

Moodic - Uma ferramenta para análise de sentimentos através de hábitos musicais

Projeto final da disciplina Processamento de Linguagem Natural (DCC/UFMG)

Hugo Araujo de Sousa
Universidade Federal de Minas Gerais
Belo Horizonte, MG
hugosousa@dcc.ufmg.br

RESUMO

Nesse trabalho é desenvolvida uma ferramenta para análise de sentimentos e emoções através dos hábitos musicais de uma pessoa. As relações íntimas entre música e as emoções do ouvinte são examinadas e, a partir desse estudo, uma modelagem é proposta para identificar, através da inspeção das músicas (com foco em suas letras) que um ouvinte escuta, suas emoções e possíveis transtornos psicológicos.

CONCEITOS

• **Processamento de Linguagem Natural** → *Análise de Sentimentos*; • **Metodologias computacionais** → Extração de informação; • **Computação aplicada** → Computação de som e música;

1 INTRODUÇÃO

Grande quantidade dos textos produzidos pelos seres humanos tem como objetivo refletirem as opiniões e sentimentos do autor, em contraste com a categoria de textos onde a preocupação é com fatos e expressões objetivas. Dentro dos textos subjetivos, aqueles de cunho artístico sempre foram exemplo mais notável de expressividade emocional [19].

Uma das formas de arte mais antigas, a música consegue alcançar níveis de expressividade emocional especialmente interessantes, combinando tanto a linguagem natural quanto sua própria linguagem através de ritmo, sons, instrumentos, entre outros elementos. Sendo dotada de tal capacidade de expressão, a música foi e continua sendo estudada por cientistas da área da psicologia [12] [20] [8]. Dentre os vários aspectos que relacionam música e emoção, três deles são particularmente instigantes para serem estudados:

- Como emoções podem influenciar a música que alguém escolhe ouvir.
- Como música pode expressar emoção e sentimentos.
- Como música pode induzir emoções no ouvinte.

Já sabemos que os hábitos musicais de uma pessoa influenciam diretamente seu humor, sentimentos e emoções [15]. Mais especificamente, existem estudos que demonstram a influência das letras de uma música no humor e comportamento do ouvinte [21]. Esses dois aspectos, aliados ao fato de que o consumo de música sofreu grande aumento com a chegada das plataformas de streaming digital e compartilhamento de arquivos [11] nos fornecem grande indicação de que podemos utilizar os hábitos musicais de um ouvinte como fonte de informações a respeito do humor e emoções do mesmo.

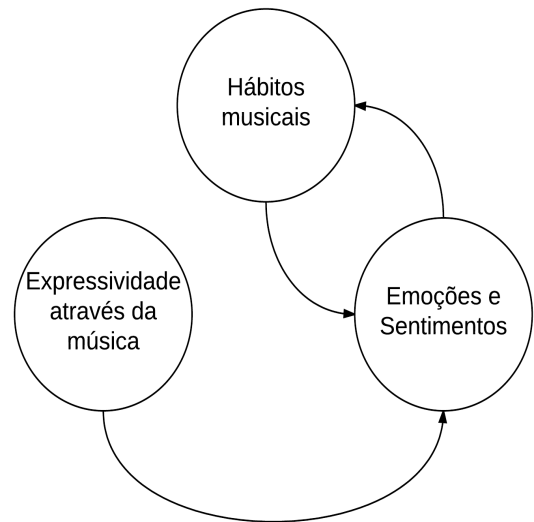


Figura 1: Relação entre expressividade de um artista através de sua música e o ciclo criado entre a indução de sentimentos no ouvinte e a escolha por músicas que correspondem a seu humor atual.

2 PROBLEMA

Diante da relação íntima entre humor e hábitos musicais, uma tarefa particularmente relevante é aquela de identificar usuários (ouvintes) cujos hábitos de consumo musical denotam possíveis transtornos psicológicos e de humor. Enquanto esse tipo de análise é muito simplista para determinar qualquer tipo de transtorno com acurácia, ele pode servir como um forte indicador de alterações de humor e, como tal, identificador de usuários que necessitam de algum tipo de suporte emocional.

