

Indice

1	Analisi dei dati di probabilità	2
1.1	Problema in esame	2
1.2	Caratteristiche degli eventi di coppia	2
1.2.1	Eventi indipendenti	3
1.2.2	Eventi dipendenti	3
1.3	Evento conosciuto ed evento indovinato	4
1.3.1	Probabilità di rispondere correttamente ad una domanda	5
1.3.2	Il piano	7
2	Rete neurale	12
2.1	Test effettuati	13
2.1.1	Configurazione della rete: 4 neuroni per ciascuno dei 2 layers	14
2.1.2	Configurazione della rete a 2 neuroni per ciascuno dei 2 layers	17
2.1.3	Configurazione della rete a 4 neuroni per 1 layer	21

1 Analisi dei dati di probabilità

Durante il periodo 20/05 - 24/05 mi sono occupata di analizzare la probabilità che ha un candidato di rispondere correttamente alle domande in fase di test; valutando le relazioni di dipendenza che possono esistere tra più domande e l'impatto che può assumere la fortuna.

1.1 Problema in esame

Test, sottoposto ad un candidato durante un colloquio, composto da *domande a tripla risposta multipla*.

Nel suddetto documento vengono analizzate le relazioni che intercorrono tra due domande, denominate A e B, a seconda se il candidato risulta in grado di rispondervi correttamente o meno.

1.2 Caratteristiche degli eventi di coppia

Tipi di eventi trattati:

- **Eventi indipendenti;**
- **Eventi dipendenti:**
 - A e B sono strettamente dipendenti;
 - A implica B.
- **Evento conosciuto ed evento indovinato.**

Struttura usata per rappresentare la probabilità degli eventi di coppia:

AB

\wedge

$A \ B$

\vee

Z

con:

- AB rappresenta la probabilità complessiva dell'evento che si verifica sempre;
- A rappresenta la probabilità che permette il verificarsi di A, ma non di B;

- B rappresenta la probabilità che permette il verificarsi di B, ma non di A;
- Z rappresenta la probabilità a zero, l'impossibilità del verificarsi dell'evento.

1.2.1 Eventi indipendenti

A e B sono due domande la quali risposte sono completamente scorrelate tra di loro.

$$\begin{array}{c}
 P(A)P(B) \\
 /\backslash \\
 P(A)(1 - P(B)) \quad (1 - P(A))P(B) \\
 \backslash/ \\
 (1 - P(A))(1 - P(B))
 \end{array}$$

Considerazioni generali La probabilità complessiva nel caso di domande indipendenti A e B viene data da $P(A)$ per $P(B)$.

Se è conosciuta dal candidato la risposta alla domanda A ma non alla domanda B la probabilità di ottenere una risposta corretta è $P(A)$, mentre la probabilità di ottenere una risposta non corretta per B vale $1 - P(B)$. Il ragionamento duale è svolto nel calcolo della probabilità per la risposta corretta alla domanda B ma non ad A.

La probabilità di non ottenere alcuna risposta corretta alle due domande viene calcolata prendendo in considerazione gli eventi contrari a quelli coinvolti. Dunque per A la probabilità che il candidato non conosca la soluzione è $1 - P(B)$, dualmente per B la probabilità è $1 - P(A)$.

1.2.2 Eventi dipendenti

A e B sono due domande fortemente correlate tra di loro se si risponde correttamente ad una delle due domande si risponde correttamente anche all'altra.

$$\begin{array}{c}
 P(A)^2 \\
 /\backslash \\
 0 \quad 0 \\
 \backslash/ \\
 (1 - P(A))^2
 \end{array}$$

Considerazioni generali La probabilità complessiva nel caso di domande dipendenti A e B viene data da $P(A)$ per $P(B)$; ma $P(A) = P(B)$ dunque $P(A)^2 = P(B)^2$.

Conseguentemente se il candidato non conosce la risposta alla domanda A non può conoscere la risposta alla domanda B per cui la probabilità di conoscere uno dei due eventi è pari a 0.

In questo caso la probabilità a 0 è $(1 - P(A))(1 - P(B)) = (1 - P(A))^2$ essendo che $A=B$.

A implica B Se si sa rispondere alla domanda A di conseguenza si è in grado di rispondere anche alla domanda B.

Tuttavia non vale il ragionamento opposto, se si sa rispondere alla domanda B non significa che si è in grado di rispondere alla domanda A.

$$P(A)$$

$$\wedge$$

$$0 \quad P(B) - P(A)$$

$$\vee$$

$$1 - P(B)$$

Considerazioni generali La probabilità complessiva nel caso di domande dipendenti A e B viene data esclusivamente da $P(A)$ in quanto la conoscenza di sia di A che di B è possibile solo se si ha piena conoscenza di A.

Dunque la probabilità che si conosca la risposta alla domanda A ma non a B è impossibile (pari a 0); mentre se si ha conoscenza della domanda B ma non di A la probabilità si stanziava a $P(B) - P(A)$.

La probabilità a zero è $1 - P(B)$ indicatore dell'impossibilità di avere la risposta corretta per A.

1.3 Evento conosciuto ed evento indovinato

Durante un test il candidato deve saper scegliere la risposta, corretta o meno, alla domanda posta. Le variabili che entrano in gioco durante l'esecuzione dell'atto non riguardano esclusivamente la conoscenza personale del singolo. La probabilità di un evento A è data dalla formula:

$$P(A) = P(A_C) + P(A_I)$$

Le variabili in uso sono:

- $P(A_C)$: probabilità che il candidato sappia rispondere alla domanda A correttamente per sua conoscenza;
- $P(A_I)$: probabilità che il candidato sappia rispondere alla domanda A correttamente indovinando.

Per quanto appena definito sopra valgono le seguenti proprietà:

1. $P(B_C|A_C) = 1$
2. $P(B_C|A_I) = P(B_C)$
3. $P(B_I|A_C) = 0$
4. $P(B_I|A_I) = P(B_I)$

1.3.1 Probabilità di rispondere correttamente ad una domanda

Variabili coinvolte:

- $P(A)$: probabilità necessaria perchè si verifichi, per la domanda A, che il candidato dia la risposta corretta. Per la legge dei grandi numeri la frequenza porta alla probabilità.
- S_0 : insieme dei casi in cui in un domanda non viene scartata alcuna risposta dal dominio delle risposte possibili;
- S_1 : insieme dei casi in cui in una domanda viene scartata una risposta dal dominio delle risposte possibili;
- S_2 : insieme dei casi in cui in una domanda vengono scartate due risposte dal dominio delle risposte possibili.
- $P(I)$: probabilità di dare la risposta corretta alla domanda A indovinando;
- $P(C)$: probabilità di dare la risposta corretta alla domanda A per conoscenza.

