

1 Rete neurale

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 trattato dalla rete è quello discusso nel precedente documento *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 0, 1 e -1 con il seguente criterio:

- **-1**: la domanda x non è stata posta al candidato;
- **0**: la domanda x è stata posta al candidato che ha risposto in maniera errata;
- **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 risposta ad una domande seguendo la formula $P(A) = \frac{1}{3} + \frac{1}{6}P(S_1) + \frac{2}{3}P(S_2)$.

1.1 Configurazione della rete

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 layer utilizzati sono due e composti da 4 neuroni ciascuno vista la numerosità dei dati di training che ho impiegato durante l'addestramento, in modo che alla rete risulti possibile imparare in modo adeguato. Attenzione se la rete presentasse un numero di neuroni troppo elevato la previsione ritornerebbe l'identità del vettore di input della stessa, come conseguenza diretta della capacità troppo elevata di immagazzinare dati.

Il numero di neuroni in regressione devono essere 6, perchè l'output che ci si

aspetta da un vettore di dimensione 6 è composto da 6 elementi. Per costruire un dataset di dati consistente che permettesse alla rete di imparare qualcosa ho costruito un grafo con lo scopo di mettere relazione degli argomenti che coinvolgono uno o più domande.

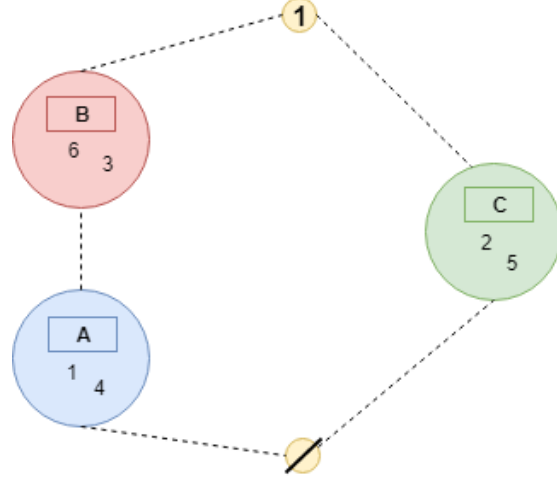


Figure 1: Grafo presentante le relazioni esistenti tra il set di domande di prova.

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. Per creare il dataset sono generati *2000* vettori di risposta in modo che sia possibile compiere in maniera esaustiva 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.

1.2 Osservazioni riscontrate

1.2.1 Training set standard

- $[0.18394862760524544, 0.5427447874383465, 0.4475798470511032, -0.2002756172921305, 0.07023832331402126, -0.38412626496750873]$

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

Gli scostamenti tra le coppie 2 e 5 sono consistenti con quelle che sono le relazioni di dipendenza fra le domande, invece per quanto concerne le coppie 2, 5 e 1, 4 sono dovuti dalla presenza di -1 all'interno dei vettori di training tuttavia tale valore supera lo 0.4 in entrambi i casi. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4; nel test in analisi questo vale per la domanda 1 e 4 in relazione con la domanda 6, negli altri casi non si è conformi a tale regola.

- $[-0.11235743604300916, -0.39879459369010783, -0.6219582601088702, 0.22754749414916, -0.3307584554090044, -0.39007701490038627]$

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

Gli scostamenti tra le coppie 3 e 6, 2 e 5 sono consistenti con quelle che sono le relazioni di dipendenza fra le domande, invece per quanto concerne la coppia 1 e 4 sono dovuti 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 sia la domanda 1 che 4 si presentano conformi alla regola.

- [-0.2399988601091234, 0.09747007794669733, 0.5093732175811206, 0.06546467766710193, 0.05567129781511258, -0.11672474718649554]

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 le coppie 3, 6 sono dovuti alla presenza di -1 all'interno dei vettori di training anche se la differenza supera lo 0.5. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4; nel test in analisi solo la domanda 4 in confronto con la domanda 6 rispetta la regola.

