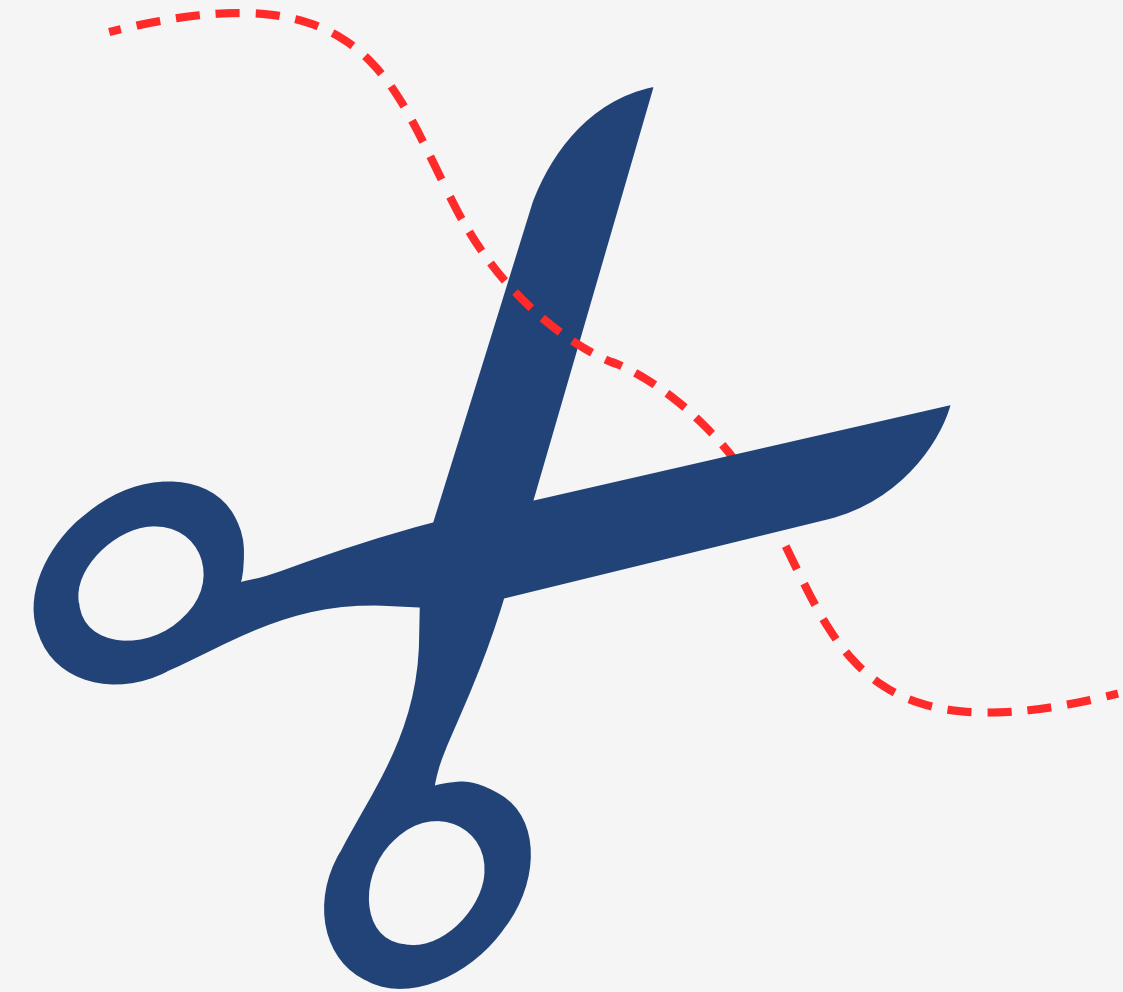


RETI NEURALI ARTIFICIALI PER LO SVILUPPO DI SHORT-FORM

Uno studio su dati simulati utilizzando
gli autoencoder



INTRODUZIONE

Negli ultimi anni, lo sviluppo di forme brevi di test è diventato una parte fondamentale della ricerca psicometrica.

Secondo una ricerca su PsycINFO riportata da Leite, Huang e Marcoulides (2008), sono stati pubblicati 46 articoli dal 2001 al 2006 che hanno sviluppato brevi forme di test.

Abbiamo effettuato una ricerca simile sulla stessa banca dati per un periodo compreso tra il 2017 e il 2022 che è risultata in 129 articoli, evidenziando la significativa crescita dello sviluppo e dell'uso delle brevi forme in contesti psicologici.



PERCHÈ ABBREVIARE I TEST PSICOMETRICI?

I test lunghi:

- sono costosi in termini di tempo e di risorse;
- spesso producono risposte di scarsa qualità inducendo nel rispondente un "effetto stanchezza"
- hanno tassi più elevati di risposte mancanti;

Inoltre, la necessità di test più brevi può riflettere la tendenza crescente all'uso più frequente di modelli complessi e approcci multivariati nella ricerca psicologica.

APPROCCIO ESPLICATIVO ALLA COSTRUZIONE DI SHORT-FORM

Gli approcci basati sulla Teoria Classica dei Test mirano a mantenere approssimativamente inalterate le proprietà psicometriche della scala originale (attendibilità e validità).

Ciò si ottiene considerando gli item con la maggior correlazione item-totale per un determinato costrutto o gli item con i loadings più elevati (in un approccio di analisi fattoriale).

In questo contesto, esistono diverse tecniche statistiche per investigare la struttura interna della misura originale e abbreviata, come PCA, Analisi Fattoriale Esplorativa e Analisi Fattoriale Confermativa.

Un'altra pratica comune è considerare le correlazioni tra gli item o misure di coerenza interna come l'Alpha di Cronbach. Sebbene l'attendibilità diminuisca quando vengono eliminati degli item (poiché essa dipende dal numero di item), l'idea è di abbreviare la misura mantenendo un livello accettabile di Alpha.

APPROCCIO ESPLICATIVO ALLA COSTRUZIONE DI SHORT-FORM

Lo shortening basato su Item Response Theory prevede la selezione di quegli item sono di con parametri di discriminazione più elevati o con una funzione informativa più grande rispetto ad altri.

L'obiettivo è quello di produrre una short form che misuri con buona precisione tutti i livelli del tratto latente d'interesse.



APPROCCIO ESPLICATIVO ALLA COSTRUZIONE DI SHORT-FORM

Negli ultimi anni, sono stati proposti vari algoritmi di ottimizzazione per automatizzare e ottimizzare lo shortening riducendo le fasi di validazione.

I criteri di ottimizzazione proposti da questi studi sono principalmente legati alla varianza spiegata o agli indici di adattamento dell'Analisi Fattoriale Confermativa.

Esempi di applicazione di questi algoritmi allo shortening includono gli Algoritmi Genetici (GA) proposti da Yarkoni (2010) e gli Ant Colony Optimization Algorithms (ACO) proposti da Leite, Huang & Marcoulides (2008).

Questi algoritmi consentono l'ottimizzazione di molteplici criteri, compresa la validità interna ed esterna, contemporaneamente.

APPROCCIO ESPLICATIVO ALLA COSTRUZIONE DI SHORT-FORM

Nonostante i suoi vantaggi, lo sviluppo tradizionale di short-form nasconde alcuni rischi:

- La costruzione di una breve forma è un problema combinatorio con un numero molto elevato di possibili soluzioni. Nella realtà, vengono considerate poche possibili combinazioni abbreviate, e ciò può portare a scegliere soluzioni non ottimali.
- Lo sviluppo di una short-form richiede al ricercatore di prendere decisioni e intervenire nel processo: ciò aumenta il rischio di decisioni errate e strumenti distorti.
- Gli item che riflettono un costrutto generale mostrano meno correlazione con quelli che definiscono un costrutto più "sottile", probabilmente perché rappresentano una piccola parte di un costrutto più generale. Pertanto, non considerare gli item con una correlazione item-totale più bassa potrebbe significare non prendere in considerazione alcuni aspetti importanti del costrutto e, di conseguenza, alterarne il significato.
- Le statistiche basate su CTT possono catturare solo relazioni lineari tra gli item. Tuttavia, poiché le relazioni presenti nel mondo reale potrebbero non essere sempre lineari, queste misure potrebbero non essere sempre metodi di analisi appropriati.

