o *MAE (Mean Absolute Error)*. Os resultados obtidos definiu: A melhor base foi o *Essentia*, com o valor de 0.79 no *Correlation Coefficient* e no *MAE* de 0.09, isso na dimensão do *Arousal*; no *Valence* o *Correlation Coefficient* foi de 0.58 e o *MAE* de 0.10; as melhores características no *Arousal* foram as de baixo nível e as de ritmo, com 0.79 no *Correlation Coefficient* e *MAE* de 0.09; enquanto que no *Valence*, as melhores características foram as de baixo nível e as tonais, com 0.56 de *Correlation Coefficient* e 0.10 de *MAE*.

A pesquisa de [Zhang et al. \(2017\)](#) utilizou a combinação de 8 características de intensidade. A escolha é justificada por serem referentes a sensação de volume do ouvido humano. Desta forma, demonstrou ser possível responder de formar pertencente as emoções do *Arousal*. Na pesquisa foi utilizado o seletor *Wrapper* e os regressores: *LR (Logistic Regression)*, *LR-FS (Logistic Regression Feature Selection)*, *Bagging-A*, *Bagging-B*, *L1-LR (Level 1 Linear Regression)* e *L2-LR (Level 2 Linear Regression)*. A taxa média de precisão foi de 83.8% no trabalho.

¹ <https://essentia.upf.edu/documentation/>

² <http://marsyas.info/downloads/datasets.html>

O trabalho proposto por [Bai et al. \(2016\)](#) extraiu 548 características da música e realizou uma seleção para determinar quais as melhores combinações. A seleção foi efetuada com as características de energia, ritmo, tempo, espectro e harmonia. Os métodos utilizados na seleção de características foram: *PCA*, que um método utilizado para testar todos os componentes possíveis e determinar quais compõem toda a base; O *Relief Attribute Eval* escolhe aleatoriamente o valor de X e realiza um cálculo com dois valores, o mais próximo e outro mais distante de X . Foram selecionadas 129 características de um dos componentes do *PCA* e 276 no *Relief*. Os regressores utilizados foram o *SVR*, *RFR* (*Random Forest Regression*) e *RNN* (*Recurrent Neural Networks*). Dentre as combinações testadas a que obteve o melhor resultado foi no *Arousal* com o regressor *SVR* e o seletor *Relief*, o valor obtido foi de 65%. No entanto, o trabalho não define quais são as melhores características.

No artigo de [Guan, Chen e Yang \(2012\)](#) é desenvolvido um modelo para extração de informações por meio do *MIDI* (*Musical Instrument Digital Interface*), do áudio e da letra e da música. Com este modelo, o trabalho contribui para a identificação de emoção dentro do modelo bidimensional, que usa *Arousal* e *Valence*. Para realização do trabalho foi necessário desenvolver uma base de dados de músicas. Os melhores resultados foram obtidos com a junção entre os meios de extração. O melhor regressor foi o *AdaBoost.RM*, que é baseado no *MultiLayer Perceptron*.

Esta monografia pretende realiza a seleção de características afim de verificar duas informações: O impacto da seleção no resultado da regressão; e quais características estão mais relacionadas ao *Arousal* e o *Valence*. A partir de 138 combinações entre seletores, regressores, quantidade de características, métricas e modelo emocional.

Tabela 1: Listagem dos Trabalhos Relacionados.

| Artigo | Seletor | Regressor | Quantidade antes da seleção | Quantidade pós seleção | Métrica de Resultado | Base Pública |
|--|--------------------------|---|-----------------------------|-------------------------|---|--------------|
| (HUQ; BELLO; ROWE, 2010) | PCA, PCA-W | Linear Regression, Regression Tree, LWR, Model Tree, SVR, SVR-RBF | 160 | Não informada | MAE | Não |
| (GREKOW, 2017) | <i>WrapperSubsetEval</i> | SMOreg, REPTree, M5P, SMOreg | 654 | Não informada | <i>Correlation Coefficient</i> e <i>MEA</i> | Sim |
| (ZHANG et al., 2017) | <i>Wrapper</i> | GML e Métodos baseados em árvores | 8 | Não informada | | Sim |
| (BAI et al., 2016) | PCA, <i>RRelief</i> | RFR, RNN, SVR | 548 | 139 - PCA, 276 - Relief | <i>Correlation Coefficient</i> | Não |
| (GUAN; CHEN; YANG, 2012) | <i>CF Subset Eval</i> | <i>AdaBoost.R2</i> , <i>AdaBoost.RM</i> <i>LinearRegression</i> e <i>SMOreg</i> | 1034 | Não informada | | Não |

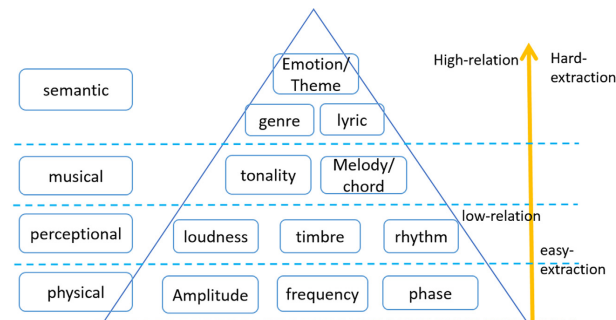
3 Referencial Teórico

Este capítulo apresenta conteúdos relevantes no desenvolvimento deste trabalho. Na Seção 3.1 é apresentado os níveis que separa as características da música. Na Seção 3.2 é relatado diferentes modelos emocionais existentes na área da psicologia, que são utilizados na área de REM. A Seção 3.3 descreve brevemente o que é a área do REM.

3.1 Caraterísticas da Música por Níveis

Os recursos das música podem ser divididos em quatro níveis: nível físico, nível de percepção, nível musical e nível semântico, como pode ser apresentado na Figura 2 (ZHANG et al., 2016). Mesmo com a separação dos níveis musicais, todos estes podem se relacionar, ou formar uma nova característica.

Figura 2: Níveis de separação das características musicais.



As propriedades físicas ou fisiológicas do som são descritas com referência especial àquelas propriedades que pertencem aos sons do coração (FEIGEN, 1971). O nosso coração tem uma amplitude a cada batida, uma frequência nas suas ondas sonoras e um faseamento, perceptível no momento que é ouvida as batidas do nosso coração, com isso é gerado um eco.

As características contidas no nível de percepção são aquelas audíveis ao ouvido humano. Algumas dessas são: timbre que é a diferença sonora de cada instrumento, ritmo usado para marcar cada mudança de acorde e volume que é a altura média de uma peça de áudio.

No nível musical contém assuntos específicos na formação de um músico, ou uma música. Uma das características que pode ser destacada é a tonalidade, esta faz parte de um sistema formal, pelo qual é criado um tipo de relação organizada entre os sons (PISTON, 1998), como também, a melodia e os acordes.

