



SWAP
researchgroup

Tesi in Metodi per il Ritrovamento dell'informazione **Sustainability of RecSys**

Relatore: Prof. Pasquale Lops
Relatore: Prof. Cataldo Musto
Laureando: Emanuele Fontana

Università degli Studi di Bari Aldo Moro

Sostenibilità e AI

- **Soddisfare bisogni senza compromettere il futuro**
- Mira a garantire un futuro migliore (es. Agenda 2030 dell'ONU).
- La **Green AI** sviluppo di modelli AI che considerano il costo computazionale e l'impatto ambientale.
- La **Red AI** sviluppo di modelli sempre più complessi senza considerare le risorse impiegate.

RecSys - Introduzione

- Software che suggerisce all'utente elementi di interesse basandosi sulle preferenze e i comportamenti passati.
- Migliorano l'esperienza utente
- Utilizzano AI



Figura: Alcuni famose piattaforme che utilizzano sistemi di raccomandazione

RecSys - Tipologie

- **Collaborative Filtering:** basato sulle preferenze degli utenti (user-user, item-item)
- **Content-based Filtering:** basato sul contenuto degli item.
- **Knowledge-based:** basato su conoscenza esterna (es. knowledge graph)
- **Hybrid:** combinazione delle precedenti.

Reserch Questions

- **RQS1:** Qual è il trade-off tra emissioni e performance dei modelli di RecSys?
- **RQS2:** Lavorare con un criterio di early-stopping basato anche sulle emissioni migliora il trade-off?
- **RQS3:** Quali criteri possono essere utilizzati per migliorare il trade-off?

Lavoro svolto

Per rispondere alle domande di ricerca sono state svolte le seguenti attività:

- **Benchmarking:** Addestramento di modelli di RecSys e misurazione delle emissioni
- **Addestramento sostenibile:** Studio del criterio di early-stopping
- **Criteri di miglioramento:** Studio di criteri per migliorare il trade-off

Sono state utilizzate le librerie **RecBole** e **CodeCarbon**.

$$emission = CI \cdot PC$$

$$CI = \sum_{s \in S} e_s \cdot p_s$$

Dataset Utilizzati

- **MovieLens**: dataset di recensioni di film (1M,10M)
- **Amazon-Books**: dataset di recensioni di libri
- **LastFM**: dataset di ascolti musicali

Modelli SOTA Utilizzati

- **Modelli di raccomandazione generali CF:** BPR, CFKG, DMF, KGNNLS, LINE, MultiDAE, LightGCN, ItemKNN
- **Modelli di raccomandazione basati su conoscenza:** CKE, KGCN, NCF, DGCF

