
Università degli Studi di Padova – Dipartimento di Ingegneria dell'Informazione
Corso di Laurea in Ingegneria dell'Informazione

Relazione per la prova finale

***«Sonic seasoning: esperimento di valutazione percettiva di un
modello computazionale per l'associazione tra musica e gusto»***

Relatore: Prof. Antonio Rodà
Correlatore: Dott. Matteo Spanio

Laureanda: Valentina Frezzato
Matricola n.: 2031105

Padova, 13/03/2026

Corrispondenze cross-modali

La mente umana stabilisce **associazioni sistematiche** tra diversi domini sensoriali

Le associazioni **musica-gusto** dipendono da caratteristiche intrinseche della musica

Sonic seasoning: uso intenzionale della musica per enfatizzare determinate caratteristiche gustative

Intelligenza artificiale

I **modelli multimodali** lavorano con dati provenienti da domini diversi e mirano a colmare il divario tra uomo e macchina

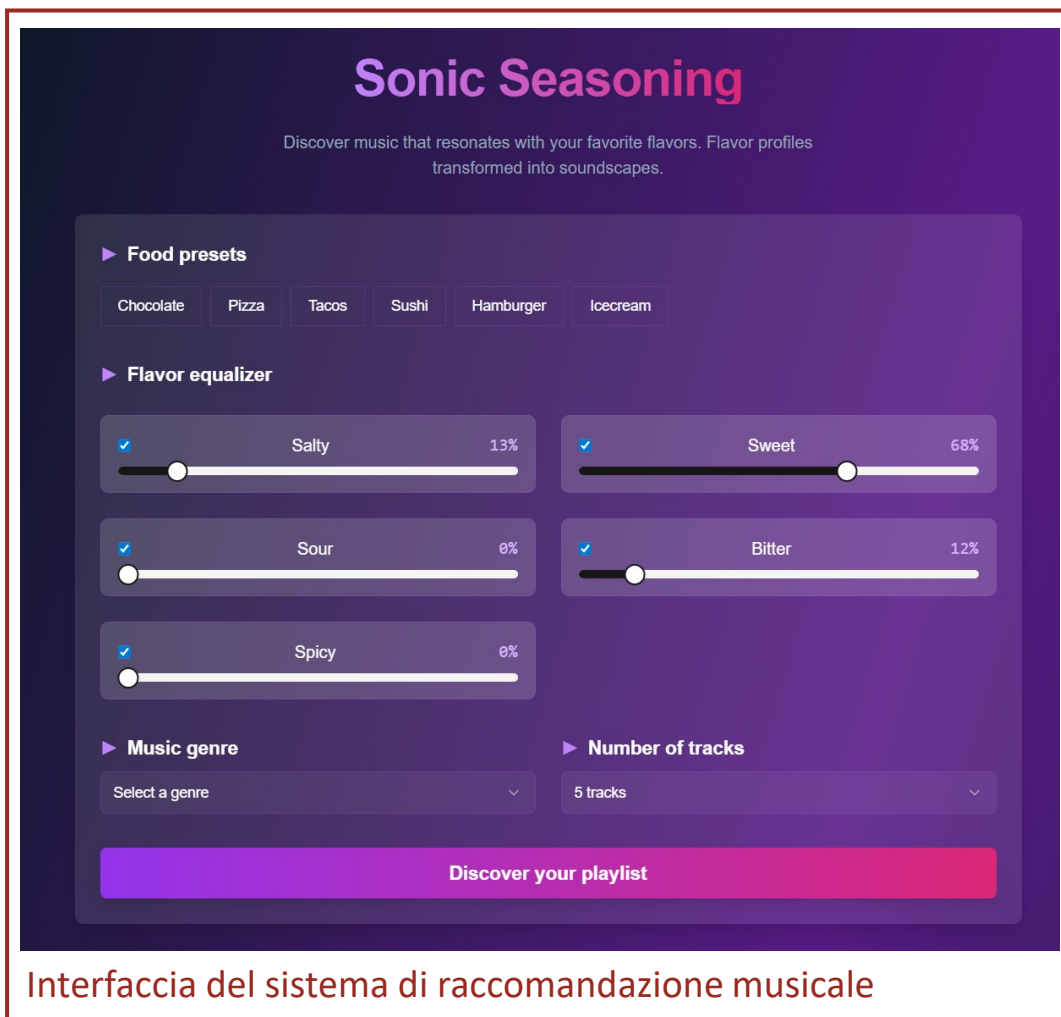
Tecniche di **ML** in grado di **estrarre informazioni dalla musica** e apprendere associazioni con altre modalità sensoriali

