prospettiveinORGANIZZAZIONE
Rivista trimestrale di Organizzazione Aziendale
ISSN: 2465-1753

ALGORITMI DI INTELLIGENZA ARTIFICIALE DI TIPO NON SUPERVISIONATO A SUPPORTO DELLE ORGANIZZAZIONI NEI PROCESSI DECISIONALI

Pubblicato il14 Aprile 2023 di Rosa Angelo



Categoria: Digital Organization

INTRODUZIONE

L'intelligenza artificiale ha completamente rivoluzionato il modo in cui le organizzazioni operano e gestiscono le proprie attività lavorative. L'IA è stata introdotta in molti settori, tra cui la sanità, la finanza, l'automotive, e viene impiegato sempre di più anche in altre aree, come la gestione delle risorse umane, il marketing, le vendite e, particolarmente, l'analisi dei dati. Infatti, negli ultimi anni i sistemi di intelligenza artificiale hanno assunto un ruolo chiave nella gestione delle risorse umane. Tali algoritmi hanno consentito alle organizzazioni di analizzare una gran mole di dati, ad esempio quelli relativi ai dipendenti, per identificare pattern e trend di non immediata interpretazione. Ciò consente alle organizzazioni di prendere decisioni più oculate nella gestione delle risorse umane, come ad esempio la pianificazione strategica delle assunzioni, la valorizzazione dei talenti e la valutazione delle performance. In questo contesto, gli algoritmi di intelligenza artificiale di tipo non supervisionato vengono impiegati per l'ottimizzazione dell'organizzazione delle risorse per il miglioramento dell'efficienza, della produttività e della soddisfazione dei dipendenti. Partendo, quindi, da grandi basi di dati, è possibile analizzare in maniera del tutto indipendente processi e contesti lavorativi per il

supporto ai decision makers organizzativi.

Alcuni degli approcci tipicamente usati per l'analisi di *dataset* con molti attributi sono i modelli di intelligenza artificiale di tipo non supervisionato [1]. In particolare, è opportuno menzionare due tipologie di algoritmi di *clustering*: il k-Means e il Fuzzy c-Means. Partendo dall'analisi dei risultati del *clustering*, si vuole descrivere il metodo deduttivo di gestione dell'organizzazione e di progettazione dei processi di gestione dell'organizzazione aziendale: l'analisi di *clustering* consente quindi di supportare i dirigenti ad un migliore gestione delle risorse umane, attivando procedure di ingegnerizzazione dei processi aziendali. Partendo dunque da un'analisi dei processi 'AS IS' si è in grado di strutturare i nuovi processi 'TO BE' automatizzando il *decision making* per mezzo dell'analisi degli algoritmi di intelligenza artificiale. Tale approccio di automazione e di ottimizzazione dei processi prende il nome di '*Process Mining*', ed è ad oggi un argomento emergente di ricerca nel campo dell'Ingegneria Gestionale con particolare applicazione ai processi aziendali di Industria 5.0 [2]. Il Process Mining è una tecnica di analisi dei processi che parte dall'analisi dei dati e consiste nell'analizzare in maniera capillare ed estremamente dettagliata i processi aziendali per come si verificano nella realtà quotidiana, con l'obiettivo di analizzarli, mapparli, scoprirne punti di forza e debolezza e scostamenti rispetto ai processi standard codificati per policy, e quindi migliorarli senza dover passare per lunghi e costosi progetti di Business Process Reengineering (BPR), che risulta essere l'approccio tradizionale alla riorganizzazione dei modelli di funzionamento delle imprese e che nel tempo, visto le turbolenti determinanti strutturali esterne, ha mostrato diversi limiti.

