Exploring Case-Based Reasoning for Multivariate Time Series Classification - Application to Music Emotion Recognition

Erwan Diebold[1,2,*,†]

[1]Université de lorraine, CNRS, LORIA, F-54000, France

Abstract

This PhD project investigates the use of Case-Based Reasoning (CBR) for Music Emotion Recognition (MER) from physiological signals. Music is known to induce emotions, and the objective is to determine whether the planned emotion is effectively felt by the listener. Classical approaches generally rely on extracting features from physiological signals, such as heart activity, electrodermal activity, body temperature, or pupil dilation, before applying machine learning classification methods. In contrast, this work proposes to reason directly on raw physiological signals in order to avoid information loss and reduce assumptions introduced by feature extraction.

To compare physiological signals, this work relies on the Time Warp Edit Distance (TWED), an elastic distance measure adapted to temporal deformations in time series. The research investigates several CBR approaches, approximation (similarity-based reasoning) and interpolation (betweenness-based reasoning). The proposed methods aim to extend these CBR approaches to complex continuous-valued temporal data structures. Preliminary results show that these CBR approaches are promising for multivariate time series classification and achieve competitive results for emotion recognition tasks.

Keywords

Case-based reasoning, Physiological Data, Music Emotion Recognition

1. Introduction

Emotions influence behavior and decision-making, and understanding them can be useful in many domains, especially in healthcare. Music is known to induce emotions [1], making it a promising tool for applications such as stress reduction [2]. However, to ensure that the desired emotion has actually been induced, it is necessary to design a system capable of recognizing the emotion felt by a person.

Music emotion recognition is commonly performed using physiological signals such as heart rate, electrodermal activity, or body temperature. Classical approaches to this task rely on the extraction of features from these signals and the application of machine learning methods for classification. For example, an approach using Case-Based Reasoning (CBR) has been proposed for stress detection [3] using extracted features from physiological signals. These methods depend on the choice of extracted features and on feature selection processes, which rely on underlying assumptions.

The objective of this work is to investigate the use of CBR directly on raw physiological signals to avoid information loss due to feature extraction, and to go beyond these assumptions. While learned representations offer an alternative to handcrafted features, they also make the reasoning process less interpretable. By reasoning directly on raw physiological signals, this work aims to preserve both the original information and the explainability of the proposed CBR approach. To achieve this, 3 CBR approaches, introduced in [4], are studied here: approximation (based on similarity), interpolation (based on the notion of betweenness) and extrapolation (based on analogy). These approaches have already been explored on simple data structures [5], on discrete sequences [6]. However, these approaches have not yet been extended to continuous multivariate time series. Due to this limitation, they cannot be used in fields like music emotion recognition and physiological signal analysis. This research aims to

ICCBR DC'26: Doctoral Consortium at ICCBR2026, August 16, 2026, Bremen, Germany

*Corresponding author.

erwan.diebold@univ-lorraine.fr (E. Diebold)

0009-0007-9973-6710 (E. Diebold)

extend CBR methods to the domain of physiological time series analysis. In particular, it investigates how similarity, betweenness, and analogy can be defined for time-series, and how these reasoning mechanisms can be efficiently applied to emotion recognition tasks. By working directly on raw signals, this work explores if CBR can offer an interpretable alternative for Music Emotion Recognition.

2. Research Plan

2.1. Research Objectives

The main objective of this research is to design a CBR system capable of recognizing emotions induced by music from physiological signals. The considered signals are heart activity, electrodermal activity, body temperature, or pupil dilation. The goal is to predict the emotion felt by a listener while listening to music, based only on the observed physiological responses. This work aims to extend CBR approaches approximation, interpolation, and extrapolation to multivariate time series with continuous values.

2.2. Approach / Methodology

The initial objective of this work is to predict the annotations of emotions in long recordings of physiological signals. However, directly modeling emotions as temporal evolutions during music listening leads to a complex reasoning problem involving both continuous signal variations and sequences of emotional states over time. To simplify the problem for this first stage of the PhD, the task has been reduced to a classification problem associating shorter physiological signals with a single emotion label.

In the dataset of music emotion recognition [7], emotions were represented as temporal annotations collected during music listening sessions. In this work, only the time intervals during which a specific emotion was reported are kept. For each interval, the corresponding physiological signals are extracted and associated with the reported emotion. Therefore, one case corresponds to physiological signals linked to a single emotion.

