Personalized musically induced emotions of not-so-popular Colombian music

Juan Sebastián Gómez-Cañón
MIRLab - Music Technology Group
Universitat Pompeu Fabra
Barcelona, Spain, 08019
juansebastian.gomez@upf.edu

Perfecto Herrera
MIRLab - Music Technology Group
Universitat Pompeu Fabra
Barcelona, Spain, 08019
perfecto.herrera@upf.edu

Estefanía Cano
Songquito UG
Erlangen, Germany, 91054
estefania.cano@songquito.com

Emilia Gómez
Joint Research Centre
European Commission
Seville, Spain, 41092
emilia.gomez-gutierrez@ec.europa.eu

Abstract

This work presents an initial proof of concept of how Music Emotion Recognition (MER) systems could be intentionally biased with respect to annotations of musically-induced emotions in a political context. In specific, we analyze traditional Colombian music containing politically-charged lyrics of two types: (1) vallenatos and social songs from the "left-wing" guerrilla Fuerzas Armadas Revolucionarias de Colombia (FARC) and (2) corridos from the "right-wing" paramilitaries Autodefensas Unidas de Colombia (AUC). We train personalized machine learning models to predict induced emotions for three users with diverse political views – we aim at identifying the songs that may induce negative emotions for a particular user, such as anger and fear. To this extent, a user's emotion judgements could be interpreted as problematizing data – subjective emotional judgments could in turn be used to influence the user in a human-centered machine learning environment. In short, highly desired "emotion regulation" applications could potentially deviate to "emotion manipulation" – the recent discredit of emotion recognition technologies might transcend ethical issues of diversity and inclusion.

1 Introduction

From the computational perspective, Music Emotion Recognition (MER) attempts to predict the emotion perceived by or induced in a particular listener [50]. Despite the criticism to MER due to subjectivity (and response diversity) [19], a growing effort has been made to produce enriched datasets of emotion judgements with more listening data to better represent the properties and context of the listener [4, 42]: demographics, cultural and individual differences, preference, familiarity, functional uses of music, physiological signals, and language. In this context, personalized and context-sensitive models that incorporate this information could more accurately predict the particular emotion judgments from a particular listener or groups of listeners [51, 16, 20]. However, in the context of evolving technologies for user profiling, the High Level Expert Group of the European Commission proposed that societal well-being is one of seven requirements to produce trustworthy artificial intelligence systems[1] – MER systems could potentially be harmful when inducing particular emotions to a listener [19].

[1] https://op.europa.eu/s/pInE

35th Conference on Neural Information Processing Systems (NeurIPS 2021), Sydney, Australia.

The aim of this paper is to attempt to understand if and under which circumstances MER algorithm can effectively be biased to recommend music that could induce negative emotions on a particular listener. Namely, a personalized MER algorithm that produces music recommendations grounded on the users' emotional judgments (an already standardized practice in music streaming platforms), could influence the user either positively or negatively. This study diverges from previous research which aims at using music to enhance memory, relieve boredom, improve concentration, promote prosociality, or aid learning [23]. The reason is that, despite the common consensus regarding beneficial uses of music, generalized misconceptions have surged historically in the perception of well-being applications of music (e.g., the Mozart effect [31], binaural beats [35], or 432 Hz music [39]). Only recent research has started to theorize and analyze music-induced harm [54, 43] – a topic that should be more widely discussed and studied by academia given the possibility that social networks, streaming platforms, and personalized advertisement companies are already studying it and potentially gaining profit from [34, 33, 55, 46].