RETI NEURALI ARTIFICIALI PER LA COSTRUZIONE E VALIDAZIONE DI TEST

Uno studio proposto da Dolce et al. (2020) utilizza reti neurali artificiali per la costruzione e la validazione di una scala.

In questo studio, gli autori hanno proposto una procedura per la selezione degli item che tiene conto sia dell'analisi esplorativa dei dati per investigare la struttura dimensionale, sia delle reti neurali artificiali per predire la diagnosi psicopatologica di soggetti clinici (un problema di classificazione).

La prospettiva esplicativa consente di ottenere approfondimenti teorici sulle caratteristiche degli item selezionati e sulla loro conformità con il quadro teorico di riferimento.

Allo stesso tempo, l'uso delle reti neurali artificiali permette la selezione di item più predittivi dei criteri esterni (la diagnosi psicopatologica).

INTEGRAZIONE DI TECNICHE ESPLICATIVE E PREDITTIVE PER LA COSTRUZIONE DI SHORT-FORM

Lo studio che discuteremo propone una procedura per automatizzare il processo di accorciamento integrando tecniche esplicative e predittive. La procedura proposta utilizza un particolare tipo di rete neurale artificiale, l'autoencoder.

Gli autoencoder permettono di integrare il punto di vista esplicativo e predittivo nello stesso metodo: infatti, grazie alle loro caratteristiche, consentono di costruire una short form in grado di prevedere le risposte del pattern della forma lunga mantenendo l'interpretabilità delle dimensioni dei dati.

Le caratteristiche dell'autoencoder possano aiutare i ricercatori a selezionare item con una maggiore validità esterna, ossia item che predicono meglio l'intero pattern di risposte della forma lunga originale.

Inoltre, a causa della loro relazione con altre tecniche di riduzione della dimensionalità, ipotizziamo che gli autoencoder preservino anche la dimensionalità della forma lunga originale.

L'idea centrale di questo studio è presentare una procedura automatizzata che ordini gli item in base alla loro importanza per la predizione dei pattern di risposta della forma lunga. L'obiettivo della procedura è fornire una guida al ricercatore per la selezione degli item da includere nella short form finale.

INTEGRAZIONE DI TECNICHE ESPLICATIVE E PREDITTIVE PER LA COSTRUZIONE DI SHORT-FORM

In particolare, lo scopo del lavoro era confrontare il funzionamento della selezione degli item di due diversi autoencoder: un autoencoder lineare che, per le sue caratteristiche, è del tutto equivalente al PCA, e un autoencoder non lineare a più strati.

Gli autoencoder sono testati su dati simulati, basati su una struttura fattoriale.

La diversa performance nella selezione degli item per la costruzione della short form è valutata considerando le caratteristiche dell'item selezionato (la sua relazione con la variabile latente), l'errore di ricostruzione e l'accuratezza nel ricostruire le risposte degli item della forma lunga originale.

I risultati della procedura sono stati confrontati con due algoritmi di ottimizzazione per lo sviluppo di brevi forme, l'ACO e il GA.

LA RIDUZIONE DELLA DIMENSIONALITÀ

In generale, una tecnica di riduzione della dimensionalità produce un'approssimazione \hat{x}_j di una variabile di partenza x_j dove j è una generica variabile delle p variabili totali e x è una singola osservazione. Quest'approssimazione è il risultato di due funzioni, f e g .

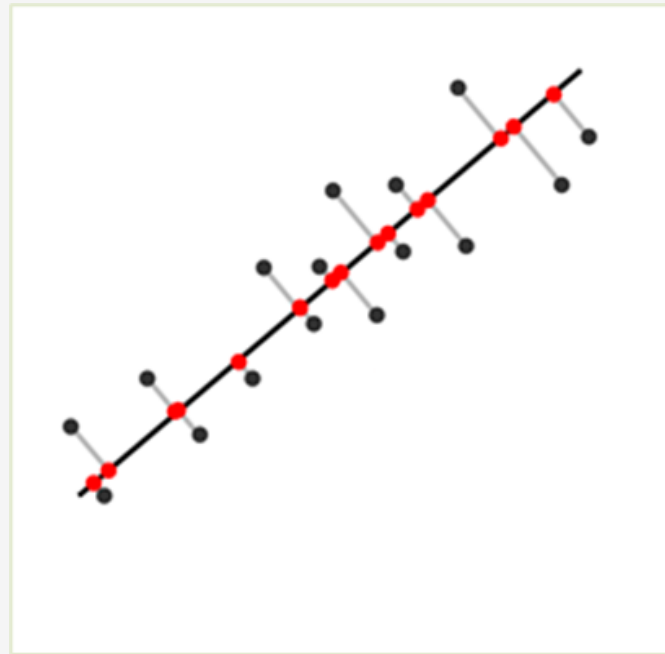
$$x_j = \hat{x}_j + \epsilon_j = g \left(f(x_j) \right) + \epsilon_j, \quad j = 1, 2, \dots, p$$

La funzione di proiezione $f : R_p \rightarrow R_z$ proietta x_j da uno spazio originario con dimensionalità p in uno sottospazio z -dimensionale a dimensionalità ridotta, mentre la funzione di espansione $g : R_z \rightarrow R_p$ espande nuovamente x_j dal sottospazio z -dimensionale allo spazio con dimensionalità iniziale p con un residuo ϵ_j .

Il problema della riduzione della dimensionalità viene risolto attraverso la determinazione delle funzioni f e g .

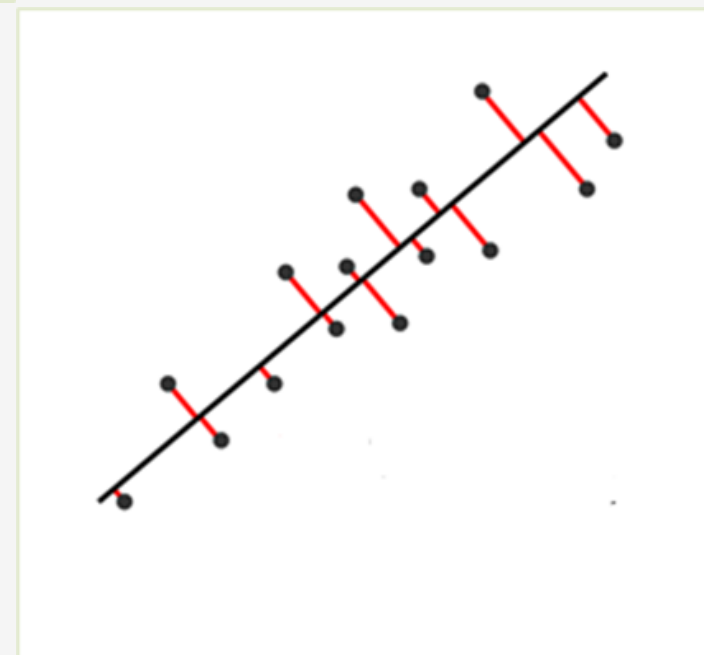
L'ANALISI DELLE COMPONENTI PRINCIPALI

Riduce la dimensionalità di un dataset con una grande quantità di variabili. Come?



Determinando le direzioni di massima varianza delle proiezioni.

Trovando le proiezioni che minimizzano l'errore di ricostruzione.