In this work, a case is represented by a problem part and a solution part. The problem is defined as a set of physiological signals, where each attribute corresponds to a time series. The solution corresponds to an emotion label. Several physiological signals are considered, including blood volume pulse (BVP), electrodermal activity (EDA), temperature (TMP), and pupil size (PupSize). Let x be a problem and a be an attribute of x; $x \cdot a$ is the value of a for x. So for the music emotion recognition task, a problem x can be defined by the following attributes:

- $x \cdot \texttt{BVP} = S^{\texttt{BVP}}$;
- $x \cdot \texttt{EDA} = S^{\texttt{EDA}}$;
- $x \cdot \texttt{TMP} = S^{\texttt{TMP}}$;
- $x \cdot \texttt{PupSize} = S^{\texttt{PupSize}}$;

where $S^{\texttt{BVP}}$, $S^{\texttt{EDA}}$, $S^{\texttt{TMP}}$, $S^{\texttt{PupSize}}$ time series represent the physiological signals presented above.

To compare two problems, a distance function between problems is defined as a weighted sum of distances between their corresponding attributes:

$$\texttt{dist}_{\mathscr{A}}(x^s, x^{tgt}) = \sum_{a:\texttt{attribute}} w_a \times \texttt{dist}_a(x^s \cdot a, x^{tgt} \cdot a)$$

.

Since each attribute is represented by a time series, the problem reduces to defining distances between time series. Among the distance functions tested, the Time Warp Edit Distance (TWED) [8] was selected for three reasons. First, like DTW [9], it is an elastic distance capable of handling temporal deformations. Second, unlike DTW, TWED is a true metric and satisfies the triangle inequality [8]. This property is particularly important because interpolation relies on betweenness relation and such a relation needs to use a distance function that fulfills triangle inegality [4]. Finally, preliminary experiments confirmed its effectiveness compared with other distances.

An important limitation of TWED is its computational cost. To address this issue, GPU-based parallelization techniques inspired by *cuTWED* [10] are investigated. In addition, a batched extension named *McuTWED* has been developed to compute several TWED distances simultaneously on the GPU. The main idea is to organize multiple pairs of time series into tensors and compute several dynamic programming matrices in parallel. This significantly reduces execution time for large-scale classification tasks.

Four CBR systems are currently explored. CBRΣA corresponds to a classical similarity-based CBR approach, where the solution of the most similar source case is reused. CBRΣI extends this idea by retrieving two source cases surrounding the target problem and predicting a solution through interpolation between their associated solutions. CBRVA applies the approximation process independently to each attribute (physiological signal), then combines the predicted solutions through a majority vote. This attribute-wise voting strategy has previously been used for CBR classification of physiological signals in a stress detection task [3], motivating its evaluation in the context of music emotion recognition.

Finally, CBRVI combines attribute-wise reasoning with interpolation, by performing interpolation independently on each signal before applying a voting process to obtain the final prediction.

The current evaluation focuses on emotion recognition tasks using physiological signals. The objective is first to assess the relevance of raw-signal CBR approaches compared with more classical feature-based methods. In a second stage, the goal will be to study how these approaches generalize to other types of continuous multivariate time series classification problem and to extend interpolation and analogy-based CBR mechanisms to this setting.

3. Progress Summary

So far, the four proposed CBR systems have been implemented in Python. A comparison with Random Forest classifiers, representing more classical machine learning approaches, has also been conducted.

In the Music Emotion Recognition task, the proposed CBR approaches obtain slightly better results than the Random Forest baseline. Experiments were also conducted on another multivariate time series classification problem that was not related to emotions. For this second task, the obtained performances are currently lower than expected. At this stage, one of the main difficulties with this classification problem concerns parameter selection. Especially for the TWED distance parameters and for the attribute weights used in the distance between problems. This issue becomes particularly challenging when dealing with high-dimensional problems containing a large number of signals.

4. Conclusion and Future Work

The current results suggest that applying CBR to multivariate time series classification problems is promising. At this stage, approximation approach obtain better performances than interpolation ones, while extrapolation approach still remains to be developed.

Future work will first focus on adaptation mechanisms for approximation. In particular, the objective is to establish adaptation rules based on edit operations performed on time series alignments. The underlying hypothesis is that edit operations observed between physiological signals may be associated with changes in emotional labels. Adaptation rules would therefore link patterns of signal transformations to transformations of solutions. Extrapolation will also be investigated, notably through analogy-based reasoning relying on similar edit operations.