Potenziale sviluppo e utilizzo di modelli cross-modali per **esperienze sensoriali** applicate alla ristorazione

Un modello computazionale è in grado di catturare corrispondenze cross-modali tra musica e gusto?

Questo lavoro ha come scopo analizzare, tramite un esperimento di valutazione percettiva, il potenziale di un modello computazionale cross-modale nel generare associazioni sistematiche tra stimoli musicali e sensazioni gustative, ed è così strutturato:





Sistema di raccomandazione musicale:

Per la selezione dei brani è stato utilizzato il sistema di raccomandazione musicale per il sonic seasoning sviluppato da Rodriguez (2025):

- Ha sviluppato e utilizzato un modello di **deep learning** che integra tecniche **MIR** per predire associazioni tra musica e gusto su un dataset di brani provenienti dal FMA
- Ogni brano del dataset è stato associato a un **vettore di valori [0,1]** che rappresenta il grado di corrispondenza con i **5 gusti**
- Un **algoritmo di raccomandazione** seleziona i brani dal dataset etichettato per generare playlist personalizzate in base al profilo gustativo selezionato dall'utente

Selezione dei brani per l'esperimento:

- Sono stati scelti 20 alimenti, ognuno associato a un profilo gustativo rappresentato dal **vettore target**
 $\mathbf{t}_i = (\text{salato}, \text{dolce}, \text{acido}, \text{amaro}, \text{piccante})$
- I vettori sono stati inseriti nel sistema generando 20 brani
- Ogni brano i è associato al vettore \mathbf{t}_i