Plataformas de streaming e recomendação de música online como Last.fm [9] e Spotify [13] possuem bases de dados massivas sobre seus usuários [18] [10], contendo informações sobre seus artistas favoritos, músicas, álbuns, além de coletar, em tempo real, informações sobre o hábito musical de um ouvinte. Dessa forma, essas plataformas a base de dados ideal para a tarefa de identificar o humor de um usuário baseado nas músicas que este ouve.

A partir desses dados, é possível coletar os hábitos musicais de um determinado usuário ao longo de um intervalo de tempo qualquer. Com esses dados em mão e uma abordagem que envolve tanto

os campos de processamento de linguagem natural - mais especificamente a área de análise de sentimentos, para tentar posicionar as letras de uma música no espectro de emoções humanas -, quanto a área da psicologia e música em si, podemos tentar extrair, a partir dos hábitos musicais de um usuário, suas emoções e sentimentos, identificando em último caso aqueles que demonstram maior necessidade de intervenção e suporte psicológico, que pode ser oferecido de várias maneiras diferentes pelas próprias plataformas.

3 MODELAGEM

O primeiro passo para resolver abordar esse problema é determinar como coletar os dados necessários para traçar um panorama sobre os hábitos musicais de um usuário. Para isso, a plataforma Last.fm foi escolhida, não só pela quantidade massiva de dados que acumula, mas também pelo fato de que todos os seus dados são exibidos publicamente na mesma, como definido nos termos de contrato do serviço [2].

Nessa etapa é importante definir a quantidade n para determinar as últimas n músicas que um usuário ouviu. Essa não é uma escolha muito simples, uma vez que, para valores muito altos de n , perdemos acurácia, no sentido de que essas músicas podem refletir um período de tempo muito longo, onde as emoções do usuário podem variar arbitrariamente. Por outro lado, é desejável escolher um valor de n suficientemente alto para reunir a maior quantidade possível de músicas que refletem um mesmo estado de sentimento/humor do usuário.

Em seguida, é preciso reunir os atributos das canções que servirão de indicadores para as emoções das mesmas e, consequentemente, dos usuários. O primeiro atributo a ser considerado são as letras das músicas. Para facilitar a implementação do trabalho, só foram levadas em consideração letras das músicas cujo idioma é o inglês (uma vez que já existe uma grande quantidade de ferramentas existentes para trabalhar com esse idioma).

De posse das letras de uma canção, podemos tentar quantificar a negatividade das palavras. Para isso, podemos usar uma base de dados que já contenha emoções associadas a cada palavra. Para esse trabalho, foi escolhida a base *NRC Word-Emotion Association Lexicon* [16]. Além disso, é necessário realizar um pré-processamento no texto, removendo as flexões das palavras e excluindo aquelas cuja classe gramatical não seja forte indicador de sentimento ou emoção, i.e., palavras como 'hi', 'him', 'there', etc. Com isso, conseguimos quantificar a porcentagem de palavras negativas nas letras de uma música, onde uma palavra é considerada negativa se, no lexicon de referência, ela ativa sentimentos considerados negativos como raiva e medo.

Um fator importante a ser considerado para a análise das letras de uma canção diz respeito ao peso ou densidade das letras em relação à canção como um todo. Isto é, é importante verificar o quanto o humor de uma canção é influenciado pelas suas letras. Imaginemos o caso de uma música de duração de 10 minutos. Caso essa canção possua apenas duas frases ao longo dos 10 minutos, é natural pensar que essas duas frases terão pouco impacto no humor geral da canção, ao passo que, caso ao longo dos 10 minutos haja poucos intervalos sem qualquer letra, então podemos pensar que o humor dessa canção é fortemente influenciado por suas letras. A

partir disso, definimos o conceito de *peso lírico*, que se refere a quão alto é o fluxo lírico em uma canção por segundo de duração.