L'obiettivo del seguente lavoro è di fornire elementi utili a comprendere come applicare in modo efficiente gli algoritmi ad apprendimento non supervisionato utilizzando metriche e analisi aggiuntive, per poter poi ipotizzare dei processi ottimizzati di gestione delle risorse umane. Il metodo proposto si basa sulla sui seguenti step:

- Partendo dalla tipologia del *dataset* disponibile, si stabilisce la tipologia di algoritmo da utilizzare (nel caso specifico, essendoci molti attributi si preferisce l'algoritmo con apprendimento non supervisionato);
- Si stabiliscono gli iper-parametri (parametri di s*etting* degli algoritmi di intelligenza artificiale) che ottimizzano le performance dell'algoritmo (nel caso specifico si valuta il coefficiente di Silhouette idoneo a definire il numero di cluster ottimale);

Si analizzano i risultati e si ipotizzano soluzioni per migliorare i processi organizzativi aziendali inerente la gestione delle risorse umane.

DATASET PER IL TESTING DEI MODELLI CLUSTERING

I dati analizzati ed utilizzati per testare i modelli degli algoritmi di *clustering* sono stato estratti da *dataset* di tipo *open* disponibili in rete [3]. Il *dataset* si riferisce ad una classificazione di lavoratori all'interno di una azienda, ipotizzando diverse tipologie di contratto e di modalità di collaborazione. Il *dataset* è caratterizzato dai seguenti principali attributi (caratteristiche del lavoratore):

- Età del lavoratore;
- Attrito (personalità con attrito verso l'altro personale);
- Predisposizione a viaggiare per business;
- Tariffa giornaliera;
- Settore di appartenenza (*Sales, Research and Development, Human Resources*);
- Km di distanza dell'ufficio dalla propria abitazione;
- Livello di istruzione (in scala da 1 a 5);
- Tipologia di istruzione (*Marketing, Life Sciences, Medical, Technical Degree, Human Resources*, altri);
- Tariffa oraria;
- Livello di coinvolgimento lavorativo (in scala da 1 a 4);
- Livello dell'attività lavorativa (in scala da 1 a 5);
- Ruolo (*Manager, Laboratory Technician, Sales Executive, Manufacturing Director, Research Scientist, Healthcare Representative*);
- Livello di soddisfazione del lavoro (in scala da 1 a 4);
- Stato matrimoniale (sposato, celibe/nubile);
- Tariffa media mensile;
- Percentuale di aumento di stipendio;

- Coefficiente di performance (in scala da 1 a 4);
- Livello d soddisfazione nelle relazioni (in scala da 1 a 4);
- Anni lavorativi totali;
- Numero di eventi formativi seguiti durante l'ultimo anno di attività;
- Livello di bilanciamento con la qualità di vita (in scala da 1 a 4);
- Anni lavorativi trascorsi nell'azienda;
- Anni lavorativi inerenti l'attuale ruolo;
- Anni trascorsi dall'ultima promozione;
- Anni trascorsi con l'attuale manager.

RISULTATI DI CLUSTERING K-MEANS E DI FUZZY C-MEANS

Un approccio utilizzato per l'esecuzione degli algoritmi di intelligenza artificiale è quello di utilizzare software ad oggetti, di alto livello, ossia in grado di poter eseguire operazioni di data processing mediante apposite interfacce grafiche senza entrare nel merito del codice, se non in piccola parte. Un metodo per la progettazione di algoritmi di intelligenza artificiale è quello di strutturate tutte le fasi del *data processing* (*data pre-processing, data processing, result visualization*) mediante un opportuno *workflow* come quello mostrato in Fig. 1, dove ogni blocco ha una specifica funzionalità. Gli algoritmi di *clustering* sono stati implementati in questo lavoro utilizzando il *tool software open source The Konstanz Information Miner* [4].