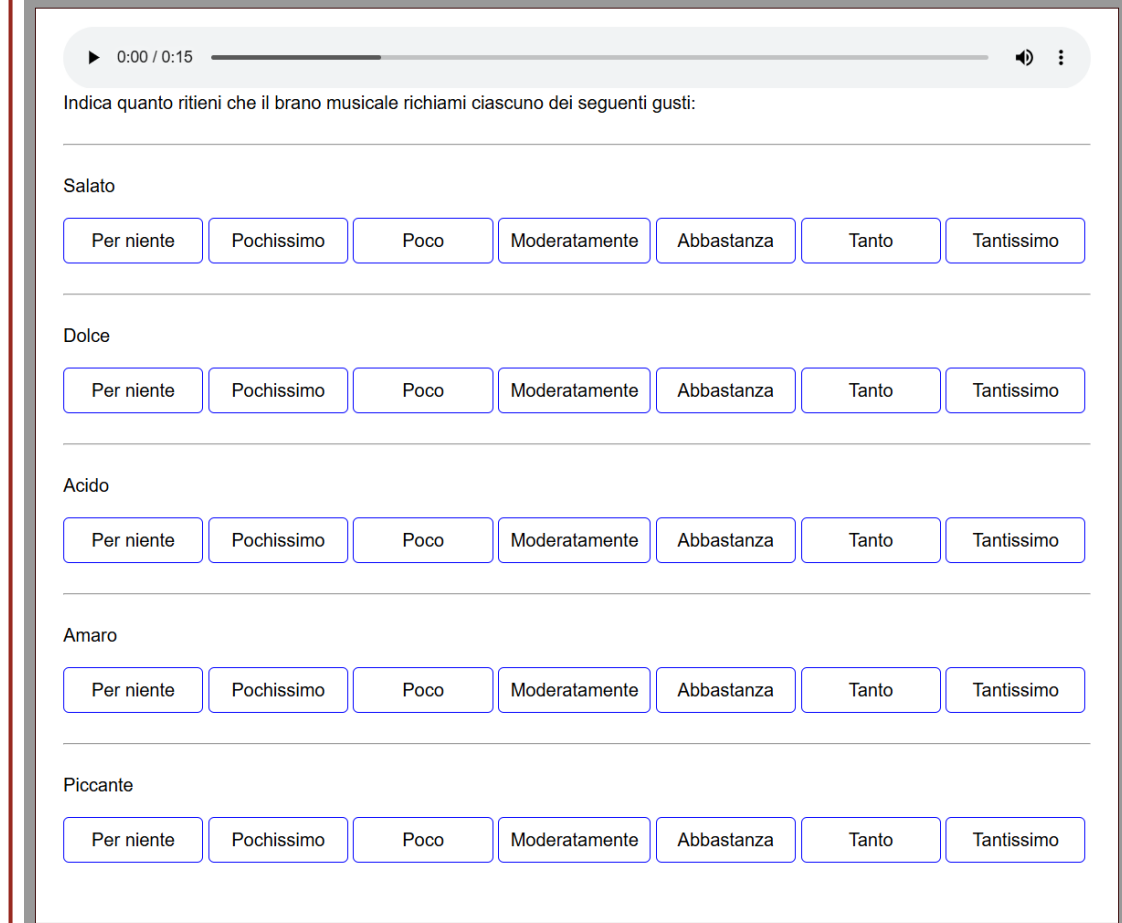
Descrizione del questionario:

- L'esperimento è stato condotto tramite un **questionario online**, progettato per analizzare le relazioni tra i brani musicali e la percezione dei gusti attraverso un **compito di ascolto**
- Il compito consisteva nell'ascolto di una traccia musicale e nella sua successiva valutazione attraverso una **scala Likert a 7 punti**
- Ogni partecipante j ha fornito un **vettore percettivo** $p_{ij} = (\text{salato}, \text{dolce}, \text{acido}, \text{amaro}, \text{piccante})$ per ciascuna traccia i ascoltata
- Al termine del questionario, ogni partecipante ha valutato un totale di 10 tracce, selezionate casualmente dall'insieme delle 20 disponibili

Descrizione del campione:

