

CLAIMS

1. A method for generating an audio stream for inducing an affective state

change in a listener, comprising:

identifying the listener's current affective state;

5 identifying the listener's target affective state;

identifying an affective trajectory from the current affective state to the target affective state;

using a trained segment identification machine learning model to identify a first audio segment likely to induce in the listener a desired affective response

10 corresponding to at least an initial portion of the affective trajectory when the first audio segment is presented to the listener as an auditory stimulus;

generating the audio stream based at least in part on the first audio segment; and

sending audio stream data based on the audio stream to a listener device.

15

2. The method of Claim 1,

further comprising, after using the trained segment identification machine learning model to identify the first audio segment:

using an affective inference process to infer an inferred new

20 affective state based on the current affective state and a set of audio feature values of the first audio segment;

identifying an updated affective trajectory from the inferred new affective state data to the target affective state; and

using the trained segment identification machine learning model to

25 identify a subsequent audio segment likely to induce in the listener a

subsequent desired affective response corresponding to at least an initial

portion of the updated affective trajectory when the subsequent audio segment is presented to the listener as an auditory stimulus; wherein the audio stream is generated based at least in part on the first audio segment and the subsequent audio segment.

5

3. The method of Claim 2, wherein:

the trained segment identification machine learning model is trained using reward data received from the affective inference process; and

the affective inference process generates the reward data by:

10 inferring an inferred affective response of the listener to a set of audio feature values of the audio stream; and

generating the reward data based on a comparison of the inferred affective response to the desired affective response.

15 4. The method of Claim 3, wherein:

the affective inference process comprises a trained affect inference machine learning model; and

the trained affect inference machine learning model is trained using training data comprising:

20 training audio feature data corresponding to a plurality of training audio segments; and

affective state data gathered from one or more human subjects in association with exposure of each human subject to each of a plurality of audio stimuli corresponding to the plurality of training audio segments.

25

5. The method of Claim 4, wherein the one or more human subjects comprises

the listener.

6. The method of Claim 5, wherein the trained segment identification machine learning model comprises a reinforcement learning model.

5

7. The method of Claim 6, wherein the trained segment identification machine learning model comprises a deep learning neural network.

8. The method of Claim 7, wherein the audio stream data comprises
10 recommendation data recommending the audio stream.

9. The method of Claim 7, wherein the audio stream data comprises the audio stream.

15 10. The method of Claim 9, further comprising, after sending the audio stream data to the listener device:

receiving updated current affective state data from the listener; and

training the trained affect inference machine learning model using runtime
training data comprising:

20 audio feature data corresponding to each of the first audio segment and
the plurality of subsequent audio segments; and
the updated current affective state data.

11. The method of Claim 1, wherein identifying the listener's target affective
25 state comprises:

receiving target affective state data from the listener via the listener

device; and

identifying the listener's target affective state based on the target affective state data.

5 12. The method of Claim 1, wherein identifying the listener's current affective state comprises:

receiving affective self-evaluation data from the listener via the listener device; and

10 identifying the listener's current affective state based on the affective self-evaluation data.

13. The method of Claim 1, wherein identifying the listener's current affective state comprises:

15 receiving physiological data correlated with or more physiological states of the listener; and

identifying the listener's current affective state based on the physiological data.

20 14. A system for generating an audio stream for inducing an affective state change in a listener, comprising:

a processor system;

a communication system; and

a memory system having stored thereon:

25 an executable trained segment identification machine learning model; and

executable instructions that, when executed by the processor

system, cause the system to:

identify the listener's current affective state based on listener state data received by the communication system;

5 identify the listener's target affective state based on target affective state data received by the communication system;

identify an affective trajectory from the current affective state to the target affective state;

10 execute the trained segment identification machine learning model to identify a first audio segment likely to induce in the listener a desired affective response corresponding to at least an initial portion of the affective trajectory in the listener when presented to the listener as an auditory stimulus;

generate the audio stream based at least in part on the first audio segment; and

15 use the communication system to send audio stream data based on the audio stream to a listener device.

15. The system of Claim 14, wherein:

20 the memory system further has stored thereon an executable trained affect inference machine learning model; and

the executable instructions, when executed by the processor system, further cause the system, after using the trained segment identification machine learning model to identify the first audio segment, to:

25 use the affective inference machine learning model to infer an inferred new affective state based on the current affective state and a set of audio feature values of the first audio segment;

identify an updated affective trajectory from the inferred new affective state data to the target affective state; and

use the trained segment identification machine learning model to identify a subsequent audio segment likely to induce in the listener a subsequent desired affective response corresponding to at least an initial portion of the updated affective trajectory when the subsequent audio segment is presented to the listener as an auditory stimulus;

wherein the audio stream is generated based at least in part on the first audio segment and the subsequent audio segment.

10

16. The system of Claim 15, wherein:

the trained segment identification machine learning model is trained using reward data received from the affective inference process; and

the affective inference process generates the reward data by:

15 inferring an inferred affective response of the listener to a set of audio feature values of the audio stream; and

generating the reward data based on a comparison of the inferred affective response to the desired affective response.

20 17. The system of Claim 16, wherein:

the trained affect inference machine learning model is trained using training data comprising:

training audio feature data corresponding to a plurality of training audio segments; and

25 affective state data gathered from one or more human subjects in association with exposure of each human subject to each of a plurality of audio

stimuli corresponding to the plurality of training audio segments.