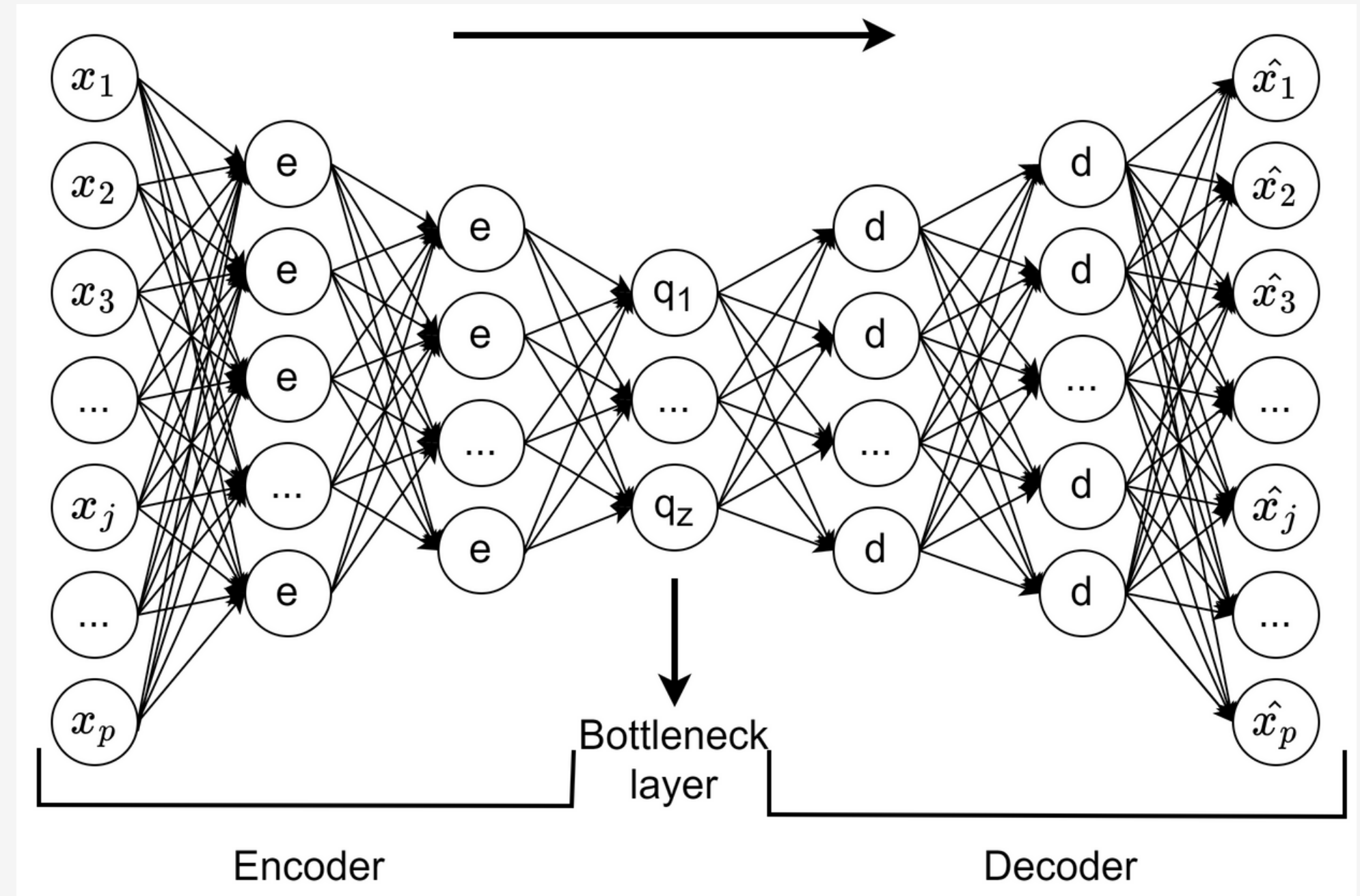


LE COMPONENTI PRINCIPALI

- Sono combinazioni lineari delle variabili originarie;
- Sono ortogonali;
- Sono ordinate in termini di varianza spiegata.

L'AUTOENCODER

- Gli autoencoder sono reti neurali artificiali che imparano a comprimere e a ricostruire i dati.
- Sono composti da due parti: una rete di codifica (encoder) e una rete di decodifica (decoder).
- Gli autoencoder vengono addestrati minimizzando la differenza tra l'input e l'output ricostruito.



AUTOENCODER E PCA

È stato dimostrato che un autoencoder con un singolo strato nascosto lineare di z unità esegue una proiezione su un sottospazio z -dimensionale, il quale è generato dalle prime z componenti principali dei dati.

In effetti, in questo caso specifico, sia l'analisi delle componenti principali che l'autoencoder eseguono una riduzione dimensionale lineare e minimizzano la stessa funzione di errore di ricostruzione.

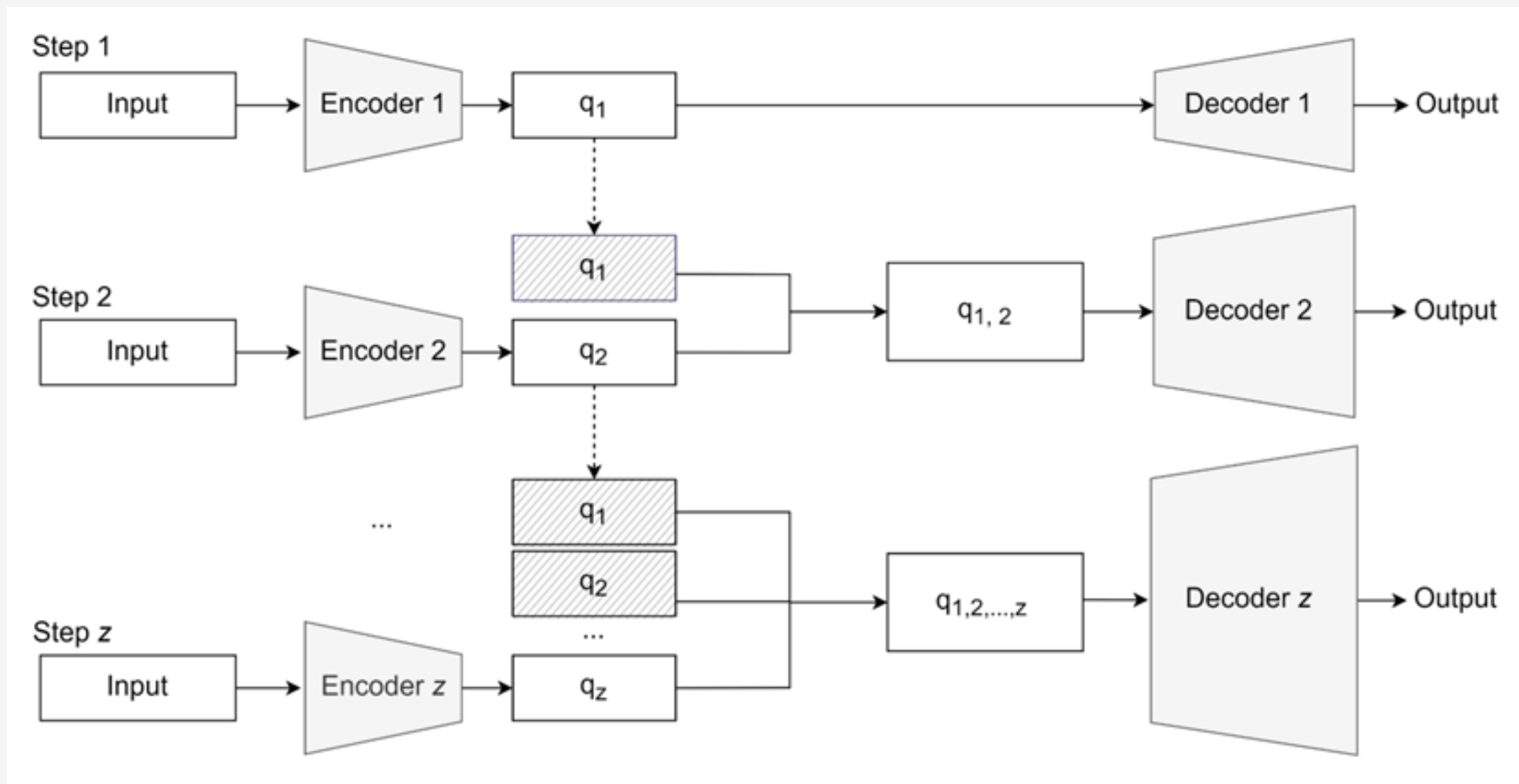
Tuttavia i nodi centrali dell'autoencoder non rappresentano esattamente le componenti principali dei dati: infatti, le componenti principali sono non correlate e sono ordinate in base alla loro rilevanza. Queste caratteristiche rendono la soluzione della PCA più interpretabile rispetto alla soluzione dell'autoencoder.

