

SPARE: SOMETHING FOR PATTERN RECOGNITION

Antonello Rizzi, Guido Del Vescovo, Lorenzo Livi, Fabio Massimo Frattale Mascioli

Dipartimento di Ingegneria dell'Informazione, Elettronica e Telecomunicazioni
Via Eudossiana 18, 00184 Roma

SPARE - Something for PAttern REcognition – è una libreria software per lo sviluppo rapido di applicazioni inerenti tecniche di Pattern Recognition ed Intelligenza Computazionale. L'idea originale nacque circa 14 anni fa, quando il gruppo di ricerca sulle reti neurali diretto dal Prof. Giuseppe Martinelli affrontò per la prima volta il problema della gestione efficace di molti distinti progetti software in cui erano implementati algoritmi di clustering, sistemi di classificazione ed approssimazione funzionale, sistemi di predizione, algoritmi di ottimizzazione evolutiva, reti neuro-fuzzy.

In particolare furono evidenziate le seguenti criticità:

1) Molti tesisti spendevano una quota significativa dei loro sforzi nel riscrivere versioni di algoritmi consolidati (come l'algoritmo di clustering K-means), librerie per la manipolazione di matrici e la gestione di dati, interfacce grafiche, etc.

2) Le versioni finali degli applicativi sviluppati durante il lavoro di tesi non erano compatibili con un comune standard per la manipolazione dei dati, creando un serio problema di interoperabilità.

3) La tracciatura e la caratterizzazione di versioni differenti di un medesimo algoritmo implementate in distinti pacchetti software era di difficile attuazione; di conseguenza la valutazione comparativa delle prestazioni di tali algoritmi non poteva essere svolta con il dovuto rigore.

Oltre queste motivazioni inerenti la progettazione del software, la ricerca metodologica sui temi di ricerca del Soft Computing evidenziò che un sistema di modellamento induttivo comunque complesso può sempre essere pensato e riorganizzato come un insieme di piccoli moduli software cooperanti tra loro che implementano algoritmi basilari caratterizzati da una comune interfaccia accuratamente definita.

Per tali ragioni si iniziò a definire un framework per lo sviluppo di applicativi software, inteso non solo come una semplice collezione di classi organizzate in una libreria, ma come un più ampio insieme di oggetti, procedure e raccomandazioni che stabilissero regole per la documentazione del software, il formato dei dati e la loro rappresentazione, le metodologie per lo sviluppo di progetti software, la definizione delle interfacce utente, le procedure di test e benchmarking, etc. Ad oggi sono stati concepiti diversi framework di questo tipo, ciascuno specializzato su diversi requisiti, come la standardizzazione dei files, la modularizzazione del software, la correttezza, la standardizzazione dell'interfaccia grafica, etc. Negli ultimi due anni è stato avviato un lavoro di unificazione, stabilendo un nuovo framework concepito per riunire le caratteristiche migliori di lavori precedenti, in cui SPARE rappresenta il nucleo di tale progetto.

SPARE è una libreria software scritta in C++ basata sulla metodologia di *template meta-programming*, avente lo scopo di fornire classi e procedure elementari per lo sviluppo rapido di applicazioni inerenti le tecniche di Intelligenza Computazionale. E' stata concepita per massimizzare la qualità del software secondo i seguenti requisiti, elencati in ordine di importanza sulla base dei pesi relativi che sono stati considerati nel prendere decisioni critiche nelle varie fasi della progettazione della libreria:

1) *Efficienza*. La progettazione della libreria ha seguito i dettami della *programmazione generica* tramite un uso sistematico di *templates*, invece di un approccio più tradizionale basato sull'ereditarietà ed il polimorfismo, in modo da consentire il *binding* (ossia la decisione circa l'attributo o il metodo da richiamare in un dato momento dell'esecuzione del programma) a tempo di compilazione, garantendo la massima efficienza per il codice eseguibile.

2) *Sinteticità*. Niente di più dello stretto necessario è stato incluso nella definizione dei concetti e delle interfacce delle classi.

3) *Portabilità*. Il codice sorgente è stato scritto rigorosamente secondo il linguaggio standard C++, utilizzando le *Standard Template Libraries* e le *Boost Libraries*. Ciò comporta la possibilità di compilare le applicazioni su diversi sistemi operativi funzionanti su varie piattaforme hardware.

4) *Consistenza*. Lo stesso stile di programmazione è stato rigorosamente adottato per tutte le classi. Librerie esterne, come quella adottata per le Support Vector Machines, sono state inglobate in opportune classi che rispettano la medesima interfaccia definita nel resto della libreria.

5) *Manutenibilità*. Tutti gli algoritmi sono stati concepiti come un insieme di blocchi software elementari ed interoperanti, ciascuno implementato in una classe sintetica e ben documentata.

Altri fattori di qualità nello sviluppo del software sono stati intenzionalmente considerati in modo marginale durante il difficile lavoro di definizione di un preciso compromesso tra la leggibilità - usabilità e l'efficienza - portabilità di una libreria.

SPARE è stata pubblicata su *SourceForge* come progetto *open-source* (<http://sourceforge.net/p/libspare/home/>), ove è possibile scaricare l'intera libreria e tutta la documentazione allegata, inclusi alcuni progetti software dimostrativi. Per tale motivo è disponibile alla comunità scientifica internazionale, fatto questo che ha imposto elevati requisiti di qualità in termini di progettazione del software e di documentazione. Attualmente la libreria SPARE implementa alcuni algoritmi di base per problemi di clustering, classificazione ed approssimazione funzionale, oltre ad una versione efficiente di un algoritmo genetico [1] ed un insieme di classi per la definizione di Fuzzy Inference Systems.

In SPARE gli algoritmi sono concepiti e implementati come meta-algoritmi in grado di operare su spazi di dati non necessariamente metrici [2]. L'attuale versione della libreria è concepita per essere utilizzata da programmatori C++ esperti. In ogni caso l'intero progetto di sviluppo del framework prevede la definizione di un appropriato meta-linguaggio con lo scopo di ridurre i requisiti del profilo dell'utilizzatore. Questo obiettivo può essere conseguito realizzando un insieme di oggetti stabiliti ad un più alto livello semantico in grado di gestire strutture dati, files, modelli, algoritmi, reports ed elementi standard di interfaccia grafica organizzati secondo un'appropriata gerarchia.

La libreria SPARE è stata sviluppata nell'ambito delle attività inerenti il Reparto "Sistemi Intelligenti di Elaborazione", presso il Polo per la Mobilità Sostenibile della Regione Lazio.

[1] G. Del Vescovo, L. Livi, A. Rizzi, and F. M. Frattale Mascioli. SPARE: Something for pattern recognition. *Journal of Machine Learning Research*, Microtome Publishing. Submitted for publication.

[2] G. Del Vescovo, L. Livi, A. Rizzi, and F. M. Frattale Mascioli. Clustering Structured Data with the SPARE Library. To appear in *Proc. of the 4th IEEE International Conference on Computer Science and Information Technology (IEEE ICCSIT 2011)*, Chengdu, China, June 10-12th, 2011.