



UNIVERSITÀ DI PARMA

Dipartimento di Ingegneria e Architettura

Corso di Laurea in Ingegneria Informatica, Elettronica e delle Telecomunicazioni

Analisi audio per la classificazione di brani musicali tramite clustering e deep learning

Audio analytics for music classification using clustering and deep learning

Relatore:

Prof. Michele Tomaiuolo

Tesi di Laurea di:

Manuel Tanzi

ANNO ACCADEMICO 2021/2022

“In che modo l’intelligenza artificiale può influire sull’esperienza musicale umana, se non fa affidamento sui dati degli *utenti* o sulla *popolarità* dei brani? Come possiamo costruire un sistema di raccomandazione musicale che catturi l’essenza dell’arte musicale senza l’aiuto di fattori esterni?”.

In questa tesi di laurea, si è condotto uno studio finalizzato alla creazione di un **sistema di raccomandazione musicale** che si basa unicamente sull’**analisi audio** della musica. Lo scopo principale del lavoro è stato quello di approfondire le tecniche di analisi audio, al fine di sviluppare un sistema che consiglia la musica in modo oggettivo, basandosi sullo stile e la qualità della musica stessa, e non sulla somiglianza con altri utenti, la popolarità della canzone o la novità che fa scalpore tramite i media. In altre parole, il sistema di raccomandazione è stato progettato per selezionare la musica in base alle sue caratteristiche intrinseche, piuttosto che a fattori esterni che potrebbero influenzare la scelta dell’utente. Nello specifico, la tesi si è concentrata sull’analisi delle tecniche di analisi audio, compresi gli algoritmi di estrazione di caratteristiche musicali, le tecniche di **clustering** e la creazione di modelli di **apprendimento automatico** per la raccomandazione musicale.

L’obiettivo finale della ricerca è stato quello di creare un sistema di raccomandazione musicale efficace e preciso, basato sull’analisi oggettiva della musica stessa, per soddisfare le esigenze degli utenti che preferiscono la selezione musicale basata sulla qualità e sullo stile.

Il processo di implementazione di un sistema di raccomandazione musicale è stato schematizzato in cinque fasi principali:

1. Consiste nella raccolta dei dati tramite le *API* di due importanti piattaforme musicali: **Spotify** e **Genius**. I dati raccolti includono informazioni riguardanti gli *artisti*, i *generi* musicali, i *testi* delle canzoni e le caratteristiche musicali come il *mood*, le *proprietà* e il *contesto* della canzone.
2. Prevede l’**analisi dei dati** raccolti nella fase precedente. In questa fase, sono state utilizzate tecniche di analisi dei dati per comprendere meglio le preferenze musicali degli utenti.
3. **Analisi audio**. In particolare, durante lo studio sono state utilizzate diverse tecniche di analisi degli spettri di frequenza delle canzoni, tra cui il **Mel-Spectrogram**, **CSS**, **ZCR**,

TH e **BH** . Queste tecniche hanno permesso di valutare in modo accurato le caratteristiche delle canzoni, come la loro energia spettrale, la loro timbrica e la loro dinamica, e di utilizzare tali informazioni per creare un sistema di raccomandazione musicale efficace e preciso.

4. Raccomandazione musicale effettiva:

- Utilizzo di **clustering** e la creazione di un **grafo delle adiacenze**, utilizzando le features fornite da Spotify. In seguito, è stato effettuato un **random walking** all'interno del grafo per ottenere le raccomandazioni musicali.
- Utilizzo di una **rete neurale Siamese** che analizza i dati ottenuti tramite l'analisi dell'audio.

5. Creazione di un **sito web** che mostrerà le raccomandazioni musicali. In questo modo, gli utenti potranno accedere alle raccomandazioni musicali e ascoltare la musica consigliata in base alle proprie preferenze.

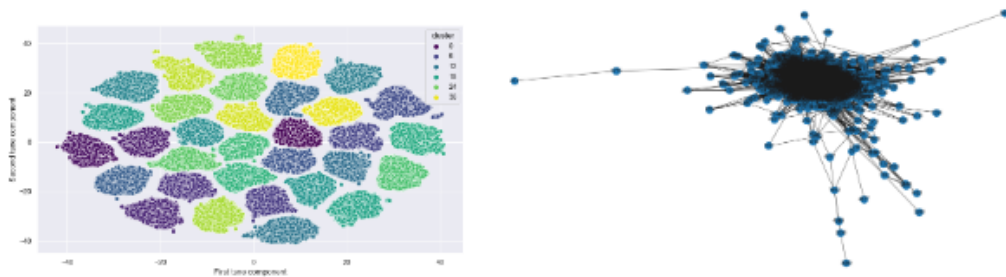


Figura 1: Rappresentazione **t-SNE** dei cluster ottenuti tramite **BIRCH** e successivamente grafo delle adiacenze del cluster 0 dove gli archi sono ottenuti tramite la **distanza di Manhattan**.

1 Conclusioni

Le metodologie analizzate offrono spunti interessanti per migliorare l'efficienza dei sistemi di raccomandazione. L'utilizzo del *grafo delle adiacenze* in combinazione con il *clustering*, infatti, può portare a una rappresentazione più compatta dei dati, con tempi di elaborazione più veloci e una ricerca più efficiente delle raccomandazioni. La bassa dimensionalità dei dati impiegati nel grafo delle adiacenze rappresenta un ulteriore vantaggio, in quanto riduce il rischio di overfitting e migliora la capacità del modello di generalizzare su nuovi dati.

L'implementazione di una *Siamese Neural Network*, grazie alla sua capacità di utilizzare dati a più alta dimensionalità, fornisce risultati più precisi e coerenti rispetto alle tecniche tradizionali. Inoltre, l'uso della **Softmax** come layer di output permette di ampliare la scelta dei brani simili, anche al di fuori di un singolo cluster.

/	Diversità	Novità	Serendipità	Inattesa	Copertura
BIRCH + Graph	68.50 %	79.90 %	37.60 %	21.20 %	3.125 %
SNN	97.20 %	84.60 %	42.10 %	22.70 %	100 %

Figura 2: Tabella dei risultati che mostra le variabili utilizzate per il calcolo e la valutazione dei sistemi di raccomandazione creati. Tale tabella è stata utilizzata per confrontare le performance dei diversi metodi di raccomandazione.