

Master Universitario in: “ANALISI DATI PER LA BUSINESS INTELLIGENCE”

A.A. 2016-2017

Titolo della tesi: Estrazione e visualizzazione di hotspot: un caso studio di data mining spaziale per il social network civico FirstLife

Autore: Stefania Buccoliero

Abstract:

Il presente lavoro ha come obiettivo lo studio dell'applicazione di una tecnica di data mining spaziale per i contenuti del progetto di FirstLife, social network civico basato sul posizionamento di contenuti geo riferiti su una mappa interattiva. Partendo dalla consapevolezza dell'immenso patrimonio informativo contenuto fra i dati della piattaforma di FirstLife e dell'esistenza di una branca del data mining dedicata ai dati spaziali, si è cercato e sviluppato un modo per indagare nell'eterogeneità dei punti mappati per evidenziare le aree di maggiore interesse statistico rispetto ad una domanda alla quale si vorrebbe rispondere attraverso tali contenuti.

FirstLife contiene i dati raccolti e collezionati da molti progetti sul territorio e differiscono per semantica e natura in molti aspetti restando quindi logicamente e strutturalmente divisi da parte dell'utilizzo dell'utente della piattaforma.

La visualizzazione attuale, inoltre, non offre la possibilità di fornire informazioni statisticamente significative rispetto alla densità dei dati nello spazio ma permette solo la possibilità di esporre il conteggio dei punti in un'area. Dopo aver quindi analizzato e confrontato tutte le possibili tecniche e metodi facenti parte il campo del data mining spaziale, il lavoro è confluito sullo studio dell'analisi degli hotspot.

I risultati ottenuti permettono di visualizzare le statistiche generate da questa tecnica di data mining spaziale attraverso un layer sovrapposto ad una mappa web e lascia la possibilità a chi utilizza lo strumento di personalizzare sia il livello di scala da evidenziare sia i contenuti spaziali da analizzare. Inoltre, con un ulteriore sviluppo ed analisi sulla libreria utilizzata per generare tali hotspot, si è arrivati all'estrazione delle geometrie prima solo disegnate sulla mappa in un formato tale da avere la possibilità di collezionarle, interrogarle o utilizzarle per ulteriori evoluzioni di data mining e statistica, oltre al visualizzarle.