



HOW CAN DEEP LEARNING MODELS IDENTIFY AND COMPARE MELODIC SIMILARITIES BETWEEN SONGS TO DETECT POTENTIAL COMPOSITIONAL COPYRIGHT INFRINGEMENT?

Arnav Karthikeyan

Research Scholars Program, Harvard Student Agencies, In collaboration with Learn with Leaders

ABSTRACT

The music industry and its evolution are closely tied to the growth of technology, including both the creation and distribution of music in the 21st century. However, such technology has yet to impact the protection of music IP. Judicial courts handle the cases of compositional copyright infringement in a subjective manner—taking in the opinions of many factfinders as if the two songs “sound” similar with regard to explicit tune, style, and authorship, to arrive at a final verdict. Thus, this research explores the extent to which deep-learning models can bring forth objectivity in the proceedings of court hearings to provide a holistic view of the entire case, therefore, increasing the certainty of a court’s verdict to be justifiable.

KEYWORDS: Deep-Learning Models, Intellectual Property, Copyright Infringement, Algorithms

INTRODUCTION

Intellectual Property or IP, refers to creations made by humans that are under legal protection. This is done to ensure that the appropriate rights holders can control the distribution of IP-protected material and the profit generated from it. There are several different types of IP protection in the music industry such as copyright, providing rights holders exclusive privilege in reproduction, distribution, performance, and display of material; trademark to protect names, brands, and logos associated with musicians and record companies; and licensing to regulate usage of copyrighted material to other parties. These laws exist to ensure the protection and monetization of musicians’ works across various aspects of the industry (Yellowbrick, 2020).

The focus of this paper will be the discussion of copyright protection laws on IPs. However, there are two types of copyright for any given song; compositional copyright and master recording. Compositional copyright refers to legal protection of the musical structures of a song such as a tune, chord progressions, harmonies, and lyrics. Master recording copyright is the legal protection of the single song recording and the sounds used in that recording. For the purpose of this research paper, the copyright in question will be compositional copyright as most prolific and publicly debated cases stem from disputes over compositional aspects of music. The most recent of such high-profile cases is the dispute of Ed Sheeran’s “Thinking Out Loud” (2014) infringing Marvin Gaye’s “Let’s Get It On” (1973)—the case starting on September 29, 2022 and ending on May 16, 2023. The case utilized a methodology very subjective in nature—raising the potential for cultural and personal bias highlighted by factfinders’ experiences.

Thus, introducing machine learning models can be used to detect similarities between songs in order to form a more objective approach when handling music copyright cases. Such

models when used in practice have the potential to bring forth better objectivity to the case proceedings. Eliminating potential bias and subjective nature of traditional methodology by using highly extensive data sets and problem-solving abilities that transcend humans. This technology has a large impact on court verdicts. While no model as such exists at this point, this research paper will look to discuss the extent to which these models’ potential utility refines current methodologies used in practice

LITERATURE REVIEW

To first understand the overview of the subject, the source “Understanding Intellectual Property in Music: A Comprehensive Overview.” allows readers to conceptualize several IP protection laws that exist to protect the IPs of musicians and all parties in the production of a song or album (Yellowbrick, 2020).

The source, “Copyright Infringement of Music: Determining Whether What Sounds Alike Is Alike”, specifically delves into the laws of copyright. It questions the current copyright laws that are in place for the recordings of music and explores the impact of these laws on the incentive to create, protect rights, and dismantle monopolies of our shared heritage. After the discussion of the above aspects, the article concludes that policies should enforce justice on individuals whose music bears “striking similarity” to another. Thus, preserving the care-free creation of music and protecting individuals’ IP (Livingston, 2013). However, the sources’ proposed methodology is prone to subjectivity from factfinders in court.

This is where the implications of the paper “Predictive Models for Music” which itself explores the mechanisms by which melodic structures and tonalities are able to be constructed by utilizing machine learning models, proposes a refined methodology. Breaking down the task into two parts;

one to analyze the rhythms of a song and the other, being a generative model, analyzes chord progressions and sequences. Thus, the implications of this paper are that a similar model can theoretically be constructed to be used to analyze the similarities between two songs and by bearing a close enough similarity within rhythms and chords, copyright laws can be enforced accordingly (Paiement, 2009).

A model that would have been useful in the dispute of compositional rights of “Thinking Out Loud” by Ed Sheeran violating “Let’s Get It On” by Marvin Gaye. The source had summarized the key events and proceedings of the entire case, including the methodology they had used in the case which involved factfinders to piece together unprotected elements of the plaintiff’s song and find similarities in authorship. Hence, being highly subjective in nature (Gosling, 2023).