- Numero di partecipanti = 49
- Età media = 41 anni

Musica e gusto | 10% della ricerca completato

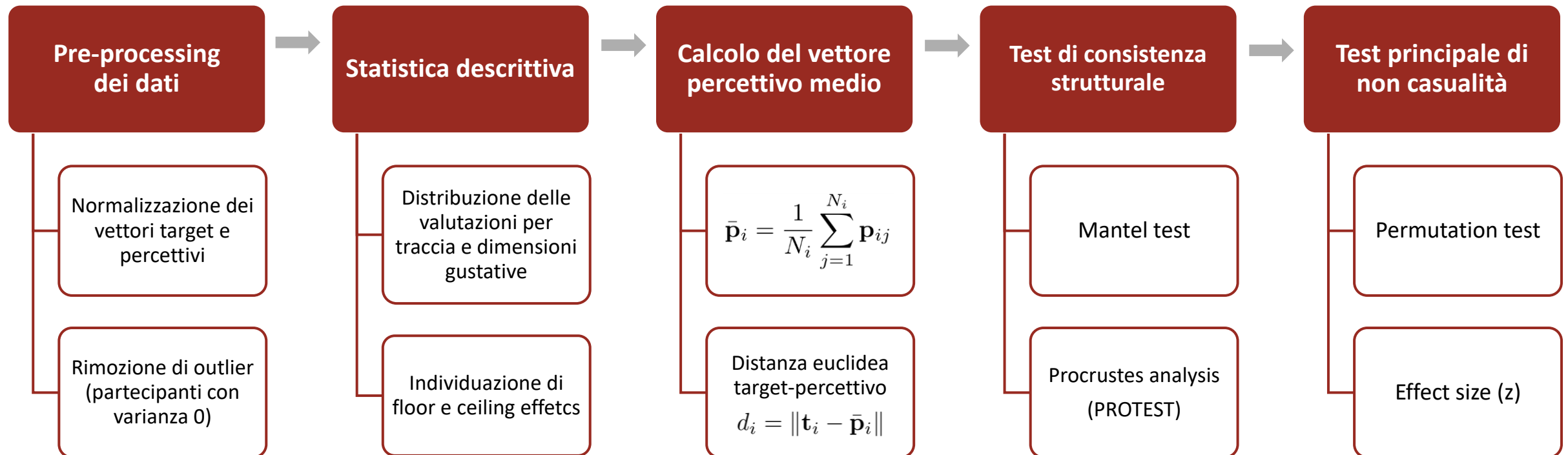


The screenshot shows a music player interface with a progress bar at 0:00 / 0:15. Below the player, the instruction reads: "Indica quanto ritieni che il brano musicale richiami ciascuno dei seguenti gusti:". There are five taste categories, each with a 7-point Likert scale:

- Salato**: Per niente, Pochissimo, Poco, Moderatamente, Abbastanza, Tanto, Tantissimo
- Dolce**: Per niente, Pochissimo, Poco, Moderatamente, Abbastanza, Tanto, Tantissimo
- Acido**: Per niente, Pochissimo, Poco, Moderatamente, Abbastanza, Tanto, Tantissimo
- Amaro**: Per niente, Pochissimo, Poco, Moderatamente, Abbastanza, Tanto, Tantissimo
- Piccante**: Per niente, Pochissimo, Poco, Moderatamente, Abbastanza, Tanto, Tantissimo