Sapendo che $P(I) = P(A) - P(C)$ logicamente vale anche $P(A) = P(I) + P(C)$.

Se un candidato non è in grado scartare alcuna risposta dalla domanda ha 1 possibilità su 3 di, indovinando, dare la risposta corretta. Se un candidato invece risulta in grado di scartare una risposta, sbagliata, alla domanda rimane con 1 possibilità su 2 di poter dare la risposta corretta. Se invece, caso ottimo, il candidato ha piena conoscenza della domanda posta risulta in grado di scartare due risposte sbagliate lasciando un'unica risposta possibile, quella esatta. Il ragionamento sopra espresso può venire espresso con la seguente espressione:

$$P(A) = P(S_0)\frac{1}{3} + P(S_1)\frac{1}{2} + P(S_2)$$

Ora individuiamo quale è la probabilità effettiva per un candidato di dare la risposta corretta ad una domanda A.

$$\begin{aligned} 1 &= S_0 + S_1 + S_2 \\ S_0 &= 1 - S_1 - S_2 \end{aligned}$$

Sostituendo:

$$\begin{aligned} P(A) &= (1 - P(S_1) - P(S_2))\frac{1}{3} + P(S_1)\frac{1}{2} + P(S_2) \\ &= \frac{1}{3} - \frac{1}{3}P(S_1) - \frac{1}{3}P(S_2) + \frac{1}{2}P(S_1) + P(S_2) \\ &= \frac{1}{3} + \frac{1}{6}P(S_1) + \frac{2}{3}P(S_2) \end{aligned}$$

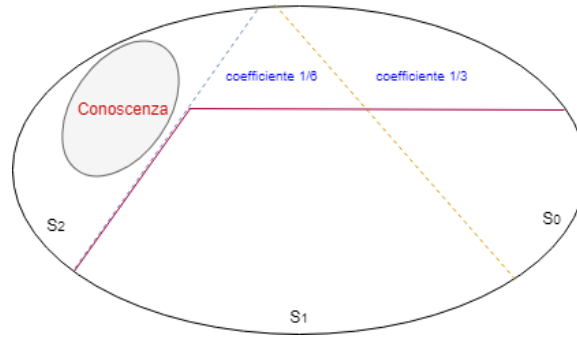


Figura 1: Rappresentazione insiemistica della probabilità di rispondere correttamente ad una domanda: $P(A)$

Considerazioni importanti

In conclusione $P(A) = \frac{1}{3} + \frac{1}{6}P(S_1) + \frac{2}{3}P(S_2)$. Ovvero la probabilità per un candidato di dare in una domanda A la risposta corretta dipende dai seguenti fattori:

- $\frac{1}{3}$: coefficiente che rappresenta la probabilità effettiva per chi non conosce la risposta alla domanda di dare la risposta corretta;
- $\frac{1}{2}P(S_1)$: coefficiente che rappresenta la probabilità effettiva di dare la risposta corretta quando il candidato è in grado di scartare una risposta sbagliata alla domanda;
- $\frac{2}{3}P(S_2)$: coefficiente che rappresenta la probabilità effettiva di dare la risposta corretta quando il candidato è in grado di scartare due risposte sbagliate alla domanda.

Dall'analisi della tipologia di eventi di coppia e dal calcolo della probabilità necessaria per poter rispondere correttamente ad una domanda, si è giunti alla valenza dei seguenti assiomi:

1. Le coppie di domande A e B devono essere fra loro disgiunte, altrimenti si genererebbero situazioni di invalidità dei risultati;
2. Per rispondere correttamente ad una domanda non è necessario che il candidato abbia piena conoscenza di tutti gli argomenti richiesti dall'esame, ma bensì ne risultano sufficienti $n - 1$;
3. La probabilità di conoscere è contenuta all'interno di S_2 , in quanto se un candidato conosce è conseguentemente in grado, da una domanda, di scartare due risposte sbagliate.

1.3.2 Il piano

La probabilità $P(A)$ che un candidato ha in gioco nel momento in cui si approccia a rispondere ad una domanda può venire rappresentata in un piano.

Di seguito viene mostrata l'immagine di un modellino, rappresentativo di $P(A)$, realizzato durante l'analisi.

TODO: foto modello

Ognuno dei tre assi cartesiani rappresenta un insieme dei casi di scarto (S_0 , S_1 , S_2). L'intersezione tra i punti del piano indica la regione accettabile contenente il range di valori assumibili da $P(A)$. Tale punto proiettato su ognuno dei tre assi permette l'individuazione esatta dei coefficienti delle variabili S_0 , S_1 , S_2 .

Ogni porzione del piano viene individuata con la seguente tecnica:

1. Per individuare ogni retta passante per S_0 , S_1 e S_2 è necessario assumere che $S_0 + S_1 + S_2 = 1$;
2. La retta passante per S_0 è rappresentabile per mezzo delle seguenti equazioni:

$$S_0 = 0 \text{ e } S_1 + S_2 = 1$$

In questo modo l'asse S_0 è fissato a 0 e estrapolando S_1 e S_2 da $S_1 = -S_2 + 1$ assumono valori tra (0,1).

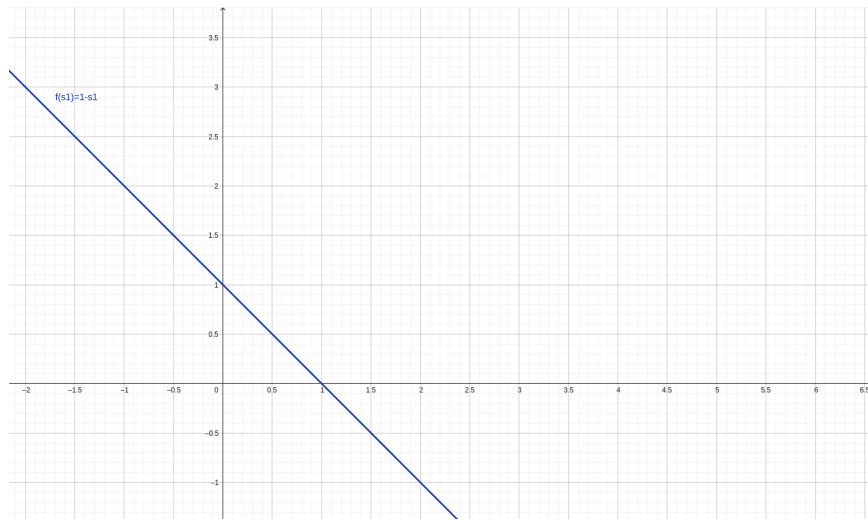


Figura 2: Rappresentazione della retta passante per $S_0 = 0$

3. Il medesimo ragionamento vale per le rette passanti per S_1 e S_2 .

$$S_1 = 0 \text{ e } S_0 + S_2 = 1$$

l'asse S_1 è fissato a 0 e S_0 e S_2 assumono valori tra (0,1).

$$S_2 = 0 \text{ e } S_1 + S_0 = 1$$

l'asse S_2 è fissato a 0 e S_0 e S_2 assumono valori tra (0,1).