No nível semântico inclui as características que não é necessário de ter conhecimento teórico de música, ou processamento de sinais. Algumas dessas são: gênero musical classi-

ficado pelo conteúdo musical que a mesma pertence, a emoção sentida por cada indivíduo e letra da música.

3.2 Modelo Emocional

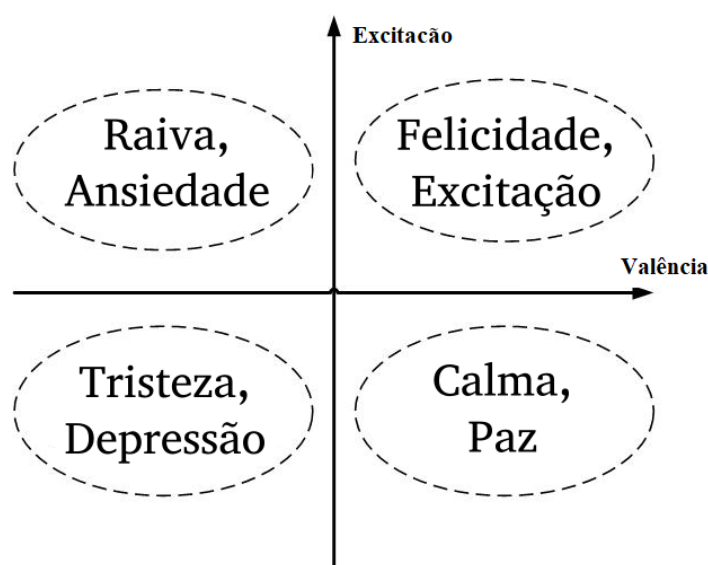
Na área da psicologia algo que não é definido de maneira clara são as emoções, por se tratar de uma área que realiza estudos empíricos (REISENZEIN, 2007), a partir das sensações das pessoas em um determinado momento. Isto se gera uma dificuldade para a definição dos modelos emocionais.

Os modelos emocionais são uma forma de relatar as emoções como uma forma exata. O primeiro estudo a tentar isso foi Hevner (1936), que desenvolveu o modelo categórico com a definição das emoções. As emoções no modelo categórico estão separados em 66 adjetivos, organizados em 8 grupos.

O modelo de Havner engloba uma grande quantidade de emoções e mesmo assim, pode não ter todas as emoções possíveis. Afim de simplificá-lo, foi desenvolvido o modelo de Russell (1980). O trabalho de Russel prova que as emoções definidas por Havner estavam inter-relacionadas, dessa forma, foi possível simplificá-las em 8 emoções básicas 3.

Atualmente o modelo emocional mais comum na área de REM é o de Theyers apresentado na Figura 3. Thayers se baseou no modelo de Russel, que definiu 4 emoções consideradas as básicas. As 4 emoções são divididas em duas dimensões chamadas de *Arousal/Excitação* e *Valence/Valência*. As dimensões cortam a imagem em duas linhas, com isso ocasiona em uma separação das emoções.

Figura 3: Modelo dimensional



Adaptado de Yeh, Lin e Chang (2009)

3.3 Reconhecimento de Emoção na Música

A área de Reconhecimento de Emoção na Música, ou *MER* se trata de um campo que visa estudar a relação entre a computação, emoções e música (AVISADO et al., 2012). Na computação através da área de RIM (Recuperação de Informações na Música). Na emoção com estudo dos diferentes modelos emocionais na psicologia. Na música através dos elementos ou características existentes.

De início os trabalhos escolhem qual modelo emocional será utilizado. Após isso, os trabalhos selecionam as músicas, que não contém restrição de direitos autorais. Com as músicas e as características escolhidas é realizada a montagem da base de dados. Para isso é escolhido características da música existentes na literatura. Com a base montada pode ser selecionada as características, mas há trabalhos que não se dedicam a esta etapa. Por fim, é classificado ou predito as emoções a partir do modelo escolhido.

4 Material e Métodos

Neste capítulo são apresentadas as técnicas e ferramentas escolhidas para as etapas descritas na Figura 1. Na Seção 4.1 é descrita a base de dados utilizada neste trabalho. Na Seção 4.2, é realizada a descrição dos métodos de seleção. Na Seção 4.3 há a descrição dos métodos de regressão utilizados neste trabalho. Na Seção 4.4, há a definição das métricas de desempenho, como também, as combinações para obter os resultados.

4.1 Base de Dados de REM

Para obter as informações das músicas este trabalho utilizou a base de dados PMEmo (ZHANG et al., 2018), na qual contém 794 músicas e 6373 características extraídas de arquivos *MP3 (MPEG Audio Layer-3)*. Das as 794 músicas, 31 foram descartas por falta de identificação emocional.

A identificação emocional foi realizada com a participação de 457 pessoas. As pessoas deram notas de 1 à 9, mas pode variar para valores como 4,6. Assim, a classificação emocional somente pode ser determinada com o uso de regressores.

4.2 Métodos de Seleção

O presente trabalho utiliza os seletores com o objetivo de obter melhores resultados na regressão e verificar quais características mais impactam no *Arousal* e *Valence*. Os seletores escolhidos foram: *Classifier Attribute Eval*, *Correlation Attribute Eval*, *Relief Attribute Eval* e *CFS Subset Eval*. Os seletores utilizados estão contidos na aplicação Weka ¹.

A escolha dos métodos de seleção *Correlation Attribute Eval*, *Relief Attribute Eval* e *CFS Subset Eval*, se da pelos resultados positivos em outros trabalhos dentro da área de REM (BAI et al., 2016; POUYANFAR; SAMETI, 2014; ALJANAKI; WIERING; VELTKAMP, 2013). O *Classifier Attribute Eval* não foi utilizado em outros trabalhos, mas foi escolhido pela sua característica de variar de acordo com o classificador escolhido no Weka.

O *Classifier Attribute Eval* avalia o valor de um atributo usando um classificador especificado pelo usuário, a validação cruzada é uma medida de avaliação de desempenho para usar na seleção de atributos (GONZÁLEZ-VIDAL; JIMÉNEZ; GÓMEZ-SKARMETA, 2019). Assim, este seletor tem como principal característica a variação

¹ <https://www.cs.waikato.ac.nz/ml/weka/>

de regressor/classificador com o mesmo método de seleção. Por conta disso, não é possível estimar uma fórmula para representar o *Classifier Attribute Eval*.