This subjective nature may also stem from an individual’s own experiences in terms of culture and understanding, and may alter the emotional or aesthetic perception of the song to factfinders. This results in cultural and perception bias, as given by the source “The Dock-in Model of Music Culture and Cross-Cultural Perception” (Fritz, 2013).

To bring forth objectivity, innovations in deep-learning models have been made such as the continuously updating model of CRNN (Convolutional Recurrent Neural Network), a hybrid of CNN (Convolutional Neural Network) and RNN (Recurrent Neural Network), an AI model that has spatial and sequential pattern recognition capabilities. In the context of this research, such models will allow the analysis of sheet music and audio drafts that musicians had during the writing process as feeding data. Thus, when comparing the two pieces of music, similarities can be highlighted (Choi et al, 2017).

Other technology that could be used to achieve the initial purpose can be seen in the source “Automated Music Recommendations Using Similarity Learning.”. This source explored a similar methodology used by Paiement—breaking down music into melody and rhythm using machine learning models to analyze both accordingly. However, the source places a greater emphasis on analyzing the similarities between the two songs’ melodic components such as harmony, patterns, and the tempo in which it is played. While the intended use of this technology was to provide users with better recommendations on streaming platforms, an extension of this technology can be used to detect similarities in songs in the context of compositional copyright disputes (Burns, 2020).

We can also consider technology that has the potential to be used in master recording disputes, such as “Musical Genre Classification of Audio Signals.”. Unlike previous sources, the methodology to highlight similarities in music is done through audio signals. Doing so will not only provide factfinders with genre classification but also realize the frequency structure of melodic components (Tzanetakis, 2002).

However, mathematical models used in these sources were used by the court of law before. Evidence of their existence

can be seen in the source “Court decisions on music plagiarism and the predictive value of similarity algorithms,” where melody similarity algorithms have been used in disputes over compositional and master recording copyrights including *Bright Tunes v Harrison* (1976) and *Selle v Gibb* (1984). The interpretations of the algorithms’ findings brought human subjectivity seen in other cases previously mentioned (Müllensiefen, 2009). Thus, highlighting the importance of deep learning models to eliminate potential bias

Despite the perceived benefits of this technology—including objectivity and depth of analysis—it has several limitations in terms of how these models are built. As they have potential for dataset bias and ethical implications to arise—the idea for a machine to influence the verdict of a court.

ANALYSIS

The findings of the literature review suggest that deep-learning models have high potential utility. Evidence to support this claim is seen in these models’ capabilities to run extensive calculations in command of their problem-solving capabilities. For example, CRNN has shown such heightened capabilities as a result of being a hybrid of two separate neural networks. Thus, establishing these models’ innate problem-solving abilities (Choi et al, 2017).

Furthermore, this ability can be stemmed from deep-learning models to process high volumes of data unparalleled to what could be accomplished by human factfinders. It is because of these two distinct capabilities that the application of this technology is likely to be high in the court of law. This can be seen in the need for such abilities in the dispute between Ed Sheeran’s “Thinking Out Loud” and “Let’s Get It On”, by Marvin Gaye, where the defendant’s main dispute lies in similarities in chord progression and harmonies. This case in particular required the plaintiff, Gaye, to show that there are several properties of unprotected parts of a song that are apparent in both. Thus, compositional copyright infringement laws protect not only the musician’s art but their authorship as well. Requiring models to piece together several data sets that attempt several combinations and permutations in order to bring forth an objective and refined verdict (Gosling, 2023).

Moreover, as there are no thresholds in the number of appearances of unprotected content on both songs, an extension of this technology can be set and agreed upon by a judge or jury with two parameters: genre and frequency of usage. Genre refers to the “aesthetic” of the song, while frequency range is more quantitative in nature, eliminating potential bias that may be seen if a human were delivering judgment (Gosling, 2017). This phenomenon was discussed as part of the literature review—the upbringing of a person affects their perceptual understanding of what music is and what kind of music depicts a particular emotion (Fritz, 2013). Thus, by automating this process using deep-learning models, thresholds can be set dynamically to the nature of each case, showing potential for high utility.

However, the limitations of these learning models are realized

when these models are unique in nature—each model has a unique set of training data. Thus, certain models that may have discrepancies or biased training data will result in poor and potentially biased verdicts being given. Moreover, machine learning models that bridge into AI such as CRNN are not standardized in behavior and output as they behave in accordance to cognitive neural networks. While this alone may not degrade the quality of verdicts—referring to minimal bias—lack of standard and consistent responses may lead to adverse effects on the overall outcome of the case, thereby limiting the utility of these models in court (Choi et al, 2017).