D'altra parte, l'autoencoder non è limitato a trasformazioni lineari, pertanto può apprendere relazioni più complesse.



PCA-AE

Uno degli autoencoder utilizzato in questo studio è stato un tipo di autoencoder (PCA-AE) che rende esplicite le proprietà della PCA attraverso una procedura particolare di training che prevede l'addestramento di un nodo centrale per volta. Se utilizzato con un singolo strato nascosto lineare, questo autoencoder è totalmente equivalente alla PCA e i suoi nodi interni sono esattamente le componenti principali dei dati.

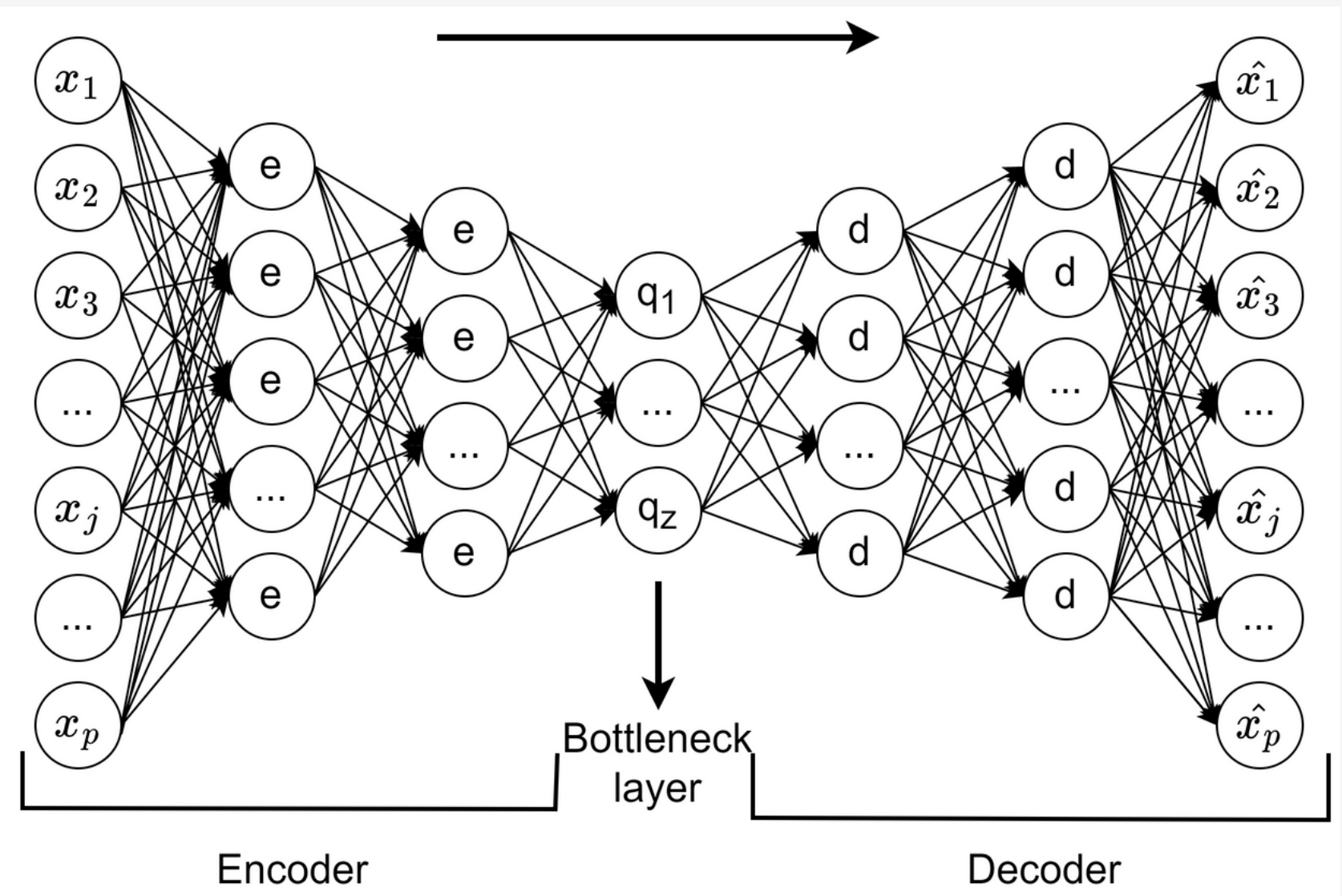


- **Architettura:** 27 input, 3 neuroni interni, 27 output
- **Attivazioni:** lineari
- **Optimizer:** Adam
- **Loss:** MSE + Covariance Term
- **Learning Rate** = 0.005
- **Training:** il training si ferma quando non ci sono miglioramenti nella loss per 40 epoche consecutive

NL-AE

Il secondo autoencoder utilizzato nello studio è un autoencoder con attivazioni non lineari e più strati nascosti.

- **Architettura:** 27*5 input, 90, 40, 15, 40, 90, 27*5 output
- **Attivazioni:** sigmoidali, tranne nello strato centrale che è lineare
- **Optimizer:** Adam
- **Loss:** Binary Cross-Entropy
- **Learning Rate** = 0.005
- **Training:** il training si ferma quando non ci sono miglioramenti nella loss per 20 epoche consecutive