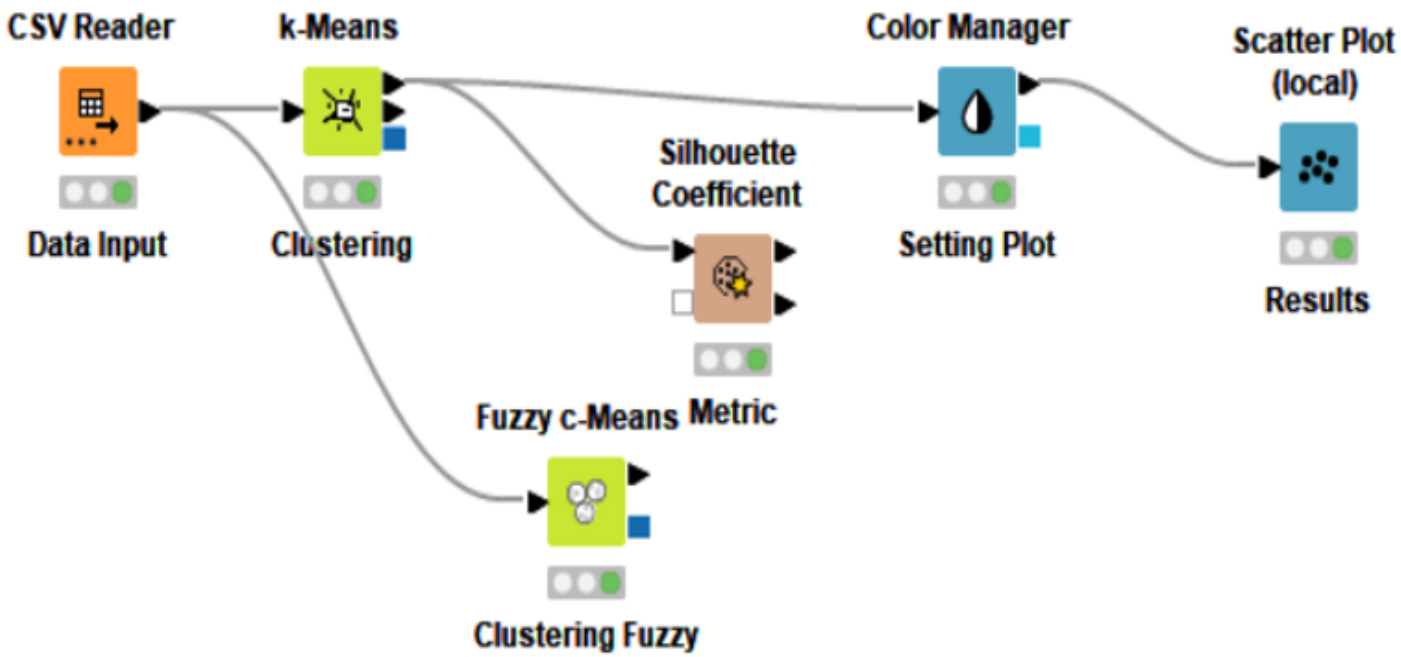


Fig.1 Workflow KNIME per il processamento dei dati mediante algoritmi non supervisionati di K-Means e di Fuzzy C-Means.

Per applicare l'algoritmo di clustering in modo efficiente, è importante stabilire il numero di cluster K che ottimizza l'analisi. Ci sono diverse tecniche per determinare il valore ottimale di K, ma una metrica comune è il coefficiente di Silhouette. Il coefficiente di Silhouette è una metrica che valuta la bontà del clustering, fornendo un valore che varia da -1 a +1, dove i valori più alti indicano un clustering migliore. Il coefficiente di Silhouette [4] valuta la qualità del clustering sulla base della distanza tra gli oggetti nello stesso cluster e la distanza degli oggetti in cluster diversi. In particolare, il coefficiente di Silhouette calcola il rapporto tra la distanza media tra un oggetto e gli altri oggetti nello stesso cluster, e la distanza media tra l'oggetto e gli oggetti del cluster più vicino (b). Un valore di Silhouette vicino a 1 indica che l'oggetto è ben assegnato al suo cluster, mentre un valore vicino a -1 indica che l'oggetto è stato assegnato erroneamente. Un valore di Silhouette vicino a 0 indica che l'oggetto potrebbe essere assegnato in modo ambiguo. Il coefficiente di Silhouette può essere stimato sia in valore medio, che fornisce una valutazione globale della qualità del clustering, sia per ogni cluster, fornendo una valutazione specifica della qualità del clustering per ogni cluster.

