RAIS RESEARCH ASSOCIATION for INTERDISCIPLINARY STUDIES
NOV. 2018

DOI: 10.5281/zenodo.1572566

Weapon-Carrying among High School Students: A Predictive Model Using Machine Learning

Yiran Fan

The Linsly School, Wheeling, WV, USA, doris.fan.620@gmail.com

ABSTRACT: This study is aimed at 1) identifying the predictors for weapon-carrying on school properties; 2) build a predictive model for parents, educators, and pediatricians for early intervention. Youth Risk Behavior Surveillance System (YRBSS) 2017 data were used for this study. Logistic regression model is used to calculate the predicted risk. Logistic regression is a part of a category of statistical models called generalized linear models, and it allows one to predict a discrete outcome from a set of variables that may be continuous, discrete, dichotomous, or a combination of these. Typically, the dependent variable is dichotomous and the independent variables are either categorical or continuous. The data is run through R program. The outcome variable is weapon-carrying based Q13 (During the past 30 days, on how many days did you carry a weapon such as a gun, knife, or club on school property?) The result identified several important predictors for carrying weapon on school properties, such as gender, alcohol use, and smoking age. This provided important information for the educators and parents for early intervention and alleviating the negative effects of weapon-carrying among teenagers.

Introduction

The massive school shooting happened at the beginning of 2018 caused teachers and parents' panic, urging them to pay more attention to the serious issue of weapon-carrying among high school students. Students felt unsafe though the schools maximized their efforts on demonstrating emergency lockdowns. In April, thousands of students participated in a walkout against school shooting, and protests in major cities followed as the gun control policy became an even more controversial topic. In some critics' view, gun control is the optimum and the only way the public could calm themselves down. However, some students claimed that "that's not the real issue here. The real issue is the people who are doing it" (Yee 2018). Indeed, it was astonishing to see 23 school shootings in the first half of 2018; but it was more anxious to face the fact that many of the suspects are teenagers, who formerly attended the school but had animosities toward it, or even to the society. To address this issue, it is better to make early intervention than to blame the policies and wait for others to dispose of the lives of children.

In fact, people have been analyzing the characteristics of the population who are likely to carry a weapon in school among teenagers. According to Child Trends, a nonprofit organization studies the development of children of all ages, 16.8 more percent male students carry weapons than female students in 2015. Considering ethnicity, more percentage of Hispanic and black students carry weapons on school properties. However, there is no indication of the reason in detail. There were studies identify the risk factors, such as the victim of extreme violence and involvement in drug or property crimes, that might explain the reasons behind those statistics (Kingery 1999). Yet, the studies were conducted in the late twentieth century and early twenty-first century. Unforeseen developments have made changes more rapid in recent years, and the trends for weapon-carrying are varying: percent of black and Hispanic students who carry weapons has continued to decline after reaching its peak in 1997, and that of white students has remained steady (Child Trend n.d.). The current situation is not as same as 1999 anymore. In 2018, there are more hypotheses concerning the reasons of weapon-carrying as specialists look into teenagers' mental health. For example, adolescent depression happens more often these years. The number of teenagers who reported an MDE (a period of at least two weeks of low mood that is present in most situations) in the previous 12 months increased 37 percent from 2004 to 2015 (Schrobsdorff 2016). Thus, a proper model built on the most current data could help to analyze the causes behind this issue and alleviate the negative effects of weapon-carrying among teenagers.

This study is aimed at 1) identifying the predictors for weapon-carrying on school properties; 2) build a predictive model for parents, educators, and pediatricians for early intervention.

Data and Method

Data:

Youth Risk Behavior Surveillance System (YRBSS) 2017 data were used for this study.

The YRBSS was developed in 1990 to monitor priority health risk behaviors that contribute markedly to the leading causes of death, disability, and social problems among youth and adults in the United States. These behaviors, often established during childhood and early adolescence, include

Behaviors that contribute to unintentional injuries and violence.