O algoritmo *Correlation Attribute Eval* avalia o valor de um atributo com uma métrica de correlação entre o atributo e a classe (BYDOVSKA; POPELÍNSKÝ, 2013). Na literatura há diferentes algoritmos de correlação, como este trabalho utiliza o Weka, o algoritmo de correlação padrão é o *Pearson*. Sua representação é dada a partir da Fórmula 4.1, na qual tem a covariância de X e Y sobre o desvio padrão de x e y .

$$CorrCoe = \frac{\text{cov}(X, Y)}{\sigma_x \sigma_y} \quad (4.1)$$

O seletor *Relief Attribute Eval* estima os atributos de acordo com a forma como os seus valores distinguem as instâncias próximas umas das outras (KONONENKO, 1994). Para isso, é determinado um vetor de peso W em cada interação é um valor de uma instancia aleatória x e seus vizinhos mais próximos, chamados de *quaseCerto* e *quaseFalha* na Equação 4.2.

$$W_i = W_i - (x_i - \text{quaseCerto}_i)^2 + (x_i - \text{quaseFalha}_i)^2 \quad (4.2)$$

O *CFS Subset Eval* avalia o subconjunto de recursos altamente correlacionados em relação à sua classe, mas baixa inter-correlação entre os recursos selecionados (CHAUHAN; BAHL, 2015). O subconjunto mais impactante é escolhido a partir de teste com todos os subconjuntos possíveis. No calculo presente na Fórmula 4.3 o valor heurístico é o Ms , k é a quantidade de característica no subconjunto, a a correlação média da correlação e j a característica média dos recursos da intercorrelação.

$$Ms = \frac{ka}{\sqrt{k + k(k-1)j}} \quad (4.3)$$

O seletor *Best First* é uma estratégia de pesquisa que explora um gráfico expandindo gradualmente o nó mais promissor escolhido de acordo com a proximidade do final de um caminho a uma solução (WANG et al., 2019). A partir dessa estratégia é selecionado apenas o subconjunto das características mais promissoras.

O método de busca *Ranker* associa as classificações aos atributos por sua avaliação individual (KHAN; SAEIDLOU; SAADAT, 2019). A avaliação é realizada de forma individual para cada atributo, que pode variar de acordo com o seletor escolhido.

A diferença entre o *Ranker* e o *Best First* se dá pela apresentação das melhores características. Os seletores que utilizam o Ranker realizam um *ranking* com as características, da mais impactante à menos impactantes. Isso os diferencia do método *Best First*, que apresenta apenas o subconjunto mais impactante para a base. Os seletores deste trabalho que seguem o *Ranker* são: *Classifier Attribute Eval*, *Correlation Attribute Eval*, *Relief Attribute Eval*. Enquanto o *CFS Subset Eval* utiliza o método *Best First*.

4.3 Métodos de Regressão

Para a obtenção dos resultados foi escolhido os regressores *MLR* (*Multiple Linear Regression*), *SVR* e *XGBooster*. A escolha do *MLR* e o *SVR* se deu pelo fato de estarem presentes no artigo oficial PMEmo. A escolha do *XGBoost* ocorre por conta de ter sido utilizado em outros trabalhos da área de REM.

O *MLR* é um método estatístico para modelar a relação entre duas variáveis, ajustando uma equação linear aos dados observados (SONG et al., 2005). Na Equação 4.4 é visto que há uma variável considerada dependente Y , uma variável explicativa X e um termo constante demonstrado no B . Também há uma relação linear entre as variáveis dependentes e as variáveis independentes.

$$Y = B_0 + B_1X_1 \dots B_n + B_nX_n \quad (4.4)$$

O *SVR* é um sistema de aprendizagem que utiliza um espaço de hipóteses de funções lineares treinado com um algoritmo de aprendizado (WANG, 2011). *SVR* tenta minimizar o limite superior no erro de generalização. Na Equação 4.5 valor de e representa o limite inferior ou superior, x é o vetor normal, y o espaço entre os valores e o b a distância em relação a origem.

$$SVR = e \leq y - x + b \leq +e \quad (4.5)$$

XGBoost é um modelo de conjunto de árvores, que gera uma soma ponderada das previsões de árvores de regressão múltipla (LI et al., 2018). Este algoritmo obtém uma implementação altamente otimizada, com isso reduz tempo na predição. Sua Equação é apresentada 4.6, onde k é o número de árvores, f é uma função no espaço funcional e F é o conjunto de todos os *CARTs* (*Classification & Regressions Trees*) possíveis.

$$Y = \sum_{k=1}^k f_k(X_i), f_k \in F \quad (4.6)$$

Para a regressão houve uma divisão dos arquivos entre treino e teste para a base formada com os rótulos de *Arousal* e *Valence*. No qual 70% foi de treino e 30% para teste. Para a execução dos regressores *SVR* e *MLR* foi usada a aplicação *Weka*, para o *XGBoost* ocorreu a utilização da linguagem de programação *Python*² com a biblioteca *XGBoost*³.

4.4 Métricas de Desempenho

Após a regressão das características é possível extrair informações referentes a esta etapa. Neste trabalho as informações são obtidas a partir das métricas *RMSE* e o *Cor-*

² <https://www.python.org>

³ https://xgboost.readthedocs.io/en/latest/python/python_intro.html

relation Coefficient. As métricas foram escolhidas por conta do seu uso na base de dados PMEmo.

O RMSE é a medida de quão bem o modelo se ajusta aos dados (YENIAY; GOKTAS, 2002). Na Fórmula 4.7, RMSE é realizada uma raiz-quadrada do somatório até as n características e a diferença entre o valor predito e o valor observado, ao quadrado, por fim uma divisão por n .

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (d_i - f_i)^2}{n}} \quad (4.7)$$

O *Correlation Coefficient* é uma medida simples e direta de linearidade entre quaisquer duas variáveis, e assim, em particular, é uma escolha óbvia para medir a linearidade de um gráfico de probabilidade normal. (FILLIBEN, 1975). Apresentado na fórmula 4.1.

Para obter os resultados foi realizada uma avaliação de desempenho a partir da junção entre quantidade de características selecionadas, método de seleção, regressor, modelo emocional e métrica. A Tabela 2 apresenta as variações entre quantidade selecionada, método de seleção e regressor. As combinações apresentadas foram implementadas para *Arousal* e *Valence* e os resultados obtido com o *RMSE* e *Correlation Coefficient*. No total foram realizadas 138 combinações.

Tabela 2: Combinações possíveis a partir da quantidade de características selecionas, método de regressão e método de seleção.

| | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 |
|----------------------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| Classifier Attribute Eval | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost |
| Corelation Attribute Eval | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost |
| Relief Attribute Eval | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost | SVR, MLR e XGBoost |
| CFS Subset Eval | SVR, MLR e XGBoost | | | SVR, MLR e XGBoost | | | |

5 Resultados

Neste capítulo será apresentado o resultado obtido com a quantidade de características selecionadas, que foram de: 100, 200, 300, 500, 1000, 1500 e 2000 características para os seletores de método de busca *Ranker*. Os valores estão definidos de forma empírica com o objetivo de verificar se uma menor quantidade de características pode obter melhores resultados.

