

Oregon State University

Can we teach AI to feel music?

Building reliable datasets for music emotion recognition

Patrick J. Donnelly, Ph.D.

October 28, 2026

soundbendor lab

Soundbendor Lab

Soundbendor Lab

Machine Understanding of Sound and Music

Our interdisciplinary research focuses on a computational understanding of sound and music using AI and deep-learning.



Music



Sound



Psycho-acoustics



Bioacoustics

www.soundbendor.org

Lab Members


Patrick Donnelly (PI)

Alex Armendariz, MS (AI)


Aidan Beery, Ph.D (CS)


Abbey Hankins, MS (CS)


Pratik Khanal, Ph.D (CS)


Derek Kwan, Ph.D (AI)


Peter McCabe, Ph.D (AI


Ian Tassin, Ph.D (CS)

Alumni


Janita Aamir
MS (2025)


Isaac Brickner
BS (2025)


Shaura Gaur
BS (2023)


Shaun Leib
MS (2026)


Shengxuan Wang
MS (2025)

\+ 28 others

soundbendor lab

Music Emotion Recognition

Music Emotion Recognition

The task of computationally estimating the emotional qualities conveyed by a piece of music using machine learning and artificial intelligence.

Music Emotion Recognition

The task of computationally estimating the emotional qualities conveyed by a piece of music using machine learning and artificial intelligence.

Modalities

- raw audio signals [1]
- spectrograms [2]
- acoustic features [3]
- metadata and context [4]
- song lyrics [5]
- symbolic music scores [6]
- physiological signals [7]

Music Emotion Recognition

The task of computationally estimating the emotional qualities conveyed by a piece of music using machine learning and artificial intelligence.

Modalities

- raw audio signals [1]
- spectrograms [2]
- acoustic features [3]
- metadata and context [4]
- song lyrics [5]
- symbolic music scores [6]
- physiological signals [7]
- **social media discourse** [8]
- **program note documents** [9]
- **LLM emotion analysis** [10]

Music Emotion Recognition

The task of computationally estimating the emotional qualities conveyed by a piece of music using machine learning and artificial intelligence.

Modalities

- raw audio signals [1]
- spectrograms [2]
- acoustic features [3]
- metadata and context [4]
- song lyrics [5]
- symbolic music scores [6]
- physiological signals [7]
- **social media discourse** [8]
- **program note documents** [9]
- **LLM emotion analysis** [10]

Applications

- emotion-based music search
- music recommendation systems
- computational musicology
- music therapy interventions
- music captioning for the deaf
- AI-assisted music composition
- adaptive film & game soundtracks
- expressive performance feedback
- dynamic ambient environments
- advertising, hospitality

Models of Emotion

Elkman's Basic Emotions [11]

- Happiness
- Sadness
- Anger
- Fear
- Disgust
- Surprise

Plutchik's Wheel of Emotions [12]

- Joy $\longleftrightarrow$ Sadness
- Anger $\longleftrightarrow$ Fear
- Trust $\longleftrightarrow$ Disgust
- Surprise $\longleftrightarrow$ Anticipation

Russell's Circumplex Model [13]

- **Valence**: unpleasant to pleasant
- **Arousal**: calm to excited



How to measure emotion?

Dataset Annotation Questions

What should we attempt to measure? [14]

1. **Perceived emotion**: *Which emotions does the music convey?*
2. **Induced emotion**: *Which emotions does the listener experience?*

Dataset Annotation Questions

What should we attempt to measure? [14]

1. **Perceived emotion**: *Which emotions does the music convey?*
2. **Induced emotion**: *Which emotions does the listener experience?*

How to conduct studies?

- By averaging user opinions?
- By consulting musical experts?
- By physiological measurements?
- By tracking eye movement?
- By measuring facial expressions?

Dataset Annotation Questions

What should we attempt to measure? [14]

1. **Perceived emotion**: *Which emotions does the music convey?*
2. **Induced emotion**: *Which emotions does the listener experience?*

How to conduct studies?

- By averaging user opinions?
- By consulting musical experts?
- By physiological measurements?
- By tracking eye movement?
- By measuring facial expressions?

Where to conduct studies?

- Formal study in a lab?
- Paid subjects working online?
- Crowd-sourced online?
- Offline questionnaires?

Dataset Annotation Questions

What should we attempt to measure? [14]

1. **Perceived emotion**: *Which emotions does the music convey?*
2. **Induced emotion**: *Which emotions does the listener experience?*

How to conduct studies?

- By averaging user opinions?
- By consulting musical experts?
- By physiological measurements?
- By tracking eye movement?
- By measuring facial expressions?

Which method to use?

- Curated human responses?
- Algorithmically generated?
- From meaning of the lyrics?
- Derived from metadata?
- Inferred from indirect sources?

Where to conduct studies?

- Formal study in a lab?
- Paid subjects working online?
- Crowd-sourced online?
- Offline questionnaires?

Dataset Annotation Questions

What should we attempt to measure? [14]

1. **Perceived emotion**: *Which emotions does the music convey?*
2. **Induced emotion**: *Which emotions does the listener experience?*

How to conduct studies?

- By averaging user opinions?
- By consulting musical experts?
- By physiological measurements?
- By tracking eye movement?
- By measuring facial expressions?

Which method to use?

- Curated human responses?
- Algorithmically generated?
- From meaning of the lyrics?
- Derived from metadata?
- Inferred from indirect sources?

