

Comparison and discussion of two Profile-based Retrieval Systems

José Antonio Ruiz Heredia
Joseph Tartivel
Álvaro Honrubia



INTRODUCTION



LinkedIn

- Domain of **Job Portals**.
- System for **Job Matching**.
- Enhances **job search** and a better **quality of applicants** for specific roles.



Coursera

- Domain of **Online Learning Platforms**.
- System for **Skill Development** based on **personalized courses**.
- Enhances access to **education** and **improves career opportunities** through specialized courses and certifications.

METHOD USED BY LINKEDIN



RETRIEVAL STAGE

A **candidate selection model** filters job listings before ranking them.

Reduce search latency:

- **Decision trees**
- **Query construction**



RANKING STAGE

Sort based on:

- **Explicit user data** (e.g., profile or past applications)
- **Implicit signals** (e.g., users' clicks.)

The system applies machine learning models like **GLMix** to score and rank jobs.



FORECASTING

LiJAR System balances job applications preventing job postings from receiving too **many** or too **few applications**.

METHOD USED BY COURSERA

Collaborative Filtering (CF)

User-based: Recommends courses based on similar users' behavior.

Item-based: Finds relationships between courses through user interactions.

Profiling

Explicit: Users manually enter their learning goals.

Implicit: The system tracks behavior and updates preferences dynamically (CBF sistem).

Content-Based Filtering (CBF)

Analysis of metadata: keywords, topics, and descriptions from courses.

Builds user profiles from past interactions to recommend relevant courses.

Hybridisation

Weighted Hybrid: Combines CF and CBF with different weights.

Switching Hybrid: Uses CF if data is available, otherwise CBF.

Feature Augmentation: One model enhances another with generated features.

COMPARISON BETWEEN THE SYSTEMS

SYSTEM	PURPOSE	METHOD	DATA	STRENGTHS	CHALLENGES
LinkedIn	Matching users to suitable jobs	Multi-tier distributed service that retrieves and ranks job postings	Past applications, user profile and employer preferences	Scalability & Efficiency Balancing Job Application High availability	Real-time personalization Unique job market constraints
	Recommending personalized courses	Hybrid content-based & collaborative filtering associated to contextual modeling	Educational & professional background, context and learning preferences	Adaptable and personalized learning experience.	Requires continuous user input

CONCLUSIONS



SUMMARY

- LinkedIn uses **ML-based ranking**.
- Coursera uses **hybrid collaborative and content-based filtering**.

CONTINUOUS IMPROVEMENT

Both systems require to **adapt in real time** to avoid over reliance in past actions.

FILTER BUBBLE

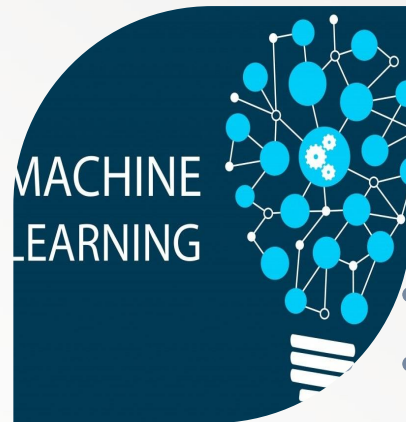
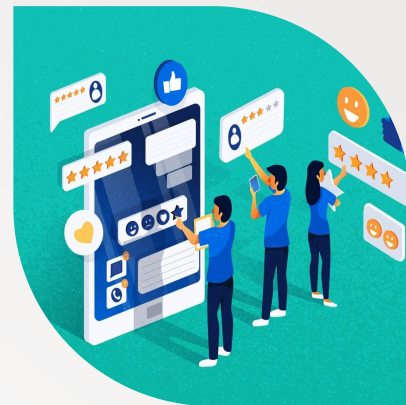
Limit users' exposure to diverse **job** or **learning opportunities**.

BALANCED APPLICATIONS

In **LinkedIn**, jobs must be **balanced** preventing some jobs from being flooded while others remain unseen.

IMPROVEMENT PROPOSALS

RESOURCE	LINKEDIN	COURSERA
Bias Mitigation	Explainable AI to provide transparency.	Adversarial Debiasing to deal with Over-reliance on past user actions.
User Feedback	NLP on user interactions (e.g., messages and posts) to infer job preferences.	NLP-based chatbots to collect more feedback.
Method	Include bidirectional recommendation to also help recruiters to find the best candidates.	Use of NLP to understand better the content and recommend not only based on titles and tags.
Personalization	Integrate external labor market data to improve rankings.	Incorporate users profiles more deeply and AI-driven adaptive learning .





**THANKS FOR YOUR
ATTENTION**