Nella tabella 1 sono riportati i valori del coefficiente di Silhouette per 5 casi di *clustering*.

Cluster	K2	K=3	K=4	K=5
Cluster_0	0,468	0,286	0,341	0,345

Cluster	K2	K=3	K=4	K=5
Cluster_1	0,456	0,529	0,284	0,297
Cluster_2	//	0,511	0,490	0,360
Clustrer_3	//	//	0,456	0,512
Cluster_4	//	//	//	0,484
Valore Medio	0,462	0,488	0,437	0,426

Tabella 1. Tabella riassuntiva dei valori dei coefficienti di Silhouette.

Considerando il solo valore medio del coefficiente di Silhouette si riporta in Fig. 2 l'andamento di tale parametro con il numero di cluster K: da tale grafiche si evince che il numero di cluster K=3 risulta ottimale per le performance dell'algoritmo di *clustering*.

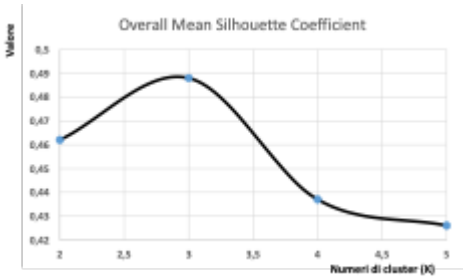


Fig.2 Parametro di medio pesato su tutti i cluster: overlall mean Silhouette coefficient in funzione del numero di cluster (K).

Fissato dunque il parametro K=3, si riportano in Fig. 3 i risultati maggiormente significativi dell'analisi di *clustering,* e nella tabella 2 le deduzioni delle analisi effettuate con relative considerazioni circa i possibili effetti e azioni sull'organizzazione aziendale.