Benchmarking - Risultati emissioni

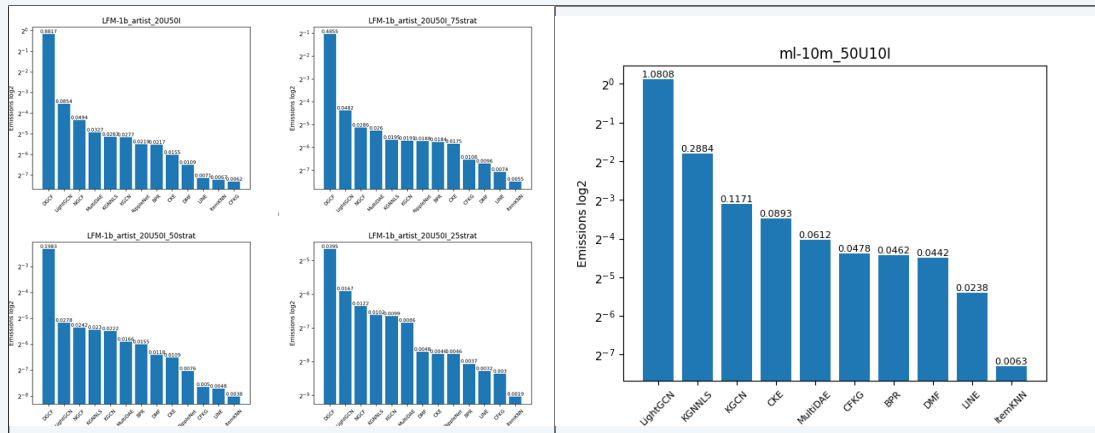


Tabella: Emissioni di CO2 per i vari modelli

Benchmarking - Trade Off

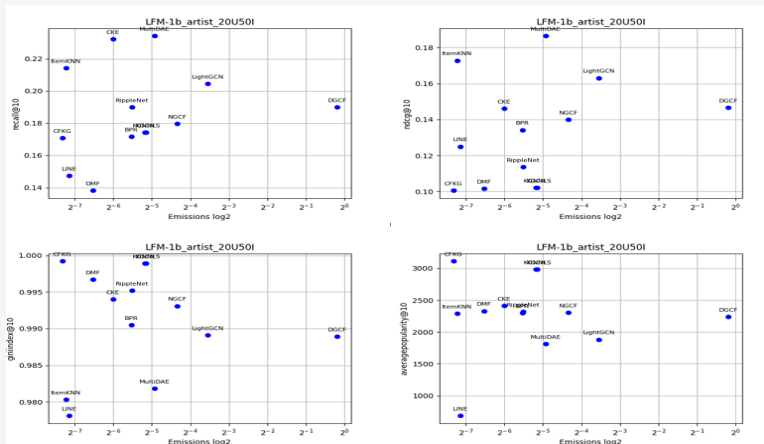
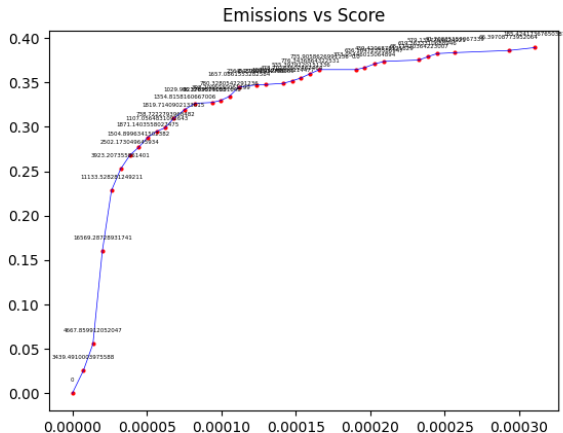


Figura: Esempio di trade-off tra emissioni e performance

Addestramento sostenibile - Introduzione

Approssimazione
della derivata della curva:

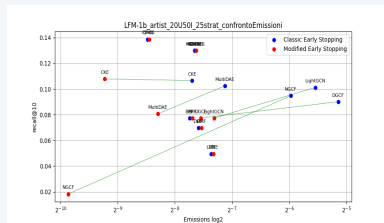
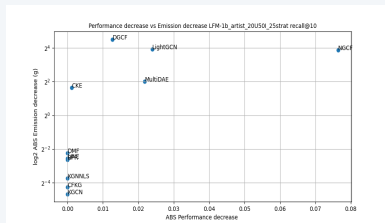
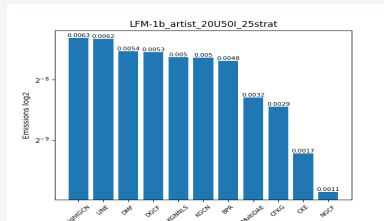
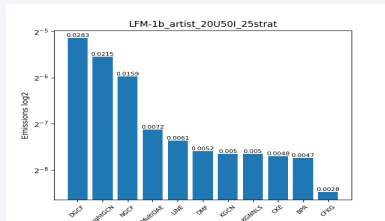
$$\frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}$$



Addestramento sostenibile - Esplorazione

- **Scopo:** capire se è possibile addestrare un modello di raccomandazione in modo sostenibile
- **Esperimento 1:** MovieLens1M con soglia 50 e 5 epoche
- **Esperimento 2:** LastFM con soglia 30 e 7 epoche
- **Esperimento 3:** Amazon_Books con soglia 40 e 6 epoche
- **Conclusioni:** Alcuni modelli (es. DGCF) sono molto sensibili al nuovo criterio, altri (es. DMF) meno