- [0.21422605841447054, -0.2636944179712092, -0.3706563171790509, 0.7764017490883244, -0.23816083562639187, -0.2524885890953481]

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

Gli scostamenti tra le coppie 1 e 4, 2 e 5, 3 e 6 sono consistenti con quelle che sono le relazioni di dipendenza fra le domande. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4; nel test in analisi tale regola viene rispettata perfettamente

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

- [0.0581156963538111, -0.1083899001264413, 0.17398947541144363, 0.19769424910168104, -0.06865983886403393, 0.06277020715810194]

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

Le domanda 1 e 4, 3 e 6 non sono più in relazione stretta; le coppie 2 e 5 rimangono invece consistenti. La domanda 1 presenta una positività inferiore rispetto alla domanda sia 3 che 6 non presentandosi conformi alla regola, invece la domanda 4 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.

- [0.07461451161172632, -0.034052968605699584, -0.12054982977145728, -0.014994452072747376, -0.02387186678202946, 0.037622445105928125]

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

Le domanda 1 e 4, 3 e 6, 2 e 5 non sono più in relazione stretta. La domanda 4 presenta una positività inferiore rispetto alla domanda sia 3 che 6 non presentandosi conforme alla regola, invece la domanda 1 risulta

conforme. L'anomalia può venire ricondotta all'uso di un set con dati "spuri" per il effettuare il training degli stessi.

- [0.009747730942870516,0.09241802945247207,-0.0925975996231829,-0.17029345029708107,-0.03732537178160159,0.01882705046994409]

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

Le domanda 1 e 4, 3 e 6, 2 e 5 non sono più in relazione stretta. La domanda 1 presenta una positività inferiore rispetto alla domanda sia 3 ma non rispetto alla domanda 6 non presentandosi parzialmente conforme alla regola, lo stesso vale per la domanda 4. L'anomalia può venire ricondotta all'uso di un set con dati "spuri" per il effettuare il training degli stessi.

- [-0.03893047230734836,-0.17335152037872065,-0.12242554699039307,-0.1449537537507802,-0.03262368658437825,0.16259494225750007]

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

Le domanda 1 ,4 e 2, 5 sono in relazione stretta, questo non vale per la coppia 3, 6. La domanda 4 presenta una positività inferiore rispetto alla domanda sia 3 che 6 non presentandosi conforme alla regola, invece la domanda 1 risulta conforme solo nel rapporto con la domanda 3. L'anomalia può venire ricondotta all'uso di un set con dati "spuri" per il effettuare il training degli stessi.

1.2.3 Osservazioni

Avendo messo 4 neuroni su una base di 2000 test, si corre il rischio che 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.

La nuova configurazione su cui ho effettuato i test è la seguente:

```
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});
```

1.2.4 Training set standard su rete a 2 neuroni

- [1.156980429249762,0.06806851158038928,0.3113362862886465,
0.17218787779201644,0.34650990282652194,-0.8215874801856704]

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

Gli scostamenti tra le coppie 1 e 4, 2 e 5 sono consistenti con quelle che sono le relazioni di dipendenza fra le domanda; invece per la coppia 3 e 6

i segni si presentano opposti tuttavia la differenza è minima e può venire associata alla presenza di valori -1 all'interno del vettore di training essendo che la differenza di valori risulta 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 nel caso della domanda 1 che supera di netto la frequenza di 3 e 6, poco male per la domanda 4 che supera esclusivamente la domanda 6 ma la differenza con la frequenza della domanda 1 sta nell'ordine di centesimi sempre da poter associare a alterazioni dovuti alla presenza di valori -1 nel vettore di training.

- [0.04982696367584444, 0.11290459035591142, -0.0696298921764785, 0.18228607258850865, 0.4

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

Gli scostamenti tra le coppie 1, 4 e 2, 5 e 3, 6 sono consistenti con quelle che sono le relazioni di dipendenza fra le domande. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4 la regola viene sia dalla domanda 6 che 3.

