Project Proposal

Teenagers with guns in school research

Alejandro González Rogel, Juraj Šušnjara and Athena Iakovidi

Problem domain

Violence in schools constitutes a serious problem in many countries, especially where weapons such as guns are involved. Although school is meant to be a secure and friendly place, many children are afraid to go there because they are often threatened and bullied by other children. Moreover, the fact that the latter children can have relatively easy access to guns can result in enormous tragedies, as seen in the past. This kind of environment is unacceptable and something has to be done in order to prevent further violence.

In this project we are focusing on some of the individual's characteristics (and the connections between them) in order to predict whether a teenager is likely to carry a gun in school.

Data

We will use data from the 2015 "Youth Risk Behavior Survey (YRBS)[1]", which analyses health-risk behaviours in teenagers in the United States. This data is public and contains the answers given by 15.624 students to 99 different questions (plus a few extra parameters about the students and their schools). There is also a "Data User's Guide" that contains additional information about the survey.

Some of the students might be discarded for our study because they didn't answer some questions or their answers were already discarded by the YRBS because they were inconsistent.

We will use the following attributes for our Bayesian network. They are all related to some question asked in the mentioned survey:

Name of the variable	Type	Related question in YRBS
Race	Discrete (1 to 8)	Raceeth (Q4 and Q5)
Age	Discrete (1 to 7)	Q1
Sex	Binary	Q2
Carry weapon in school	Binary	Q15
Feel unsecure at school	Binary	Q16
Participate in a fight last year	Discrete (1 to 4)	Q20
Bullied in school	Binary	Q24
Suffered depression last year	Binary	Q26
Alcohol use last month	Discrete (1 to 4)	Q43
Sports practiced	Discrete (1 to 4)	Q80
Grades	Discrete (1 to 4)	Q89

Implementation plan

We will implement our Bayesian Network using probability tables and any of the Python's libraries suggested (*pgmpy* or *libpgm*).

Inference problem

Our main goal would be to determine the importance of the selected attributes in predicting how likely is for a given individual to carry some kind of weapon in school.

---

[1] http://www.cdc.gov/healthyyouth/data/yrbs/index.htm