

Big Network Visualizzazione Tool for iNSIdEnano

Luigi Giugliano

Università degli studi di Salerno

27 gennaio 2016

OVERVIEW

iNSIdEnano

Problema

Soluzione

OVERVIEW

iNSIdEnano

Problema

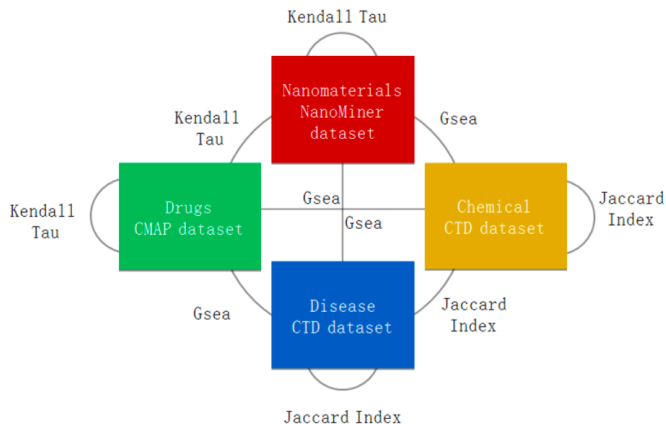
Soluzione

iNSIdENANO

iNSIdEnano è un tool grafico che mette in evidenza le connessioni tra entità fenotipiche del tipo:

- ▶ Esposizione ai nanomateriali
- ▶ Trattamenti farmaceutici
- ▶ Esposizione ad agenti chimici
- ▶ Malattie

L' interazione tra queste entità è valutata in base al loro effetto sull'espressione dei geni.



E' stata calcolata la distanza per ogni coppia di entità. Sono poi state normalizzate tra -1 e 1 per renderle confrontabili.

Per ogni entità fenotipica nel dataset, è assegnata una lista di geni. In particolare un'insieme di geni è associato a ogni malattia e ogni agente chimico, invece per ogni farmaco e per ogni nanomateriale è associata una lista ordinata di geni.

Quindi per costruire una network di similarità tra entità fenotipiche è stato necessario calcolare la similarità a coppie per ogni entità.

INSIEME DI GENI VS INSIEME DI GENI

Il Jaccard index è stato utilizzato per calcolare la similarità tra due malattie, tra due agenti chimici o tra un agente chimico e una malattia.

Dati due insiemi A e B l'indice di Jaccard è dato dalla dimensione della loro intersezione diviso la dimensione della loro unione.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

Questa misura è zero se i due insiemi non condividono neanche un gene, mentre 1 se sono esattamente uguali.

Per ogni agente chimico vengono considerati due set di geni: quelli che sono up-regolati da quell'agente chimico e quelli che sono down-regolati. Per quelli down-regolati il Jaccard index è calcolato con il segno negativo.

INSIEME DI GENI VS INSIEME DI GENI

La distanza Kendall Tau è stata utilizzata per calcolare la similarità tra nanomateriali e nanomateriali, tra farmaci e farmaci e tra nanomateriali e farmaci, basata sulla lista ordinata dei geni. La distanza Kendall Tau tra due liste T_1 e T_2 è definita come segue:

$$K(T_1, T_2) = |(i, j) : i < j, (T_1(i) < T_1(i) \wedge T_2(i) > T_2(j)) \vee (T_1(i) > T_1(i) \wedge T_2(i) < T_2(j))| \quad (2)$$

questa distanza è compresa tra 0 e $n(n-1)$, dove n è la lunghezza della lista. Il valore significa che gli elementi nella lista sono nello stesso ordine, mentre il valore $n(n-1)$, indica che gli elementi sono in ordine opposto

OVERVIEW

iNSIdEnano

Problema

Soluzione

.

OVERVIEW

iNSIdEnano

Problema

Soluzione

.