**Caratteristiche**

**Tipo Traccia**: Applicazione Machine Learning

**Nome gruppo**: AGCM

**Componenti**: Alex Lombardi, Mattia Mascia, Giacomo Vecchio, Cosimo Zaccaria

**Data**: 19/09/2024

**Descrizione**: L'obiettivo è applicare diversi modelli di Machine learning per associare ad un utente, in base ai suoi interessi, i prodotti a lui più affini tra quelli presenti nel catalogo di una piattaforma di streaming. Questo verrà fatto allenando la macchina su database pubblici disponibili su Kaggle che associano le caratteristiche dei film o dei prodotti alle valutazioni fornite dagli utenti.

Ciò consentirà di creare una profilazione degli utenti in base alle valutazioni dei titoli visionati e prevedere quali siano i nuovi prodotti più adatti a lui.   
Successivamente l’utente potrà fornire una valutazione dei titoli consigliati e questa verrà aggiunta al database, potenzialmente migliorando le previsioni successive, così da creare un modello evolutivo.

**Passi da seguire**:

1. Raccolta dati da Kaggle
2. Comprensione dei dati
3. Pulizia e riorganizzazione dati
4. Eventuale rimozione outlier
5. Feature engineering (creazione di nuove variabili, come media delle valutazioni per genere o frequenza di visione per utente)
6. Pre-processing dei dati
7. Divisione dataset in training e test set
8. Scelta del modello e sviluppo
9. Addestramento del modello e scelta dei parametri
10. Test sul modello per verificare over/underfitting, accuratezza
11. Implementazione tramite le opinioni di nuovi utenti
12. Creazione di un repository per l’utente finale

* Organizzazione del codice e della documentazione
* Documentazione dettagliata della guida all’uso del modello
* Versionamento e controllo del codice
* Utilizzo di GitHub per la distribuzioni online

**Input richiesti**: Database contenenti informazioni e caratteristiche dei prodotti disponibili su piattaforme di streaming e valutazioni fornite da un campione di utenti. I Dataset che utilizzeremo principalmente saranno: The Movie Dataset e Netflix Prize Data.

**Output attesi**: Accuratezza nel fitting tra utente e prodotto più adeguato ai suoi interessi.

**Note aggiuntive**: Un possibile sviluppo futuro è adattare il modello altri tipi di piattaforme e servizi.