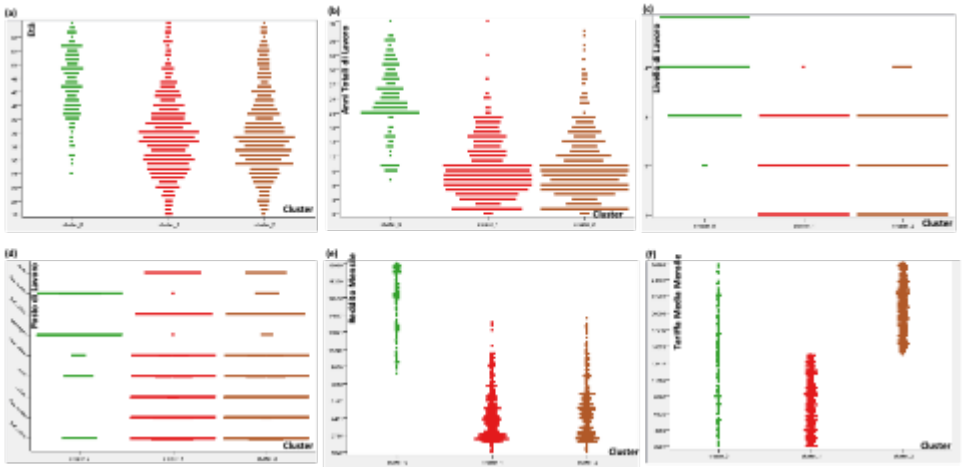


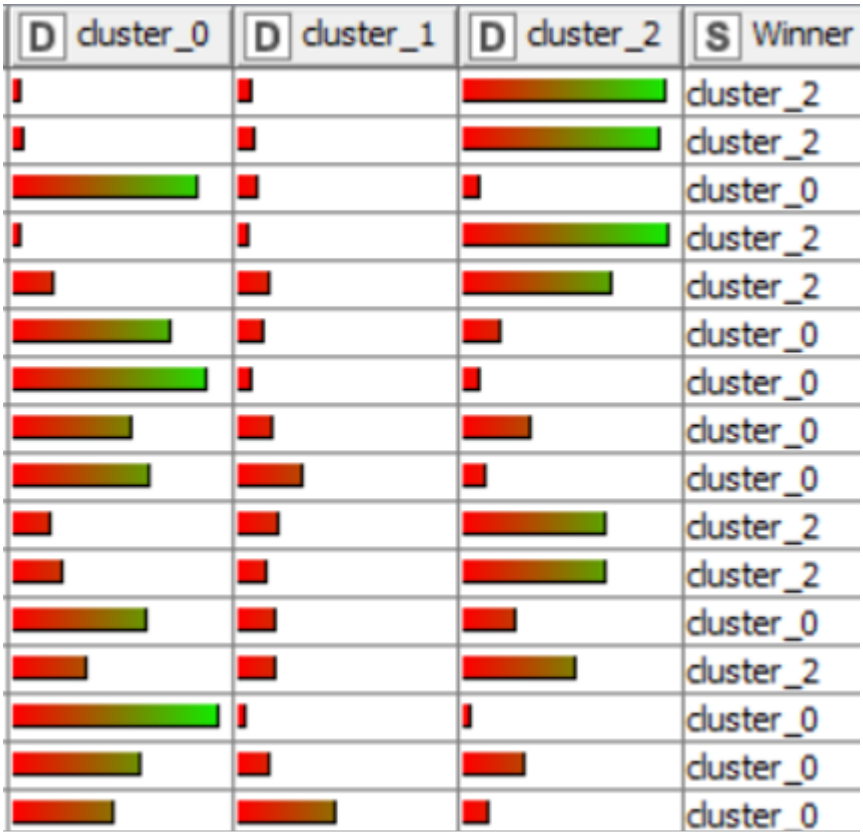
Fig.3 Risultati principali dell'algoritmo di K-Means: cluster in funzione dell'età dei lavoratori (a), degli anni totali di lavoro (b), del livello del lavoro (c), della tipologia di posto di lavoro (d), del reddito mensile (e), e della tariffa media mensile (f).

Cluster	Lettura dei Risultati (Caratteristiche del Cluster)	Possibili Effetti/Azioni sull'Organizzazione Aziendale
Cluster_0	I lavoratori appartenenti a tale cluster sono: - Minori in numero; - Lavorano da più anni nello stesso ambiente lavorativo; - Rivestono un livello di lavoro più alto; - Sono tendenzialmente dei manager e dei ricercatori (settore R&D); - Hanno un reddito mensile più alto; - Hanno una tariffa media mensile variabile per tipologia di attività svolta;	Tendenzialmente tale cluster di lavoratori lavora maggiormente rispetto agli altri. Il loro utilizzo è necessario solo per i livelli maggiormente alti (4 e 5), mentre per le mansioni del livello 3 si potrebbero utilizzare i lavoratori del cluster_1 e cluster_2 incrementando l'efficienza e diminuendo i costi associati all' impiego delle risorse umane. Essendo caratterizzati una maggiore anzianità lavorativa, potrebbero formare i lavoratori degli altri cluster per creare una nuova classe di manager (futura sostituzione della vecchia classe dirigente).
Cluster_1	Sono caratterizzati da una tariffa media mensile bassa per tipologia di attività svolta, e hanno un reddito basso (normale/alto utilizzo mensile di tale cluster).	Il livello di lavoro medio-basso e il basso valore di tariffa mensile, suggerisce di utilizzare tali lavoratori per più mansioni aumentando in generale l'efficienza dell'azienda abbattendo i costi.
Cluster_2	Pur avendo una tariffa media mensile più alta per tipologia di attività svolta hanno un reddito basso (poco utilizzo mensile di tale cluster).	Il livello di lavoro medio-basso e l'alto valore tariffa mensile, suggerisce di utilizzare tali lavoratori per meno mansioni ottimizzando in costi aziendali. L'impiego di tali lavoratori potrebbe essere richiesto nei soli casi di necessità o in periodi particolari di alto carico di lavoro.

Tabella 2. Tabella riepilogativa dei risultati dei modelli applicati al dataset *open data*.