2 Related work

2.1 Music and Emotions

Despite longstanding debates around the emotions that music can evoke, the appealing nature of human emotion studies has sparked what has been recently referred to as "the rise of affectivism" [9]: methodological and technical advancements of the emotion studies have surged to gain deeper understanding of behavioral and cognitive processes. With respect to music, *induced* emotions concern the arousal of psycho-physiological responses to a particular stimuli [29]. We refer the reader to [32, 28, 24, 26, 11, 10, 47] for extended theory and research on this topic. We use a discretized model of emotion based on Russell's circumplex model [40] and recent work on MER [37], which conceptualizes emotions in two dimensions (i.e., arousal and valence) and four distinct categories/quadrants of emotion: Q_1 (positive valence and arousal), Q_2 (positive arousal and negative valence), Q_3 (negative valence and arousal), Q_4 (negative arousal and positive valence). Arousal refers to energy or activation and valence relates to pleasantness or positiveness of an emotion: Q_1 refers to emotions such as joy, wonder, and power; Q_2 refers to emotions as tension, anger, and fear; Q_3 refers to emotions as sadness, bitterness; Q_4 refers to emotions as tenderness, peacefulness, and transcendence [53]. For this particular study, we focus on music that can induce Q_2-emotions for a particular user.

2.2 Use case: Polarization in Colombian music

The contextualization of the political landscape in Colombia escapes the scope of this paper, thus we refer the reader to [8, 52, 5, 44, 2, 13, 30] for deeper analysis regarding the history of violence in Colombia. As context, the "biblical holocaust" of Colombian violence – portrayed by the writer Gabriel García Márquez – has resulted in more than 420,000 violent deaths over the last 70 years, more than 11 million Colombians leaving the country or internally displaced, and one of the most unequal distribution of income in the continent [30]. Diverse sources of inequalities (e.g., agrarian capitalism, socioeconomic exclusion, decolonization processes, the war on drugs, illegal economies, and exploitation of natural resources) are the cause of the formation of illegal armies fueled by political ideologies [17]: "left-wing" Fuerzas Armadas Revolucionarias de Colombia (FARC) and "right-wing" Autodefensas Unidas de Colombia (AUC), amongst several other illegal groups. As an oversimplification of Colombia's historical process (and reflecting the generalized trend in the world), polarization arose over whether and how to pursue peace in the country, producing negative relationships between political discourses and everyday life [14]. Polarization results in fragile societies, cooperation within citizens becomes complicated, and collective action for problem solving becomes impossible [46]. We argue that the music selection, described as follows, is *not-so-popular* Colombian music since the political content of the lyrics may arouse arguments from different political views.

3 Methodology

3.1 Music selection

We refer the reader to studies by Quishpe [38], Barbosa Caro and Suavita [3], and Katz-Rosene [27] with respect to historical, functional, and lyrical analysis from the two types of music used: (1) FARC-songs (mainly in the style of *vallenato* and *canción social*) and (2) AUC-songs (in the style of *corridos*). These musical styles make part of traditional Colombian (and Latin-american) music, yet they have distinctive sonorities, structures, and instrumentation. It must be noted that music with politically motivated lyrics from both types have used similar styles of music as well (e.g., hip-hop and rock), but this study only considers this reduced range of styles. Additionally, FARC-songs have been typically created by active members from the guerrilla as a mechanism of identity confirmation and propaganda [38], while AUC-songs have been typically produced by sympathizers of the paramilitaries as promotion to their deeds and open crtiticism to the FARC [3] – the functionality of the music and the target listener are different.

We remark that humans frequently listen to music *without* feeling any emotion at all [28, 26], but we assume that music *might* trigger episodic memories to particular individuals [25, 10]. The potential induction of emotions from this music is due mainly to the semantic content of the lyrics (see 4.2) – inducing different emotions to listeners with different political views. Despite the importance of lyrics to the induced emotions, we argue that the acoustic features are useful to provide a content-based contrast among the different styles of music: (1) FARC-songs typically use less instruments and might include only voice and guitar, and (2) AUC-songs are more heavily orchestrated with faster tempo. We use 50 music excerpts with lyrics from each music type (30 seconds long) and extracted 260 emotionally relevant acoustic features (mean and standard deviation of 65 low-level music descriptors and their first order derivatives) from segments of 1 second [1], with 50% overlap, and standardize across features – using the IS13 ComParE feature set [48] and OpenSMILE toolbox [12]. Namely, the machine learning models should be able to differentiate between the types of music – the interesting element is to attempt to understand which users will provide problematizing labels (i.e., music that induces subjective emotions of anger or fear) that can bias the algorithm towards a particular class.

