



Progetto ELDM : Electronic Logbook Data Mining

Applicazione di tecniche di Educational Data Mining ai dati dei registri elettronici: approccio "data-driven" per l'analisi delle dinamiche didattiche

Descrizione del Data-Set di interesse

Il dataset completo che si vuole costruire consta di 6 tipologie di informazioni necessarie per gli scopi del progetto.

Nel dettaglio queste informazioni sono:

Dati relativi alla scuola aderente al progetto

Della scuola è fondamentale essere a conoscenza del codice meccanografico dal quale è possibile recuperare la tipologia di istituto e la dislocazione territoriale.

Questa informazione è centrale per l'analisi comparata a livello geografico e di tipologia di istituti.

Dati relativi ai corsi di studio attivi nella scuola aderente al progetto

Necessari, per analisi comparate con altri istituti che hanno attivato lo stesso corso di studio, è l'informazione sulla tipologia di corsi presenti nella scuola che partecipa al progetto di ricerca.

E' sufficiente specificare il codice ministeriale del corso o, in alternativa, indicare la descrizione del corso stesso (Liceo Scientifico, Liceo Classico, Istituto Tecnico Informatico, etc..)

Dati relativi alle materie dei corsi di studio

Altra informazione centrale è la descrizione della disciplina che viene insegnata nella scuola. Fondamentale la descrizione della materia (Italiano, Inglese, Matematica, etc...), opzionale l'area disciplinare di appartenenza della materia stessa (Area umanistica, Area scientifica, Area dei linguaggi, etc..)

Dati relativi agli studenti

Nucleo fondante di tutto il progetto di ricerca sono i dati relativi agli studenti.

Le informazioni relative all'andamento scolastico verranno fornite in maniera anonimizzata dalle scuole. Ogni studente deve essere caratterizzato da un ID univoco (che preferibilmente non coincida con la matricola interna).

Gli altri dati relativi ad uno studente saranno:

- **anno di corso**
- **classe:** *preferibilmente non la sezione di appartenenza ma un numero che sia associato tutti gli studenti dello stesso gruppo-classe.*
Questo dato permette di confrontare l'andamento di uno studente con quello dei suoi compagni di classe.
- **anno scolastico**
- **codice meccanografico della scuola**
- **id del corso:** *indica il corso seguito dallo studente*
- **sesso:** *dato non obbligatorio*
- **voto di italiano** del 1 e 2 quadrimestre
- **voto di matematica** del 1 e 2 quadrimestre
- **numero di insufficienze** del 1 e 2 quadrimestre
- **media finale** del 1 e 2 quadrimestre
- **esito dello scrutinio** di giugno
- **eventuale esito** dello scrutinio di **settembre**

Dati relativi alle misurazioni e alle valutazioni

Fondamentali anche i dati base delle misurazioni e delle valutazioni.

Questo dato, fornito anch'esso in maniera anonimizzata in quanto fa riferimento al generico ID dello studente, ha come livello di dettaglio l'associazione materia/giorno/tipo di dato

Un esempio potrebbe essere

STUD001, INGLESE, 6, 2016-03-01, ORALE

che indica che lo studente con codice STUD0001 ha avuto una valutazione di 6/10 in una prova ORALE di INGLESE il giorno 1 Marzo 2016.

Dati relativi alle assenze

In modo del tutto analogo al punto precedente

Questo dato, fornito anch'esso in maniera anonimizzata in quanto fa riferimento al generico ID dello studente, ha come livello di dettaglio solo il giorno

Un esempio potrebbe essere

STUD001, 2016-03-01

che indica che lo studente con codice STUD0001 è stato assente il giorno 1 Marzo 2016.