Where to conduct studies?

- Formal study in a lab?
- Paid subjects working online?
- Crowd-sourced online?
- Offline questionnaires?

Which paradigm to use?

- Discrete or continuous scale?
- Single or multiple labels?
- Static or dynamic timing?
- Excerpt or whole song?

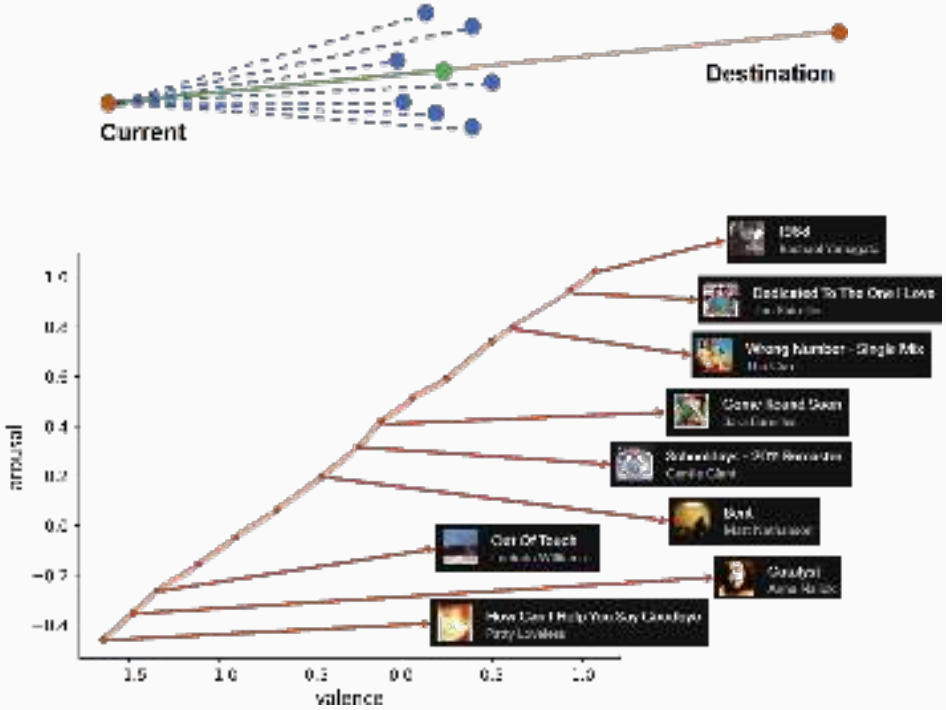
MER for Therapy Interventions

Mood-Dynamic Playlists

TASK: Generate mood-dynamic playlists for use in music therapy

Approach

- User chooses:
  - starting emotion
  - target emotion
  - number of songs
  - perferred genre
- KNN algorithm
- Vector difference
  - smoothness
  - evenness
- Playlist traverses songs in an emotional space

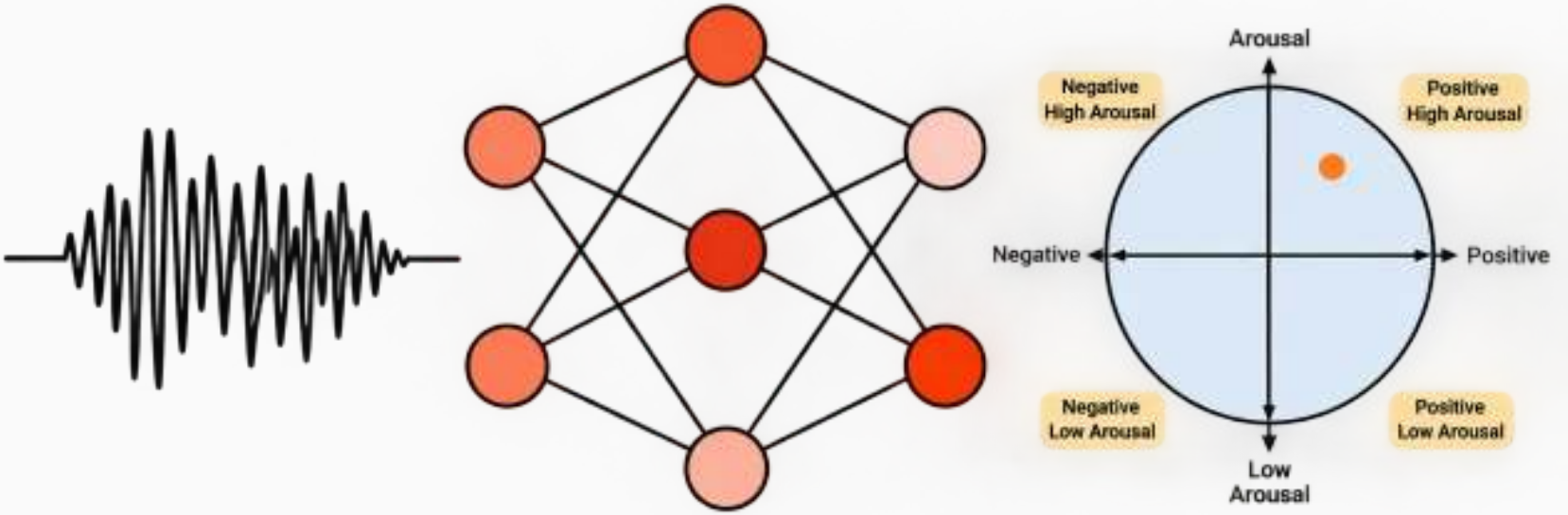


[15] S. Gaur and P.J. Donnelly. **"Generating Smooth Mood-Dynamic Playlists with Audio Features and KNN"**. EvoMUSART 2024.