3.2 Classification strategy

We use the "machine consensus" MER personalization strategy presented by Gómez-Cañón et al. [20]: consensus entropy for active learning. This strategy uses a committee of classifiers to analyze their output agreement and queries each user for instances with the highest uncertainty. A committee of classifiers (15 independent Extreme Gradient Boosting models [6]) has previously been pre-trained on separate cross-validation splits of the DEAM dataset, the benchmark dataset for MER [1]. In order to select uncertain data to be labeled, classifiers predict the output probabilities for the pool of excerpts. We then perform the consensus entropy strategy by analyzing the disagreement across classifiers. For example, full disagreement from a committee of four classifiers results when each one predicts a different class/quadrant with 100% probability. This yields average probabilities per quadrant $p_{avg} = \{Q_1 : 0.25, Q_2 : 0.25, Q_3 : 0.25, Q_4 : 0.25\}$ and high inter-class entropy/uncertainty of 1.386. Excerpts with highest uncertainty are then queried to the oracle (i.e., each user) to be annotated. Initially, we randomly draw 5 excerpts from each type of music (10 excerpts for the first annotation iteration), retrain our classifiers with the annotations provided by each user, identify the excerpts to be annotated for the next iteration, and present the new batch of music to be annotated. Given the low amount of available music, we perform only three iterations for a total of 30 annotations per user – past research has shown that only 20-30 annotations are needed in order to reach personalization [45, 7]. Please refer to [20] for additional information of the methodology.

4 Preliminary results and Discussion

We test our initial models with three users who have reported their political opinions: one with "left-wing", one with "right-wing", and one with "center" political views [14]. We ask our users to report their possible vote in a forced choice question – if the elections would happen now for whom would they vote? They were able to choose between the "left-wing" candidate from the Colombia Humana party (Gustavo Petro), a "right-wing" candidate from the Centro Democrático government party (there is no clear candidate yet), or cast a blank vote.

4.1 Q_2 music prediction

Given that this study is not centered on improving the accuracy of models but rather understanding how the models are progressively biased towards a particular category, we test the finalized models on the music which has not been annotated by the user. Namely, each user annotates 30 excerpts (3 iterations of 10 annotations each) and we test the personalized model on the remaining testing data (70 excerpts), which are different for each user. In particular, our interest is to study if a personalized model trained for a user with a particular political view, can effectively identify new music that can potentially induce negative emotions (i.e., music belonging to Q_2). We obtain output probabilities for the testing data, sort the highest probabilities for Q_2, and count the amount of songs that belong to each political view from the top 10 predictions. Additionally, we report the mean output probability p_{avg} of each type of music belonging to Q_2 (see Table 1).

Table 1: Preliminary results of personalized models.

	Top 20		Output p_{avg}	
User type	FARC-songs	AUC-songs	FARC-songs	AUC-songs
“Left-wing”	10%	90%	93.37%	71.02%
“Center”	30%	70%	90.57%	34.49%
“Right-wing”	70%	30%	92.60%	72.76%

The initial remarks are that, for this reduced number of users, the personalized models for users with distinctive political views (i.e., “right-wing” and “left-wing”) appear to capture that music from the opposite political perspective will have more likelihood to produce Q_2 emotions. While the amount of participants is not indicative of generalizable trends, we argue that these personalized models could indeed be “breached” with problematizing subjective data – testing data was completely unseen by the users, relying exclusively on acoustic features which are not the reason for inducing emotions. In fact, further analysis with p_{avg} reveals that music of both types is very likely to produce Q_2-emotions – in fact, the assumption that the music selection is *not-so-popular* might hold. As a limitation, we identify that this methodology can be randomly determined by the selection of songs for the initial iteration (see 4.2).