4. In questo modo l'unione di tutte le rette passanti per gli assi creano la regione accettabile dei valori di $P(A)$.

Avendo rappresentato il piano si ottiene nei punti di intersezioni fra le tre rette la regione accettabile per $P(A)$. Inoltre è possibile, ora, individuare il fascio di rette che tangenti il piano permettono di affermare se una specifica domanda è, in base alla sua frequenza, ha difficoltà bassa, media, alta per un candidato.

- Se una domanda ha una difficoltà bassa la retta si situa passante per i punti $0 < S_2 \leq 1$ (molto vicino a 1) e $(S_0, S_1) < 0$ (tendenti a 0);
- Se una domanda ha una difficoltà alta la retta si situa passante per i punti $S_2 \leq 0$ (molto vicino a 0), $S_1 < 1$ e $S_0 \leq 1$ (tendente a non scartare alcuna risposta);

- Se una domanda ha una difficoltà media la retta si situa nella parte centrale della regione accettabile, passante per i punti $0 \leq (S_0, S_1, S_2) \leq 1$.

Rappresentazione di P(A)

Vediamo alcuni casi di come le domande possono venire rappresentate sul piano:

La funzione di partenza è:

$$F = \frac{1}{3} + \frac{1}{6}S_1 + \frac{2}{3}S_2$$

Va esplicitato S_1 , i passaggi utili da fare sono i seguenti:

$$\frac{-1}{6}S_1 = \frac{1}{3} + \frac{2}{3}S_2 - F \rightarrow S_1 = -4S_2 - 2 + 6F$$

Essendo che $0 \leq S_2 \leq 1$ usando $S_1 = 1$ e $S_2 = 0$ allora si ottiene che $F = \frac{1}{2} = 0.5$

Quanto appena calcolato può venire rappresentato graficamente impiegando la retta $S_1 = 1 - S_2$ (responsabile di definire una porzione del piano in base alle variabili coinvolte) e mediante la retta $S_1 = -4S_2 - 2 + 6F$ (che permette di calcolare il fascio di rette tangenti alla prima retta).

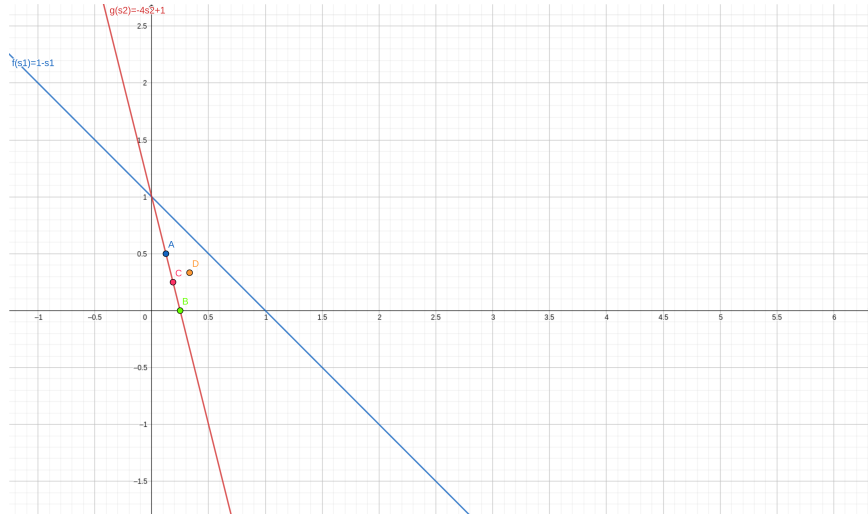


Figura 3: Rappresentazione di P(A) per una frequenza 0.5 proiettata su assi $S_0 = 0$, S_1 e S_2 .

Nella figura sopra sono rappresentati i seguenti significati:

- La linea azzurra rappresenta $S_2 = 1 - S_1$;
- La linea rosa rappresenta la retta tangente $S_1 = -4S_2 + 1$;
- Punto A (blu):

$$S_1 = 0.5 = \frac{1}{2}$$

$$\frac{1}{2} = -4S_2 + 1 \rightarrow 4S_2 = 1 - \frac{1}{2} \rightarrow S_2 = \frac{1}{8}$$

$$S_0 = 1 - \frac{1}{2} - \frac{1}{8} = \frac{3}{8}$$

Ovvero metà dei candidati sottoposti alla domanda sa scartare una delle risposte, lo 0.16% sa dare la risposta corretta e lo 0.36% non sa scartare alcune delle risposte possibili.

- Punto B (verde):

$$S_1 = 0$$

$$S_2 = \frac{1}{4}$$

$$S_0 = 1 - \frac{1}{4} = \frac{3}{4}$$

Ovvero nessun dei candidati sottoposti alla domanda sa scartare una delle risposte, lo 0.25% sa dare la risposta corretta e lo 0.75% non sa scartare alcune delle risposte possibili.

- Punto C (fucsia):

$$S_1 = \frac{1}{4}$$

$$S_2 = \frac{3}{16}$$

$$S_0 = 1 - \frac{1}{4} - \frac{3}{16} = \frac{9}{16}$$

Ovvero lo 0.25% dei candidati sottoposti alla domanda sa scartare una delle risposte, lo 0.19% sa dare la risposta corretta e lo 0.56% non sa scartare alcune delle risposte possibili.

- Punto D (arancione):

$$S_1 = \frac{1}{3}$$

$$S_2 = \frac{1}{3}$$

$$S_0 = 1 - \frac{1}{3} - \frac{1}{3} = \frac{1}{3}$$

Osserviamo che il punto in esame fuoriesce dalla regione delimitata dalla retta tangente di frequenza 0.5 ($S_1 = -4S_2 + 1$). Conseguenza diretta data dall'impossibilità di ottenere una probabilità del 50% sulla domanda con $\frac{1}{3}$ di candidati che sa scartare 2 risposte, $\frac{1}{3}$ che ne sa scartare 1 e $\frac{1}{3}$ nessuna.