O seletor *CFS Subset Eval* utiliza o método *Best First*, por conta disso foi selecionada 92 características para *Arousal* e 96 para *Valence*. A quantidade diferente de características ocorre por conta do seletor avaliar todos os subconjunto possíveis e definir o que obtém o melhor resultado.

Este capítulo está dividido em duas seções. Na Seção 5.1 é demonstrado os resultados com as diferentes combinações e comparado com a coluna Alvo, que compara o resultado deste trabalho com o obtido na base PMEmo. Na Seção 5.2 tem o resultado das características de maior impacto no *Arousal* e *Valence*.

5.1 Resultado das Combinações

Nas Tabelas 3, 4, 5, 6, 7 e 8 são demonstrados os resultados do seletor *Classifier Attribute Eval* com os regressores *MLR*, *SVR* e *XGBoost* respectivamente para as métricas de *RMSE* e *Correlation Coefficient*. Em ambas as Tabelas não é obtido resultados expressivos em comparação com o Alvo. Isso ocorreu por conta de não ter sido escolhido no Weka um classificador comum na área de REM para selecionar os valores no *Classifier Attribute Eval*. O classificador utilizado foi o *ZeroR*.

Tabela 3: Resultado do *RMSE* na combinação do *Classifier Attribute Eval MLR*.

| RMSE | | | | | | | | |
|--|--------|--------|--------|--------|--------|--------|--------|---------------|
| <i>Classifier Attribute Eval + MLR</i> | | | | | | | | <i>Alvo</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,1443 | 0,1556 | 0,1858 | 0,2519 | 0,1895 | 0,1900 | 0,1908 | 0,1110 |
| <i>Valence</i> | 0,1517 | 0,1958 | 0,1883 | 0,4008 | 0,1570 | 0,1643 | 0,1659 | 0,1360 |

Tabela 4: Resultado do *Correlation Coefficient* na combinação do *Classifier Attribute Eval* mais *MLR*.

| <i>Correlation Coefficient</i> | | | | | | | | |
|--|--------|--------|--------|--------|--------|--------|------|---------------|
| <i>Classifier Attribute Eval + MLR</i> | | | | | | | | <i>Alvo</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,3393 | 0,5686 | 0,4082 | 0,2872 | 0,1895 | 0,0000 | 0 | 0,5460 |
| <i>Valence</i> | 0,4157 | 0,4311 | 0,3396 | 0,2189 | 0,0000 | 0,0000 | 0 | 0,7190 |

Tabela 5: Resultado do *RMSE* na combinação do *Classifier Attribute Eval* mais *SVR*

| <i>RMSE</i> | | | | | | | | |
|--|--------|--------|--------|--------|--------|--------|--------|---------------|
| <i>Classifier Attribute Eval + SVR</i> | | | | | | | | <i>Alvo</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,1871 | 0,1526 | 0,1840 | 0,1581 | 0,1887 | 0,1893 | 0,1728 | 0,1110 |
| <i>Valence</i> | 0,1565 | 0,163 | 0,1546 | 0,1645 | 0,1566 | 0,1642 | 0,1653 | 0,1360 |

Tabela 6: Resultado do *Classifier Coefficient* na combinação com o *Correlation Attribute Eval* mais *SVR*.

| <i>Correlation Coefficient</i> | | | | | | | | |
|--|---------|--------|--------|--------|--------|--------|--------|---------------|
| <i>Classifier Attribute Eval + SVR</i> | | | | | | | | <i>Alvo</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,02887 | 0,5814 | 0,77 | 0,104 | 0,0943 | 0,0894 | 0,1634 | 0,5460 |
| <i>Valence</i> | 0,1748 | 0,1186 | 0,6570 | 0,1645 | 0,0898 | 0,0737 | 0,092 | 0,7190 |

Tabela 7: Resultado do *RMSE* na combinação com o *Classifier Attribute Eval* mais *XGBoost*.

| <i>RMSE</i> | | | | | | | | |
|--|--------|--------|--------|--------|--------|--------|--------|---------------|
| <i>Classifier Attribute Eval + XGBoost</i> | | | | | | | | <i>ALVO</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,1659 | 0,1268 | 0,1416 | 0,1296 | 0,1049 | 0,1009 | 0,1065 | 0,1110 |
| <i>Valence</i> | 0,1396 | 0,1229 | 0,1196 | 0,1286 | 0,1120 | 0,1183 | 0,112 | 0,1360 |

Tabela 8: Resultado do com a combinação do *Classifier Attribute Eval* mais o *XGBoost*.

| <i>Correlation Coefficient</i> | | | | | | | | |
|--|--------|--------|--------|--------|--------|--------|--------|---------------|
| <i>Classifier Attribute Eval + XGBoost</i> | | | | | | | | <i>ALVO</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,4628 | 0,7384 | 0,6393 | 0,6047 | 0,8339 | 0,8486 | 0,7949 | 0,5460 |
| <i>Valence</i> | 0,4832 | 0,6109 | 0,6329 | 0,6292 | 0,7015 | 0,6874 | 0,7383 | 0,7190 |

As Tabelas 9, 10, 11 e 12, é possível verificar os resultados do seletor *Correlation Attribute Eval* com os regressores *MRL* e *SRV* no *RMSE* e *Correlation Coefficient*. Pode ser percebido que há poucos resultados melhores em comparação ao alvo. Também é visto

que os melhores resultados são com baixa quantidade de características selecionadas. Isso ocorre pela característica do *Correlation Coefficient* de se comportar melhor com baixa quantidade de elementos selecionados. As Tabelas 13 e 14 são exceções do *Correlation Coefficient*, por conta do alto desempenho do XGBoost independente da quantidade de características selecionadas.