Estimating Music Emotion

GOAL

Predict perceived music emotion directly from raw audio signals



This would enable the automatic labeling of large music repositories.

Predicting Emotion from Audio

TASK: Predict music emotion using curated acoustic features

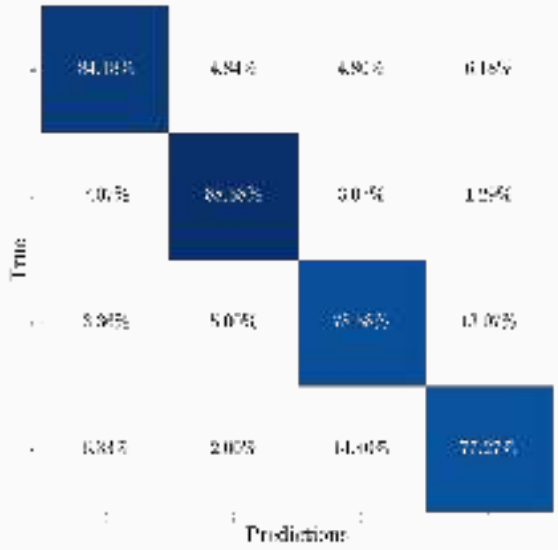
Dataset

- 4Q dataset [16], n=900
- balanced by emotion
- Tags from AllMusic API
- V/A quadrants

Features

- Generated 3339 features
  - selected best 300
- reduced features helped
- neural network was best
- outperformed [16]
- most difficulty with:
  - Q3: sad, sleepy
  - Q4: calm, peaceful

Model	[16]: F_1	Our: F_1
SVM	0.768	**0.786**
RF	0.717	**0.731**
MLP	0.748	**0.799**



[17] Shaun Leib. **"Interpreting emotion from sound: A comparison of features and models for music emotion classification"**. Undergraduate Honors Thesis, **Oregon State University**, 2025.

Comparing Static and Dynamic Features

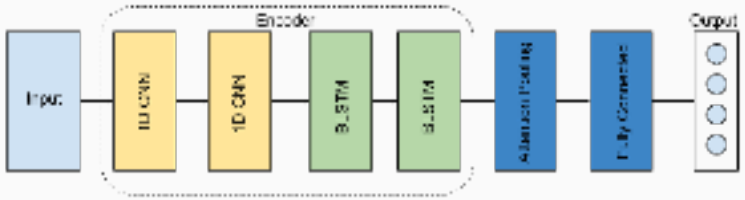
TASK: Predict music emotion using static and dynamic annotations

Dataset

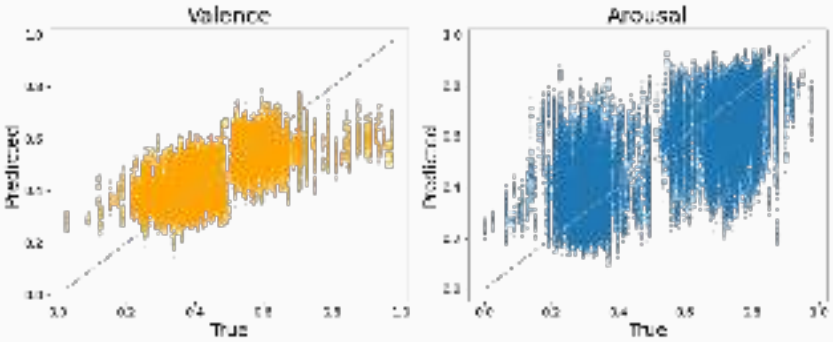
- MERGE [18], n=3554
- 30-second excerpts
- V/A quadrants

Features

- n=710 features
- Static and dynamic
- Acoustic features
- Model embeddings
  - Wav2Vec 2.0 [19]
  - MERT [20]
  - MAEST [21]



Feature	Pearson's r V	A
Static	**0.741**	**0.720**
Dynamic	0.733	0.711



[22] Shaun Leib. **"Time is not on your side: Assessing the value of dynamic audio features for static music emotion labels"**. MS Thesis, **Oregon State University**, 2026.

soundbendor lab

Chinese Opera Sentiment

TASK: Detect positive or negative sentiment in Chinese Opera songs

Dataset [23]

- 139 recordings
- 61 songs
- 28 singers

Feature Modality

- Spectrograms
- MFCCs
- Raw waveforms
- Embeddings
- Lyrics
- Audio+Lyrics



	Features	Method	F_1
Acoustic	Spectrogram	CNN-LSTM	0.65
	MFCCs	CNN	0.59
	Raw Audio	LSTM	0.64
	Embeddings	FC MLP	0.63
Textual	Lyrics	BERT	**0.71**
Fusion	Audio+Lyrics	FC MLP	0.54

[24] Shengxuan Wang. **"Comparative analysis of sentiment in Chinese Opera with different features."**. MS Thesis Project, **Oregon State University**, 2025.

