# Do You Want To Dance?

BY VEENA CALAMBUR, DANIEL HANGGI, JAVIER ORTIZ & LA VESHA PARKER

# BACKGROUND & SETUP

# -PROJECT MOTIVATION ----

MUSIC CAN BE QUANTIFIED NUMERICALLY IN TERMS OF VARIABLES SUCH AS VALENCE, ENERGY, TEMPO, BUT TO WHAT EXTENT CAN WE USE THESE VARIABLES TO LEARN MORE COMPLICATED CLASSIFICATIONS THAT HINGE ON HUMAN PERCEPTIONS OF MUSIC?

#### -IMPLEMENTATION-

1. FEATURE MAPPING APPLICATION

METHOD: WITH R'S LINEAR FIT METHODS, WE DERIVED A MODEL BASED VARIABLES WITH HIGH CORRELATIONS AMONGST EACH OTHER

## PROBLEM STATEMENT-

FOCUSING ON THE FOLLOWING ATTRIBUTES

**DANCEABILITY:** A COMBINATION OF ENERGY, RHYTHM, AND



- DANCEABILITY =  $-(3.279225 * 10^{-6})$  VALENCE \* TEMPO<sup>2</sup>
- + -(8.160651 \* 10<sup>-4</sup>) ENERGY \* LOUDNESS<sup>2</sup>
- + (3.218223 \* 10<sup>-1</sup>) ACOUSTICNESS \* ENERGY<sup>2</sup>
- + -(3.27 9225 \* 10<sup>-6</sup>) LIVENESS \* SPEECHINESS
- + (3.793979 \* 10<sup>-4</sup>) TEMPO \* ENERGY
- + -(7.492498 \* 10<sup>-3</sup>) LOUDNESS \*ACOUSTICNESS
- + (7.001209 \* 10<sup>-2</sup>) INSTRUMENTALNESS \* VALENCE
- + -(6.132415 \* 10<sup>-4</sup>) VALENCE \* LOUDNESS

TEMPO APPROXIMATED USING ALGORITHMIC ESTIMATION. (VALUES IN RANGE: [0.0,1.0])

**ENERGY:** ENERGY FROM LISTENER POINT OF VIEW (VALUES IN RANGE: [0.0,1.0])

**VALENCE**: MEASURE OF THE EMOTIONAL CONTENT OF A SONG (VALUES IN RANGE: [0.0, 1.0])

WE WANT TO RESEARCH THE QUESTIONS:

- 1. CAN WE DERIVE THE VALUES OF THE SUBJECTIVE ATTRIBUTES FOR A GIVEN SONG?
- 2. CAN WE LEARN UNDERLYING GENRES OF SONGS BASED ON THESE GIVEN MUSIC ATTRIBUTES THAT WE DERIVED IN SUPERVISED LEARNING?

-0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 -3.279225e-06 \* valence \* I(tempo^2) + -0.0008160651 \* energy \* I(loudness^2) + 0.3218223 \* acousticness \* I(energy^2) +

+  $(1.359087 * 10^{-2})$  \* ENERGY \* LOUDNESS \* VALENCE<sup>2</sup>)

# 2. Supervised: K-Nearest Neighbors

#### REGRESSION

PREDICTED A SONG'S DANCEABILITY, ENERGY, & VALENCE, BASED ON EUCLIDEAN DISTANCE METRIC ON OUR TRANSFORMED DATA

# 3. Unsupervised: Hierarchical Agglomerative Clustering

EMPLOYED GROUP-LINK HAC FROM MATLAB & JUDGED CLUSTER PURITY BY LOOKING AT GENRE DISTRIBUTION WITHIN EACH CLUSTER

# ANALYSIS & CONCLUSION

## -KNN RESULTS -





# -TAKEAWAY FROM RESULTS

TRANSFORMATION HELPED US LEARN RELATIONSHIPS AMONG DATA ATTRIBUTES

RELATIVELY FEW "PURE" CLUSTERS WITH DISTINCT GENRES: WE ARE NOT CONSIDERING ALL ATTRIBUTES THAT DEFINE GENRE

# -CHALLENGES FACED.

ATTRIBUTES WERE NOT PROPERLY DEFINED AND WERE DISCOVERED TO HAVE ALL BEEN LEARNED THROUGH MACHINE LEARNING ALGORITHMS

PROVIDED MOTIVATION TO CHANGE OUR PROJECT TO RE-DERIVING ECHO NEST, TO UNDERSTAND OUR DATA



8.70%





#### - Future Work-

ATTEMPT TO FIND THE MISSING ATTRIBUTE FOR BETTER CLUSTERING

CURRENT IDEAS: LYRIC OR SENTIMENT ANALYSIS TO "INTERACT" WITH OUR CURRENT ATTRIBUTES

CONSIDER USING WEIGHTED KNN TO FACTOR IN SONG DISTANCES FOR PREDICTIONS