Naturalmente, canções instrumentais também devem ser identificadas e tratadas de forma especial, de modo que também possam ser extraídas informações importantes das mesmas. De fato, o essas canções recebem no trabalho o mesmo tratamento de canções cujo idioma é diferente do inglês, uma vez que a análise da letra dessas também não foi incluída, como mencionado anteriormente.

Para uma análise mais razoável e completa, é necessário obter uma medida de negatividade para as canções instrumentais e aquelas cujas letras não podem ser analisadas. Para isso, a decisão foi de incluir uma medida já definida pela própria plataforma de streaming Spotify [6], chamada valência (*valence*). De acordo com a própria documentação da API do Spotify, essa medida é definida como:

Uma medida de 0.0 a 1.0 que descreve a positividade musical transmitida por uma faixa. Faixas com valência alta soam mais positivas (e.g. felizes, alegres, eufóricas), enquanto faixas com baixa valência soam mais negativas (e.g. tristes, deprimentes, irritadas). - (tradução livre)

Dessa forma, temos três medidas para construir o índice de negatividade de uma canção, seu peso lírico, a negatividade lírica e a valência. Essas três medidas serão usadas para definir o índice de negatividade de cada canção que um determinado usuário ouviu ao longo do tempo, obtendo-se assim uma medida indicadora do próprio humor do usuário.

4 IMPLEMENTAÇÃO

Nessa seção são discutidas todas as decisões de implementação tomadas durante o desenvolvimento do trabalho em questão.

4.1 Linguagem e Bibliotecas

A partir da modelagem descrita na Seção 3, foi desenvolvida uma ferramenta para análise dos sentimentos de um usuário através de seus hábitos musicais. A ferramenta foi apelidada de *Moodic* (um jogo de palavras com *mood* e *music* - humor e música em inglês).

A ferramenta foi desenvolvida utilizando a linguagem Python 3 [5]. Essa linguagem foi escolhida pela grande quantidade de ferramentas para processamento de linguagem natural disponíveis para uso gratuito através da mesma.

Para a coleta de letras das canções, por exemplo, foi utilizada a biblioteca *lyricwikia* [4]. Para essa tarefa, não houve qualquer fator decisivo na escolha, porém o ideal, para trabalhos futuros, seria a escolha de uma biblioteca com o maior banco de dados possível, a fim de evitar que um grande número de canções, mesmo aquelas que possuem letras, não tenham suas letras encontradas pelo programa e, como tal, acabem sem uma medida de negatividade lírica.

Com relação à coleta de dados da plataforma Last.fm, foi identificado um *wrapper* em Python para a própria API da plataforma. Essa biblioteca, a *pyLast* [1] fornece acesso aos dados da plataforma de forma fácil, com ela, é possível coletar dados sobre canções, álbuns e músicas escutadas por um determinado usuário. Para esse trabalho, somente foi necessária a utilização da mesma para coletar as n canções ouvidas por um usuário.

De forma similar, a biblioteca *Spotipy* [7] fornece uma interface em Python para a API Web da plataforma Spotify. Com essa biblioteca, o programador consegue obter os dados de um artista, canções, álbuns, imagens, etc. O interesse para o trabalho em questão, entretanto, é obter algumas informações interessantes sobre as faixas ouvidas por um usuário, como duração (necessária para o cálculo de peso lírico) e valência.

Também é importante mencionar a biblioteca *langdetect* [3], que foi responsável por detectar a linguagem de cada uma das letras das músicas, permitindo que o programa apenas trabalhasse com letras em inglês.

Por fim, a biblioteca *NLTK* [17] foi de grande utilidade para realizar o pré-processamento do texto das canções, como descrito na Seção 4.4, além da tarefa de filtragem das palavras cujas classes gramaticais eram mais relevantes para o programa, como descrito na Seção 4.5.