4.2 Future work and broader impact

Our final interest is to understand whether the use of this music can potentially have an impact on decision-making – studies regarding the impact of musically-induced emotions on decision making have been studied since 2006, as reviewed by Palazzi et al. [36]. It has been argued that music and persuasion have indirect relationships and are never strictly causal – as the persuasion/manipulation of a person is already a difficult task, it can only be “helped” or “promoted” by music. As mentioned by Herrera [22], music can only contribute as a “persuading factor” to an induced emotional state, which can be associated between music and a particular person or message, contributing to re-evaluating attitudes and actions. In fact, the author stresses the fact that music containing a message within the lyrics can “make the complete brain work in a coordinated manner”: text and lyrics will activate more the left hemisphere, while music will activate the right one. Given that the main emotion-inducing mechanism is the semantic content of the lyrics for this study, we plan to add multi-modality to this approach by using sentiment analysis models on the lyrics. In principle, sentiment analysis models could also be re-trained with new annotations that could bias the ensemble model to identify this problematic data. Furthermore, the variability of the initial annotations for personalization and the impact of the initial iteration on concept drift [49], must still be evaluated and analyzed.

In summary, it is very likely that any type of stimuli that produces strong political responses can be somewhat captured by the personal annotations used as input to a machine learning model. It has been argued that music is a powerful and engaging stimulus that can promote prosociality [41], impact customer behaviors [21], and influence processes of decision-making and risk-aversion [15, 18]. We want to understand how the use of growing emotion recognition technologies can have a direct impact on ethical issues, mainly societal well-being – human-centered technologies must be evaluated for both beneficial and harmful use cases.

Acknowledgments and Disclosure of Funding

The research work conducted at the Universitat Pompeu Fabra is partially supported by the European Commission under the TROMPA project (H2020 770376) and the Project Musical AI - PID2019-111403GB-I00/AEI/10.13039/501100011033 funded by the Spanish Ministerio de Ciencia, Innovación y Universidades (MCIU) and the Agencia Estatal de Investigación (AEI).

References

[1] A. Aljanaki, Y.-H. Yang, and M. Soleymani. Developing a benchmark for emotional analysis of music. *PLoS One*, pages 1–22, 2017. doi: 10.1371/journal.pone.0173392.

[2] J. Arocha R. et al. *Colombia: Violencia y Democracia*. Universidad Nacional de Colombia, 1988.

[3] E. Barbosa Caro and J. R. Suavita. Paramilitarism and music in Colombia. *Journal of Language and Politics*, 18(4):541–559, 2019. doi: 10.1075/jlp.19019.bar.

[4] M. Barthet, G. Fazekas, and M. Sandler. Music emotion recognition: From content- to context-based models. In *From Sounds to Music and Emotions*, pages 228–252, Berlin, Heidelberg, 2013. Springer.

[5] C. W. Berquist. *Coffee and Conflict in Colombia, 1886-1910*. Duke University Press, 1978.

[6] T. Chen and C. Guestrin. XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, Aug 2016.

[7] Y.-A. Chen, J.-C. Wang, Y.-H. Yang, and H. H. Chen. Component tying for mixture model adaptation in personalization of music emotion recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25(7):1409–1420, 2017.

[8] N. Chomsky. On colombia. In D. Stokes, editor, *Americas's Other War: Terrorizing Colombia*, chapter On Colombia. Zed Books, 2004.

[9] D. Dukes, K. Abrams, R. Adolphs, et al. The rise of affectivism. *Nature Human Behaviour*, 5(7):816–820, Jul 2021. doi: 10.1038/s41562-021-01130-8.

[10] T. Eerola. Music and emotion. In R. Bader and S. Koelsch, editors, *Handbook of Systematic Musicology*, chapter Music and Emotion, pages 539–556. Springer, 2018.

[11] T. Eerola and J. K. Vuoskoski. A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music Perception: An Interdisciplinary Journal*, 30(3):307–340, 2013.

