



Simone Cipriani &lt;sim.cipr@gmail.com&gt;

---

**Re: Info su Progetto DMO**

---

**Donatella Merlini** <donatella.merlini@unifi.it>

4 gennaio 2018 09:54

A: Simone Cipriani &lt;sim.cipr@gmail.com&gt;

Cc: Sabrina Innocenti &lt;sabrina.innocenti01@gmail.com&gt;

Nessun disturbo, ci mancherebbe. Quello che mi aspetto è un progettinio di data mining completo tramite una relazione che descriva prima le caratteristiche del dataset con calcolo di misure di similarità per vedere se tra i dati esistono particolari correlazioni, eventualmente utilizzando anche tecniche di visualizzazione (in tal caso probabilmente dovreste usare R, come fatto all'inizio del corso). Sarà probabilmente necessario poi un preprocessing per poter utilizzare gli algoritmi di clustering che userete. Ovviamente tutte queste fasi dovranno essere descritte nella relazione, inserendo i grafici che ritenete più interessanti e commentando i risultati. Potreste anche prevedere una fase di postprocessing per tener conto in qualche modo dei risultati trovati nel dataset. Come ultima fase mi aspetto un'analisi dei risultati che avete trovato: vengono fuori delle regolarità interessanti dai dati? Che significato hanno e che ripercussioni potrebbero avere i vostri risultati sul CdL Informatica? Mi aspetto di avere anche i file con i dati rielaborati in modo da poter rieseguire le vostre prove.

Ovviamente ci dovete mettere un po' di fantasia e tirar fuori la vostra anima data scientist!

Buon lavoro

Donatella Merlini

Il Giovedì 04/01/2018 09:07 Simone Cipriani ha scritto:

Buongiorno prof. Merlini,

mi dispiace disturbarla subito all'inizio dell'anno, ma avremmo un dubbio riguardo alla presentazione dei risultati del clustering.

In pratica, cosa le interessa vedere? Una relazione, un qualche output del software Weka, ...

Abbia pazienza per la domanda banale e auguri di un buon 2018,  
Simone

----- Messaggio inoltrato -----

Da: **Sabrina Innocenti** <sabrina.innocenti01@gmail.com>

Date: 3 gennaio 2018 17:20

Oggetto: Re: Info su Progetto DMO

A: Simone Cipriani <sim.cipr@gmail.com>

Bella domanda...

ancora non mi ero posta il problema perché non c'ho messo mano. Non mi viene in mente niente che ci abbia detto in proposito ma così a senso direi che dovrebbe essere una relazione corredata con figure per visualizzare i risultati. Magari chiediamo per conferma.

Sabrina

Il giorno 3 gennaio 2018 17:08, Simone Cipriani <sim.cipr@gmail.com> ha scritto:

Ciao,

ho cominciato ora a fare questo benedetto progetto. Sai per caso se ci sono indicazioni su come presentare i risultati del clustering? Nel senso, una relazione con screenshot e spiegazioni, una presentazione, ...?

Simone

----- Messaggio inoltrato -----

Da: **Donatella Merlini** <[donatella.merlini@unifi.it](mailto:donatella.merlini@unifi.it)>

Date: 22 dicembre 2017 17:12

Oggetto: Re: Info su Progetto DMO

A: Simone Cipriani <[sim.cipr@gmail.com](mailto:sim.cipr@gmail.com)>

Cc: Sabrina Innocenti <[sabrina.innocenti01@gmail.com](mailto:sabrina.innocenti01@gmail.com)>

Facciamo così,

visto che anche Sabrina Innocenti (che metto in cc) ha la necessità di fare il progetto da sola, potrebbe scegliere anche lei il progetto

"Coorti 2010- 2013 studenti Informatica - esami I, II e III anno clustering"

ma poi ognuno di voi lo svolge in modo autonomo, ovviamente con la possibilità di potervi consultare se necessario. In questo modo risparmiamo un progetto, altrimenti c'è il rischio che le tracce non siano sufficienti per tutti.

Se va bene, invio i dati anche a lei.

Cordiali saluti

Donatella Merlini

Il 22/12sa se la disturbo (di nuovo), ma ho visto la scelta del progetto sul Moodle ed avrei bisogno di chiederle una conferma.

Come le avevo accennato l'altro giorno alla fine della prova intermedia, io avrei bisogno di fare il progetto da solo perché il lavoro impegna la maggior parte delle mie giornate, quindi i miei orari di studio sono piuttosto "inconsueti" per coordinarmi con altri e svolgere un progetto insieme.

E' possibile? Che devo fare? Se scelgo un dataset ed una tecnica che ha già scelto qualcun'altro, mi ritrovo in un gruppo?

La ringrazio per la pazienza,  
Simone

--

=====

Donatella Merlini  
Dipartimento di Statistica, Informatica, Applicazioni "G. Parenti"  
Viale Morgagni 65, 50134 Firenze  
tel. +39 055 2751509 - fax. +39 055 4237436  
[donatella.merlini@unifi.it](mailto:donatella.merlini@unifi.it)  
<http://local.disia.unifi.it/merlini/>

=====