4.2 NRC Word-Emotion Association Lexicon

O processo de quantificação de negatividade nas letras de uma determinada canção começa no problema de como associar emoções a palavras. Há várias abordagens que podem ser utilizadas nesse problema, seja levando em consideração palavras individualmente ou sentenças de k palavras. A fim de facilitar implementação, a escolha para esse trabalho foi a de analisar cada palavra do texto de uma música de forma individual, isto é, sem considerar as palavras que cercam uma determinada palavra (seu contexto).

Para isso, foi utilizado o NRC Emolex, que consiste em uma grande lista de palavras, seguidas de 10 possíveis emoções, seguidas de um número (0 ou 1), que indica se aquela palavra "ativa" tal emoção ou não, sendo 0 representante de emoção não ativada e 1 de emoções ativadas. O conceito de ativação aqui indica simplesmente associação entre a palavra e a emoção que transmite, de forma isolada, sem informações de contexto.

Palavra	Emoção	Ativação
suicide	anger	1
suicide	anticipation	0
suicide	disgust	0
suicide	fear	1
suicide	joy	0
suicide	negative	1
suicide	positive	0
suicide	sadness	1
suicide	surprise	0
suicide	trust	0

Tabela 1: Exemplo de dados do NRC Emolex para a palavra "suicide".

4.3 Emoções negativas

Através do lexicon apresentado na Seção 4.2, foi necessário definir quais dentre as 10 emoções teriam influência no índice de negatividade de uma canção. De maneira geral, as 10 emoções presentes no lexicon são muito bem definidas como negativas ou positivas, sendo a separação feita como mostrado na Tabela 2.

Polaridade	Emoções
Positivo	Joy Trust
Neutro	Anticipation Surprise
Negativo	Anger Disgust Fear Negative Sadness

Tabela 2: Separação das emoções do NRC Emolex em positivas e negativas.

Após essa separação de polaridades, o foco do trabalho foi nas emoções negativas. Dessa forma, ao processar uma palavra no texto de uma canção, o programa verifica se a mesma ativa alguma das emoções negativas, como definido na Tabela 2. Caso isso aconteça, um contador de palavras negativas é incrementado.

Essa abordagem é simplista e poderia ser melhorada para incluir pesos às diferentes emoções dentro da polaridade negativa. Por exemplo, poderia ser interessante definir que a emoção *disgust* é mais *anger*, ou vice-versa. Essas decisões devem levar em conta o propósito da aplicação. Para o trabalho em questão, como o objetivo é simplesmente determinar um índice de negatividade, a abordagem não levou em consideração pesos.

4.4 Pré-processamento de texto

Uma tarefa fundamental em processamento de linguagem natural diz respeito ao pré-processamento do texto a ser analisado, uma vez que esse, em sua maioria, não se apresenta de forma estruturada, isto é, com o formato necessário para ser usado como entrada para os diversos algoritmos existentes [14].

Para o trabalho em questão, foi necessário coletar as letras das canções em formato bruto e transformá-las em uma lista de palavras no mesmo formato em que se apresentam as palavras no NRC Emolex. Para isso, o primeiro passo foi realizar um processo de *tokenização* onde a letra de uma canção, inicialmente uma longa string, é quebrada em tokens. Para cada um desses tokens, aplicamos um algoritmo de *lematização*, removendo flexões.

Cada token também foi transformado em seu equivalente apenas com letras minúsculas. Além disso, é importante remover do texto original aqueles tokens que não representam palavras. Esses podem ser tanto numerais quanto símbolos de pontuação. Por fim, foram removidas do texto aquelas palavras que denotam conectores sentenciais, como "if", "into", "in", etc, uma vez que essas palavras não denotam emoções e, portanto, não são relevantes para a nossa análise.