Existing Music Emotion Datasets

Motivation

Problem

There is a severe lack of large, reliable, and diverse datasets for MER.

Motivation

Problem

There is a severe lack of large, reliable, and diverse datasets for MER.

Difficulties:

- task is difficult to define [25]
- expensive to conduct [26]
- time-consuming process [27]
- sampling selection bias [28]
- temporal structure of music [29]
- listener subjectivity [30]
- conflicting studies [31]

Motivation

Problem

There is a severe lack of large, reliable, and diverse datasets for MER.

Difficulties:

- task is difficult to define [25]
- expensive to conduct [26]
- time-consuming process [27]
- sampling selection bias [28]
- temporal structure of music [29]
- listener subjectivity [30]
- conflicting studies [31]

Limitations:

- only short excerpts [32]
- Western-centric genres [33]
- bias in song selection [34]
- crowd-sourced processes [35]
- collapse to the average [36]
- rater annotation fatigue [37]
- copyright concerns [38]

Summary of Existing MER Datasets

Dataset		Count	Method	Target
AMG1608	[39]	1608	Online	Perceived
DEAM	[27]	1820	Online	Perceived
DEAP	[7]	120	Lab	Induced
Deezer	[40]	18,644	Algorithmic	Derived
PMEmo	[41]	794	Lab	Induced

Correlations between Human Datasets

	AMG	DEAM	DEAP	Deezer	PmEmo
AMG		— (0)	-0.353 (3)	0.639 (246)	-0.254 (4)
DEAM	— (0)		— (0)	0.576 (3)	— (0)
DEAP	0.901 (3)	— (0)		0.705 (29)	— (0)
Deezer	0.571 (246)	-0.374 (3)	0.484 (29)		-0.014 (9)
PMEmo	0.257 (4)	— (0)	— (0)	-0.001 (9)	

Correlations of Arousal (left of diagonal) and Valence (right)

[10] J. Aamir and P.J. Donnelly. **"Addressing dataset scarcity in music emotion recognition with LLMs"**. EvoMUSART 2026

Exploring Alternate Approaches

Exploring Alternate Approaches

Research Question

Can we find new ways to generate pseudo-annotations for music emotion?

Exploring Alternate Approaches

Research Question

Can we find new ways to generate pseudo-annotations for music emotion?

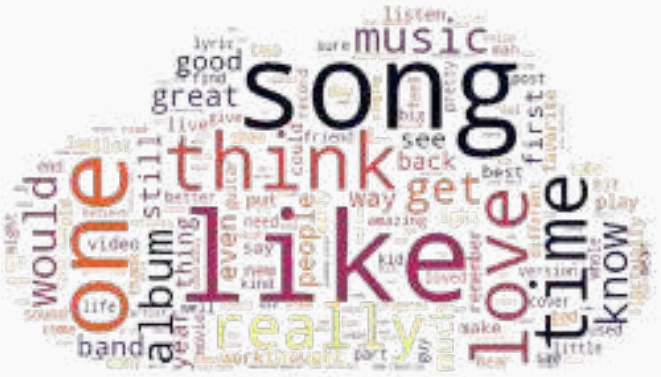
And can we do so *quickly*, *cheaply*, and *automatically* on a large scale?

Social Media Discourse

TASK: Estimate music emotion directly from social media comments

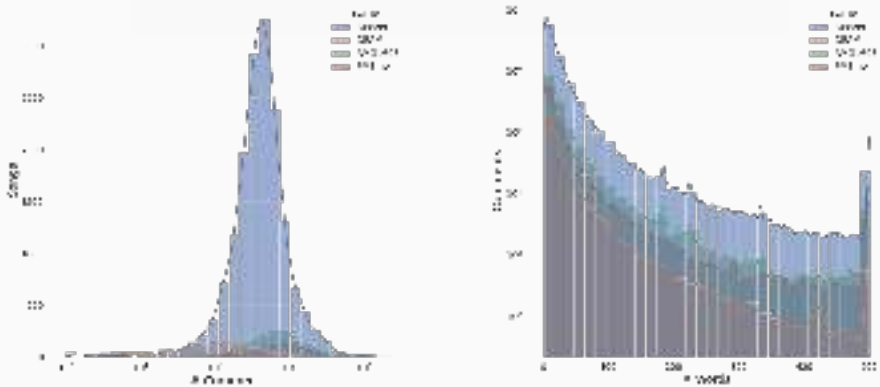
Dataset

- 19,627 songs
- 11.8M comments
  - YouTube
  - Reddit
  - Twitter



Approach

- Harvest comments
- Annotate emotion [42]
- Train BERT model
  - Valence r=0.80
  - Arousal r=0.79



[8] A. Beery and P.J. Donnelly. **"Learning Affective Responses to Music from Social Media Discourse"**. Practical solutions for Diverse Real-World NLP Application. Signals and Communication Technology, 2024.