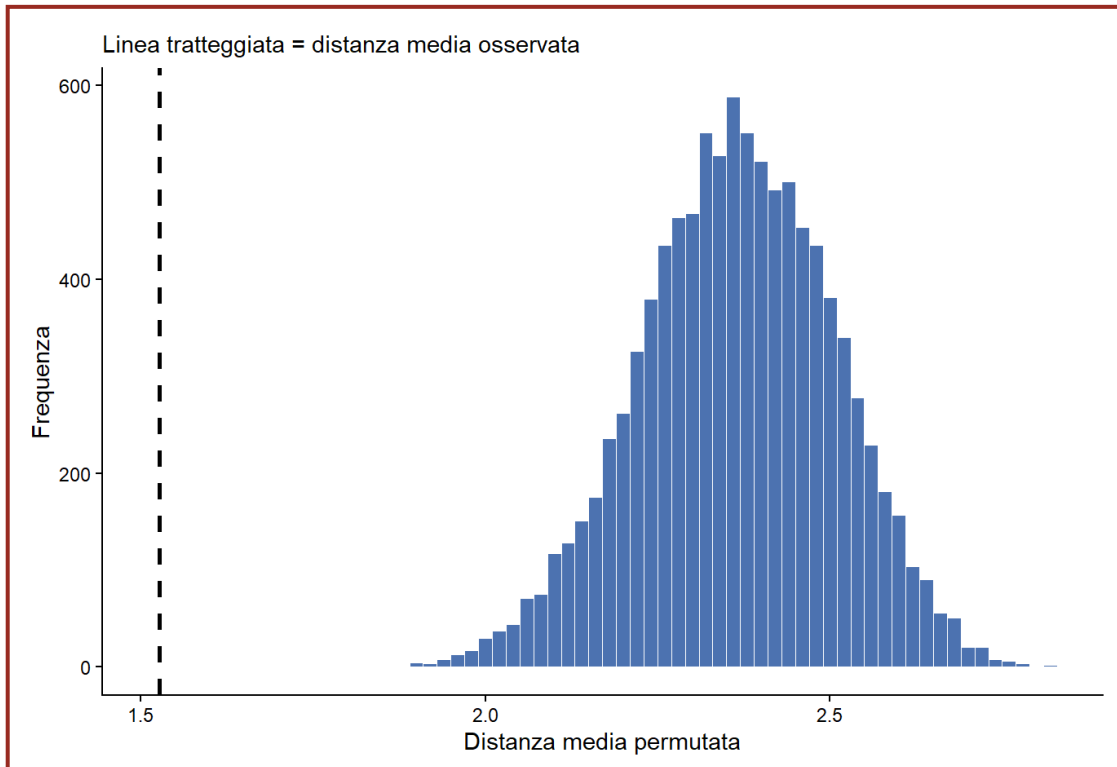
Interfaccia del questionario online

L'obiettivo dell'analisi è verificare l'esistenza di una **relazione sistematica** tra **vettori target** e **vettori percettivi**



Test	Risultati	Commento
Permutation test	$p \ll 0.001$, $d_{\text{obs}} < d_{\text{perm}}$	La similarità osservata tra spazio target e percettivo è maggiore di quella attesa per caso
Effect size (z)	ES = 5.913	ES elevato - la distanza media osservata si colloca a diverse DS al di sotto della media della distribuzione nulla delle distanze permutate
Mantel test	$r = 0.4519$, $p = 0.0001$	Correlazione positiva e statisticamente significativa tra le matrici di distanza nei due spazi
Procrustes (PROTEST)	$m^2 = 0.5113$, $p = 0.0001$	Sovrapposizione significativa, ma parziale, tra le due configurazioni spaziali

Distribuzione nulla delle distanze medie permutate



TEST PRINCIPALE DI NON CASUALITÀ

Permutation test:

- È stata calcolata una distribuzione nulla delle distanze medie permutate
- Il p-value è stato calcolato come la proporzione di distanze medie permutate che sono minori o uguali alla distanza media osservata