18. A non-transitory processor-readable medium containing instructions for executing one or more of the methods of Claims 1 to 14.

5

19. A non-transitory storage medium containing the audio stream generated by one or more of the methods of Claims 1 to 14.

20. A method for training a machine learning model to predict human affective
10 responses to musical features, comprising:
 presenting a listener with music having a set of musical features;
 obtaining affective response data from the listener indicating the listener's
affective response to presentation of the music;
 labelling the musical features of the music with the affective response
15 data to generate labelled musical feature data; and
 using the labelled musical feature data as training data to train the
machine learning model to predict the affective response data based on the
musical feature data.

20 21. A system for predicting human response to music, comprising:
a processor system; and
 a memory system having stored thereon:
 a machine learning model trained according to the method of claim
20.

25

22. The system of claim 21, wherein the machine learning model is a generative

model for:

receiving desired affective response data indicating a desired affective response; and

generating music having musical features predicted by the machine

5 learning model to induce the desired affective response in a listener.

23. The system of Claim 22, wherein generating the music comprises:

using the generative model to process the desired affective response data to generate musical feature data indicating the musical features; and

10 generating the music based on the musical feature data.

24. The system of Claim 23, wherein the musical feature data is music information retrieval (MIR) data.

15 25. The system of Claim 24, wherein the MIR data is a MIR blueprint for the music.

26. The system of Claim 23, wherein the generative model is a generative adversarial network (GAN).

20

27. The system of Claim 26, wherein the GAN comprises a generator network, a probability network, and a control network.

28. The system of Claim 27, wherein the GAN comprises a conditional GAN.

25

29. The system of Claim 28, wherein the generator network comprises a

generator neural network.

30. The system of Claim 29, wherein the probability network comprises a discriminator neural network.

5

31. The system of Claim 30, wherein the generator neural network and discriminator neural network each comprise a recurrent neural network (RNN) with long short-term memory (LSTM).

10 32. The system of Claim 23, wherein generating the music based on the musical feature data comprises:

generating a score based on the musical feature data; and

generating the music based on the score.

15 33. The system of Claim 32, wherein the score is a musical instrument digital interface (MIDI) score.

34. The system of Claim 32, wherein generating the score comprises:

receiving composition intention information; and

20 generating the score based on the musical feature data and the composition intention information.

35. The system of Claim 34, wherein the composition intention information comprises one or more of: score type information, instrumentation information,

25 and score length information.

36. The system of Claim 34, wherein the composition intention information is indicated by composition intention user input received from a user.

37. The system of Claim 32, wherein:

5 the memory further stores a score generation machine learning model;
and

 generating the score comprises using the score generation machine learning model to process the musical feature data to generate the score.

10 38. The system of Claim 37, wherein the score generation machine learning model is a score generation generative adversarial network (GAN).

39. The system of Claim 38, wherein the score generation GAN comprises a generator network, a probability network, and a control network.

15

40. The system of Claim 39, wherein the score generation GAN comprises a conditional GAN.

41. The system of Claim 40, wherein the generator network comprises a
20 generator neural network.

42. The system of Claim 41, wherein the probability network comprises a discriminator neural network.

25 43. The system of Claim 42, wherein the generator neural network and discriminator neural network each comprise a recurrent neural network (RNN)

with long short-term memory (LSTM).

44. The system of Claim 32, wherein generating the music based on the score comprises:

- 5 presenting the score to a user;
- receiving rough mix user input from the user;
- generating a rough mix based on the rough mix user input; and
- generating the music based on the rough mix.

10 45. The system of Claim 23, wherein generating the music based on the generating the music based on the musical feature data comprises:

- generating a composition lead sheet based on the musical feature data;
- presenting the composition lead sheet to the user;
- receiving rough mix user input from the user;
- 15 generating a rough mix based on the rough mix user input; and
- generating the music based on the rough mix.

46. The system of Claim 44 or 45, wherein generating the music based on the rough mix comprises:

- 20 generating a production lead sheet based on the musical feature data and the rough mix;
- presenting the composition lead sheet to the user;
- receiving final mix user input from the user;
- generating a final mix based on the final mix user input; and
- 25 generating the music based on the final mix.

47. The system of Claim 46, wherein:

the memory further stores a mastering machine learning model; and
generating the final mix comprises using the mastering machine learning
model to process the musical feature data and the final mix to generate the
5 music.

48. The system of claim 47, wherein the mastering machine learning model is
further configured to:

receive an existing piece of music; and
10 process the musical feature data and the existing piece of music to
generate the music having the musical features.

49. The system of claim 48, wherein the mastering machine learning model is
15 further configured to generate the music having the musical features based on a
plurality of existing musical stems.

50. The system of claim 47, wherein the mastering machine learning model
comprises a recurrent deep Q network (DQN).

20 51. The system of claim 50, wherein the mastering machine learning model
comprises a branching recurrent DQN.

52. The system of claim 51, wherein the mastering machine learning model
25 includes a long short term memory (LSTM).

53. A non-transitory storage medium containing the audio stream generated by the system of any one of Claims 20 to 52.

54. The system of claim 21, wherein the machine learning model is further
- 5 configured to:
- receive a plurality of existing pieces of music;
 - receive the desired affective response data; and
 - identify an existing piece of music of the plurality of existing pieces of music likely to induce the desired affective response in a listener.

10