Vediamo ulteriori due esempi che permettono di valutare cosa accade nel piano nel caso di una frequenza:

1. Quasi in prossimità di 1;

2. Pari alla soglia minima dell'indovinato.

Il grafico è il seguente:

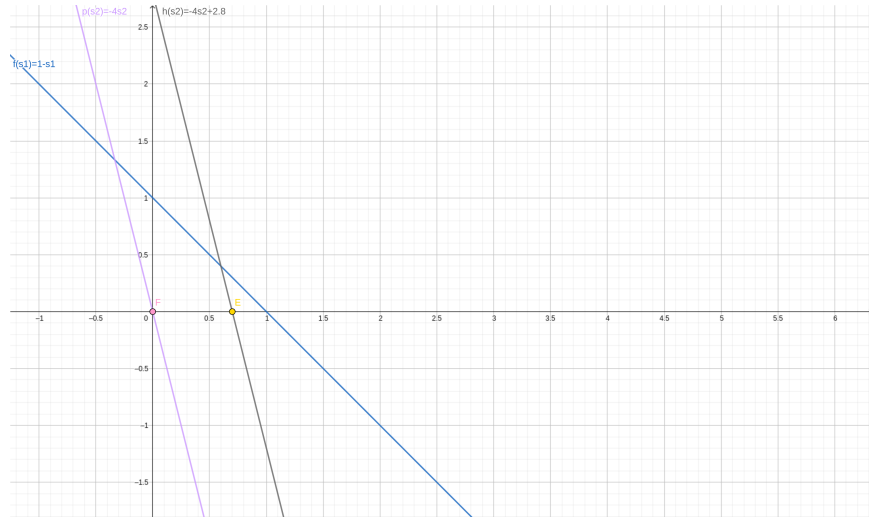


Figura 4: Rappresentazione di $P(A)$ per una frequenza 0.33 e 0.8 proiettate su assi $S_0 = 0$, S_1 e S_2 .

- La linea azzurra mostra la retta tangente con frequenza 0.80%. In questa abbiamo calcolato il punto E (giallo):

$$\begin{aligned} S_1 &= 0 \\ S_2 &= \frac{14}{20} \\ S_0 &= 0 \end{aligned}$$

Quasi la totalità dei candidati ha la conoscenza per poter scartare tutte le risposte sbagliate e dare la risposta giusta alla domanda.

- La linea viola mostra la retta tangente con frequenza 0.33%. In questa abbiamo calcolato il punto F (rosa):

$$\begin{aligned} S_1 &= 0 \\ S_2 &= 0 \\ S_0 &= 1 \end{aligned}$$

Ovvero nessuno dei candidati ha la conoscenza per poter scartare nè una nè due risposte, per cui l'unica possibilità per un candidato di rispondere alla domanda è indovinare. È evidente come se un candidato non sa la risposta ad una domanda ha una probabilità dello 0.33% di poter indovinare la risposta corretta.

2 Rete neurale

La libreria utilizzata per sviluppare la Rete neurale è stata *ConvNetJS*. L'aspetto positivo di tale scelta è stata la semplicità nell'utilizzo del linguaggio javascript; l'aspetto negativo ha riguardato la totale mancanza di mantenibilità della libreria stessa che comporta la scarsità di esempi applicativi, oltre alla documentazione ufficiale, che costringono lo sviluppatore ad una ricerca approfondita personale in un ambiente ove le nozioni si presentano scarse e a continue prove per verificare la validità del codice prodotto.

Questions test - Prevision Neural Net

CONFIGURAZIONE RETE NEURALE

```
layer_defs = [];  
layer_defs.push({type:'input', out_sz:1, out_sy:1, out_depth:6});  
layer_defs.push({type:'fc', num_neurons:2, activation: 'tanh'});  
layer_defs.push({type:'fc', num_neurons:2, activation: 'tanh'});  
layer_defs.push({type:'regression', num_neurons:6});  
  
net = new convnetjs.Net();  
net.makeLayers(layer_defs);  
  
trainer = new convnetjs.SGDTrainer(net, {learning_rate:0.01, momentum:0.1, batch_size:10, l2_decay:0.001});
```

OUTPUT DELLA RETE

```
Richiesta di previsione inoltrata alla rete ...  
Richiesta di previsione accettata  
Il vettore [0,1,0,0,0,0] ha previsione calcolata di  
[-0.14178025275470646,-0.014639237878831693,-0.00131295482446743,0.11790470015357472,0.14410273493990078,0.1329311955165462]  
Rete Neurale in attesa ... Inserire risultati del test...  
Inserimento risposte alle domande del test andato a buon fine  
Ricapitolazione dati inseriti: Risposte ottenute [0,-1,0,0,0,0]  
Inizio allenamento della rete ...  
Richiesta di previsione inoltrata alla rete ...  
Richiesta di previsione accettata  
Il vettore [0,1,1,0,0,0] ha previsione calcolata di  
[-0.9793841414997074,-0.7824502014273779,0.34748891066895563,-0.8448991676224333,-0.5825080715009475,0.29369850449980084]  
Rete Neurale in attesa ...
```

INSERIMENTO DATI DI ADDESTRAMENTO

Number of fields (max 9999): [Insert fields](#)

Input1 Input2 Input3 Input4 Input5 Input6

[save](#)

[start trainer](#)

Vettore di previsione

[prevision trainer](#)

Figura 5: Interfaccia utente della Rete neurale di prova.

Durante il periodo 24/05 - 31/05 mi sono occupata dello sviluppo di una Rete neurale in grado di ricevere in input un training set di dimensione 6 e di restituire una previsione sui dati di apprendimento ricevuti. Il problema che la rete mira ad analizzare è quello discusso nel precedente capitolo *Analisi dei dati di probabilità*

Per agevolare l'apprendimento della rete, ed ottenere delle previsioni stabili mi sono occupata di implementare due metodi di generazione randomica di dati in modo da far apprendere massicciamente la stessa. Il dato prodotto consiste in un vettore di 6 elementi, composto da -1, 0 e 1 con il seguente criterio:

- **-1**: la domanda x è stata posta al candidato che ha risposto in maniera errata;
- **0**: la domanda x non è stata posta al candidato;
- **1**: la domanda x è stata posta al candidato che ha saputo rispondere correttamente.