Tabela 9: Resultado do *RMSE* na combinação do *Correlation Attribute Eval* mais *MLR*.

| RMSE | | | | | | | | |
|---|---------------|--------|--------|--------|--------|--------|--------|---------------|
| <i>Correlation Attribute Eval + MLR</i> | | | | | | | | <i>Alvo</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,1031 | 0,1451 | 0,1215 | 0,1931 | 0,1779 | 0,1830 | 0,1824 | 0,1110 |
| <i>Valence</i> | 0,1411 | 0,1968 | 0,1416 | 0,1275 | 0,1629 | 0,1575 | 0,1824 | 0,1360 |

Tabela 10: Resultado do *Correlation Coefficient* na combinação do *Correlation Attribute Eval* mais *MLR*.

| <i>Correlation Coefficient</i> | | | | | | | | |
|---|--------------|--------------|---------------|---------------|--------|--------|------|---------------|
| <i>Correlation Attribute Eval + MLR</i> | | | | | | | | <i>Alvo</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,838 | 0,738 | 0,7802 | 0,6082 | 0,2226 | 0,0000 | 0 | 0,5460 |
| <i>Valence</i> | 0,6038 | 0,482 | 0,5454 | 0,6096 | 0,0000 | 0,0000 | 0 | 0,7190 |

Tabela 11: Resultado do *RMSE* na combinação com *Correlation Attribute Eval* mais *SVR*.

| RMSE | | | | | | | | |
|---|--------|--------|--------|--------|--------|--------|--------|---------------|
| <i>Correlation Attribute Eval + SVR</i> | | | | | | | | <i>Alvo</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,1193 | 0,1843 | 0,1809 | 0,1843 | 0,1809 | 0,1827 | 0,1822 | 0,1110 |
| <i>Valence</i> | 0,1612 | 0,1634 | 0,161 | 0,1611 | 0,1610 | 0,1565 | 0,1623 | 0,1360 |

Tabela 12: Resultado *Correlation Coefficient* na combinação com o *Correlation Attribute Eval* mais *SVR*.

| <i>Correlation Coefficient</i> | | | | | | | | |
|---|--------|--------|--------|--------|--------|--------|--------|---------------|
| <i>Correlation Attribute Eval + SVR</i> | | | | | | | | <i>Alvo</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,1882 | 0,0016 | 0,174 | 0,1223 | 0,1149 | 0,0665 | 0,052 | 0,5460 |
| <i>Valence</i> | 0,0367 | 0,0472 | 0,0068 | 0,0108 | 0,1482 | 0,1194 | 0,1195 | 0,7190 |

Tabela 13: Resultado do *RMSE* com o *Correlation Attribute Eval* mais o *XGBoost*.

| RMSE | | | | | | | | |
|---|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| <i>Correlation Attribute Eval + XGBoost</i> | | | | | | | | <i>Alvo</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,1024 | 0,1117 | 0,1003 | 0,1039 | 0,0994 | 0,0984 | 0,0966 | 0,1110 |
| <i>Valence</i> | 0,1262 | 0,1180 | 0,1194 | 0,114 | 0,1063 | 0,1059 | 0,1183 | 0,1360 |

Tabela 14: Resultado do *Correlation Coefficient* na combinação com *Correlation Attribute Eval* mais o *XGBoost*.

| <i>Correlation Coefficient</i> | | | | | | | | |
|---|---------------|--------------|---------------|---------------|---------------|---------------|---------------|---------------|
| <i>Correlation Attribute Eval + XGBoost</i> | | | | | | | | |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,8419 | 0,798 | 0,8376 | 0,8296 | 0,8357 | 0,8471 | 0,8488 | 0,5460 |
| <i>Valence</i> | 0,6373 | 0,6903 | 0,6727 | 0,7058 | 0,7587 | 0,7378 | 0,687 | 0,7190 |

Nas Tabelas 15, 16, 17 e 18 contém as combinações entre Relief Attribute Eval com os seletores *MLR* e *SRV* nas métricas de *RMSE* e *Correlation Coefficient*. Nestas combinações é verificado poucos resultados melhores que o alvo. O resultado negativo ocorre pela aleatoriedade na escolha do valor de X nas N repetições do seletor, além do baixo impacto dos regressores com o conjunto de dados deste trabalho. As Tabelas 19 20 demonstram uma exceção com os testes realizados com o *XGBoost*, que é o regressor com melhores resultados neste trabalho.

Tabela 15: Resultado do *RMSE* na combinação do *Relief Attribute Eval* mais *MLR*.

| RMSE | | | | | | | | |
|----------------------------------|---------------|--------|--------|--------|--------|---------------|--------|---------------|
| <i>Relief Attribute Eval MLR</i> | | | | | | | | <i>Alvo</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,1025 | 0,1145 | 0,1374 | 0,195 | 0,196 | 0,196 | 0,1258 | 0,1110 |
| <i>Valence</i> | 0,1112 | 0,1409 | 0,1549 | 0,2207 | 0,1574 | 0,1279 | 0,1629 | 0,1360 |

Tabela 16: Resultado do *Correlation Coefficient* na combinação do *Relief Attribute Eval* mais *MLR*.

| <i>Correlation Coefficient</i> | | | | | | | | |
|----------------------------------|--------|--------|--------|--------|--------|--------|--------|---------------|
| <i>Relief Attribute Eval MLR</i> | | | | | | | | <i>Alvo</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,8261 | 0,7906 | 0,7268 | 0,6103 | 0,5583 | 0,6608 | 0,7292 | 0,5460 |
| <i>Valence</i> | 0,73 | 0,5949 | 0,5662 | 0,4139 | 0 | 0 | 0 | 0,7190 |

Tabela 17: Resultado do *RMSE* na combinação com o *Relief Attribute Eval* mais *SVR*.

| RMSE | | | | | | | | |
|------------------------------------|--------|--------|--------|--------|--------|--------|--------|---------------|
| <i>Relief Attribute Eval + SVR</i> | | | | | | | | <i>Alvo</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,1669 | 0,1862 | 0,1793 | 0,1792 | 0,1852 | 0,1893 | 0,1826 | 0,1110 |
| <i>Valence</i> | 0,1849 | 0,1628 | 0,1614 | 0,1567 | 0,1575 | 0,1628 | 0,1617 | 0,1360 |

Tabela 18: Resultado do *Correlation Coefficient* na combinação com o *Relief Attribute Eval* mais *SVR*.

| <i>Correlation Coefficient</i> | | | | | | | | |
|------------------------------------|--------|---------|--------|--------|--------|--------|--------|---------------|
| <i>Relief Attribute Eval + SVR</i> | | | | | | | | <i>Alvo</i> |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,5057 | -0,0297 | 0,0845 | 0,1773 | 0 | 0,0783 | 0,1098 | 0,5460 |
| <i>Valence</i> | 0 | 0,0528 | 0,1294 | 0,1603 | 0,0865 | 0,1349 | 0,1308 | 0,7190 |

Tabela 19: Resultado do *Relief Coeficient* na combinação com *Correlation Attribute Eval* mais o *XGBoost*.

| RMSE | | | | | | | | |
|--------------------------------------|---------------|---------------|---------------|---------------|---------------|--------------|---------------|---------------|
| <i>Relief Attribute Eval+XGBoost</i> | | | | | | | | |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,0954 | 0,0974 | 0,1027 | 0,0971 | 0,1071 | 0,105 | 0,1097 | 0,1110 |
| <i>Valence</i> | 0,0996 | 0,1142 | 0,1121 | 0,1165 | 0,1089 | 0,109 | 0,1197 | 0,1360 |