Addestramento sostenibile - Esempi di risultati



Addestramento sostenibile - Confronto criteri

- **Scopo:** confrontare i criteri di addestramento sostenibile (dataset MovieLens1M)
- **Esperimento 1:** Soglia 40, 5 epoche
- **Esperimento 2:** Soglia 30, 5 epoche
- **Esperimento 3:** Soglia 40, 6 epoche
- **Esperimento 4:** Soglia 30, 6 epoche
- **Esperimento 5:** Soglia 40, 7 epoche
- **Esperimento 6:** Soglia 30, 7 epoche

Addestramento sostenibile - Esempio sensibilità

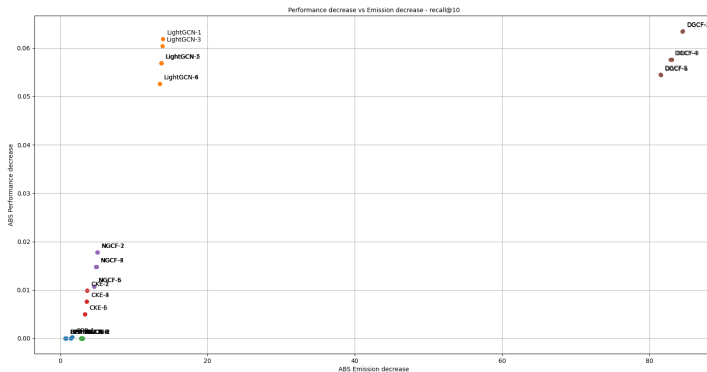


Figura: Sensibilità dei parametri con metrica Recall@10

Addestramento sostenibile - Risultati confronto criteri

Modello	Parametro più impattante	Migliori risultati
BPR	Soglia	Soglia 40 e 6 epoche
CFKG	Soglia	Soglia 40 e 6 epoche
CKE	Epoche consecutive	Soglia 40 e 6 epoche
DMF	Nessuno predominante	Soglia 40 e 7 epoche
KGCN	Epoche consecutive	Soglia 40 e 5 epoche
KGNNLS	Soglia	Soglia 40 e 5 epoche
LINE	Soglia	Soglia 40 e 7 epoche
MultiDAE	Soglia	Soglia 40 e 7 epoche
LightGCN	Soglia	Soglia 40 e 6 epoche
NGCF	Epoche consecutive	Soglia 40 e 5 epoche
DGCF	Epoche consecutive	Soglia 40 e 6 epoche

Tabella: Parametri più impattanti e migliori risultati per ciascun modello

Tipo di Modello	Parametro predominante	Numero di Modelli	Modelli
Collaborative Filtering	Soglia	5	BPR, DMF, LightGCN, MultiDAE, LINE
Collaborative Filtering	Epoche	2	NGCF, DGCF
Knowledge Aware	Soglia	2	CFKG, KGNNLS
Knowledge Aware	Epoche	2	CKE, KGCN

Conclusioni

Benchmarking

Si dimostra come spesso i modelli più complessi hanno emissioni maggiori non giustificate da un miglioramento delle performance elevato.

Addestramento sostenibile

E' possibile ridurre le emissioni di un modello di raccomandazione senza perdere in modo significativo di performance

Sviluppi futuri

Benchmarking

E' necessario effettuare più esperimenti variando dataset, modelli e hardware per avere una visione più completa del problema.

Addestramento sostenibile

Eseguire più esperimenti con altri dataset e altri hardware per confermare o meno i risultati ottenuti.

Iperparametri

Tutti gli esperimenti sono stati effettuati con iperparametri di default. Dunque tutta la fase di benchmarking e di addestramento sostenibile potrebbe essere rivista anche in termini di ricerca degli iperparametri migliori.

Grazie per l'attenzione!

Relatore: Prof. Pasquale Lops
Relatore: Prof. Cataldo Musto
Laureando: Emanuele Fontana

Università degli Studi di Bari Aldo Moro