A possible research direction would also be to study how approximation, interpolation, and extrapolation can interact within the same reasoning framework. For instance, one objective could be to determine which reasoning mechanism is the most appropriate depending on the target problem and the structure of the retrieved cases. Another direction would be to combine the solutions proposed by the three approaches in order to improve the overall quality and robustness of the final prediction.

Another important direction concerns the acceleration of the retrieval process. Several optimizations are planned, such as integrating MAC/FAC [11] retrieval mechanisms and reducing the computation

cost of TWED using constraints like the Sakoe–Chiba band [12], which limits the number of explored cells in the dynamic programming matrix.

Finally, the long-term objective is to return to the original and more complex problem: predicting emotional annotations over complete listening sessions, where several emotions may appear at different moments in time. Additional experiments on newly collected data are also planned in order to evaluate the proposed approaches in more realistic settings.

Declaration on Generative AI

During the preparation of this work, the authors used a local version of Gemma4 and Writefull. Writefull was used for grammar and spelling checks, and a local version of Gemma4 assisted with phrasing and style improvements. All suggestions were reviewed and adapted by the authors, who take full responsibility for the final content.

References

[1] V. J. Konecni, Does music induce emotion? a theoretical and methodological analysis., Psychology of Aesthetics, Creativity, and the Arts 2 (2008) 115–129. URL: https://api.semanticscholar.org/CorpusID:264237034.

[2] S. Lall, S. Gupta, Music genres and stress reduction: A comparative experimental study with university students, International Journal For Multidisciplinary Research (2025). URL: https://api.semanticscholar.org/CorpusID:282086313.

[3] S. Begum, S. Barua, R. Filla, M. U. Ahmed, Classification of Physiological Signals for Wheel Loader Operators Using Multi-scale Entropy Analysis and Case-Based Reasoning, Expert Systems with Applications 41 (2014) 295–305. doi:10.1016/j.eswa.2013.05.068.

[4] J. Lieber, E. Nauer, H. Prade, G. Richard, Making the Best of Cases by Approximation, Interpolation and Extrapolation, in: M. T. Cox, P. Funk, S. Begum (Eds.), Case-Based Reasoning Research and Development, volume 11156, Springer International Publishing, Cham, 2018, pp. 580–596. doi:10.1007/978-3-030-01081-2_38.

[5] É. Astier, H. Iopeti, J. Lieber, H. Mathieu Steinbach, L. Yvoz, Case-Based Cleaning of Text Images, in: S. Massie, S. Chakraborti (Eds.), Case-Based Reasoning Research and Development, volume 14141, Springer Nature Switzerland, Cham, 2023, pp. 344–358. doi:10.1007/978-3-031-40177-0_22.

[6] E. Diebold, J. Lieber, N. Merah, A. Bannay, Case-Based Identification of Anti-Coagulation Therapy for Pregnant Women Using Approximation and Interpolation of Sequences, in: Proceedings of the 34th International Conference on Case-Based Reasoning (ICCBR 2026), Lecture Notes in Computer Science, Springer, Bremen, Germany, 2026.

[7] L. E. Bishop, G. Bonnin, Analysing Physiological Data Collected During Music Listening: An Introduction, in: 24th International Society for Music Information Retrieval Conference, Milan, Italy, 2023. URL: https://inria.hal.science/hal-04407724.

[8] P.-F. Marteau, Time Warp Edit Distance with Stiffness Adjustment for Time Series Matching, IEEE Transactions on Pattern Analysis and Machine Intelligence 31 (2009) 306–318. URL: https://hal.science/hal-00135473. doi:10.1109/TPAMI.2008.76.

[9] Cory S. Myers and Lawrence R. Rabiner and Aaron E. Rosenberg, Performance tradeoffs in dynamic time warping algorithms for isolated word recognition, IEEE Transactions on Acoustics, Speech, and Signal Processing 28 (1980) 623–635. URL: https://api.semanticscholar.org/CorpusID:14297130.

[10] Wright, G., Improved Time Warp Edit Distance – A Parallel Dynamic Program in Linear Memory, 2020. URL: https://arxiv.org/abs/2007.16135. arXiv:2007.16135.

[11] K. D. Forbus, D. Gentner, K. Law, MAC/FAC: A Model of Similarity-Based Retrieval, Cognitive Science 19 (1995) 141–205.

[12] H. Sakoe, S. Chiba, Dynamic programming algorithm optimization for spoken word recognition, IEEE Transactions on Acoustics, Speech, and Signal Processing 26 (1978) 43–49. doi:10.1109/TASSP.1978.1163055.