Tabela 20: Resultado do *Relief Coeficient* na combinação com *Correlation Attribute Eval* mais o *XGBoost*.

| <i>Correlation Coefficient</i> | | | | | | | | |
|--------------------------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| <i>Relief Attribute Eval+XGBoost</i> | | | | | | | | |
| Número de características | 100 | 200 | 300 | 500 | 1000 | 1500 | 2000 | 6373 |
| <i>Arousal</i> | 0,8914 | 0,8533 | 0,827 | 0,8466 | 0,8173 | 0,8331 | 0,8224 | 0,5460 |
| <i>Valence</i> | 0,8013 | 0,7122 | 0,7228 | 0,7338 | 0,7225 | 0,7485 | 0,6793 | 0,7190 |

A tabela 24 apresenta os resultados do *CFS Subset Eval*, com o regressor *SVR* na métrica *Correlation Coefficient*. Na combinação com o *Valence* foi obtido um resultado abaixo do alvo. Isso ocorre pelo fato do alto valor no Alvo na dimensão do *Valence* no *Correlation Coefficient*, que ser considera este valor fortemente relacionado. Nas Tabelas 21, 22, 23, 25 e 26 é visto os resultados do *CFS Subset Eval*, com os regressores *MLR*, *SVR* e *XGBoost*, nas métricas de *RMSE* e *Correlation Coefficient*. Todos estes obtiveram uma quantidade expressiva de resultados positivos com relação o Alvo. O resultado positivo do *CFS Subset Eval* é por conta do seletor escolher o subconjunto mais impactante para a base inserida.

Tabela 21: Resultado do *RMSE* na combinação do *CFS Subset Eval* mais *MLR*.

| RMSE | | |
|------------------------------|---------------|---------------|
| CFS Subset Eval + MLR | | <i>Alvo</i> |
| Número de características | - | 6373 |
| <i>Arousal</i> (92) | 0,1064 | 0,1110 |
| <i>Valence</i> (96) | 0,1181 | 0,1360 |

Tabela 22: Resultado do *Classifier Coeficient* na combinação com o *CFS Subset Eval* mais *MLR*.

| Correlation Coeficient | | |
|-------------------------------|---------------|---------------|
| CFS Subset Eval + MLR | | <i>Alvo</i> |
| Número de características | - | 6373 |
| <i>Arousal</i> (92) | 0,8206 | 0,5460 |
| <i>Valence</i> (96) | 0,7313 | 0,7190 |

Tabela 23: Resultado do *RMSE* na combinação *CFS Subset Eval* mais *SVR*.

| RMSE | | |
|------------------------------|---------------|---------------|
| CFS SUBSET EVAL + SVR | | <i>Alvo</i> |
| Número de características | - | 6373 |
| <i>Arousal</i> (92) | 0,1015 | 0,1110 |
| <i>Valence</i> (96) | 0,1366 | 0,1360 |

Tabela 24: Resultado do *Correlation Coeficient* na combinação *CFS Subset Eval* mais *SVR*.

| Correlation Coeficient | | |
|-------------------------------|---------------|---------------|
| CFS SUB SET EVAL + SVR | | <i>Alvo</i> |
| Número de características | - | 6373 |
| <i>Arousal</i> (92) | 0,8259 | 0,5460 |
| <i>Valence</i> (96) | 0,6202 | 0,7190 |

Tabela 25: Resultado do *RMSE* na combinação com *CFS Subset Eval* mais o *XGBoost*.

| RMSE | | |
|----------------------------------|---------------|---------------|
| CFS SUBSET EVAL + XGBoost | | <i>Alvo</i> |
| Número de características | - | 6373 |
| <i>Arousal</i> (92) | 0,1037 | 0,1110 |
| <i>Valence</i> (96) | 0,1168 | 0,1360 |

Tabela 26: Resultado do *Correlation Coeficient Eval* na combinação com *CFS Subset Eval* mais o *XG-Boost*.

| Correlation Coeficient | | |
|----------------------------------|---------------|---------------|
| CFS SUBSET EVAL + XGBoost | | <i>Alvo</i> |
| Número de características | - | 6373 |
| <i>Arousal</i> (92) | 0,8151 | 0,5460 |
| <i>Valence</i> (96) | 0,7419 | 0,7190 |

Um dos objetivos secundários deste trabalho é verificar os diferentes métodos de seleção no resultado de predicação. Com os resultados é percebido que os seletores do método de busca *Ranker* não obtém resultados expressivos com relação ao Alvo.

Com os resultados apresentados nas Tabelas 7, 8, 13, 14, 19, 20, 26 e 25 é percebido que o *XGBoost* é o melhor regressor deste trabalho. A expressiva quantidade de resultados melhor que Alvo, ocorre por conta deste regressor ter a capacidade de trabalhar com pequena ou grande quantidade de dados.

5.2 Impacto no *Arousal* e *Valence*

Este trabalho também tem como objetivo verificar quais características podem influenciar no *Arousal* e no *Valence*. Para obter estas informações foi realizada uma de seleção de características com o método *CFS Subset Eval*, que foi o melhor dentre os testados neste trabalho.

Nas Tabelas 27 e 28 são apresentadas a quantidade de sub-características do *Arousal* e *Valence* que obtiveram maior impacto nas emoções. As características com os melhores resultados são: *MFCC* (*Mel Frequency Cepstral Coefficient*), *PCM* (*Pulse-code modulation*) e *audspec*. Este resultado ocorreu por conta da quantidade de aparições na base de dados. Somadas todas as suas sub-características foi encontrada 5900 ocorrências. A frequência de ocorrências ocorreu porque a base PMEmo se baseou no trabalho (WENINGER et al., 2013), que provou funcionar efetivamente para a tarefa de regressão de humor de música, segundo Zhang et al. (2018).

Tabela 27: Resultado das melhores características para *Arousal*

| Característica | Quantidade |
|----------------|------------|
| Audspec | 34 |
| PCM | 21 |
| MFCC | 20 |
| F0 | 7 |
| Jitter | 3 |
| Log HNR | 3 |
| Shimmer Local | 2 |
| Voicing | 1 |

Tabela 28: Resultado das melhores características para *Valence*

| Característica | Quantidade |
|----------------|------------|
| Audspec | 38 |
| PCM | 25 |
| MFCC | 25 |
| F0 | 3 |
| Shimmer Local | 2 |
| Log HNR | 2 |
| Voicing | 1 |

A característica com a maior quantidade de sub-características é o *audspec*. O *audspec* faz a conversão de Hertz para mel e prossegue a filtragem de frequência de mel em referência a um banco de filtros de frequência que leva a uma redução da dimensionalidade (SUEUR, 2018). O *MFCC* é uma característica de que representa um padrão de recurso de curta duração (TINGLE; KIM; TURNBULL, 2010). O *PCM*, é uma característica baseada na amplitude do sinal e utilizada na etapa de pre-processamento do áudio (PAO; YEH; CHEN, 2008; NALINI; CHAKRABORTY, 2013).