[12] F. Eyben, F. Weninger, F. Gross, and B. Schuller. Recent Developments in OpenSMILE, the Munich Open-source Multimedia Feature Extractor. In *Proceedings of the 21st ACM International Conference on Multimedia*, pages 835–838, New York, NY, USA, 2013.

[13] O. Fals Borda et al. *Movimientos Sociales, Estado y Democracia en Colombia*. Universidad Nacional de Colombia, 2001.

[14] A. E. Feldmann. *Colombia's Polarizing Peace Efforts*, pages 153–176. The Global Challenge of Political Polarization. Brookings Institution Press, 2019. doi: 10.7864/j.ctvbd8j2p.9.

[15] P. Fischer and T. Greitemeyer. Music and aggression: The impact of sexual-aggressive song lyrics on aggression-related thoughts, emotions, and behavior toward the same and the opposite sex. *Personality and Social Psychology Bulletin*, 32(9):1165–1176, 2006. doi: 10.1177/0146167206288670.

[16] J. S. Gómez-Cañón, E. Cano, P. Herrera, and E. Gómez. Joyful for you and tender for us: the influence of individual characteristics and language on emotion labeling and classification. In *Proceedings of the 21st International Society for Music Information Retrieval Conference*, pages 853–860, Montréal, Canada, 2020.

[17] J. Grajales. *Agrarian Capitalism, War and peace in Colombia*. Routledge, 2021.

[18] T. Greitemeyer. Exposure to music with prosocial lyrics reduces aggression: First evidence and test of the underlying mechanism. *Journal of Experimental Social Psychology*, 47(1):28–36, 2011. doi: 10.1016/j.jesp.2010.08.005.

[19] J. S. Gómez-Cañón, E. Cano, T. Eerola, P. Herrera, X. Hu, Y.-H. Yang, and E. Gómez. Music Emotion Recognition: towards new robust standards in personalized and context-sensitive applications. *IEEE Signal Processing Magazine*, 38, 2021.

[20] J. S. Gómez-Cañón, E. Cano, Y.-H. Yang, P. Herrera, and E. Gómez. Let's agree to disagree: Consensus entropy active learning for personalized music emotion recognition. In *Proceedings of the 22nd International Society for Music Information Retrieval Conference (ISMIR)*, 2021.

[21] F. Hansen and S. Christensen. *Emotions, Advertising and Consumer Choice*. Copenhagen Business School Press, 2007.

[22] P. Herrera. Música y persuasión. In J. Ayats et al., editors, *La música y su reflejo en la sociedad*, chapter Música y persuasión, pages 27–38. Indigestió Musical, 2009.

[23] X. Hu, J. Chen, and Y. Wang. University students' use of music for learning and well-being: A qualitative study and design implications. *Information Processing and Management*, 58(1):1–14, 2021.

[24] P. N. Juslin. *Handbook of Music and Emotion: Theory, Research, Applications*. Oxford University Press, Oxford, 2010. ISBN 9780199230143.

[25] P. N. Juslin. From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, 10:235–266, 2013. doi: 10.1016/j.plrev.2013.05.008.

[26] P. N. Juslin. *Musical Emotions Explained*. Oxford University Press, Oxford, 2019.

[27] J. Katz-Rosene. *From Protest Song to Social Song: Music and Politics in Colombia 1966-2016*. PhD thesis, City University of New York, 2017.

[28] P. Kivy. *Music alone: reflections on a purely musical experience*. Cornell University Press, 1990.

[29] C. L. Krumhansl. An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 51(4):336, 1997.

[30] R. D. Mahoney. *Colombia: what everybody needs to know*. Oxford University Press, 2020.

[31] S. A. Mehr, A. Schachner, R. C. Katz, and E. S. Spelke. Two randomized trials provide no consistent evidence for nonmusical cognitive benefits of brief preschool music enrichment. *PLOS ONE*, 8(12):1–12, 2013.

[32] L. B. Meyer. *Emotion and meaning in music*. Chicago University Press, 1956.