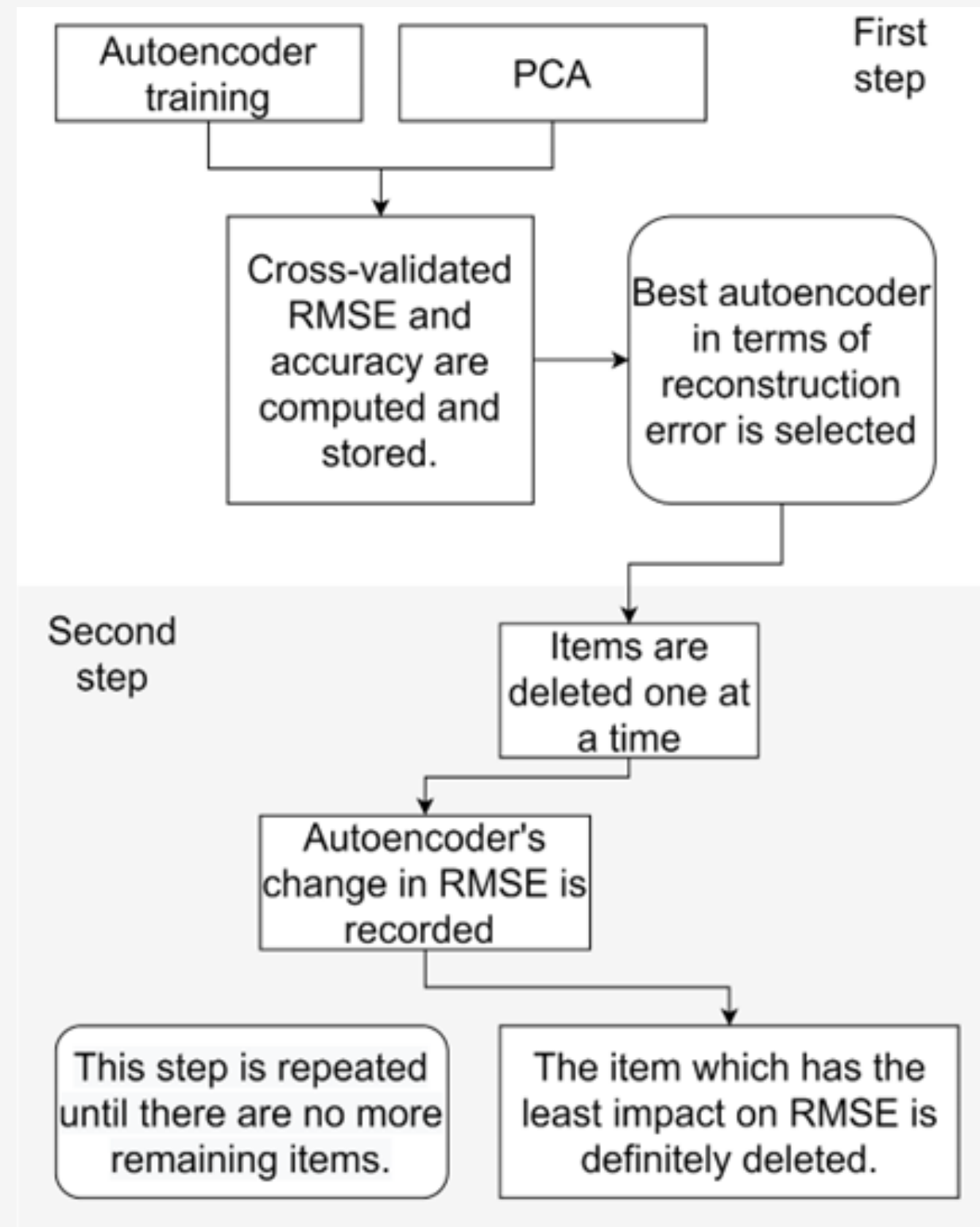


ONE-HOT ENCODING

- Il One-Hot Encoding è una tecnica utilizzata nella fase di preparazione dei dati.
- Si tratta di una trasformazione che converte variabili categoriche in un formato numerico, permettendo ai modelli di apprendimento di elaborare tali informazioni correttamente.
- Un attributo binario viene impostato su 1 quando la categoria è presente per una determinata osservazione e a 0 quando non è presente.
- Questo procedimento crea una matrice di dimensione $M \times N$, dove M è il numero di osservazioni e N è il numero di categorie della variabile.

Original response categories		One-hot encoded response categories
1	→	10000
2		01000
3		00100
4		00010
5		00001

SELEZIONE DEGLI ITEM



Nella prima fase, l'autoencoder viene addestrato sulla forma lunga. Viene eseguita cross-validation per selezionare l'autoencoder con l'errore quadratico medio (RMSE) inferiore.

Per capire quali item fossero più importanti per la ricostruzione della forma lunga del test, è stata valutata la variazione nell'errore di ricostruzione dell'autoencoder rimuovendo in sequenza gli item, uno per volta.

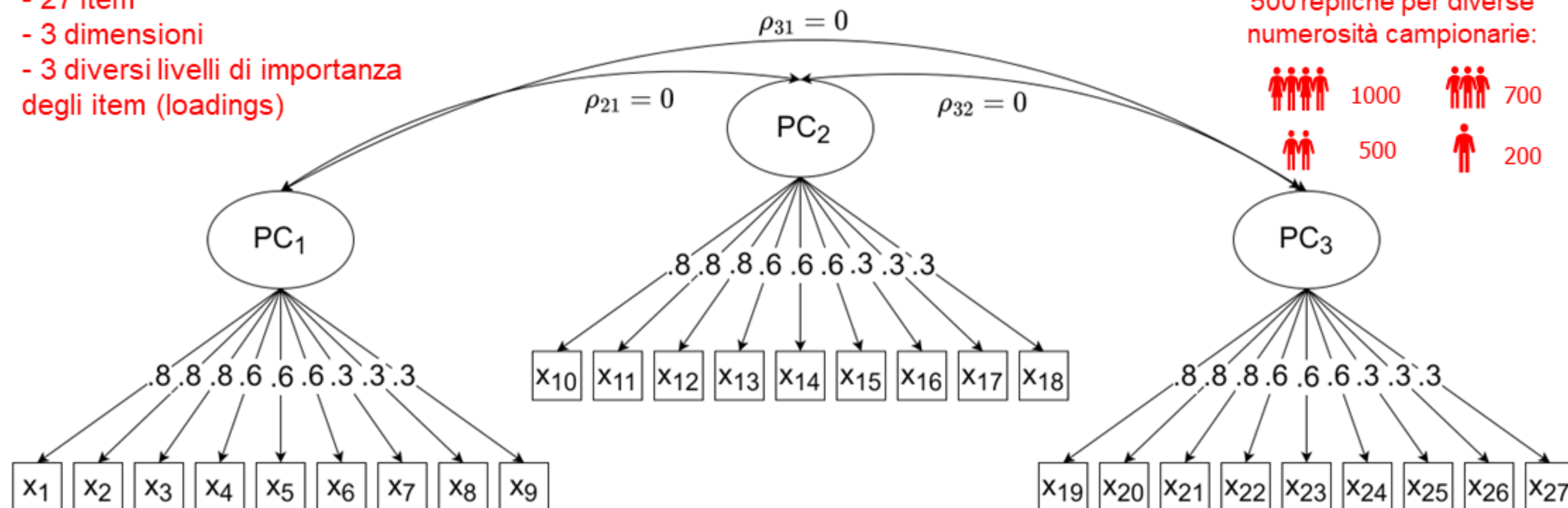
Una volta individuato l'item con il minor impatto sull'RMSE, questo è stato definitivamente rimosso (ovvero l'input della rete neurale che codificava tale item è stato impostato a 0) e la procedura è stata ripetuta finché il numero di item non si è esaurito.

Ipotizziamo che la variazione dell'RMSE per ogni item rimosso sia un'indicazione della sua importanza relativa: in particolare, l'idea è che quando un item non è molto importante nella ricostruzione della long-form, dopo la sua eliminazione l'errore di ricostruzione ha un incremento minimo.

A questo punto, infatti, è stato possibile definire un ordine di importanza dove gli ultimi item selezionati dalla procedura sono quelli che meglio ricostruiscono la long-form.

