An idea about automatic setting the weight-value between 2 features.

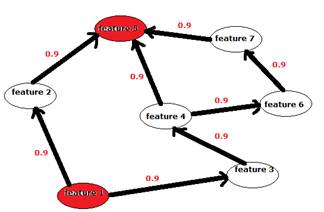
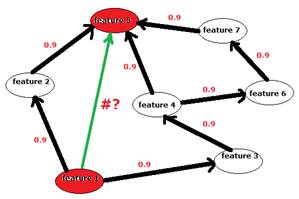
Present Challenge:

* How to indicate the exact weight-value, e.g. 0.7 or 0.8? the criterion(0.0~1.0)?

Existing Resource:

* The relationship(relative/irrelative) between 2 features, especially the heaviest dependence, indicate by 0.9.

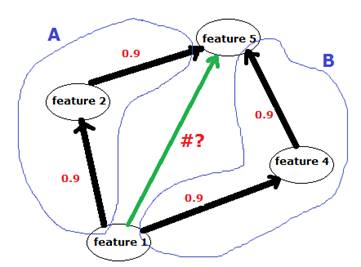
Expected result:

From existing data to knowledge, “#” means the weight-value between feature 1 and feature 5, as P(f5|f1).

f1 is the event “feature 1” , f5 is the event “feature 5”.

So now the question is how to calculate P(f5|f1), simplify the prototype as below:



P(f5|f1)=P(f5\*f1)/P(f1), P(f1) is a constant 1.

* P(f5\*f1)
* 1- (1-P(A))\*(1-P(B)), P(A)=P(f1)\*P(f2), P(B)=P(f1)\*P(f4), both them are independent events.
* 1- (1-0.9\*0.9)\*(1-0.9\*0.9)
* 0.9639

Then we can see “#” is a quite high value, means the feature 5 is heavily dependent on feature 1.

For a poor and sparse matrix, e.g. only some “0.9” marked manually, the method can mine more detailed and distant dependences.

NOTE:

* Absolutely, there are some parts should be improved, please correct it freely.
* Although the method can create weight-values automatically, it increase Features-Network’s complexity, what’s your suggestion?