Il primo metodo che ho sviluppato si occupa di generare un vettore di dati di apprendimento basandosi esclusivamente su come le domande sono interconnesse tra di loro (grazie all'uso di un grafo della conoscenza costruito ad hoc); il secondo metodo ripropone quanto perseguito dal primo metodo con il valore aggiunto di generazione di un profilo randomico di un candidato, che tiene conto della probabilità di risposta ad una domande seguendo la formula $P(A) = \frac{1}{3} + \frac{1}{6}P(S_1) + \frac{2}{3}P(S_2)$.

2.1 Test effettuati

Alcune decisioni che ho preso durante la configurazione della rete riguardano i seguenti settori:

1. Una rete neurale non deve, per fornire dei dati attendibili, possedere un numero di neuroni troppo elevato rispetto al trainset effettuato; altrimenti la previsione ritornerebbe l'identità del vettore di input della stessa, come conseguenza diretta della capacità troppo elevata di immagazzinare dati.
2. I layers, ho deciso, di allenarli mediante tecnica di regressione, che permette l'inserimento in input di una funzione obiettivo e l'ottenimento di un risultato, in output, anche in virgola mobile e composto di tanti elementi quanti sono i neuroni di regressione dichiarati. Per la mia rete di prova è necessario dichiarare 6 neuroni in regressione perchè l'output, appunto, che ci si aspetta dal sistema è di 6 elementi.
3. Per costruire un dataset di dati consistente che permettesse alla rete di imparare qualcosa ho costruito un grafo della conoscenza con lo scopo di mettere in relazione degli argomenti che coinvolgono uno o più domande.

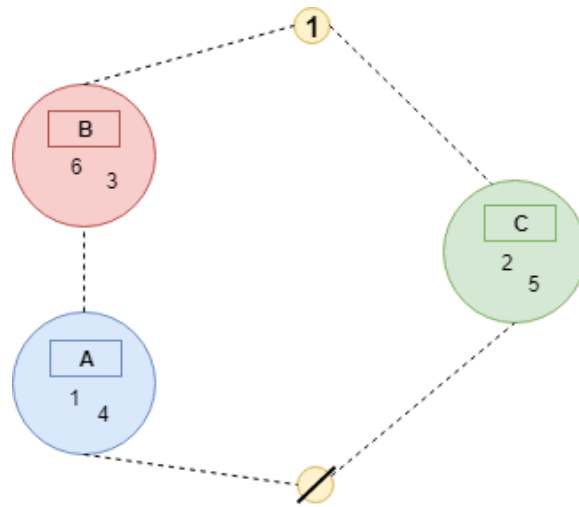


Figura 6: Grafo rappresentante le relazioni esistenti tra il set di domande di prova.

Per svolgere l'apprendimento ogni vettore, facente parte del dataset, viene dato in pasto alla rete che a sua volta provvede alla sua assimilazione come conoscenza mediante la tecnica dell'autoencoder, ovvero la rete impara il vettore riducendone lo spazio occupato.

4. Per creare il dataset ho ritenuto sufficiente generare *2000* vettori di risposta in modo da compiere in maniera esaustivo l'apprendimento della rete.

Il vettore passato in input per svolgere le previsioni è $[0,1,0,0,0,0]$. Di seguito riporto quanto è stato rilevato in fase di test.

2.1.1 Configurazione della rete: 4 neuroni per ciascuno dei 2 layers

Configurazione della rete utilizzata:

```
layer_defs = [];
layer_defs.push({type:'input', out_sx:1, out_sy:1, out_depth:6});
layer_defs.push({type:'fc', num_neurons:4, activation: 'tanh'});
layer_defs.push({type:'fc', num_neurons:4, activation: 'tanh'});
layer_defs.push({type:'regression', num_neurons:6});

net = new convnetjs.Net();
net.makeLayers(layer_defs);
```

```
trainer = new convnetjs.SGDTrainer(net, {learning_rate:0.01,
momentum:0.1, batch_size:10, l2_decay:0.001});
```

I layers utilizzati sono 2 e composti da 4 neuroni.

Training set standard a 4 neuroni per ciascuno dei 2 layers

- [0.06185960218322828,-0.27133263709888233,0.11724865743521667,-0.03165246440684366,0.0930428691265781,-0.4358467270595998]

Appaiono in relazione le domande 1, 3, 5 e 2, 4, 6.

Gli scostamenti tra le coppie 2 e 5, 1 e 4, 3 e 6 non sono consistenti con quelle che sono le relazioni di dipendenza, è una differenza che parte da qualche millesimo fino 0.5 circa; che mi sembra troppo per venire associata solamente alla presenza di valori -1 all'interno del vettore di training. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4; nel test in analisi questo viene rispettato solo dalle coppie 1, 6 e 4, 6 per una differenza trascurabile attorno allo 0.1.

- [0.0591737964385515,0.04700554500974672,0.1770953961599532,-0.032873052126447984,-0.21556028214336714,0.05438007273096417]

Appaiono in relazione le domande 1, 2, 3, 6 e 4, 5.

Gli scostamenti tra la coppia 3, 6 sono consistenti con quelle che sono le relazioni di dipendenza fra le domande; invece per quanto concerne le coppie 1, 4 e 2, 5 hanno una differenza che oscilla da qualche millesimo fino a 0.2, che mi sembra adeguata dalla presenza di -1 all'interno dei vettori di training. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4; nel test in analisi questo vale per le coppie 1, 6; in tutti gli altri casi vi è una differenza trascurabile di qualche millesimo.

- [0.24761793901633827,-0.006279637560350819,-0.05180076990401846,0.1813232184712338,-0.057742845674898,0.07468235957894961]

Appaiono in relazione le domande 1, 4, 6 e 2, 3, 5.

Gli scostamenti tra le coppie 1, 4 e 2, 5 sono consistenti con quelle che sono le relazioni di dipendenza fra le domande; invece per quanto concerne la coppia 3, 6 vi è una differenza attorno allo 0.012 circa; che mi sembra adeguata alla presenza di valori -1 all'interno del vettore di training. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4; nel test in analisi solo la regola viene rispettata pienamente.

- [-0.41386243742641304,0.2293601640299243,-0.15172273330068514,-0.08993270051292439,-0.16696727087969201,-0.030153227547984836]

Appaiono in relazione 1, 3, 4, 5, 6 e 2 (a parte).

