Predict and Prevent Bullying Victimization among High School Students:
A Large-Scale Machine Learning Approach

Danrui Wang
Columbia University
Quantitative Methods in the Social Science
Master's Thesis 2024

Table of Contents

Abstract

Bullying is a serious global public health problem that needs more attention to prevent. This thesis first built 6 machine learning algorithms (Naive Bayes, K-Nearest Neighbors, Logistic Regression, Decision Tree, Random Forest, and Light Gradient Boosting Machine) to predict whether an adolescent from 12 to 18 might be bullied or not in the U.S. I used 24 features which are from 4 domains: individual, family, peers, and school. Random Forest and LightGBM attained similar high accuracy, and they are both suitable precautionary models. Then, I compared the most important features of the two models and the combined model for both female and male adolescents. The results show that even though the forms of bullying and most important features are different for both genders, factors from all 4 domains are determinant and most of the features have the same direction of influence on the prediction.

As a call to action, this research encourages stakeholders to collaborate in creating a supportive and empowering environment for adolescents. By adopting a proactive stance against bullying, society can work towards a future where adolescents can thrive without the pervasive threat of victimization. This thesis seeks to contribute to the ongoing dialogue on bullying prevention, offering insights and recommendations to guide the development of effective, sustainable interventions for the betterment of adolescent well-being.

Introduction

Bullying is a type of violence that threatens the wellness of school-aged individuals in schools and communities. Bullying victimization is a serious global public health problem and is prevalent in both developed and developing countries (Biswas et al., 2020). Globally approximately one-third of adolescents have been bullied by their peers at school (UNESCO, 2018). According to the Centers For Disease Control and Prevention (CDC, 2023), bullying is unwanted aggressive behavior done by another individual or a group who are not siblings or current dating partners; the action is intentional and can be captured by evaluating the purpose of the harmful behavior (Gladden et al., 2014). The power imbalance is the priori of bullying where the perpetrators specifically choose targets that are less physically powerful, less in numbers, have lower social status, or are one of the minority so that they can control the victims' behavior and limit their ability to resist (Nelson et al., 2019). Usually, bullying is repetitive or is highly likely to be repetitive over time. Much research has established that bullying may inflict long-lasting devastating effects on the targets including educational, social, physical, and psychological harm (Schoeler, 2018).

Bullying can be direct and indirect and can be divided into physical, verbal, relational, and property damage. Physical and verbal bullying are both direct forms that are overt expressions of power (Shetgiri et al., 2013). Physical abuse is the use of physical force against the victims such as hitting, pushing, punching, kicking, choking, spitting, tripping, and forcefully taking something from the victim (Ferguson et al., 2007; Lamb et al., 2009; Gladden et al., 2014). Verbal bullying includes oral and written harassment in the form of mean taunting, malicious teasing, name-calling, psychological intimidation, inappropriate sexual comments, and offensive written words or gestures (Merrell et al., 2008; Gladden et al., 2014). Relational bullying attacks the victim's reputation and relationships and can be both direct and indirect. Direct relational aggression involves ignoring or isolating the targets from interacting with peers, and indirect relational aggression includes spreading false or harmful rumors, convincing peers to exclude the target, publicly writing derogatory comments, and posting humiliating images in a physical or electronic environment without the target's knowledge or permission. Bullying can also take the form of vandalization which includes stealing, alteration, and damaging the target's property. Deleting one's electronic information also belongs to this category (Gladden et al., 2014).

Bullying can occur in different contexts and usually in circumstances that lack adult supervision, either in the physical world like school, school events, bus, and neighborhoods or on the internet which is known as cyberbullying (Smokowski, 2005). Bullying can be one-to-one or many-to-one. Individual bullying is typically defined as a perpetrator acting in a certain way to obtain power over the victim (Valerie, 1989). However, most of the time bullying tends to be a group behavior (mobbing). The bully may have one or more followers. Some of these followers serve as assistants to the bully who encourages and sometimes joins in the bullying behavior. Others might not directly get involved in the conduct but reinforce the bullying behavior through positive feedback like laughing or cheering during or after the incident (Salmivalli, 2014). Some bystanders witness but stay out of the scene and choose not to get involved as they either are not friends with the victim or are afraid of being bullied themselves (Thornberg, 2012).

Bullying victimization might induce physical injuries, social and emotional distress, and self-harm. Those who are persistently subjected to abusive behavior might develop long-term psychological and behavioral problems, including depression, anxiety, eating disorders, sleep difficulties, body dysmorphia, lower academic achievement, and absence from school (CDC, 2023). According to a self-reported study for 9th to 12th graders, bullying victims have more psychological issues than those who have never been bullied, and the depressive symptoms might last even many years later (Klomek, 2007; Kaltiala-Heino et al., 2009). Bullying might also increase the victim's susceptibility to illness and lead to low self-esteem (Kipling et al., 2013). More seriously, these negative effects especially depression might eventually cause the victims to commit suicide (Kim & Leventhal, 2013). Studies have shown that 23.4% of the bullied victims have seriously considered suicide (UNESCO, 2018), and 4.37% of the victims have attempted suicide (Zygo et al., 2019). Thus, bullying victimization is truly a serious issue that needs more action and attention.

Bullying culture may exist in any environment where humans interact with each other, including family, neighborhood, school, and workspace (Williams,2011). Rationalization of the abuse usually consists of differences in race, religion, gender, sexual orientation, appearance, personality, physical strength, social class, or ability (Ericson, 2001; Meyer, 2016). However, bullying victimization is more devastating for adolescents as it is the transitional period that the influence of parents diminishes and social interactions with peers become more important (De Goede et al., 2009), and it is also the time of rapid cognitive and social development (Lansford, 2023). The interactions, on the one hand, provide opportunities for students to make friends and build social social, but on the other hand, increase the possibilities of interpersonal conflicts between peers. These conflicts are inevitable parts of youth development and how students react may promote or impair their academic achievements, wellness, and moral and emotional growth (Dirks et al., 2018; Legkauskas & Magelinskaitė-Legkauskien, 2019; Valente et al. 2022).