4.5 Classes gramaticais não expressivas

Um fator importante observado durante a implementação do trabalho foi o de que o NRC Emolex contém somente determinados tipos de palavras. Esses tipos são definidos pelas suas classes gramaticais. Logo, foi importante filtrar dos tokens obtidos através das letras de uma canção aqueles cuja classe gramatical não faça parte do

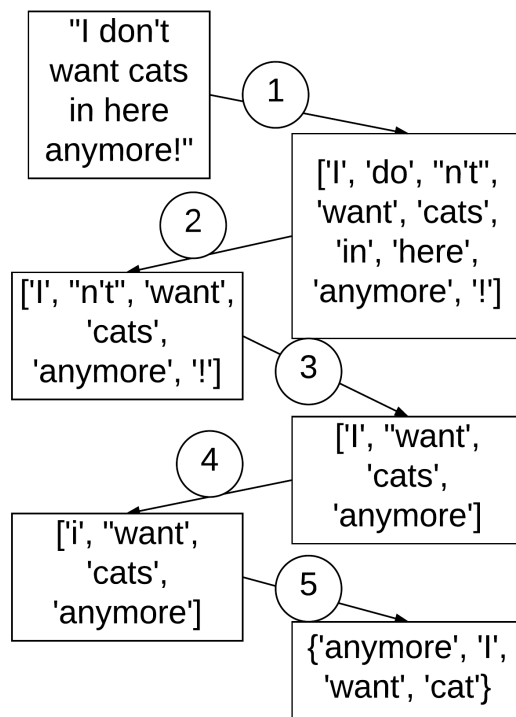


Figura 2: Pré-processamento de texto realizado para cada texto de canção encontrada pelo programa. 1: A string que se refere à letra de uma canção é quebrada em tokens. 2: Remove-se conectores sentenciais dos tokens. 3: Remove-se tokens não formados exclusivamente por letras. 4: Todas as letras dos tokens são transformadas para minúsculas. 5: Remove-se as flexões dos tokens. Observamos que o token "i" volta a apresentar letras maiúsculas. Isso se deve ao fato de que a palavra "I" inglês sempre tem grafia em maiúsculo.

Emolex. Com isso, aumentamos a precisão da análise que conta as palavras negativas no conjunto de tokens.

Essa filtragem que mantém apenas as classes gramaticais emocionalmente expressivas foi feita de forma empírica, observando, de todas as classes gramaticais possíveis dentro da biblioteca NLTK, aquelas que não indicam necessariamente algum tipo de humor ou sentimento. A lista das classes gramaticais definidas como não expressivas é apresentada na Tabela 3.

4.6 Peso lírico

4.7 Índice de negatividade

5 ESTUDO DE CASO

6 CONCLUSÃO

REFERÊNCIAS

- [1] [n. d.]. GitHub pyLast: A Python interface to Last.fm and Libre.fm. <https://github.com/pylast/pylast>. ([n. d.]). Acessado em: 04-12-2017.
- [2] [n. d.]. LastFM Terms of Use. <https://www.last.fm/legal/terms>. ([n. d.]). Acessado em: 04-12-2017.

Classe Gramatical na NLTK	Descrição
CC	Conjunção coordenativa
CD	Dígito cardinal
DT	Determinador
EX	"There"existencial: "there is"
IN	Preposição / Conjunção subordinativa
LS	Marcador de lista: 1)
NNP	Nome próprio
POS	Terminações possessivas: parent's
PRP	Pronome pessoal
PRP\$	Pronome possessivo
TO	"To"quando usado como preposição
UH	Interjeições
WDT	Determinador wh: "which"
WP	Pronomes wh: "who", "what"
WP\$	Pronome possessivo wh: "whose"
WRB	Advérbios wh: "where", "when"

Tabela 3: Relação das classes gramaticais na biblioteca NLTK definidas como não expressivas emocionalmente no NRC Emolex.