Sexual behaviors related to unintended pregnancy and sexually transmitted infections, including HIV infection.

Alcohol and other drug use.

Tobacco use.

Unhealthy dietary behaviors.

Inadequate physical activity.

In addition, the YRBSS monitors the prevalence of obesity and asthma and other priority health-related behaviors plus sexual identity and sex of sexual contacts. From 1991 through 2017, the YRBSS has collected data from more than 3.8 million high school students in more than 1,100 separate surveys.

Method:

Logistic regression model is used to calculate the predicted risk in this study. Logistic regression is a part of a category of statistical models called generalized linear models, and it allows one to predict a discrete outcome from a set of variables that may be continuous, discrete, dichotomous, or a combination of these. Typically, the dependent variable is dichotomous and the independent variables are either categorical or continuous.

Variable:

The outcome variable is weapon-carrying based Q13 (During the past 30 days, on how many days did you carry a weapon such as a gun, knife, or club on school property?). Students who answered all questions as '0 days' were identified as no involvement of weapon, and the remaining were identified some involvement of weapon-carrying.

Results

About 2.8%% of 7008 students carried weapon on school property.

The logistic regression is run in the R program. The significant factors are listed in the table.

Table 1. Logistic Regression Output

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-9.37535	1.75770	-5.334	9.61e-08	***
Q1	0.12952	0.13947	0.929	0.35308	
factor(Q2)2	0.97665	0.23687	4.123	3.74e-05	***
factor(Q3)2	0.14689	0.30484	0.482	0.62991	
factor(Q3)3	0.51590	0.36776	1.403	0.16067	
factor(Q3)4	0.07103	0.46949	0.151	0.87975	
factor(Q3)5	-12.11497	354.97985	-0.034	0.97277	
factor(Q4)2	0.02310	0.21877	0.106	0.91592	
factor(Race)2	-0.84379	0.70699	-1.193	0.23267	
factor(Race)3	-0.51127	0.44958	-1.137	0.25545	
factor(Race)4	0.53702	0.59353	0.905	0.36557	

factor(Race)5	0.20784	0.39582	0.525	0.59951	
Q6	1.77125	1.04570	1.694	0.09029	.
factor(Q31)2	0.54454	0.37754	1.442	0.14921	
factor(Q31)3	0.88055	0.35579	2.475	0.01333	*
factor(Q31)4	0.61666	0.30101	2.049	0.04050	*
factor(Q31)5	0.57814	0.26122	2.213	0.02688	*
factor(Q31)6	0.55206	0.24466	2.256	0.02404	*
factor(Q31)7	-0.50232	0.62024	-0.810	0.41802	
factor(Q41)2	1.45688	0.33935	4.293	1.76e-05	***
factor(Q41)3	1.42178	0.34639	4.105	4.05e-05	***
factor(Q41)4	1.08056	0.33390	3.236	0.00121	**
factor(Q41)5	1.20209	0.27685	4.342	1.41e-05	***
factor(Q41)6	0.69639	0.27707	2.513	0.01196	*
factor(Q41)7	0.47804	0.46009	1.039	0.29880	
factor(Q47)2	1.24638	0.49783	2.504	0.01229	*
factor(Q47)3	-0.01022	0.49227	-0.021	0.98344	
factor(Q47)4	0.56245	0.30263	1.859	0.06309	.
factor(Q47)5	-0.41185	0.26326	-1.564	0.11772	
factor(Q47)6	-0.06984	0.23244	-0.300	0.76380	
factor(Q47)7	0.11994	0.42187	0.284	0.77617	
factor(Q60)2	0.78031	0.65551	1.190	0.23389	
factor(Q60)3	1.61850	0.68568	2.360	0.01825	*
factor(Q60)4	1.46586	0.63143	2.321	0.02026	*
factor(Q60)5	1.03143	0.64433	1.601	0.10943	
factor(Q60)6	1.07412	0.62640	1.715	0.08639	.
factor(Q60)7	1.73568	0.63316	2.741	0.00612	**
factor(Q60)8	1.71941	0.69141	2.487	0.01289	*
factor(Q61)2	-0.70611	0.39731	-1.777	0.07553	.
factor(Q61)3	-0.35833	0.37880	-0.946	0.34416	
factor(Q61)4	-0.26278	0.37626	-0.698	0.48493	
factor(Q61)5	-0.50508	0.42964	-1.176	0.23977	
factor(Q61)6	-0.53923	0.54529	-0.989	0.32272	
factor(Q61)7	NA	NA	NA	NA	
factor(Q62)2	-0.48141	0.65334	-0.737	0.46121	
factor(Q62)3	-0.54960	0.62291	-0.882	0.37761	
factor(Q62)4	-0.22185	0.64189	-0.346	0.72963	
factor(Q62)5	-0.42958	0.69667	-0.617	0.53749	
factor(Q62)6	0.25154	0.80182	0.314	0.75375	
factor(Q62)7	0.20336	0.99236	0.205	0.83763	
factor(Q62)8	NA	NA	NA	NA	
factor(Q64)2	-0.18753	0.19971	-0.939	0.34774	
factor(Q64)3	NA	NA	NA	NA	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The most significant factors are factor(Q2)2 (male sex), factor(Q41)2 (drinking alcohol 1-2 days in past 30 days), factor(Q41)3 (drinking alcohol 3-5 days in past 30 days), and factor (Q41)5 (drinking alcohol 6-9 days in past 30 days). In addition, Q60 (How old were you when you had sexual intercourse for the first time?), Q31 (How old were you when you first tried cigarette smoking, even one or two puffs?), Q47 (How old were you when you tried marijuana for the first time?), and Q61 (During your life, with how many people have you had sexual intercourse?) all have some significance concerning weapon-carrying students.