Despite the apparent limitations when applying such models, it is likely the development of systems used in the court of law are ones developed by the country in which the trial is occurring. Therefore, the only way the counterclaim stands is if the country is biased towards either the plaintiff or defendant—instances such as those are rare. Even so, it is important to assume all courts and their respective governments who may utilize such learning models to detect potential compositional copyright infringement are principled. This is done to showcase the extent to which the applications of deep-learning models achieve the highest utility.

CONCLUSION

In conclusion, based on the findings of both the literature review and the analysis of its findings, the extent to which the learning models will be used in court will likely be hybrid in nature. Utilizing the subjective intuition of human factfinders and the objective nature of deep-learning models will provide a balanced set of evidence and verdicts to deliver a comprehensive final judgment.

This particular conclusion is in tandem with the level of technology and infrastructure available in most developed or developing countries. To effectively put models such as these into practice, standardization of source code and construction methods of these programs must be open source to all interested governments. In addition, the model must have information regarding copyright law in the country it is being used to ensure universal application of the same program. Moreover, the current deep learning models that are able to detect similarities in music are being used in streaming platform services and musicians to make music, but also extend its utility in the court of law by achieving the objective first set out, thereby creating positive effects on artists' creative freedom and liberty. However, it would require high capital investment to make such advancements in technology with the objective of reaching full autonomy. A key limitation of this study is that the development of these models is highly specific to detect similarities in copyright infringement rather than the other protection laws as discussed in the introduction.

REFERENCES

1. Burns, Jamie, and Terence L. van Zyl. "Automated Music Recommendations Using Similarity Learning." *SACAIR 2020* (2020): 288. Accessed 7 Sep. 2024.
2. Choi, Keunwoo, et al. "Convolutional Recurrent Neural Networks for Music Classification." 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),

- 2017, pp. 2392–96. IEEE Xplore, <https://doi.org/10.1109/ICASSP.2017.7952585>. Accessed 7 Sep. 2024.
3. Fosler-Lussier, Danielle. "Copyright, Surveillance, and the Ownership of Music." *Music on the Move*, University of Michigan Press, 2020, pp. 180–201. JSTOR, <https://www.jstor.org/stable/10.3998/mpub.9853855.16>. Accessed 5 Sep. 2024.
4. Gosling, Jake. Edward Christopher Sheeran, p/k/a Sheeran, Sony/ATV Music Publishing, LLC, Atlantic Recording Corporation d/b/a Atlantic Records, BDI Music Ltd., Bucks Music Group Ltd., The Royalty Network, Inc., David Platz Music (USA) Inc., Amy Wadge. Accessed 1 Sep. 2024.
5. Livingston, Margit, and Joseph Urbinato. *Copyright Infringement of Music: Determining Whether What Sounds Alike Is Alike*. Winter 2013. Accessed 3 Aug. 2024.
6. Müllensiefen, Daniel, and Marc Pendzich. "Court Decisions on Music Plagiarism and the Predictive Value of Similarity Algorithms." *Musicae Scientiae*, vol. Discussion Forum 4B, 2009, pp. 257–95. APA PsycNet, <https://doi.org/10.1177/102986490901300111>. Accessed 7 Sep. 2024.
7. Paiement, Jean-François, et al. "Predictive Models for Music." *Connection Science*, vol. 21, no. 2–3, Sept. 2009, pp. 253–72. tandfonline.com (Atypon), <https://doi.org/10.1080/09540090902733806>. Accessed 7 Sep. 2024.
8. Tzanetakis, G., and P. Cook. "Musical Genre Classification of Audio Signals." *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 5, July 2002, pp. 293–302. IEEE Xplore, <https://doi.org/10.1109/TSA.2002.800560>. Accessed 7 Aug. 2024.
9. Yellowbrick. "Understanding Intellectual Property in Music: A Comprehensive Overview." Yellowbrick, 27 Aug. 2023, [https://www.yellowbrick.co/blog/entertainment/understanding-intellectual-property-in-music-a-comprehensive-overview#:~:text=Intellectual%20property%20\(IP\)%20in%20music%20refers%20to%20the%20legal%20protection,%2C%20perform%2C%20and%20display%20works](https://www.yellowbrick.co/blog/entertainment/understanding-intellectual-property-in-music-a-comprehensive-overview#:~:text=Intellectual%20property%20(IP)%20in%20music%20refers%20to%20the%20legal%20protection,%2C%20perform%2C%20and%20display%20works). Accessed 6 Sep. 2024.