Gli scostamenti tra le coppie 2 e 5 non sono consistenti con quelle che sono le relazioni di dipendenza fra le domande per una differenza attorno allo 0.2 circa. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4; nel test in analisi solo dalle coppie 4, 3 e 4, 6; in tutti gli altri casi la differenza oscilla da uno 0.25 ad uno 0.4 circa che mi sembra troppo da associare esclusivamente alla presenza di valori -1 nel set di training.

Training set con generazione del profilo di un candidato e calcolo delle probabilità di risposta a 4 neuroni per ciascuno dei 2 layers

- [0.12233318607288908,0.11271870424991579,0.03946474842257475,0.010689268243140052,-0.2212488419725651,0.08871808059256903]

Appaiono in relazione le domande 1, 2, 3, 4, 6 e 5 (a parte).

La coppia 2, 5 non è in relazione stretta con una differenza dello 0.3 circa. La domanda 4 presenta con una positività inferiore rispetto alla domanda sia 3 che 6 non mostrandosi così conforme alla regola per una differenza trascurabile attorno a qualche millesimo, invece la domanda 1 risulta conforme. L'anomalia può venire ricondotta all'uso di un set con dati "spuri", calcolati mediante la probabilità che un candidato ha di rispondere correttamente o meno ad una i-esima domanda (tale formula ha fatto venire meno la validità parziale delle relazioni che intercorrono tra le domande) che alla presenza dei valori -1 del vettore di training. Il secondo fattore però ha sicuramente un'influenza inferiore rispetto al primo sui risultati ottenuti.

- [0.10964014479802053,0.1858249376842786,-0.5006601985527853,-0.0334676581944329,-0.09971574849174258,-0.3801105216529717]

Appaiono in relazione le domande 1, 2 e 3, 4, 5, 6.

Le coppie 2, 5 e 1, 4 non è più in relazione stretta ma con una differenza 0.2 - 0.5; questo non vale per la coppia 3, 6 che rimane conforme alla regola. Sia la domanda 1 che 4 si presentano in rapporto con la domanda 3 e 6 con positività superiore conforme alla regola. Le anomalie riscontrate possono venire ricondotte all'uso di un set con dati "spuri", calcolati mediante la probabilità che un candidato ha di rispondere correttamente o meno ad una i-esima domanda (tale formula ha fatto venire meno la validità parziale delle relazioni che intercorrono tra le domande) che alla presenza dei valori -1 del vettore di training. Il secondo fattore però ha sicuramente un'influenza inferiore rispetto al primo sui risultati ottenuti.

- [0.6928150525498127,-0.12346483511365285,0.3519927766897261,-0.29213616574371004,-0.2664118985214089,-0.4278610241686793]

Appaiono in relazione le domande 1, 3 e 2, 4, 5, 6.

La coppia 1, 4 non è più in relazione stretta ma con una differenza importante che oscilla attorno allo 0.8 circa. La domanda 1 presenta una positività superiore rispetto alla domanda alla 3 e 6, mostrandosi conforme alla regola; invece la domanda 4 ha una differenza attorno allo 0.6 - 0.8. L'anomalia può venire ricondotta all'uso di un set con dati "spuri" usati per effettuare il training degli stessi.

- [-0.36082231867252107,-0.04413191737526341,0.5586488173865226,0.7496419035838204,-0.3774536487859032,0.20900746632808334]

Appaiono in relazione le domande 1, 2, 5 e 3, 4, 6.

Le domanda 2, 5 e 3, 6 sono in relazione stretta, questo non vale per la coppia 1, 4 che ha una differenza che supera l'unità. La domanda 4 presenta una positività superiore rispetto alla domanda sia 3 che 6 dimostrandosi conforme alla regola; negli altri casi la differenza si attesta attorno con una differenza attorno 0.6 - 0.8. L'anomalia può venire ricondotta all'uso di un set con dati "spuri" usati per il effettuare il training degli stessi.

Osservazioni

La configurazione testata si compone di 4 neuroni a layer su una base di 2000 test correndo il rischio di avere una rete che apprende troppo e come effetto negativo "veda" addirittura cose che non esistono.

Per fare un 'ulteriore verifica del sistema da me sviluppato ne ho mutato la configurazione riducendo il numero di neuroni presenti in ciascun layers e/o il numero di layers presenti.

Le nuove configurazione su cui ho effettuato i test sono esposte nei paragrafi seguenti.

2.1.2 Configurazione della rete a 2 neuroni per ciascuno dei 2 layers

Configurazione della rete utilizzata:

```
layer_defs = [];
layer_defs.push({'type':'input', out_sx:1, out_sy:1, out_depth:6});
layer_defs.push({'type':'fc', num_neurons:2, activation: 'tanh'});
layer_defs.push({'type':'fc', num_neurons:2, activation: 'tanh'});
layer_defs.push({'type':'regression', num_neurons:6});
```

```
net = new convnetjs.Net();
net.makeLayers(layer_defs);

trainer = new convnetjs.SGDTrainer(net, {learning_rate:0.01,
momentum:0.1, batch_size:10, l2_decay:0.001});
```

I layers utilizzati sono 2 composti da 2 neuroni.

Training set standard su rete a 2 neuroni per ciascuno dei 2 layers

- [0.21244167155728633,0.3163244589382085,0.5553080475586194,0.9592565134409513,-0.4523485201932427,0.05451735176218134]

Appaiono in relazione le domande 1, 2, 3, 4, 6 e 5 (a parte).

Gli scostamenti tra le coppie 1, 4 e 3, 6 sono consistenti con quelle che sono le relazioni di dipendenza fra le domande; invece per la coppia 2, 5 i segni sono opposti con una differenza dello 0.8, che posso far ricondurre alla presenza dei valori -1 nel vettore di training. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4; in questo test la regola viene rispettata ad esclusione del rapporto tra le coppie 3, 1 per una differenza attorno allo 0.3. L'anomalia può venire ricondotta all'uso di un set con dati "spuri" usati per il effettuare il training degli stessi.

- [-0.04983929591704967,-0.028242217295170846,-0.08053647611979069,-0.1306487555784668,0.06466493103049097,-0.10238617176530473]

Appaiono in relazione le domande 1, 2, 3 4, 6 e 5 (a parte)

Gli scostamenti tra le coppia 2 e 5 non sono consistenti con quelle che sono le relazioni di dipendenza fra le domande; la differenza è trascurabile oscilla tuttavia nell'ordine di qualche millesimo Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4, la regola non viene rispettata esclusivamente nel rapporto tra le domande 3 e 6 per una differenza trascurabile di qualche millesimo.