According to the logistic regression, the male is more likely to carry weapon on school properties than female. Students who tried cigarette smoking before are more likely to carry a weapon in the past 30 days in school. The more often a student drink, the more likely for him or her to carry weapon on school properties. Marijuana use at an early age is another positive predictor, so does sexual intercourse experience. Race, on the other hand, does not identify as a significant factor in this case.

Conclusions

In this study, the result identified several important predictors for carrying weapon on school properties, such as gender, alcohol use and smoking age. This provided important information for the educators and parents for early intervention and alleviating the negative effects of weapon-carrying among teenagers. Based on the most recent data, we built a predictive model using logistic regression to provide a tool for early identification of students at higher risk to carry weapon in schools, especially high schools. Moreover, future studies could use outside data and test the performance of the outputs from the model in this study. This study also provides a reference for future experiments and studies.

References

Centers for Disease Control and Prevention. (2018, June 14). *Adolescent and School Health*. Retrieved from https://www.cdc.gov/healthyyouth/data/yrbs/index.htm.

Child Trend. (n.d.). *High School Students Carrying Weapons*. Retrieved September 2, 2018, from https://www.childtrends.org/indicators/high-school-students-carrying-weapons.

Kingery, P. M., Coggeshall, M. B., & Alford, A. A. 1999. "Weapon carrying by youth: Risk factors and prevention." *Education and Urban Society* 31(3): 309-333.

Schrobsdorff, S. (2016, November 16). *Teen Depression on the Rise Says New Pediatrics Study*. Retrieved September 2, 2018, from http://time.com/4572593/increase-depression-teens-teenage-mental-health/.

Stepman, J. (2018, April 21). To reduce mass shootings, we must treat the mentally ill first | Opinion. Retrieved September 2, 2018, from https://www.newsweek.com/mass-shootings-gun-rights-gun-laws-guns-mentally-ill-opinion-892001.

Yee, V., & Blinder, A. (2018, March 14). National School Walkout: Thousands Protest Against Gun Violence Across the U.S. Retrieved September 2, 2018, from https://www.nytimes.com/2018/03/14/us/school-walkout.html.