PIANO DI SIMULAZIONE E GENERAZIONE DEI DATI

- 27 item
- 3 dimensioni
- 3 diversi livelli di importanza degli item (loadings)



LA GENERAZIONE DEI DATI

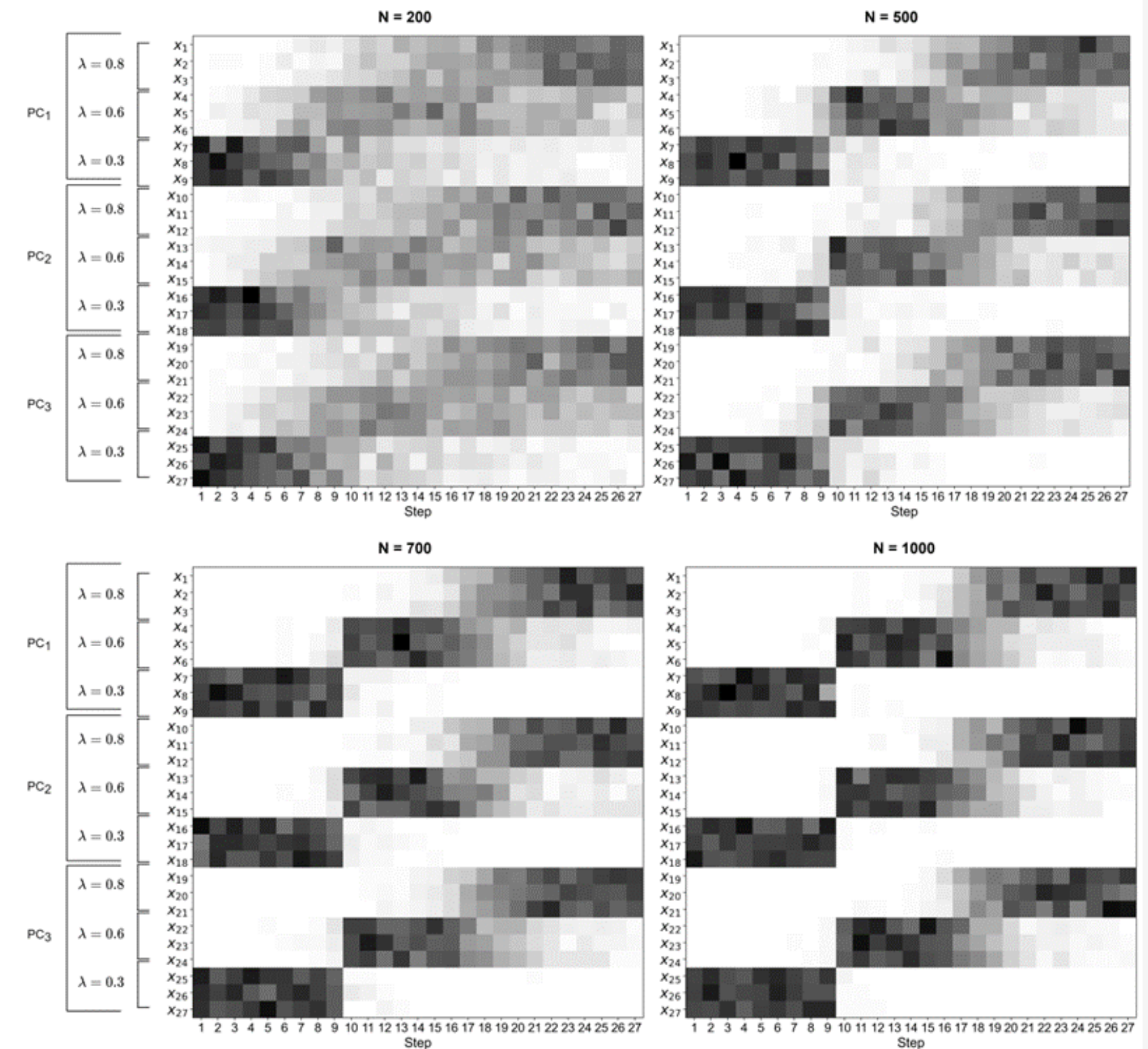
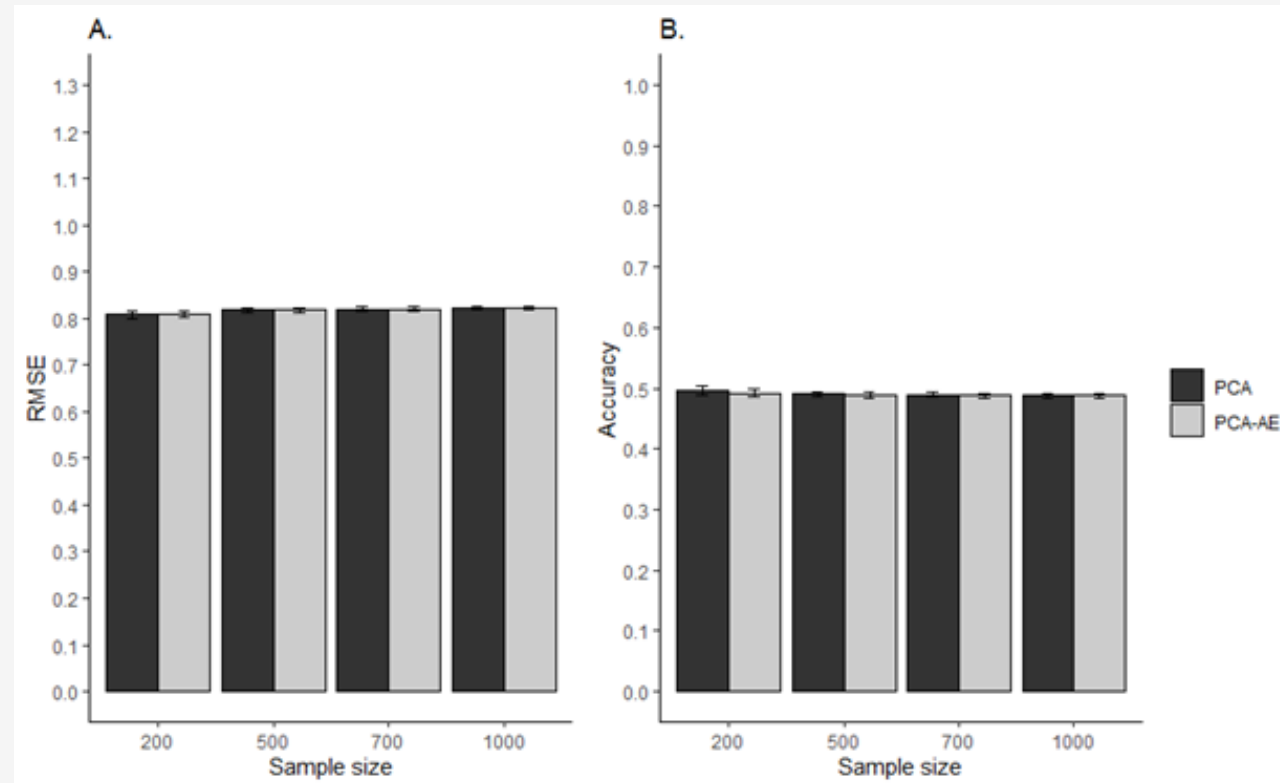
In particolare, nell'ambito dei modelli di equazioni strutturali di cui l'analisi fattoriale confermativa rappresenta un caso particolare, è possibile simulare dei dati che abbiano una struttura fattoriale definita a priori dal ricercatore. In sintesi, dal momento che a partire dai parametri del modello specificato si può riscrivere la matrice di varianza e covarianza secondo le seguenti relazioni:

$$Var(x_j) = \sum_{k=1}^q \lambda_{jk}^2 + Var(\delta_j)$$

$$Cov(x_i, x_j) = \sum_{k=1}^q \lambda_{ik} \lambda_{jk}$$

In seguito, tramite una decomposizione di Cholesky si può ricavare una matrice di dati grezzi simulata. È possibile simulare i dati in maniera automatica utilizzando il software R e in particolare il pacchetto LAVAAN.