- [-0.20255330492560975,0.08174389617093263,0.3793856082770166,-0.09440896178066316,-0.2459101239568775,-0.05389443129944521]

Appaiono in relazione le domande 1, 4, 5, 6 e 2, 3.

Gli scostamenti tra la coppia 1, 4 sono consistenti con quelle che sono le relazioni di dipendenza fra le domande; invece per le coppie 3, 6 e 2, 5 i segni sono opposti con una differenza che oscilla attorno allo 0.2 - 0.3. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4, la regola non viene rispettata; ma per una differenza trascurabile che oscilla da pochi millesimi fino ad un valore

attorno allo 0.3; l'uguaglianza con quanto rilevato appena sopra mi fa poter associare tali anomalie ad alterazioni dovute alla presenza di valori -1 nel vettore di training.

- [-0.27852676930010534, 0.5313659297438961, -0.18141922931321913, -0.2725908891309912, -0.24799393899350472, 0.12846505297906688]

Appaiono in relazione le domande 1, 3, 4, 5 e 2, 6.

Gli scostamenti tra le coppie 3 e 6, 2 e 5 non sono consistenti con quelle che sono le relazioni di dipendenza fra le domande; la differenza si attesta nei valori 0.4 - 0.7 circa. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4, la regola non rispettata da nessuna delle coppie; ma in ogni caso la differenza è marginale nell'ordine di qualche millesimi che posso sempre ricondurla ad anomalie ad alterazioni dovute alla presenza di valori -1 nel vettore di training.

Training set con generazione del profilo di un candidato e calcolo delle probabilità di risposta a 2 neuroni per ciascuno dei 2 layers

- [0.0014939473221869328, 0.00008866264558449942, -0.0051039572299997425, -0.020202729539546355, 0.009480113116598188, 0.0018674039920875888]

Appaiono in relazione le domande 1, 2, 5, 6 e 3, 4.

Gli scostamenti tra la coppia 2 e 5 sono consistenti con quelle che sono le relazioni di dipendenza fra le domande; invece per la coppia 1, 4 e 3, 5 i segni sono opposti con una differenza tuttavia trascurabile inferiore al millesimo. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4, la regola viene rispettata nel caso della domanda 1 che supera la frequenza di 3 e 4, poco male perchè la differenza è trascurabile inferiore al millesimo, da poter associare ad alterazioni dovute alla presenza di valori -1 nel vettore di training.

- [0.544465180879028, 0.30741483494259525, -0.05978298930333867, 0.07618785219787755, 0.20828694965540176, 0.28254321048406633]

Appaiono in relazione le domande 1, 2, 4, 5, 6 e 3 (a parte)

Gli scostamenti tra le coppie 1, 4 e 2, 5 sono consistenti con quelle che sono le relazioni di dipendenza fra le domande; invece per la coppia 3, 6 i segni sono opposti con una differenza dello 0.3, che posso far ricondurre alla presenza dei valori -1 nel vettore di training. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4, la regola non viene rispettata dalla domanda 4 in rapporto con 3; ma la differenza sta sempre nell'ordine dei millesimi da poter associare ad alterazioni dovute alla presenza di valori -1 nel vettore di training.

- [0.1910638695115555,0.000591429229778241,0.03066702296109937,0.08222365813985216,0.007067502425449214,-0.013488907175148557]

Appaiono in relazione le domande 1, 4, 2, 3, 5 e 6 (a parte).

Gli scostamenti tra le coppie 1, 4 e 2, 5 sono consistenti con quelle che sono le relazioni di dipendenza fra le domande; invece per la coppia 3, 6 i segni sono opposti con una differenza tuttavia trascurabile che oscilla attorno all'ordine dei millesimi, che posso far ricondurre alla presenza dei valori -1. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4, la regola viene rispettata pienamente.

- [0.06749554602043151,-0.5960274542125068,-0.29293270738073995,-0.887419364280586,1.0148867009493885,0.8353342224531083]

Appaiono in relazione le domande 1, 5, 6 e 2, 3, 4.

Gli scostamenti tra le coppie 1 e 4, 2 e 5, 3 e 6 non sono consistenti con quelle che sono le relazioni di dipendenza fra le domande i segni sono opposti con una differenza trascurabile attorno a qualche millesimo fino a 0.7. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4, la regola viene rispettata esclusivamente nel rapporto tra le domande 3 e 4; con differenza è che va dallo 0.7 fino ad 1 unità. In questo caso mi sembra adeguato far risalire tali anomalie alla presenza di valori -1 nel vettore di training; anche se le differenze riscontrate appaiono molto elevate.

Osservazioni

Confrontando i risultati ottenuti dalla rete con i layers impostati a 4 neuroni con quanto emerso dai dati risultanti dalla rete a 2 neuroni, posso dire che sia nel caso di Training set standard che con generazione di profilo del candidato la situazioni, rispetto ai valori attesi, nel secondo gruppo di test sono leggermente migliori (entambe le configurazioni standard presentano una situazione critica, 1-3-1-0 contro 0-2-1-2 di non consistenza delle coppie tra la seconda e prima).

Emerge nel training standard una previsione che rispecchia uniformemente il grafo della conoscenza utilizzato a meno di errori di millesimi trascurabili (raramente viene superato lo 0.4. di differenza tra il valore reale e quello atteso). Tale effetto è meno evidente quando al set viene applicata la formula della probabilità di una domanda perchè i dati, rispetto al grafo, vengono "sporcati"; ma comunque le coppie che risultano ancora tali e la frequenza che vincola le domande dell'insieme A con quelle dell'insieme B rimangono di una precisione lievemente superiore rispetto alla prima configurazione della rete.

2.1.3 Configurazione della rete a 4 neuroni per 1 layer

Configurazione della rete utilizzata:

```
layer_defs = [];  
layer_defs.push({type:'input', out_sx:1, out_sy:1, out_depth:6});  
layer_defs.push({type:'fc', num_neurons:4, activation: 'tanh'});  
layer_defs.push({type:'regression', num_neurons:6});  
  
net = new convnetjs.Net();  
net.makeLayers(layer_defs);  
  
trainer = new convnetjs.SGDTrainer(net, {learning_rate:0.01,  
    momentum:0.1, batch_size:10, l2_decay:0.001});
```

Viene utilizzato un unico layer da 4 neuroni.

Training set standard su rete a 4 neuroni per 1 layer

- [-0.33962352034516724,0.15766643526150442,0.5311946439432398,
-0.8983729798246516,-0.07735861910229852,0.045229619914774866]

Appaiono in relazione le domande 1, 4, 5 e 2, 3, 6.