As demais características tem uma menor importância na identificação de emoção, com base na triagem realizada pelos trabalhos, as características são: Zhang et al. (2018) e (WENINGER et al., 2013). As características com menor impacto foram: *Jitter*, *Shimmer Local*, *Log HNR*, *F0* e *Voicing*. O *Jitter* e o *Shimmer* são variações de características linguísticas do comprimento e amplitudes (respectivamente) da frequência fundamental para sons harmônicos, segundo Weninger et al. (2013). O *Log HNR* são características de baixo nível fornecidas a partir de 64 descritores (WENINGER et al., 2013). As sub-características de *F0* e *Voicing* são mais eficazes no *Arousal* e suas informações são extraídas da fala (WENINGER et al., 2013).

Outro resultado verificado nas Tabelas 27 e 28 é a presença 3 sub-características do *Jitter*, que influenciam o *Arousal* e nenhuma no *Valence*. Isso ocorre pelo maior impacto e listagem na bibliografia da integração do *Jitter* dentro do *Arousal*, como no trabalho Kirke et al. (2018), que fez um mapeamento das características no *Arousal*.

6 Conclusão

Esta monografia realizou a seleção de características em uma base de dados pública da área de REM. O intuito da seleção de características é verificar o impacto desta etapa para a predição das emoções e identificar quais características tem uma maior relação com o *Arousal* e o *Valence*.

Com a utilização da seleção de características foi perceptível o impacto desta etapa no predição emocional. Pode ser concluído que os seletores com o método de busca *Ranker* causam baixo impacto nos resultados das métricas de predição. No entanto, as combinações com o seletor *CFS Subset Eval*, conseguiram os resultados positivos com relação ao Alvo.

Também pode ser concluído, que a escolha do regressor ocasiona em um impacto positivo ou negativo nos resultados de predição. Dos regressores escolhidos, o *SRV*, *MLR* obtiveram os resultados inferiores comparado ao Alvo. Entretanto, o *XGBoost* alcançou os melhores resultados na predição emocional. Com os resultados do *XGBoost* e do método de busca *Best First* pode ser concluído que algoritmos baseados em árvore ou grafos se comportam melhor com dados contínuos da área de REM.

Em relação as características relacionadas com o *Arousal* e o *Valence*, as que mais se destacaram foram: *audspec*, *PCM* e *MFCC*. As três características obtiveram impacto no *Arousal* e no *Valence*. Mas das três características principais, o *audspec* pode ser definida como a mais influente, por obter a maior quantidade de sub-características que as outras.

Como trabalhos futuros é pretendido analisar mais bases de dados públicas, a partir disso, pode ser obtido um resultado mais abrangente sobre a seleção de características dentro da área de REM. Como também, pode ser obtido uma relação mais ampla das características que impactam no *Arousal* e *Valence*. Assim, uma melhorara na predição das emoções.

Referências

- ABDUL, A. et al. An emotion-aware personalized music recommendation system using a convolutional neural networks approach. *Applied Sciences*, Multidisciplinary Digital Publishing Institute, v. 8, n. 7, p. 1103, 2018. Citado na página 13.
- ALJANAKI, A.; WIERING, F.; VELTKAMP, R. C. Mirutrecht participation in mediaeval 2013: Emotion in music task. In: CITESEER. *MediaEval*. [S.l.], 2013. Citado na página 21.
- ALJANAKI, A.; WIERING, F.; VELTKAMP, R. C. Studying emotion induced by music through a crowdsourcing game. *Information Processing & Management*, Elsevier, v. 52, n. 1, p. 115–128, 2016. Citado na página 13.
- ALJANAKI, A.; YANG, Y.-H.; SOLEYMANI, M. Developing a benchmark for emotional analysis of music. *PloS one*, Public Library of Science, v. 12, n. 3, p. e0173392, 2017. Citado na página 13.
- AVISADO, H. G. et al. Analysis of music timbre features for the construction of user-specific affect model. In: *Theory and Practice of Computation*. [S.l.]: Springer, 2012. p. 28–35. Citado na página 20.
- BAI, J. et al. Dimensional music emotion recognition by valence-arousal regression. In: IEEE. *2016 IEEE 15th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)*. [S.l.], 2016. p. 42–49. Citado 2 vezes nas páginas 17 e 21.
- BYDOVSKA, H.; POPELÍNSKÝ, L. Predicting student performance in higher education. In: IEEE. *2013 24th International workshop on database and expert systems applications*. [S.l.], 2013. p. 141–145. Citado na página 22.
- CHANEL, G.; ANSARI-ASL, K.; PUN, T. Valence-arousal evaluation using physiological signals in an emotion recall paradigm. In: IEEE. *2007 IEEE International Conference on Systems, Man and Cybernetics*. [S.l.], 2007. p. 2662–2667. Citado na página 13.
- CHAUHAN, N.; BAHL, S. Performance analysis of dimension reduction techniques with classifier combination for intrusion detection system. In: IEEE. *2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom)*. [S.l.], 2015. p. 1084–1089. Citado na página 22.
- CHEN, T. et al. Emotion recognition from eeg signals enhanced by user’s profile. In: ACM. *Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval*. [S.l.], 2016. p. 277–280. Citado na página 13.
- CHEN, Y.-A. et al. The amg1608 dataset for music emotion recognition. In: IEEE. *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. [S.l.], 2015. p. 693–697. Citado na página 13.
- FEIGEN, L. P. Physical characteristics of sound and hearing. *American journal of Cardiology*, Elsevier, v. 28, n. 2, p. 130–133, 1971. Citado na página 18.