RISULTATI PCA-AE

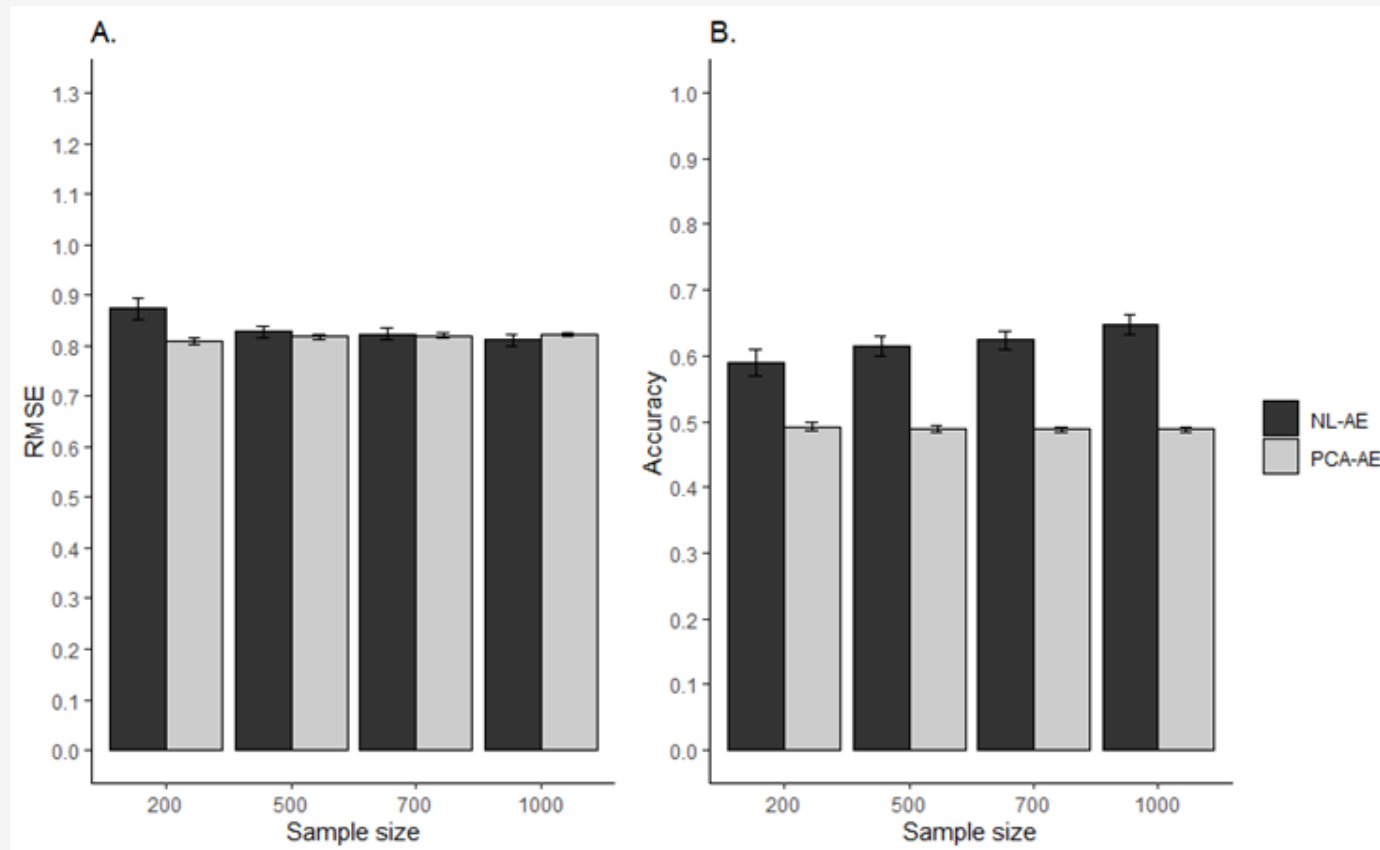


PCA-AE e PCA hanno un'accuracy e un RMSE molto simili, in media, per tutte le numerosità campionarie. La figura a destra mostra in sfumature di grigio quante volte su 500 repliche un singolo item viene rimosso dalla procedura di selezione iterativa in quel preciso step.

Dai risultati emerge che nelle prime nove iterazioni, gli item più selezionati sono quelli con i loadings pari a 0.3, mentre nelle iterazioni successive gli item rimossi sono quelli con loadings pari a 0.6. Solo nelle ultime iterazioni vengono rimossi gli item con loadings pari a 0.8.

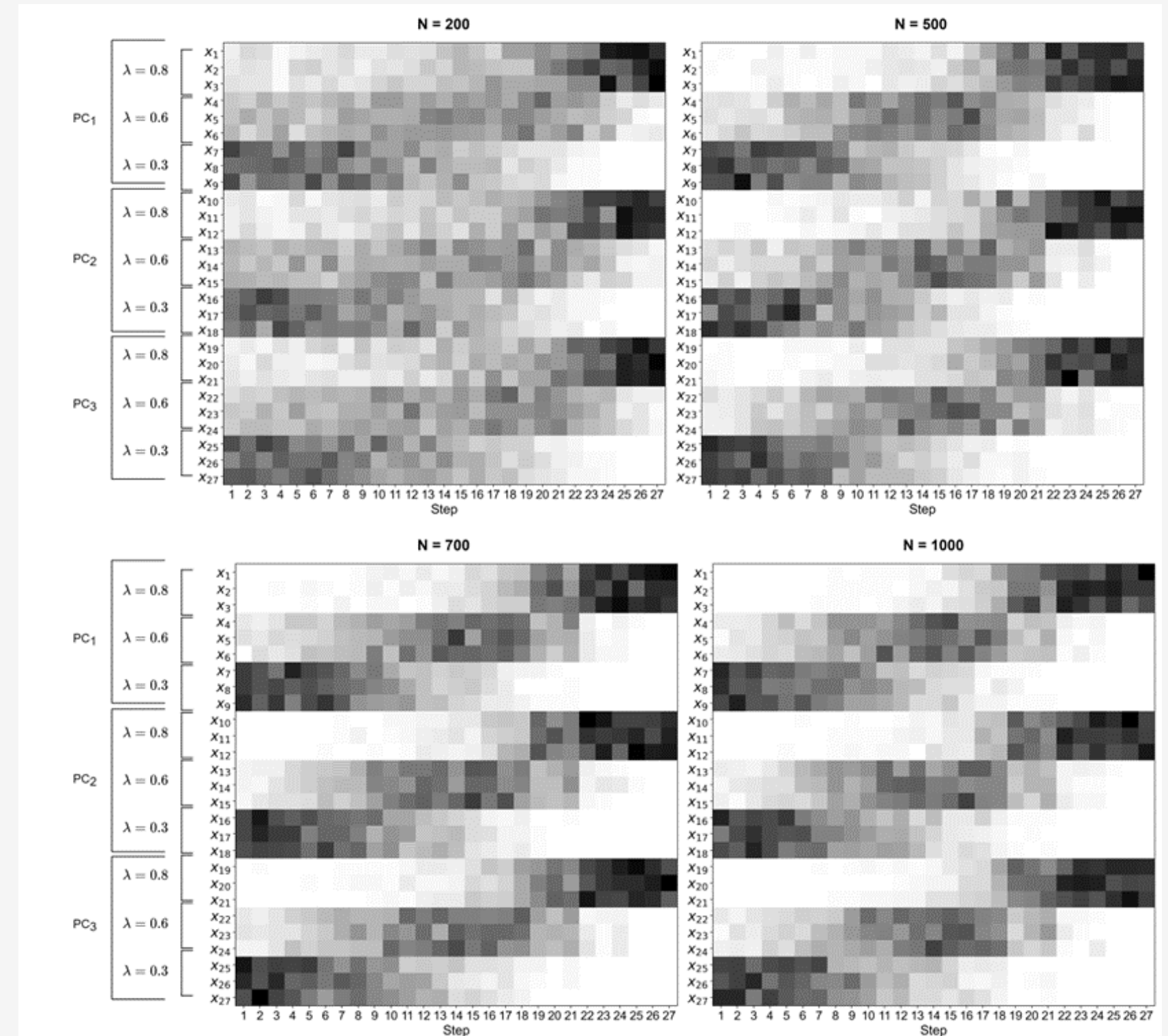
Il pattern di PCA-AE nella selezione degli item è meno definito quando la numerosità campionaria è pari a 200.