Classical Music Programs

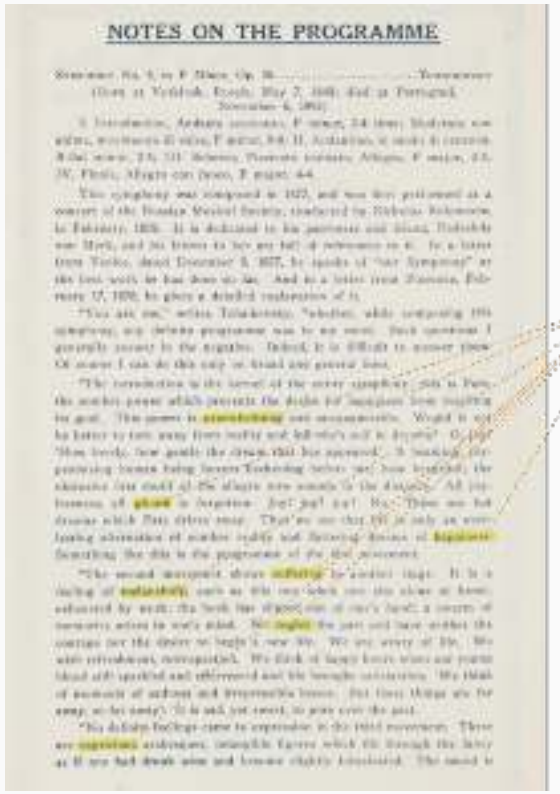
TASK: Estimate emotion from historical classical music programs

Dataset

- NY Philharmonic
- 14,743 booklets
- 31,966 notes
- from 1842-2024

Approach

- Perform OCR
- Clean notes with LLM
- Annotate emotion [42]
- Train model to predict
  - Valence r=0.899
  - Arousal r=0.860



Word	Valence	Arousal
overwhelming	0.361	0.527
gloom	0.224	0.372
happiness	0.968	0.792
suffering	0.109	0.466
melancholy	0.341	0.409
regret	0.296	0.533
capricious	0.514	0.620

[9] P. Khanal and P.J. Donnelly. **"EmotioNotes dataset: Decoding emotions in classical music through concert program notes"**. EvoMUSART 2025

Trends in Billboard Charts

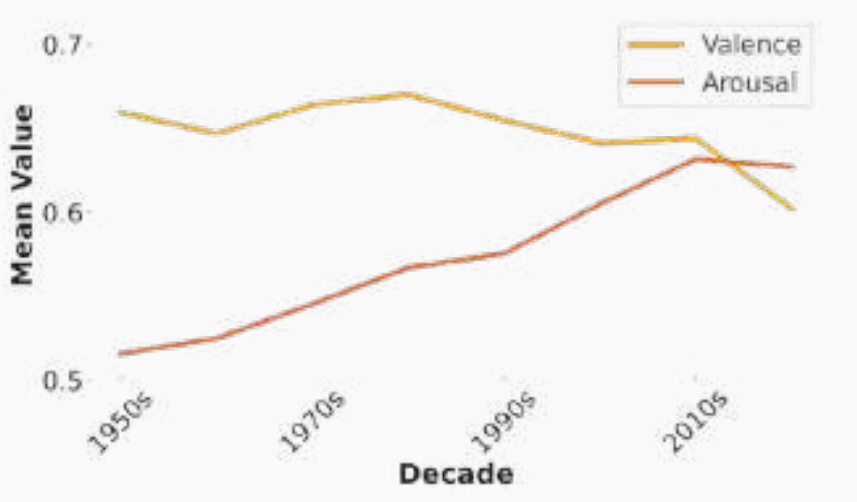
TASK: Identify emotion trends in Billboard Top-100 tracks

Dataset

- Billboard Top-100
- from 1958-2024
- 31,596 distinct songs
- 10,954 different artists

Approach

- Prompt ChatGPT
- Annotate emotion [42]
- Analyze changing trends
  - Small decrease in valence
  - Large increase in arousal



Issac Brickner. **"Analysis of Emotion Prediction in Human and LLM Annotated Datasets"**.
Cascades Research Fellowship, 2024-2025

LLM Annotation Generation

TASK: Generate LLM annotations for music emotion recognition

Human Annotations

- AMG1608, n=1608
- DEAM, n=1802
- DEAP, n=120
- Deezer, n=18,644
- PMEmo, n=794

Approach

- ChatGPT 3.5-turbo
- Compare prompt types
  - direct, numeric
  - five emotion words
  - longform description
- Publish new dataset

	Direct	
	Arousal	Valence
AMG1608	0.582	0.635
DEAM	0.185	0.201
DEAP	0.560	0.688
Deezer	0.208	0.446
PMEmo	-0.091	0.044

Pearson's r: Human vs. ChatGPT

[10] J. Aamir and P.J. Donnelly. **"Addressing dataset scarcity in music emotion recognition with LLMs"**. EvoMUSART, 2026