- [3] [n. d.]. Python Packages: langdetect 1.0.7. <https://pypi.python.org/pypi/langdetect/>. ([n. d.]). Acessado em: 06-12-2017.
- [4] [n. d.]. Python Packages: lyricwikia 0.1.4. <https://pypi.python.org/pypi/lyricwikia>. ([n. d.]). Acessado em: 06-12-2017.
- [5] [n. d.]. Python: Python 3.0 release. <https://www.python.org/download/releases/3.0/>. ([n. d.]). Acessado em: 06-12-2017.
- [6] [n. d.]. Spotify Developer: Get audio features for a track. <https://developer.spotify.com/web-api/get-audio-features/>. ([n. d.]). Acessado em: 06-12-2017.
- [7] [n. d.]. Spotipy: spotipy 2.0 documentation. <http://spotipy.readthedocs.io/en/latest/>. ([n. d.]). Acessado em: 06-12-2017.
- [8] Elvira Brattico, V. Alluri, B. Bogert, T. Jacobsen, Nuutti Vartiainen, Sirke Nieminen, and M. Tervaniemi. 2011. A functional MRI study of happy and sad emotions in music with and without lyrics. (2011), 1–16. <https://doi.org/10.3389/fpsyg.2011.00308>
- [9] Jon Haupt. 2009. Last.fm: People-Powered Online Radio. *Music Reference Services Quarterly* 12, 1-2 (2009), 23–24. <https://doi.org/10.1080/10588160902816702> arXiv:<https://doi.org/10.1080/10588160902816702>
- [10] V. Henning and J. Reichelt. 2008. Mendeley - A Last.fm For Research?. In *2008 IEEE Fourth International Conference on eScience*. 327–328. <https://doi.org/10.1109/eScience.2008.128>
- [11] Chun-Yao Huang. 2005. File Sharing as a Form of Music Consumption. *International Journal of Electronic Commerce* 9, 4 (2005), 37–55.
- [12] Patrik N Juslin and John A Sloboda. 2001. *Music and emotion: Theory and research*. Oxford University Press.
- [13] G. Kreitz and F. Niemela. 2010. Spotify – Large Scale, Low Latency, P2P Music-on-Demand Streaming. In *2010 IEEE Tenth International Conference on Peer-to-Peer Computing (P2P)*. 1–10. <https://doi.org/10.1109/P2P.2010.5569963>
- [14] Christopher D. Manning and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*. MIT Press, Cambridge, MA, USA.
- [15] Rollin McCraty, B Barrios-Choplin, M Atkinson, and Dana Tomasino. 1998. The effects of different types of music on mood, tension, and mental clarity. 4 (02 1998), 75–84.
- [16] Saif M. Mohammad and Peter D. Turney. 2013. Crowdsourcing a Word-Emotion Association Lexicon. 29, 3 (2013), 436–465.
- [17] Jacob Perkins. 2014. *Python 3 text processing with NLTK 3 cookbook*. Packt Publ., Birmingham. <http://cbs.cern.ch/record/1953497>
- [18] Martin Pichl, Eva Zangerle, and Günther Specht. 2014. Combining Spotify and Twitter Data for Generating a Recent and Public Dataset for Music Recommendation. In *Grundlagen von Datenbanken*. 35–40.
- [19] Jenefer Robinson. 2007. Expression and Expressiveness in Art. *Postgraduate Journal of Aesthetics* 4, 2 (2007). <http://pjaesthetics.org/index.php/pjaesthetics/article/view/57>
- [20] Thomas Schäfer, Peter Sedlmeier, Christine Städtler, and David Huron. 2013. The psychological functions of music listening. *Frontiers in psychology* 4 (2013). <https://doi.org/10.3389/fpsyg.2013.00511>

- [21] Valerie N. Stratton and Annette H. Zalanowski. 1994. Affective Impact of Music Vs. Lyrics. *Empirical Studies of the Arts* 12, 2 (1994), 173–184. <https://doi.org/10.2190/35T0-U4DT-N09Q-LQHW> arXiv:<https://doi.org/10.2190/35T0-U4DT-N09Q-LQHW>