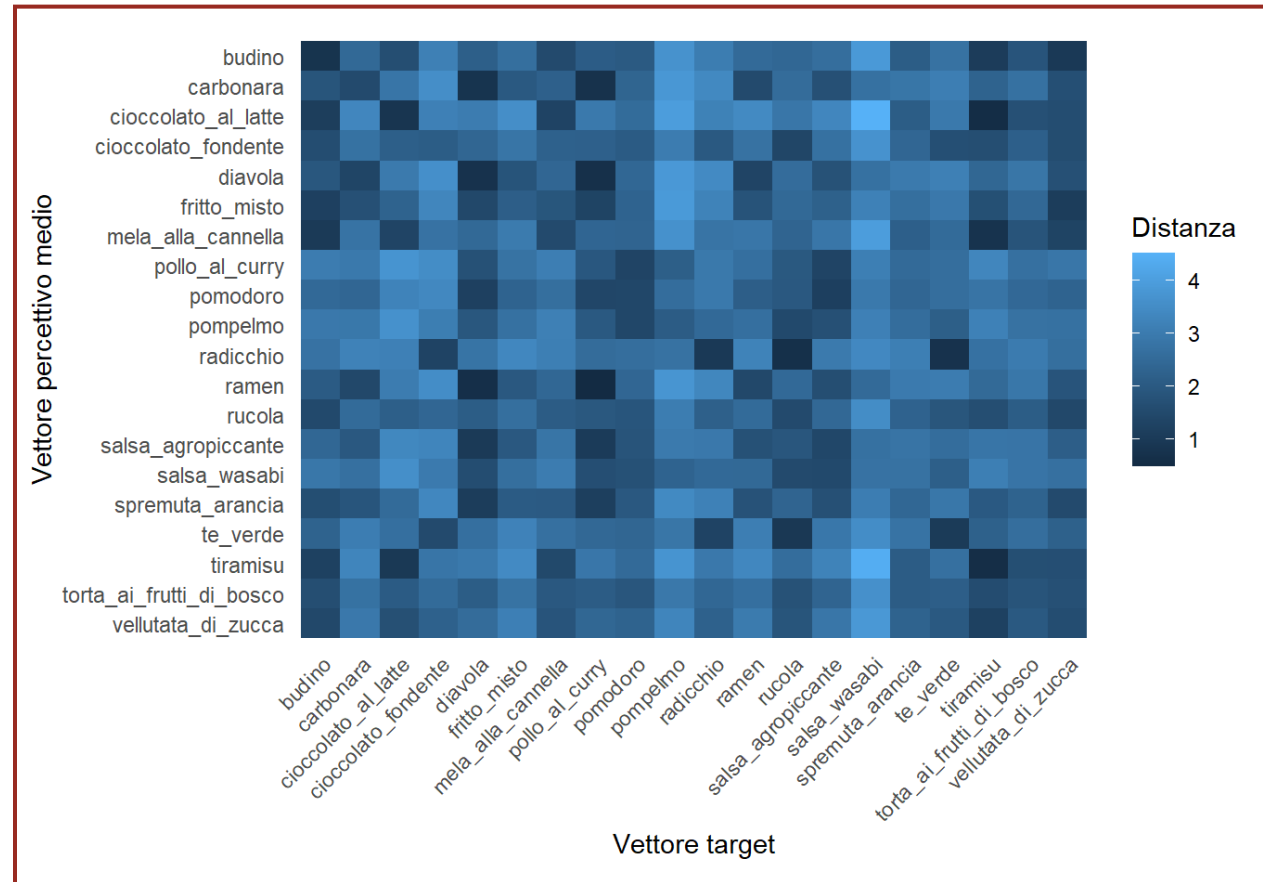
Cosa suggeriscono i risultati?

La distanza media osservata è significativamente minore di quella ottenuta sotto l'ipotesi di associazione casuale

=> Esiste una relazione sistematica tra i vettori target e i vettori percettivi medi

Distanza media osservata	Distanza media permutata
1.5271	2.3643

Heatmap delle distanze tra vettori percettivi medi e target:
 la diagonale principale indica la distanza tra vettore percettivo e il rispettivo target



TEST DI CONSISTENZA STRUTTURALE

- Verifica che la struttura geometrica dei brani sia preservata nello spazio target e percettivo
- La relazione tra gli spazi è visualizzabile tramite la **heatmap** delle distanze, che conferma i risultati di Mantel test e Procrustes analysis

Cosa suggeriscono i risultati?

I match sulla diagonale della heatmap sono parziali:

- alcuni vettori percettivi hanno una distanza minima coi rispettivi target
- ci sono dei mismatch tra vari brani con profili gustativi simili

⇒ **C'è una struttura condivisa tra lo spazio target e percettivo, seppur senza una corrispondenza precisa a livello dei singoli brani**

Il sistema di raccomandazione musicale utilizzato è in grado di apprendere e predire associazioni sistematiche tra musica e gusto, seppur senza una corrispondenza precisa a livello dei singoli brani.



Questo disallineamento potrebbe essere dovuto all'utilizzo di un **dataset limitato**.
Sviluppi futuri in questo campo potrebbero concentrarsi sull'**ampliamento dei dataset musicali**, che descrivano in maniera più approfondita le associazioni tra musica e gusto.



Nel complesso, i risultati supportano l'idea che i modelli computazionali siano in grado di catturare corrispondenze cross-modali tra musica e gusto.

Ciò rafforza il potenziale applicativo dell'intelligenza artificiale in contesti reali, come nel sonic seasoning, e apre prospettive interessanti per la sua integrazione nel settore gastronomico.