- FILLIBEN, J. J. The probability plot correlation coefficient test for normality. *Technometrics*, Taylor & Francis Group, v. 17, n. 1, p. 111–117, 1975. Citado na página 24.
- GONZÁLEZ-VIDAL, A.; JIMÉNEZ, F.; GÓMEZ-SKARMETA, A. F. A methodology for energy multivariate time series forecasting in smart buildings based on feature selection. *Energy and Buildings*, Elsevier, v. 196, p. 71–82, 2019. Citado na página 21.
- GREKOW, J. Audio features dedicated to the detection of arousal and valence in music recordings. In: IEEE. *2017 IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*. [S.l.], 2017. p. 40–44. Citado 2 vezes nas páginas 16 e 17.
- GUAN, D.; CHEN, X.; YANG, D. Music emotion regression based on multi-modal features. In: *Proc. International Symposium on Computer Music Modeling and Retrieval*. [S.l.: s.n.], 2012. p. 70–77. Citado na página 17.
- HAN, B.-j. et al. Smers: Music emotion recognition using support vector regression. In: *ISMIR*. [S.l.: s.n.], 2009. p. 651–656. Citado na página 13.
- HEVNER, K. Experimental studies of the elements of expression in music. *American journal of Psychology*, v. 48, n. 2, p. 246–268, 1936. Citado na página 19.
- HUQ, A.; BELLO, J. P.; ROWE, R. Automated music emotion recognition: A systematic evaluation. *Journal of New Music Research*, Taylor & Francis, v. 39, n. 3, p. 227–244, 2010. Citado 2 vezes nas páginas 16 e 17.
- KHAN, Z. M. A.; SAEIDLOU, S.; SAADAT, M. Ontology-based decision tree model for prediction in a manufacturing network. *Production & Manufacturing Research*, Taylor & Francis, v. 7, n. 1, p. 335–349, 2019. Citado na página 22.
- KIRKE, A. et al. Unconsciously interactive films in a cinema environment—a demonstrative case study. *Digital Creativity*, Taylor & Francis, v. 29, n. 2-3, p. 165–181, 2018. Citado na página 32.
- KONONENKO, I. Estimating attributes: analysis and extensions of relief. In: SPRINGER. *European conference on machine learning*. [S.l.], 1994. p. 171–182. Citado na página 22.
- LI, T. R. et al. Sentiment-based prediction of alternative cryptocurrency price fluctuations using gradient boosting tree model. *arXiv preprint arXiv:1805.00558*, 2018. Citado na página 23.
- LOPES, P. S. et al. The use of the convolutional neural network as an emotion classifier in a music recommendation system. In: ACM. *Proceedings of the XIV Brazilian Symposium on Information Systems*. [S.l.], 2018. p. 41. Citado na página 13.
- NALINI, N.; CHAKRABORTY, A. Speech emotion recognition using mfcc and aann. In: *Proc. International Conference on Engineering and Technology*. [S.l.: s.n.], 2013. p. 223–225. Citado na página 32.
- PAO, T.-L.; YEH, J.-H.; CHEN, Y.-T. Emotion recognition via continuous mandarin speech. In: *Advances in human computer interaction*. [S.l.]: IntechOpen, 2008. Citado na página 32.

- PISTON, W. *Armonía, España, Edit.* [S.l.]: Span Press–Universitaria, 1998. Citado na página 18.
- POUYANFAR, S.; SAMETI, H. Music emotion recognition using two level classification. In: IEEE. *2014 Iranian Conference on Intelligent Systems (ICIS)*. [S.l.], 2014. p. 1–6. Citado na página 21.
- REISENZEIN, R. What is a definition of emotion? and are emotions mental-behavioral processes? *Social Science Information*, Sage Publications Sage UK: London, England, v. 46, n. 3, p. 424–428, 2007. Citado na página 19.
- RUSSELL, J. A. A circumplex model of affect. *Journal of personality and social psychology*, American Psychological Association, v. 39, n. 6, p. 1161, 1980. Citado na página 19.
- SEO, Y.-S.; HUH, J.-H. Automatic emotion-based music classification for supporting intelligent iot applications. *Electronics*, Multidisciplinary Digital Publishing Institute, v. 8, n. 2, p. 164, 2019. Citado na página 13.
- SONG, K.-B. et al. Short-term load forecasting for the holidays using fuzzy linear regression method. *IEEE transactions on power systems*, IEEE, v. 20, n. 1, p. 96–101, 2005. Citado na página 23.
- SUEUR, J. Mel-frequency cepstral and linear predictive coefficients. In: *Sound Analysis and Synthesis with R*. [S.l.]: Springer, 2018. p. 381–398. Citado na página 32.
- TINGLE, D.; KIM, Y. E.; TURNBULL, D. Exploring automatic music annotation with acoustically-objective tags. In: ACM. *Proceedings of the international conference on Multimedia information retrieval*. [S.l.], 2010. p. 55–62. Citado na página 32.
- WANG, K.-C. A hybrid kansei engineering design expert system based on grey system theory and support vector regression. *Expert Systems with Applications*, Elsevier, v. 38, n. 7, p. 8738–8750, 2011. Citado na página 23.
- WANG, Y. et al. Best-first search guided multistage mass spectrometry-based glycan identification. *Bioinformatics*, 2019. Citado na página 22.
- WENINGER, F. et al. On the acoustics of emotion in audio: what speech, music, and sound have in common. *Frontiers in psychology*, Frontiers, v. 4, p. 292, 2013. Citado 2 vezes nas páginas 31 e 32.
- YEH, C.-H.; LIN, H.-H.; CHANG, H.-T. An efficient emotion detection scheme for popular music. In: IEEE. *Circuits and Systems, 2009. ISCAS 2009. IEEE International Symposium on*. [S.l.], 2009. p. 1799–1802. Citado na página 19.
- YENIAY, O.; GOKTAS, A. A comparison of partial least squares regression with other prediction methods. *Hacettepe Journal of Mathematics and Statistics*, v. 31, n. 99, p. 99–101, 2002. Citado na página 24.
- ZHANG, J. et al. Bridge the semantic gap between pop music acoustic feature and emotion: Build an interpretable model. *Neurocomputing*, Elsevier, v. 208, p. 333–341, 2016. Citado na página 18.

ZHANG, J. L. et al. Feature selection and feature learning in arousal dimension of music emotion by using shrinkage methods. *Multimedia Systems*, Springer, v. 23, n. 2, p. 251–264, 2017. Citado 2 vezes nas páginas 16 e 17.

ZHANG, K. et al. The pmemo dataset for music emotion recognition. In: ACM. *Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval*. [S.l.], 2018. p. 135–142. Citado 3 vezes nas páginas 21, 31 e 32.



**TERMO DE AUTORIZAÇÃO PARA PUBLICAÇÃO DIGITAL NA BIBLIOTECA
"JOSÉ ALBANO DE MACEDO"**

Identificação do Tipo de Documento

- () Tese
() Dissertação
(X) Monografia
() Artigo

Eu, Douglas Siqueira Gonçalves,
autorizo com base na Lei Federal nº 9.610 de 19 de Fevereiro de 1998 e na Lei nº 10.973 de
02 de dezembro de 2004, a biblioteca da Universidade Federal do Piauí a divulgar,
gratuitamente, sem ressarcimento de direitos autorais, o texto integral da publicação
Seleção de Características para Predição da Emoção na Música
de minha autoria, em formato PDF, para fins de leitura e/ou impressão, pela internet a título
de divulgação da produção científica gerada pela Universidade.

Picos-PI 01 de agosto de 2019.

Douglas Siqueira Gonçalves
Assinatura