RISULTATI NL-AE



Per quanto riguarda i risultati di NL-AE, l'accuratezza media di addestramento calcolata su tutti i dataset è maggiore per NL-AE rispetto a PCA-AE. L'RMSE medio è quasi lo stesso per NL-AE e PCA-AE, tranne che per la numerosità campionaria più bassa.

La figura a destra mostra i risultati della procedura di selezione degli item: anche in questo caso, nelle prime iterazioni, gli item prevalentemente selezionati sono quelli con loadings pari a 0.3, mentre gli item più preservati sono quelli con loadings pari a 0.8. Tuttavia, questo autoencoder mostra pattern più sfumati nelle iterazioni intermedie.

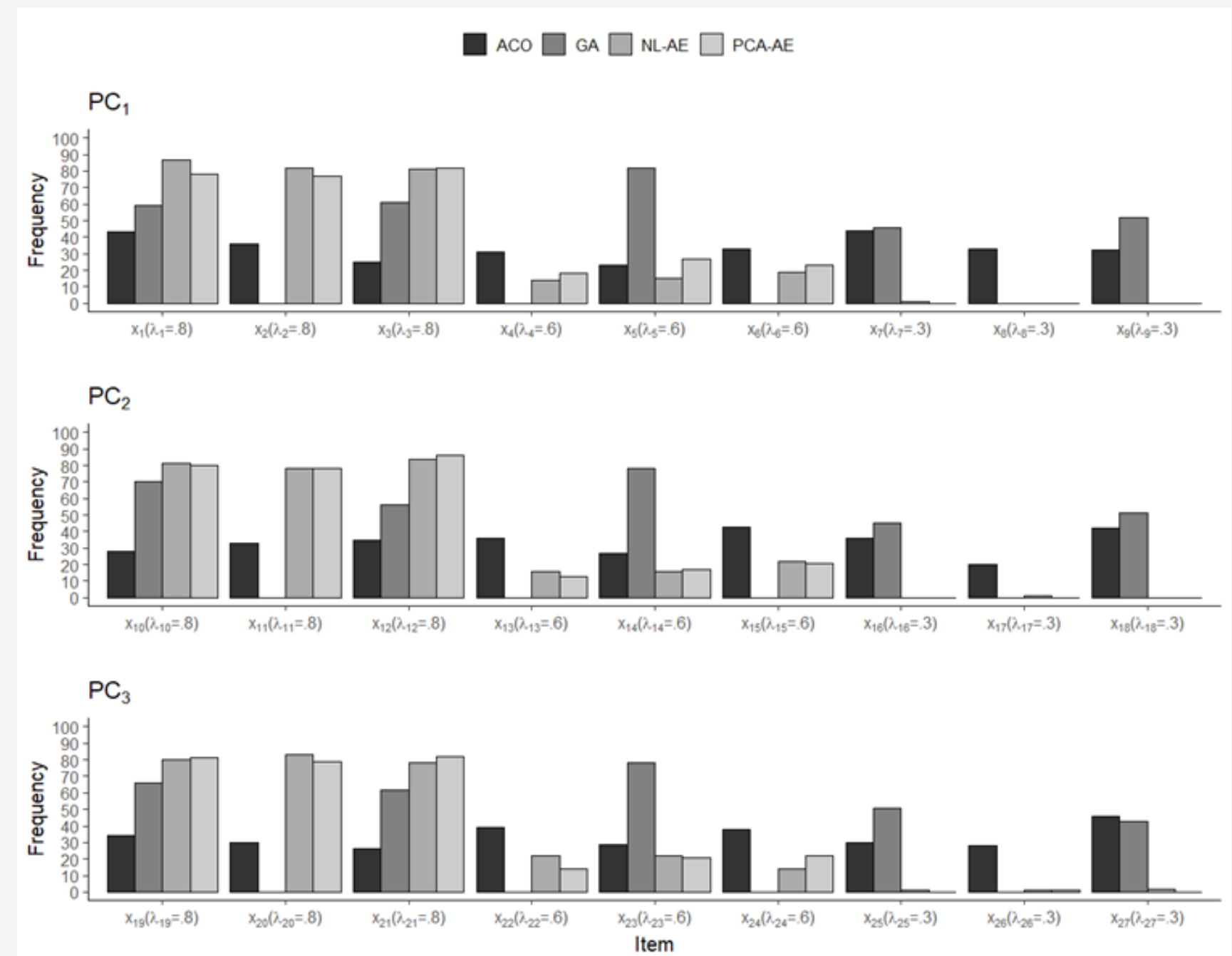


RISULTATI CONFRONTO CON ACO E GA

Abbiamo scelto di testare la costruzione di una short form a 9 item per confrontare i due autoencoder con due algoritmi di ottimizzazione (ACO e GA).

La figura a destra riporta per ogni dimensione simulata quante volte ciascun item viene selezionato dai diversi algoritmi testati in questo studio.

PCA-AE e NL-AE convergono verso soluzioni molto simili, in cui gli item più selezionati sono quelli con i loadings più alti. Le soluzioni di GA e ACO presentano un pattern meno chiaro nella selezione degli item, con un campionamento omogeneo su tutti gli item.



CONCLUSIONI

I risultati mostrano che entrambi gli autoencoder preservano la dimensionalità del test originale e mantengono gli item più importanti fino agli ultimi step, dando la possibilità di costruire una short-form che massimizzi la predizione della long form di partenza, integrando l'ottica esplicativa e predittiva.

Nonostante gli autoencoder non conoscano a priori la struttura fattoriale dei dati, essa emerge dalla procedura.

L'approccio presentato fornisce al ricercatore una classifica di importanza degli item contenuti, consentendo di determinare il numero di item da mantenere nella forma breve, anche sulla base di considerazioni relative al contenuto degli item.

Bisogna, tuttavia tenere a mente alcuni limiti e possibili implementazioni future:

I dati utilizzati in questo studio sono simulati con relazioni lineari, pertanto la performance del NL-AE non può essere troppo diversa dalle altre tecniche.

In questa forma, l'autoencoder non lineare è poco interpretabile.

BIBLIOGRAFIA

Casella, M., Dolce, P., Ponticorvo, M., Milano, N., & Marocco, D. (2023). Artificial Neural Networks for Short-Form Development of Psychometric Tests: A Study on Synthetic Populations Using Autoencoders. Educational and Psychological Measurement, 0(0). <https://doi.org/10.1177/00131644231164363>

Dolce, P., Marocco, D., Maldonato, M. N., & Sperandeo, R. (2020). Toward a Machine Learning Predictive-Oriented Approach to Complement Explanatory Modeling. An Application for Evaluating Psychopathological Traits Based on Affective Neurosciences and Phenomenology. Frontiers in Psychology, 11, 446.
<https://doi.org/10.3389/fpsyg.2020.00446>

Gonzalez, O. (2020). Psychometric and Machine Learning Approaches to Reduce the Length of Scales. Multivariate Behavioral Research. <https://doi.org/10.1080/00273171.2020.1781585>

Raborn, A. W., Leite, W. L., & Marcoulides, K. M. (2020). A comparison of metaheuristic optimization algorithms for scale short-form development. Educational and Psychological Measurement, 80(5), 910-931.

ESERCITAZIONE

In questa esercitazione, utilizzando un dataset che vi sarà fornito, dovrete implementare:

1) un autoencoder che minimizza l'errore di ricostruzione con 5 strati così organizzati:

- input: numero degli item
- I interno : 5 (sigmoidale)
- II interno: 3 (lineare)
- III interno: 5 (sigmoidale)
- output : numero degli item (lineare)

2) la procedura di selezione degli item così com'è stata presentata, in particolare:

- al primo step, dovrete annullare tutte le variabili una per volta e calcolare l'errore di ricostruzione (RMSE) ogni volta.
- Dovrete annullare la variabile che quando eliminata produce l'RMSE minore
- Dovrete ripetere questa procedura per tutti gli item rimasti e finché non annullerete tutti gli item.