

UNIVERSITA' DEGLI STUDI DI TORINO

DIPARTIMENTO DI CULTURE, POLITICA E SOCIETÀ

DIPARTIMENTO DI INFORMATICA

DIPARTIMENTO DI MATEMATICA "GIUSEPPE PEANO"

DIPARTIMENTO DI ECONOMIA E STATISTICA "COGNETTI DE MARTIIS"

MASTER DI I LIVELLO IN ANALISI DATI PER LA BUSINESS INTELLIGENCE E DATA SCIENCE

ANNO ACCADEMICO 2016/2017

ABSTRACT

Candidato: Roberta Zola

Il fenomeno del bullismo: la situazione del territorio torinese attraverso i dati della Polizia Municipale

Il presente lavoro vi racconterà come un processo di trasformazione qualitativa di un dato abbia portato a ottenere informazioni e considerazioni utili per la descrizione e il monitoraggio di un fenomeno quale il bullismo in città.

In particolare, partendo dalla standardizzazione e trasformazione delle basi dati che la Polizia Municipale di Torino usa per la protocollazione delle proprie attività di prossimità, l'informazione è stata classificata per tema e arricchita di indicatori sulle scuole, cosicché si è potuta svolgere un'analisi di relazione tra il fenomeno del bullismo e gli interventi formativi presso le scuole o quelli riparatori, previsti come giustizia riparatrice dopo la conferma del reato.

Esso vuole essere di spunto per un progetto più ampio di razionalizzazione della raccolta delle informazioni presso gli enti che collaborano sul progetto della prevenzione del bullismo in città o più in generale del disagio giovanile.

La tesi completa il momento di formazione che mi sono concessa frequentando questo Master e che mi ha permesso sì, di “scontrarmi”, ma anche di approfondire alcuni temi sui quali non avevo alcun tipo di conoscenza e di strutturare un lavoro di data mining, pur senza un background sui linguaggi di programmazione.

Il lavoro è stato svolto lavorando su Mac, affrontando qualche limitazione nella disponibilità di tool open source e con interfaccia grafica con codice precompilato, soprattutto per quanto riguarda la creazione di grafici dinamici (PowerBI e QlikView).

- Il processo di esplorazione dei dati e arricchimento con l'analisi del testo e l'uso di algoritmi di classificazione supervisionata, è stato svolto essenzialmente con Knime e supportato in Excel quando poteva velocizzare il lavoro.
- Le analisi di correlazione sono state fatte con Sas Studio, adattando i comandi imparati durante l'anno accademico in classe.
- Per alcuni grafici non supportati da Excel, mi sono avvalsa di R-Studio e Raw Graphs del Density Design Research Lab del Politecnico di Milano.
- Infine per la visualizzazione georeferenziata, ho usato Qgis.

Il lavoro sottostante la classificazione delle varie attività protocollate nel database della Polizia Municipale è derivata da un processo di analisi del testo sulle parole ricorrenti e non da una classificazione nativa del database, presente ma non sempre disponibile né sufficiente per l'analisi.

La numerosità del campione di eventi ascrivibili al bullismo non è alta se circostanziata su 1 anno, in quanto limitata ai casi conosciuti dalla Polizia Municipale che non è l'unico ente ad intervenire per indagare e prevenire il fenomeno. Pertanto l'analisi di correlazione è stata condotta sui dati aggregati negli 8 anni e si è osservato l'evolvere del fenomeno georeferenziandolo sulla mappa della città.