Gli scostamenti tra la coppia 2, 5 e 1, 4 sono consistenti con quelle che sono le relazioni di dipendenza fra le domande; invece per la coppia 2 e 5 i segni si presentano con una differenza di circa uno 0.15. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4, la regola non viene rispettata per nessuna delle coppie per una differenza molto elevata che supera 0.5 fino a superare l'unità stessa, ciò mi sembra eccessiva se fatta risalire solo alla presenza di valori -1 nel vettore di training.

- [0.4689181585312483,-0.12738526421179722,-0.020252601645710966,
0.3690665851248379,0.14974405282488307,0.09966055983535718]

Appaiono in relazione le domande 1, 4, 5, 6 e 2, 3.

Gli scostamenti tra la coppia 1 e 4 è consistente con quelle che sono le relazioni di dipendenza fra le domande; invece le coppie 2 e 5, 3 e 6 presentano una differenza che va da qualche millesimo fino a 0.25 circa. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4, la regola in questo test viene rispettata pienamente.

- [-0.00792331014831936,-0.5620727426105633,0.050227369930150756,
0.06293398881662707,-0.013272174547818288,-0.3428331763255027]

Appaiono in relazione le domande 1, 2, 5, 6 e 3, 4.

Gli scostamenti tra la coppia 2 e 5 sono consistenti con quelle che sono

le relazioni di dipendenza fra le domanda; invece per le coppie 3, 6 e 1, 4 presenta dei segni con una differenza tocca qualche millesimo fino a 0.3. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4, la regola viene rispettata nel caso delle coppie 6, 1 e 6, 4; in tutti gli altri casi la differenza è però trascurabile nell'ordine di millesimi. In questo caso mi sembra adeguato far risalire tali anomalie esclusivamente alla presenza di valori -1 nel vettore di training.

- i [0.16382419484909602, -0.4843321857549283, 0.0798110372950237, 0.1468091810283858, 0.09370650530018387, -0.6027027216393346]

Appaiono in relazione le domande 1, 3, 4, 5 e 2, 6.

Gli scostamenti tra la coppia 1, 4 sono consistenti con quelle che sono le relazioni di dipendenza fra le domanda; invece per le coppia 3, 6 e 2, 5 i segni si presentano con una differenza che va da 0.4 ad 0.6, che mi sembra troppo per venire associata solamente alla presenza di valori -1 all'interno del vettore di training. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4, la regola viene rispettata pienamente; tuttavia lo sfasamento iniziale tra i valori delle coppie mi sembra eccessivo se fatta risalire solo alla presenza di valori -1 nel vettore di training.

Training set con generazione del profilo di un candidato e calcolo delle probabilità di risposta a 4 neuroni per 1 layer

- [0.12770865743682994, -0.24336167359634645, 0.3227039332157453, -0.03629205324980521, 0.0025701762607406864, -0.0979149988715754]

Appaiono in relazione le domande 1, 3, 5 e 2, 4, 6.

Gli scostamenti tra la coppia 2 e 5, 3 e 6, 1 e 4 non sono consistenti con quelle che sono le relazioni di dipendenza fra le domande per una differenza che va da qualche millesimo fino a 0.2 circa. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4, la regola non viene rispettata nel caso delle coppie 1, 3 e 4, 3 con una differenza da qualche millesimo fino attorno allo 0.2, che mi sembra possa essere fatta risalire alla presenza di valori -1 nel vettore di training.

- [-0.3753246129012504, -0.32489290432521356, 0.11638579769117009, 0.1637997248566981, 0.3963980492869333, 0.08221072964296722]

Appaiono in relazione le domande 1, 2 e 3, 4, 5, 6.

Gli scostamenti tra le coppie 1, 4 e 2, 5 non sono consistenti con quelle che sono le relazioni di dipendenza fra le domanda con una differenza attorno allo 0.6 - 0.7. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4, la regola non

viene rispettata dalle coppie 6, 1 e 4, 1 con una differenza attorno allo 0.4 circa. In questo caso mi sembra eccessivo far risalire tali anomalie esclusivamente alla presenza di valori -1 nel vettore di training.

- [-0.7376658352719108, 0.039807798764236374, 0.15698050207879827, 0.11018118060896148, 0.47406083711443714, 0.28114954244803336]

Appaiono in relazione le domande 1 e 2, 3, 4, 5, 6.

Gli scostamenti tra la coppia 2, 5 e 3, 6 sono consistenti con quelle che sono le relazioni di dipendenza fra le domanda; invece la coppia 1, 4 presenta dei segni con una differenza attorno allo 0.8 circa. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4; la regola non viene rispettata da alcuna delle coppie con una differenza che va da qualche millesimo fino ad 0.9 circa. In questo caso mi sembra adeguato far risalire tali anomalie alla presenza di valori -1 nel vettore di training; anche se le differenze riscontrate appaiono molto elevate.

- [-0.23242901300784663, -0.06707019901013181, 0.10105686814296164, -0.031588675400]

Appaiono in relazione le domande 1, 2, 4, 5, 6 e 3 (a parte).

Gli scostamenti tra le coppie 1, 4 e 2, 5 sono consistenti con quelle che sono le relazioni di dipendenza fra le domanda; invece per la coppia 3, 6 i segni si presentano con una differenza inferiore allo 0.4. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4, la regola viene rispettata dalle coppie 6, 1 e 6, 4; negli altri casi la differenza si attesta attorno ai valori 0.1 - 0.4. In questo caso mi sembra adeguato far risalire tali anomalie alla presenza di valori -1 nel vettore di training;

Osservazioni

Rispetto a quanto osservato nei casi precedenti, mi risulta che sia nel caso di utilizzo di training set standard che con profilo di candidato la tecnica dei risultati buoni ma inferiori rispetto all'impiego di due layers con 2 neuroni ognuno (1-2-2-2 contro 0-2-1-2 del training set standard). Se con il set di dati "spuro" potrei sorvolare di più nelle oscillazioni delle previsioni, questo non vale per il primo set che presenta delle anomalie, anche se ridotte, rispetto a quanto presentato dalla precedente configurazione.

*Fino ad ora la configurazione di rete che ha dato i maggiori risultati di previsione risulta essere a **2 layers con 2 neuroni ognuno**.* Tale conclusione ha senso in quanto il grafo della conoscenza che ho usato come base per costruire i vettori di apprendimento è composto da 3 nodi (A, B, C) indicanti 3 neuroni; il quarto può venire valutato come un nodo della rete utile per parametri in entrata e in uscita.