LLM Inter-rater Reliability

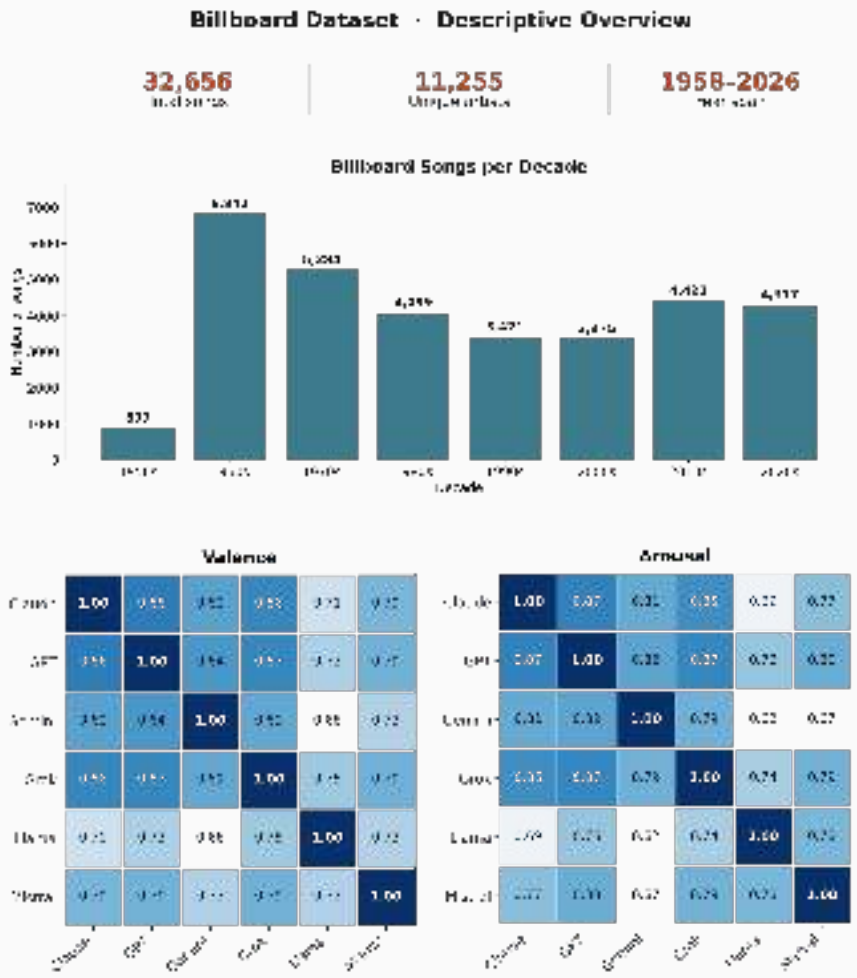
TASK: Compare LLM annotations for music emotion recognition

Dataset

- Billboard Top-100
- from 1958-2026
- 32,656 distinct songs
- 11,255 artists

Approach

- Prompt LLM models
  - Claude
  - ChatGPT
  - Gemini
  - Grok
  - LLaMa
  - Mistral
- Inter-rater reliability
- Publish new dataset



P. Khanal and P.J. Donnelly. **"LLMs as music emotion annotators: A multi-LLM annotation framework for music emotion recognition"**. In Preparation, 2027

soundbendor lab

Next Steps

Next Steps

Better human annotations

- new annotation framework
  - best match
  - pick all
  - compare two
  - opposites
  - circumplex 2D
  - matching
  - temporal
- CS Capstone 2026-2027

Next Steps

Better human annotations

- new annotation framework
  - best match
  - pick all
  - compare two
  - opposites
  - circumplex 2D
  - matching
  - temporal
- CS Capstone 2026-2027

Novel pseudo-annotations

- album cover art
- critic reviews
- playlist analysis
- music scores
- facial expressions

Next Steps

Better human annotations

- new annotation framework
  - best match
  - pick all
  - compare two
  - opposites
  - circumplex 2D
  - matching
  - temporal
- CS Capstone 2026-2027

Novel pseudo-annotations

- album cover art
- critic reviews
- playlist analysis
- music scores
- facial expressions

Predictions from audio

- large-scale dataset
- collect Billboard audio
- build multimodal models
  - raw audio
  - extracted audio features
  - spectrograms
  - song lyrics
  - metadata & tags
  - comments & discourse

Next Steps

Better human annotations

- new annotation framework
  - best match
  - pick all
  - compare two
  - opposites
  - circumplex 2D
  - matching
  - temporal
- CS Capstone 2026-2027

Predictions from audio

- large-scale dataset
- collect Billboard audio
- build multimodal models
  - raw audio
  - extracted audio features
  - spectrograms
  - song lyrics
  - metadata & tags
  - comments & discourse

Novel pseudo-annotations

- album cover art
- critic reviews
- playlist analysis
- music scores
- facial expressions

Better problem understanding

- perceived vs induced
- discrete vs continuous
- static vs dynamic
- single vs multi-label

Thank You!

References i

[1] N. He and S. Ferguson, **"Multi-view neural networks for raw audio-based music emotion recognition,"** in *2020 IEEE International Symposium on Multimedia (ISM)*, 2020, pp. 168–172. DOI: 10.1109/ISM.2020.00037

[2] X. Jia, **"A music emotion classification model based on the improved convolutional neural network,"** *Computational Intelligence and Neuroscience*, vol. 2022, p. 6 749 622, 2022. DOI: 10.1155/2022/6749622

[3] R. Panda, R. Malheiro, and R. P. Paiva, **"Audio features for music emotion recognition: A survey,"** *IEEE Transactions on Affective Computing*, vol. 14, no. 1, pp. 68–88, 2023. DOI: 10.1109/TAFFC.2020.3032373

[4] E. Cano and M. Morisio, **"Music mood dataset creation based on last fm tags,"** in *Proceedings of the Fourth International Conference on Artificial Intelligence and Applications (AIAP 2017)*, 2017, pp. 15–26. DOI: 10.5121/csit.2017.70603

[5] E. Çano and M. Morisio, **"Moodylyrics: A sentiment annotated lyrics dataset,"** in *Proceedings of the 2017 International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence (ISMSI)*, Hong Kong, Hong Kong: Association for Computing Machinery, 2017, pp. 118–124, ISBN: 9781450347983. DOI: 10.1145/3059336.3059340

