



UNIVERSITÀ  
DEGLI STUDI DI BARI  
ALDO MORO

**SWAP**  
researchgroup

**Tesi in Metodi per il Ritrovamento dell'informazione**

**ANALISI DELLA SOSTENIBILITÀ DEI SISTEMI DI RACCOMANDAZIONE A STATO  
DELL'ARTE E PROPOSTA DI UN ADDESTRAMENTO SOSTENIBILE**

**Relatore: Prof. Pasquale Lops**

**Relatore: Prof. Cataldo Musto**

**Correlatore: Dott. Giuseppe Spillo**

**Laureando: Emanuele Fontana**

Università degli Studi di Bari Aldo Moro

# Sostenibilità e AI

- **Sostenibilità:** Soddisfare bisogni senza compromettere il futuro
- **Sostenibilità ambientale:** mantenimento del capitale naturale
- **Green AI:** considera l'impatto ambientale.
- **Red AI:** non considerara l'impatto ambientale (massime prestazioni)

# Recommender Systems - Introduzione

- Software che suggerisce all'utente elementi di interesse basandosi sulle preferenze e i comportamenti passati.
- Basati su Intelligenza Artificiale.



Figura: Alcuni famose piattaforme che utilizzano sistemi di raccomandazione

# Recommender Systems - Tipologie

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

# Recommender Systems – Alcuni problemi

- Cold Start: difficoltà nel suggerire item a nuovi utenti.
- More of the same: suggerire sempre item simili.
- **Sostenibilità: Ad oggi i sistemi di raccomandazione sono dei modelli Red AI**

## DOMANDA

E' possibile migliorare la sostenibilità di un sistema di raccomandazione?

# Domande di ricerca e lavoro svolto

- **RQ1:** Qual è il trade-off tra emissioni e performance dei modelli di raccomandazione a stato dell'arte<sup>1</sup>? -> **Benchmarking**
- **RQ2:** E' possibile usare un criterio di early-stopping basato sulle emissioni per migliorare il trade-off tra emissioni e performance dei modelli di raccomandazione a stato dell'arte? -> **Addestramento sostenibile**
- **RQ3:** Quali parametri possono essere utilizzati in questi criteri per migliorare il trade-off? -> **Parametri di miglioramento**

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

---

<sup>1</sup>Modelli classici a cui fare riferimento

# Dataset e modelli utilizzati

## Dataset

- **MovieLens**: dataset di recensioni di film
- **Amazon-Books**: dataset di recensioni di libri
- **LastFM**: dataset di ascolti musicali

## Modelli

- **Modelli di raccomandazione Collaborative Filtering**: BPR, DMF, LINE, MultiDAE, LightGCN, ItemKNN, NCF, DGCF
- **Modelli di raccomandazione Knowledge Aware**: CKE, KGCM, KGNNLS, CFKG

# Benchmarking - Emissioni

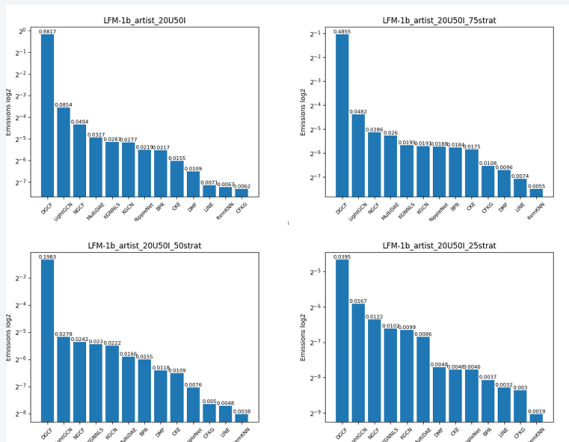


Figura: Emissioni di CO2 per i vari modelli



# Benchmarking - Trade-off

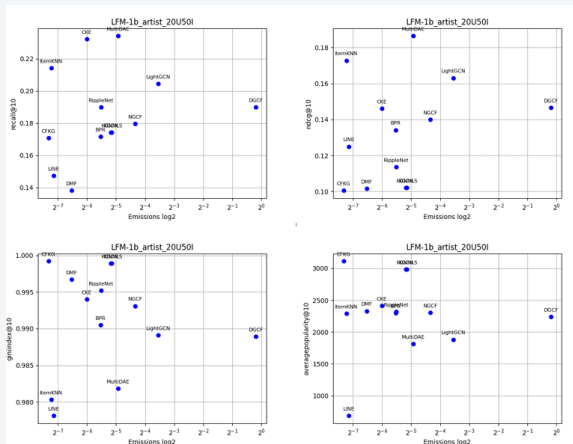


Figura: Trade-off tra emissioni e performance

# Addestramento sostenibile - Introduzione

Sull'asse delle x troviamo le emissioni, sull'asse delle y lo score. Quando la derivata è al di sotto di una certa soglia **S** per un certo numero di epoche consecutive **E** l'addestramento termina (comportamento asintotico). Approssimazione della derivata della curva:

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

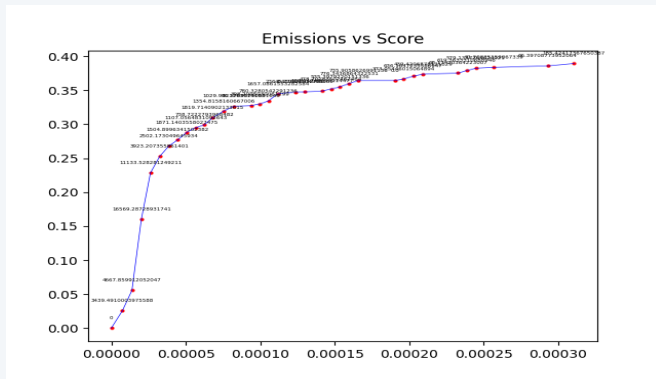


Figura: Andamento score e emissioni

# Addestramento sostenibile - Esplorazione

Esperimento	Dataset	Soglia	Epoche
Esperimento 1	MovieLens1M	50	5
Esperimento 2	LastFM	30	7
Esperimento 3	Amazon_Books	40	6

Tabella: Parametri degli Esperimenti

- **Conclusioni:** Alcuni modelli (es. DGCF) sono molto sensibili al nuovo criterio, altri (es. DMF) meno

# Addestramento sostenibile - Confronto criteri

Fissato il dataset MovieLens1M, sono stati eseguiti 6 esperimenti variando i parametri:

<b>Esperimento</b>	<b>Soglia</b>	<b>Epoche</b>
Esperimento 1	40	5
Esperimento 2	30	5
Esperimento 3	40	6
Esperimento 4	30	6
Esperimento 5	40	7
Esperimento 6	30	7

Tabella: Parametri degli Esperimenti

# 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

# Addestramento sostenibile - Risultati confronto criteri

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

Tabella: Riassunto dei parametri dominanti per tipo di modello

# 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**

**Correlatore: Dott. Giuseppe Spillo**

**Laureando: Emanuele Fontana**

Università degli Studi di Bari Aldo Moro