Il modello di Fuzzy C-Means (FCM) è un algoritmo di clustering basato sulla teoria degli insiemi fuzzy. Rispetto al K-Means, il FCM è un metodo "rafforzativo" in quanto consente di ottenere una valutazione più precisa dell'appartenenza di ogni singolo record a un cluster specifico. Nel K-Means, ogni record viene assegnato in modo esclusivo a un singolo cluster, sulla base della distanza dall'oggetto centroide. Ciò significa che un record può appartenere solo a un cluster, anche se presenta caratteristiche comuni con gli oggetti di altri cluster. Al contrario, nel FCM, ogni record viene associato a tutti i cluster in modo proporzionale alla sua vicinanza con i centroidi dei singoli cluster. (vedi Fig. 4). In Fig. 5 si riassume il metodo utilizzato per implementare algoritmi di *clustering* di *decision making* all'interno di un processo decisionale manageriale. Tale *workflow* è in notazione

grafica di tipo *Business Processs Modelling and Notation* (BPMN) [2], notazione tipica per la progettazione dei processi e per la reingegnerizzazione degli stessi.



D cluster_0	D cluster_1	D cluster_2	S Winner
			cluster_2
			cluster_2
			cluster_0
			cluster_2
			cluster_2
			cluster_0
			cluster_0
			cluster_0
			cluster_0
			cluster_2
			cluster_2
			cluster_0
			cluster_2
			cluster_0
			cluster_0
			cluster_0

Fig.4 Risultati del metodo Fuzzy C-Means. Si osserva che l'ultimo valore dell'elenco appartiene al cluster_0 anche se vi è un'alta probabilità che appartenga anche al cluter_1.

RISULTATI E CONCLUSIONI

Questo contributo evidenzia la modalità di integrare gli algoritmi di intelligenza artificiale nei processi decisionali, al fine di ottimizzare l'organizzazione delle risorse umane all'interno delle nuove organizzazioni. Il metodo deduttivo suggerisce dunque di spostare il personale su diverse attività in base all'esigenza e alle competenze, con l'obiettivo di massimizzare l'efficienza dell'operato e di ridurre i costi aziendali (Tabella 1).

L'analisi per cluster (Tabella 2), combinata con le tecniche di K-means e di Fuzzy C-means (Figure 2,3,4) ha evidenziato come i lavoratori appartenenti al Cluster_0 sono caratterizzati da una maggiore presenza sul posto di lavoro rispetto agli altri. Questo suggerisce una maggiore dedizione al lavoro e un'attitudine positiva verso l'impiego delle proprie competenze. Inoltre, essi vengono principalmente impiegati in mansioni di livello più

elevato, mentre per le mansioni di livello 3 si potrebbe utilizzare il personale degli altri cluster, in modo da ottimizzare l'efficienza e ridurre i costi associati all'impiego delle risorse umane. In più, i lavoratori di questo cluster hanno una maggiore anzianità lavorativa, il che indica un'esperienza consolidata nel loro settore professionale. Ciò potrebbe essere impiegato per formare i lavoratori degli altri cluster, creando una nuova classe di manager in grado di sostituire la vecchia classe dirigente ove necessario. In generale, questi risultati suggeriscono che le aziende potrebbero beneficiare di una gestione più attenta delle risorse umane, che tenga conto delle caratteristiche specifiche dei lavoratori e dei loro cluster di appartenenza. In questo modo, sarebbe possibile migliorare l'efficienza aziendale, ridurre i costi e prepararsi al futuro sostituendo in modo graduale la vecchia classe dirigente con professionisti di nuova formazione.