[6] H.-T. Hung, J. Ching, S. Doh, N. Kim, J. Nam, and Y.-H. Yang, **"EMOPIA: A multi-modal pop piano dataset for emotion recognition and emotion-based music generation,"** in *Proceedings of the 22nd International Society for Music Information Retrieval Conference (ISMIR)*, 2021, pp. 318–325. [Online]. Available: https://arxiv.org/abs/2108.01374

[7] S. Koelstra et al., **"DEAP: Database for emotion analysis using physiological signals,"** *IEEE transactions on affective computing*, vol. 3, no. 1, pp. 18–31, 2011. DOI: 10.1109/T-AFFC.2011.15

[8] A. Beery and P. J. Donnelly, **"Learning affective responses to music from social media discourse,"** in *Practical Solutions for Diverse Real-World NLP Applications*, M. Abbas, Ed., Cham: Springer International Publishing, 2024, pp. 93–119, ISBN: 978-3-031-44260-5. DOI: 10.1007/978-3-031-44260-5_6

[9] P. Khanal and P. J. Donnelly, **"Emotionotes dataset: Decoding emotions in classical music through concert program notes,"** in *International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar)*, Springer, 2025, pp. 339–356. DOI: 10.1007/978-3-031-90167-6_23

[10] J. Aamir and P. J. Donnelly, **"Addressing dataset scarcity in music emotion recognition with LLMs,"** in *International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar)*, Springer, 2026, pp. 3–20. DOI: 10.1007/978-3-032-24350-8_1

[11] P. Ekman, **"An argument for basic emotions,"** *Cognition and Emotion*, vol. 6, no. 3–4, pp. 169–200, 1992. DOI: 10.1080/02699939208411068

References ii

[12] R. Plutchik, **"A general psychoevolutionary theory of emotion,"** in *Theories of Emotion*, R. Plutchik and H. Kellerman, Eds., Academic Press, 1980, pp. 3–33, ISBN: 978-0-12-558701-3. DOI: 10.1016/B978-0-12-558701-3.50007-7

[13] J. A. Russell, **"A circumplex model of affect,"** *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980. DOI: 10.1037/h0077714

[14] A. Gabrielsson, **"Emotion perceived and emotion felt: Same or different?"** *Musicae Scientiae*, vol. 5, no. 1_suppl, pp. 123–147, 2001. DOI: 10.1177/10298649020050S105

[15] S. Gaur and P. J. Donnelly, **"Generating smooth mood-dynamic playlists with audio features and KNN,"** in *Proceedings of the 13th International Conference of Evolutionary and Biologically Inspired Music, Sound, Art and Design (EvoMUSART)*, C. Johnson, S. M. Rebelo, and I. Santos, Eds., vol. 14633, Springer, 2024, pp. 162–178, ISBN: 978-3-031-56992-0. DOI: 10.1007/978-3-031-56992-0_11

[16] R. Panda, R. Malheiro, and R. P. Paiva, **"Novel audio features for music emotion recognition,"** *IEEE Transactions on Affective Computing*, vol. 11, no. 4, pp. 614–626, 2020. DOI: 10.1109/TAFFC.2018.2820691

[17] S. Leib, **"Interpreting emotion from sound: A comparison of features and models for music emotion classification,"** Undergraduate honors thesis, Oregon State University, 2025.

[18] P. L. Louro, H. Redinho, R. Santos, R. Malheiro, R. Panda, and R. P. Paiva, **Merge – a bimodal audio-lyrics dataset for static music emotion recognition,** 2025. arXiv: 2407.06060 [cs.SD]. [Online]. Available: https://arxiv.org/abs/2407.06060

[19] A. Baevski, H. Zhou, A. Mohamed, and M. Auli, **Wav2vec 2.0: A framework for self-supervised learning of speech representations,** 2020. arXiv: 2006.11477 [cs.CL]. [Online]. Available: https://arxiv.org/abs/2006.11477

[20] Y. Li et al., **Mert: Acoustic music understanding model with large-scale self-supervised training,** 2024. arXiv: 2306.00107 [cs.SD]. [Online]. Available: https://arxiv.org/abs/2306.00107

[21] P. Alonso-Jiménez, X. Serra, and D. Bogdanov, **Efficient supervised training of audio transformers for music representation learning,** 2023. arXiv: 2309.16418 [cs.SD]. [Online]. Available: https://arxiv.org/abs/2309.16418

[22] S. Leib, **"Time is not on your side: Assessing the value of dynamic audio features for static music emotion labels,"** Master's thesis, Oregon State University, 2026.

[23] D. A. A. Black, M. Li, and M. Tian, **"Automatic identification of emotional cues in Chinese opera singing,"** *ICMPC, Seoul, South Korea*, pp. 250–255, 2014.

References iii

[24] S. Wang, **"Comparative analysis of sentiment in chinese opera with different features,"** Master's thesis project, Oregon State University, 2025.

