

Cognome e Nome _____

Matricola _____

Appello del 12 febbraio 2026

Esercizio 1 [8 punti]

Si progettino (eventualmente utilizzando più step di computazione tra loro concatenati) le funzioni *Map* e *Reduce* per generare tutti i **cicli di lunghezza 4 con il loro relativo peso** in un **grafo diretto pesato** $G=(V,E,w)$. Si assuma che un arco del grafo di input, che va dal nodo i al nodo j ed ha peso x , sia codificato come la tripla in input (i, j, x) e che il peso di un ciclo sia dato **dalla somma dei pesi** degli archi che lo compongono.

Si chiede di dare in output i cicli **senza ripetizioni**: ad esempio, il ciclo formato dai nodi $2 \rightarrow 4 \rightarrow 3 \rightarrow 10$ non dovrà comparire anche sottoforma di $4 \rightarrow 3 \rightarrow 10 \rightarrow 2$, $3 \rightarrow 10 \rightarrow 2 \rightarrow 4$ o $10 \rightarrow 2 \rightarrow 4 \rightarrow 3$.

Suggerimento: prima trovare i path di 2 archi (quindi con 3 nodi) e successivamente trovare i cicli di 4 archi analizzando come si possono concatenare i path di 2 archi trovati al primo step.

Esercizio alternativo [valutazione massima 4 punti]: generare, invece dei cicli di 4 nodi, tutti i path di 2 archi (e quindi con 3 nodi) con il relativo peso (ottenuto sempre come somma dei pesi degli archi).

Esercizio 2 [9 punti]

Si considerino gli insiemi descritti nella tabella qui sotto.

- **[3 punti]** Si scriva la colonna dell'insieme S_2 in modo che la Jaccard Similarity tra S_1 e S_2 sia $2/5$. Si calcoli quindi la Jaccard Similarity tra S_2 e S_3 e tra S_2 e S_4 .
- **[4 punti]** Si calcoli, applicando il metodo visto a lezione ed evidenziando tutti i passaggi fatti riga per riga, la matrice delle firme **minhash** degli insiemi S_1, S_2, S_3 e S_4 secondo le permutazioni indotte dalle 3 funzioni hash evidenziate in tabella.
- **[2 punti]** Sfruttando la matrice delle firme appena calcolata, si approssimi la Jaccard Similarity tra S_1 e S_2 , tra S_2 e S_3 e tra S_2 e S_4 .

Riga	S_1	S_2	S_3	S_4	$(x+3) \bmod 7$	$2x \bmod 7$	$(x-1) \bmod 7$
0	1		1	0			
1	0		1	1			
2	0		0	0			
3	1		0	1			
4	0		1	0			
5	0		0	1			
6	1		1	1			

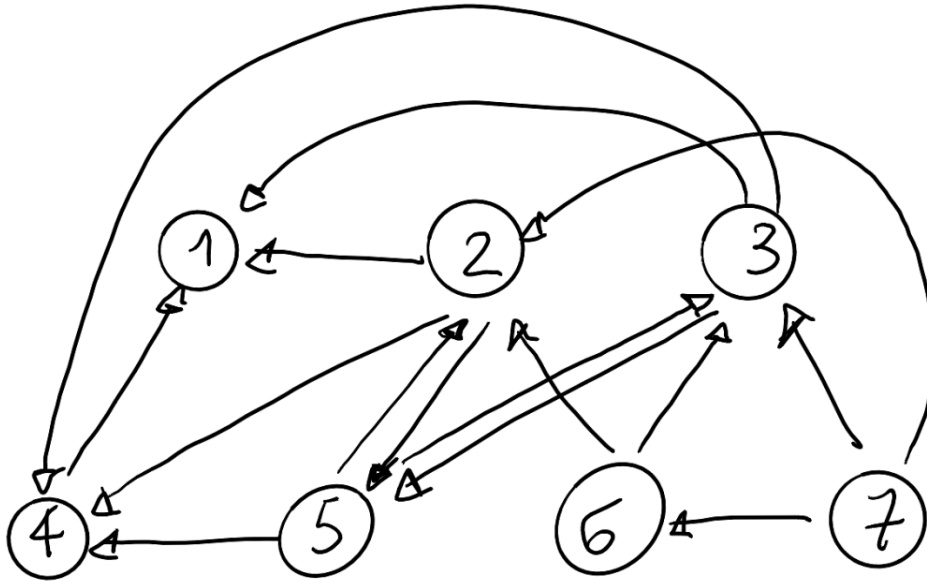
Esercizio 3 [8 punti]

Si consideri il problema dei **Frequent Itemset**.

- Si definisca il **bordo negativo** e si indichi come è possibile calcolarlo
- Si descriva il funzionamento dell'**algoritmo di Tovoinen**
- Si definiscano i **falsi negativi** e i **falsi positivi**, e si discuta come vengono gestiti dall'algoritmo di Tovoinen

Esercizio 4 [9 punti]

Si consideri la rete sociale in figura.



Applicando l'algoritmo A-priori, **si individui** i cluster di almeno 5 nodi indotti da soglia di supporto uguale almeno a 2. Si mostri ogni passo della computazione effettuata.

Regole per lo svolgimento della prova scritta:

- Per svolgere il compito si hanno a disposizione **90** minuti (**70** minuti se non si svolge l'Esercizio 1)
- Scrivere **subito** nome, cognome, matricola su **OGNI FOGLIO (compreso questo)**.
- Durante la prova scritta **non** è possibile abbandonare l'aula.
- Non è ammesso **per nessun motivo** comunicare in qualsiasi modo con altre persone
- Non è possibile consultare appunti, libri e dispense.
- Qualsiasi strumento elettronico di calcolo o comunicazione (telefoni cellulari, calcolatrici, tablet, computer, smartwatch, etc...) deve essere **completamente disattivato e depositato in vista sulla cattedra**
- Mettere in vista sul banco un valido documento di identità.