- [0.19504797225824305, -0.2295101910352556, -0.028760245350834636, 0.007144898117011814, -0.056011983451495176, -0.1803934455401963]

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

Gli scostamenti tra le coppie 1, 4 e 2, 5 e 3, 6 sono consistenti con quelle che sono le relazioni di dipendenza fra le domande. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4 la regola viene non rispettata pienamente sia dalla domanda 3 che 6.

- [0.08339384459022353, -0.15782370764467343, 0.0005080967853213457, -0.020723132326211795, -0.03207289911077578, 0.004139946640153436]

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

Gli scostamenti tra le coppie 3, 6 e 2, 5 sono consistenti con quelle che sono le relazioni di dipendenza fra le domande; invece per la coppia 1 e 4 i segni sono opposti posso farli ricondurre alla presenza di valori -1 nel vettore di training essendo che la differenza di valore è inferiore delle 0.18. Le domande 3 e 6 si dovrebbero presentare con una positività inferiore rispetto a 1 e 4 la regola viene non rispettata dalla domanda 6 nei confronti della domanda 4 e dalla domanda 3 nei confronti con la domanda 4 ma in ogni caso la differenza è marginale e posso sempre ricondurli a oscillazioni della rete.

1.2.5 Training set con generazione del profilo di un candidato e calcolo delle probabilità di risposta a 2 neuroni

- [-0.1035508475297289, 0.0633805382071866, -0.04231133040694016, 0.037670465236242075, 0.11461540750490316, -0.04359674755310973]

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

La domanda 1, 4 non è più in relazione stretta; le coppie 2, 5 e 3, 6 rimangono invece consistenti. La domanda 1 presenta una positività inferiore rispetto alla domanda sia 3 che 6 non presentandosi conformi alla regola, invece la domanda 4 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.

- $[-0.03694505953408343, -0.11462792478438069, 0.11455401234234787, -0.04130903688780185, -0.10017417464274732, 0.23137329348272218]$

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

Le domande 1 e 4, 3 e 6 e 2, 5 non sono più in relazione stretta. La domanda 4 presenta una positività inferiore rispetto alla domanda sia 3 che 6 non presentandosi conforme alla regola, lo stesso vale per la domanda 1. L'anomalia può venire ricondotta all'uso di un set con dati "spuri" per il effettuare il training degli stessi.

- $[0.09835734578250267, 0.2783003702917536, -0.5203071370347077, 0.21074141492570442, 0.0874911093414027, 0.22885850762943963]$

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

Le domande 3 e 6 non sono più in relazione stretta; le coppie 2, 5 e 1, 4 rimangono invece consistenti. La domanda sia 1 che 4 si presentano con una positività superiore rispetto alla regola. L'utilizzo di un set con dati "spuri" per il effettuare il training degli stessi in questo caso ha impattato marginalmente sui valori dati dal grafo della conoscenza,

- $[0.02267051416257774, 0.005474349898744291, 0.016535957156990452, 0.0034334957553278657, 0.016359614225486967, -0.009108496312618647]$

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

Le domande 1 e 4 e 2, 5 sono in relazione stretta, questo non vale per la coppia 3, 6. La domanda sia 1 che 4 si presentano una positività superiore rispetto alla domanda sia 3 che 6 presentandosi conforme alla regola, invece la domanda 1 risulta conforme solo nel rapporto con la domanda 3. Come sopra l'utilizzo di un set con dati "spuri" per il effettuare il training degli stessi in questo caso ha impattato marginalmente sui valori dati dal grafo della conoscenza,.

1.2.6 Osservazioni

Confrontando i risultati ottenuti dalla rete con i layers impostati a 4 neuroni con quanto emerso dai dati risultanti dalla rete con 2 neuroni, la layers, posso dire che sia nel caso di Training set standard che con generazione del profilo del candidato la situazioni rispetto ai valori attesi nel secondo gruppo di test sembra essere migliore.

Emerge nel training standard una previsione che rispecchia più uniformemente il grafo della conoscenza utilizzato, risultano addirittura dei set dei dati completamente conformi con le aspettative. 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 superiore rispetto alla prima configurazione della rete.