[25] Y.-H. Yang and H. H. Chen, *Machine Recognition of Music Emotion: A Review*. Boca Raton, Florida: CRC Press, 2012. DOI: 10.1201/b12252

[26] N. Gutiérrez-Páez, J. Gómez-Cañón, L. Porcaro, P. Santos, D. Hernández-Leo, and E. Gómez, **"Emotion annotation of music: A citizen science approach,"** in *Collaboration Technologies and Social Computing*, Springer, 2021, pp. 51–66. DOI: 10.1007/978-3-030-85071-5_4

[27] A. Aljanaki, Y.-H. Yang, and M. Soleymani, **"Developing a benchmark for emotional analysis of music,"** *PloS one*, vol. 12, no. 3, pp. 1–22, 2017. DOI: 10.1371/journal.pone.0173392

[28] J. Ching and G. Widmer, **"A study on the data distribution gap in music emotion recognition,"** in *Proceedings of the 17th International Symposium on Computer Music Multidisciplinary Research (CMMR)*, 2025. [Online]. Available: https://arxiv.org/abs/2510.04688

[29] M. Soleymani et al., **"Emotional analysis of music: A comparison of methods,"** in *Proceedings of the ACM International Conference on Multimedia*, Association for Computing Machinery, 2014, pp. 1161–1164. DOI: 10.1145/2647868.2655019

[30] M. Schedl, E. Gómez, E. S. Trent, M. Tkalčič, H. Eghbal-Zadeh, and A. Martorell, **"On the interrelation between listener characteristics and the perception of emotions in classical orchestra music,"** *IEEE Transactions on Affective Computing*, vol. 9, no. 4, pp. 507–525, 2018. DOI: 10.1109/TAFFC.2017.2663421

[31] Y. Zhang, X. Ding, and N. Gu, **"Understanding fatigue and its impact in crowdsourcing,"** in *22nd International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, IEEE, 2018, pp. 57–62. DOI: 10.1109/CSCWD.2018.8465305

[32] E. Bigand, S. Filipic, and P. Lalitte, **"The time course of emotional responses to music,"** *Annals of the New York Academy of Sciences*, vol. 1060, no. 1, pp. 429–437, 2005. DOI: 10.1196/annals.1360.036 [Online]. Available: https://doi.org/10.1196/annals.1360.036

[33] H. Lee et al., **"GlobalMood: A cross-cultural benchmark for music emotion recognition,"** in *Proceedings of the 26th International Society for Music Information Retrieval Conference*, 2025. DOI: 10.48550/arXiv.2505.09539 [Online]. Available: https://doi.org/10.48550/arXiv.2505.09539

References iv

[34] H. Strauss et al., **"The emotion-to-music mapping atlas (EMMA): A systematically organized online database of emotionally evocative music excerpts,"** *Behavior Research Methods*, vol. 56, pp. 3560–3577, 2024. DOI: 10.3758/s13428-024-02336-0 [Online]. Available: https://doi.org/10.3758/s13428-024-02336-0

[35] I. P. Samiotis, S. Qiu, C. Lofi, J. Yang, U. Gadiraju, and A. Bozzon, **"An analysis of music perception skills on crowdsourcing platforms,"** *Frontiers in Artificial Intelligence*, vol. 5, p. 828 733, 2022. DOI: 10.3389/frai.2022.828733 [Online]. Available: https://doi.org/10.3389/frai.2022.828733

[36] E. Gómez-Sánchez, C. Guedes, P. Herrera, and E. Gómez, **"Music emotion recognition: Toward new, robust standards,"** *IEEE Signal Processing Magazine*, vol. 38, no. 6, pp. 65–75, 2021. DOI: 10.1109/MSP.2021.3106232

[37] A. Wiafe, S. Sieranoja, A. Bhuiyan, and P. Fränti, **"Emotional response to music: The Emotify+ dataset,"** *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2025, no. 1, p. 31, 2025. DOI: 10.1186/s13636-025-00419-0 [Online]. Available: https://doi.org/10.1186/s13636-025-00419-0

[38] R. M. Bittner, M. Fuentes, D. Rubinstein, A. Jansson, K. Choi, and T. Kell, **"Mirdata: Software for reproducible usage of datasets,"** in *Proceedings of the 20th International Society for Music Information Retrieval Conference*, 2019.

[39] Y.-A. Chen, Y.-H. Yang, J.-C. Wang, and H. Chen, **"The AMG1608 dataset for music emotion recognition,"** en, in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, Apr. 2015, pp. 693–697, ISBN: 978-1-4673-6997-8. DOI: 10.1109/ICASSP.2015.7178058

[40] R. Delbouys, R. Hennequin, F. Piccoli, J. Royo-Letelier, and M. Moussallam, **"Music mood detection based on audio and lyrics with deep neural net,"** en, in *Proceedings of the 19th International Society for Music Information Retrieval Conference (ISMIR)*, ISMIR, 2018, pp. 370–375. DOI: 10.5281/zenodo.1492427

[41] K. Zhang, H. Zhang, S. Li, C. Yang, and L. Sun, **"The PMEmo dataset for music emotion recognition,"** en, in *Proceedings of the 2018 International Conference on Multimedia Retrieval*, Yokohama Japan: ACM, Jun. 2018, pp. 135–142, ISBN: 978-1-4503-5046-4. DOI: 10.1145/3206025.3206037

[42] A. B. Warriner, V. Kuperman, and M. Brysbaert, **"Norms of valence, arousal, and dominance for 13,915 English lemmas,"** *Behavior Research Methods*, vol. 45, pp. 1191–1207, 2013. DOI: 10.3758/s13428-012-0314-x