[33] S. U. Noble. *Algorithms of oppression: How search engines reinforce racism*. New York University Press, 2018.

[34] C. O'Neil. *Weapons of Math Destruction*. Crown, 2016.

[35] H. D. Orozco Perez, G. Dumas, and A. Lehmann. Binaural beats through the auditory pathway: From brainstem to connectivity patterns. *eNeuro*, 7(2), 2020.

[36] A. Palazzi, B. Wagner Fritzen, and G. Gauer. Music-induced emotion effects on decision-making. *Psychology of Music*, 47(5):621–643, 2019. doi: 10.1177/0305735618779224.

[37] R. Panda, R. M. Rui, and P. Paiva. Musical texture and expressivity features for music emotion recognition. In *Proceedings of the 19th International Society for Music Information Retrieval Conference*, Paris, France, 2018.

[38] R. Quishpe. Corcheas insurgentes: usos y funciones de la música de las FARC-EP durante el conflicto armado en Colombia. *Izquierdas*, 49:554–579, 2020. doi: 10.4067/s0718-50492020000100231.

[39] R. E. Rosenberg. Perfect pitch: 432 hz music and the promise of frequency. *Journal of Popular Music Studies*, 33(1):137–154, 03 2021. doi: 10.1525/jpms.2021.33.1.137.

[40] J. A. Russell. A circumplex model of affect. *Personality and Social Psychology*, 39(6):1161–1178, 1980.

[41] N. Ruth. "They don't really care...": Effects of music with prosocial content and corresponding media coverage on prosocial behavior. *Musicae Scientiae*, 22(3):415–433, 2018.

[42] M. Schedl, A. Flexer, and J. Urbano. The neglected user in music information retrieval research. *Journal of Intelligent Information Systems*, 41:523–539, 2013.

[43] M. J. Silverman, L. F. Gooding, and O. Yinger. It's...complicated: A theoretical model of music-induced harm. *Journal of Music Therapy*, 57(3):251–281, 2020. doi: 10.1093/jmt/thaa008.

[44] D. Stokes. *America's other war: Terrorizing Colombia*. Zed Books, 2005.

[45] D. Su and P. Fung. Personalized music emotion classification via active learning. In *Proceedings of the Second International ACM Workshop on Music Information Retrieval with User-Centered and Multimodal Strategies*, page 57–62, New York, NY, USA, 2012.

[46] C. Véliz. *Privacy Is Power: Why and How You Should Take Back Control of Your Data*. Bantam Press, 2020.

[47] L. A. Warrenburg. Comparing musical and psychological emotion theories. *Psychomusicology: Music, Mind, and Brain*, 30(1):1–19, 2020.

[48] F. Weninger, F. Eyben, B. W. Schuller, M. Mortillaro, K. R. Scherer, and J. Krajewski. On the acoustics of emotion in audio: what speech, music, and sound have in common. *Frontiers in Psychology*, 4:1–12, 2013. doi: 10.3389/fpsyg.2013.00292.

[49] G. Widmer and M. Kubat. Learning in the Presence of Concept Drift and Hidden Contexts. *Machine Learning*, pages 69–101, 1996. doi: 10.1023/A:1018046501280.

[50] Y.-H. Yang and H. H. Chen. *Music Emotion Recognition*. CRC Press, 2011.

[51] Y.-H. Yang, Y.-F. Su, Y.-C. Lin, and H. H. Chen. Music Emotion Recognition: The Role of Individuality. In *Proceedings of the International Workshop on Human-centered Multimedia*, pages 13–22, 2007.

[52] L. Zamosc. *The agrarian question and the peasant movement in Colombia*. Cambridge University Press, 1986.

[53] M. Zentner, D. Grandjean, and K. R. Scherer. Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, 8(4):494–521, 2008.

[54] N. Ziv. Music and compliance: Can good music make us do bad things? *Psychology of Music*, 44(5): 953–966, 2016. doi: 10.1177/0305735615598855.

[55] S. Zuboff. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Public Affairs, 2019.