Viceversa, i lavoratori appartenenti al Cluster_1 hanno in media una tariffa mensile bassa per tipologia di attività svolta, e un reddito mensile basso o nella media, nonostante l'utilizzo di tale cluster sia alto o molto alto. Questo suggerisce che potrebbe essere conveniente utilizzare tali professionisti per più mansioni, in modo da aumentare l'efficienza dell'azienda e ridurre i costi associati all'impiego delle risorse umane. Poiché tali lavoratori sono caratterizzati da un livello di complessità professionale medio-bassa, essi potrebbero essere utilizzati per mansioni che richiedono meno competenze, lasciando le mansioni più specializzate a quelli degli altri cluster. Tale strategia potrebbe anche consentire ai lavoratori di acquisire nuove competenze e di crescere professionalmente, migliorando così la loro carriera all'interno dell'azienda. Inoltre, data la propria esperienza lavorativa maggiore rispetto ad altri cluster, potrebbero anche essere destinati a programmi di *mentorship* e formazione per i nuovi dipendenti o per i lavoratori dei cluster meno esperti.

Il Cluster_2 mostra un profilo lavorativo simile a quello del Cluster_2, ma con una tariffa media mensile più alta. Tuttavia, a differenza degli altri cluster, questi lavoratori hanno un impiego mensile più basso, il che indica che ricevono una tariffa elevata rispetto a poche ore di lavoro. Questo profilo lavorativo suggerisce che tali dipendenti potrebbero essere utilizzati solo in modo selettivo e solo quando necessario, per evitare di aumentare i costi aziendali. Inoltre, uno scarso utilizzo mensile potrebbe indicare una mancanza di assegnazione adeguata di compiti o un sottoutilizzo delle loro competenze. Ciò suggerisce che l'organizzazione potrebbe avere bisogno di rivedere la pianificazione delle attività per utilizzare meglio questi lavoratori e massimizzare la loro efficienza, riducendo così i costi aziendali (Figura 5).

Il workflow risultante dall'analisi dei cluster e dal modello Fuzzy C-Means (Fig.5) ricostruisce un processo organizzativo dinamico che fa perno sull'analisi dei dati attraverso la scelta dell'algoritmo di analisi e che, attraverso l'interpretazione dei risultati dei cluster, offra un supporto decisionale ai manager sia in termini di riallocazione della risorsa che di conferma dell'andamento della performance di quel processo.

L'approccio di interazione tra algoritmi, seppur riportato in maniera esplicativa, mostra come gli algoritmi di apprendimento non supervisionato possano portate a diverse ipotesi di riorganizzazione virtuosa dei processi di gestione delle risorse umane [5] Tale approccio si innesta perfettamente all'interno del trend di ricerca in forte crescita sull'ottimizzazione *AI-based* dello Human Resource Management [6; 7], e vuole rappresentare un primo sforzo esplorativo negli studi organizzativi, nonché un inquadramento in prospettiva per le

organizzazioni di oggi e del domani.

Bibliografia

1] Massaro, A. *Electronic in Advanced Research Industry: From Industry 4.0 to Industry 5.0 Advances*; Wiley: Hoboken, NJ, USA; IEEE: New York, NY, USA, 2021; ISBN 9781119716877.

[2]Massaro, A. Advanced Control Systems in Industry 5.0 Enabling Process Mining. *Sensors* **2022**, *22*, 8677. https://doi.org/10.3390/s22228677

[3] https://www.kaggle.com/datasets/varunbarath/human-resources?datasetId=942845&sortBy=dateRun&tab=profile

[4] Massaro, A. Multi-Level Decision Support System in Production and Safety Management. Knowledge 2022, 2, 682-701. https://doi.org/10.3390/knowledge2040039

[5] Rosa, A., & Martinez, M. (2022). Modelli Organizzativi Agili per la Pubblica Amministrazione. Franco Angeli.

[6] Gu, J. (2022). Image Model and Algorithm of Human Resource Optimal Configuration Based on FPGA and Microsystem Analysis. Wireless Communications and Mobile Computing, 2022. https://doi.org/10.1155/2022/7911419

[7] Keegan, A., & Meijerink, J. (2023). Dynamism and realignment in the HR architecture: Online labor platform ecosystems and the key role of contractors. *Human Resource Management, 62(1),* 15-29. https://doi.org/10.1002/hrm.22120