

Organizzazione e contenuti della relazione di mini-progetto

Filippo Chiandotto¹, Federico Andreas Dotti², and Simone Tosato³

¹ Corso di laurea magistrale in Scienze Statistiche, matricola 1234567
`primo.autore@studenti.unipd.it`

² Corso di laurea magistrale in Scienze Statistiche, matricola 1234567
`primo.autore@studenti.unipd.it`

³ Corso di laurea magistrale in Scienze Statistiche, matricola 1234567
`primo.autore@studenti.unipd.it`

Sommario La relazione si prefigge l'obiettivo di approfondire e (affinare?) metodi di classificazione binaria quali perceptron ed SVM, implementando delle tecniche volte alla riduzione del tempo di esecuzione degli algoritmi e alla loro scalabilità. Partendo dalla descrizione del capitolo 12 del libro di testo (Riferimento in bibliografia) e dagli approfondimenti basati sugli articoli è stata proposta una implementazione in Python utilizzando le strutture dati e le operazioni matriciali forniti dal pacchetto numpy (riferimento in bibliografia).

Keywords: SVM · Perceptrone · MapReduce · Center Distance Ratio Method · Vettori di supporto

1 Introduzione

Il lettore della relazione è lo studente medio del corso di laurea magistrale in Scienze statistiche al quale la relazione deve dare tutti gli strumenti per comprendere il contenuto. Ci si metta nei suoi panni e si scriva tutto ciò e solo ciò che serve. Chiedersi qual è il messaggio che lo studente deve “portarsi a casa”, esplicitarlo in questo paragrafo e concentrarsi su quello nel resto della relazione.

L'introduzione della relazione dovrà anche illustrare

- i motivi per i quali il problema affrontato è significativo e
- i metodi utilizzati sono *efficaci* ed *efficienti*.

Per significatività s'intende o il mero interesse teorico-scientifico o l'impatto industriale che avrebbe la soluzione del problema. Per efficacia s'intende la capacità di un metodo a risolvere il problema, mentre l'efficienza è relativa al costo computazionale necessario.

La relazione consiste di quattro paragrafi principali dopo questa introduzione e prima della bibliografia. Se il mini-progetto è descrittivo i paragrafi 2 e 3 devono essere particolarmente accurati ed estesi. Per la bibliografia si suggerisce BibTeX se si scrive con L^AT_EX.

2 Base di partenza

La base di partenza è formata dai metodi documentati nei libri di testo o negli articoli scientifici. Variazioni significativi di essi vanno tra i metodi proposti. In caso di dubbio si chieda al docente.

Per quanto riguarda gli strumenti *software*, si tenga presente che l'insegnamento è di Informatica e che quindi si preferisce lo studio, il progetto e l'implementazione di strutture di dati e di algoritmi “fatti in casa” piuttosto che l'utilizzo di ambienti di sviluppo, come ad esempio Pandas, Julia, R e simili, in cui le strutture e i programmi sono “nascosti” e disponibili come scatole chiuse. A lezione sono stati fatti parecchi esempi di cosa s'intende.

Per capire quando dover sviluppare propri algoritmi o strutture di dati, si provi ad esempio a vedere se i propri metodi sono già stati realizzati in qualche ambiente di sviluppo e se hanno tutte le caratteristiche desiderate, come ad esempio scalabilità e tempi di esecuzione in ordine polinomiale medio-basso. Se c'è un qualche “pezzo” di un metodo di base che non soddisfa almeno un requisito, come ad esempio la sparsità di una matrice o il tipo di dati accettati, allora è un “problema” da affrontare con propri algoritmi e strutture.

3 Problema affrontato

Il problema affrontato dal gruppo deve essere illustrato nel dettaglio in questo paragrafo. Si tenga presente che questo insegnamento è di Informatica e quindi sono gli aspetti informatici da mettere in evidenza. Durante le lezioni ne sono stati messi in evidenza parecchi se non addirittura tutti:

- la dimensione dei dati,
- il tipo dei dati,
- il tempo di esecuzione e l'ordine di complessità,
- lo spazio di memoria di lavoro e quello di disco.

4 Metodi proposti

La metodologia è l'insieme coordinato dei metodi che formano la base di partenza o che sono stati proposti dal gruppo e che quindi sono da ritenersi un contributo originale. Non è necessario che il gruppo proponga nuovi metodi, ma se il gruppo ritiene che ci siano, la relazione deve evidenziarli e dire in cosa si differiscono dalla base di partenza.

Se il mini-progetto è teorico o metodologico allora i risultati teorici e le metodologie vanno descritti in questo paragrafo. La descrizione deve essere rigorosa, non necessariamente matematica. Si faccia attenzione alle definizioni. Se ci sono teoremi già riportati in letteratura, è sufficiente l'enunciato ed è necessaria la citazione ai lavori scientifici. I propri teoremi richiedono una propria dimostrazione che spesso esaurisce il mini-progetto.

In ogni caso la descrizione deve essere completa e precisa; poiché la relazione deve essere di 12 pagine al massimo è importante che si scelga con cura cosa scrivere.

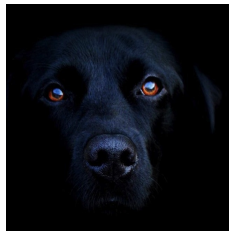
5 Esperimenti

Se il mini-progetto è sperimentale o empirico gli studi di questo tipo vanno descritti in questo paragrafo. Si descrivano le collezioni di dati in termini di dimensione e tipo ed i risultati dei confronti tra metodi di base o quelli proposti. Cruciale è la descrizione accurata degli esperimenti: essa deve permetterne la replicazione.

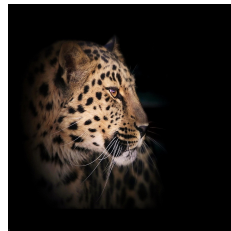
Il *software*, sia quello a titolo esemplificativo dei concetti che quello che realizza i propri metodi, va caricato su moodle nella cartella messa a disposizione più avanti. Nella relazione va scritta la descrizione dell'algoritmo in modo preciso e completo da permetterne la riproduzione.

Riferimenti bibliografici

1. Autore, F.: Article title. Journal **2**(5), 99–110 (2016)
2. Autore, F., Autore, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) CONFERENCE 2016, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg (2016). <https://doi.org/10.1007/1234567890>
3. Autore, F., Autore, S., Autore, T.: Book title. 2nd edn. Publisher, Location (1999)
4. Autore, A.-B.: Contribution title. In: 9th International Proceedings on Proceedings, pp. 1–2. Publisher, Location (2010)
5. LNCS Homepage, <http://www.springer.com/lncs>. Last accessed 4 Oct 2017



(a) Primo Autore



(b) Secondo Autore



(c) Terzo Autore