Some may suggest that conflict management skills might help prevent bullying victimization and reduce the negative effects on psychological aspects (Zapf, 2001; Mura, 2010, Bickmore, 2010), while others disagree that the skills are only effective for everyday disputes (Olweus, 1993). Bullying is fundamentally different from peer conflicts. In everyday conflicts, both conflictual parties can make their interests count (Rahim, 2002), while as defined above, there is always a clear power advantage of one group over the other in bullying victimization. There is nothing much the victims can do to reverse the bullying situation. To stop bullying victimization, adults must intervene, especially teachers as they're the grown-ups who spend most of the time with students in the school environment and are also obligated to protect every student from harm (Burger, 2022). How bystanders react is also very crucial as they are the audience of 85% of instances of bullying, and if they take action, bullying should stop (Padgett & Notar, 2013). As for those ongoing bullying cases, it is the responsibility of parents, teachers, school administrators, healthcare providers, and policymakers to make efforts to stop and reduce the harm. Certainly, it is also important to be alert to those potential targets such that people around can intervene and provide support to prevent bullying before harm is done. Thus, I would like to use machine learning algorithms to develop an early warning system for adolescents in the United States that can predict potential instances of bullying before they occur. Also, much research has indicated that bullying behaviors differentiate between genders, so I would also like to compare what are the different top characteristics of victims between genders.

This research could have important implications for improving the safety and well-being of students in schools, and provide more understanding of the portraits of victims between different genders.

Data

Background

In this research, I'm going to use the data *National Crime Victimization Survey: School Crime Supplement* (ICPSR 37816) collected in the United States in 2019. The National Crime Victimization Survey (NCVS) is a nationally representative household survey authorized by the Bureau of Justice Statistics (BJS). The survey is conducted at the Inter-university Consortium for Political and Social Research in Ann Arbor and funded by the Bureau of Justice Statistics. The NCVS is a major statistical source of the nature and extent of crime, both reported and not reported, in the United States. In 2019, the household completion rate was 72.7%, and student completion rate was 48.8%.

Along with the NCVS, there are also different series related to specific issues with the crime. The School Crime Supplement (SCS) is one of the surveys sponsored by the National Center for Education Statistics (NCES). As the abuse at school poses an obvious threat to the safety and well-being of students and acts as a significant barrier to the education process, NCES and BJS co-designed SCS to help policymakers, practitioners, and researchers learn the relationship between school-related victimization in the school and monitor changes in students' experiences of victimization so that they could make better decisions. The School Crime

Supplement was conducted after the NCVS interview from January to June 2019, with a response rate of 35.5%.

Participants

The data contain 14273 cases (female = 48.2%, male = 51.8%, original sex) with each member aged 12 to 18 years who were in primary or secondary schools leading to a high school diploma. The mean age was 15 years old. Early dropped-out or expelled students were also valid participants as long as they were attending school at any time during the six months before the month of the interview. Home school students were not included. Each eligible respondent was randomly selected by the U.S. Census Bureau, and interviewed by computer-assisted telephone and personal interviewing. 25.1% of the respondents are Hispanic, 54.2% are non-Hispanic white-only, 11.6% are non-Hispanic black-only, and 9.1% are Non-Hispanic other races. Participants are from different regions in the United States. 16.7% are from the Northeast, 21.6% are from the Midwest, 37.5% are from the South, and 24.3% are from the West. Among these regions, 30.1% are in the Metropolitan Statistical Area (MSA) and principle city, 56.5% are in MSA not principle city, and 13.4% are outside MSA. The survey contains questions asking about the experiences and perceptions of students on crime and safety at school, including participation in afterschool activities, preventive measures implemented by the school, students' perception of these rules, the presence of alcohol, drugs, weapons, and gangs in school, bullying, and hate-related violence.

Bullying Victimization Demographic

Table 1: Students Reported being Bullied at School Characteristic Distribution

Gender	**Female** 25.5 (1.18)	**Male** 19.1 (1.04)					
Control of School	**Public** 22.7 (0.90)	**Private** 21.7 (3.86)					
School Locale	**City** 22.4 (1.80)	**Suburban** 20.5 (1.18)	**Town** 21.7 (2.37)	**Rural** 27.7 (1.89)			
Race	**White** 24.6 (1.16)	**Black** 22.2 (2.48)	**Hispanic** 18.0 (1.32)	**Asian** 13.5 (2.71)	**Two or more races** 37.1 (5.96)		
Grade	**6th** 28.1 (2.68)	**7th** 28.0 (2.36)	**8th** 26.7 (2.14)	**9th** 18.9 (1.95)	**10th** 18.7 (1.72)	**11th** 21.7 (2.08)	**12th** 15.8 (1.90)

According to SCS, around 22.2% of the students reported they were being bullied in 2019, with 19.1% male participants and 25.5% female participants. Most of the students being bullied are of two or more races (37.1%), white (24.6%) and black (22.2%) percentages are similar, Hispanic students reported slightly less (18.0%), and Asians are the least reporting being bullied (13.5%). Most of the students being bullied occur in the 6th (28.1%) and 7th (28.0%) grades and the lowest percentage being reported during middle school was in the 9th grade (18.9%). During the high school period, the highest percentage occurred in the 11th grade

Figure 1: Percentage of female and male being bullied by different types in 2019

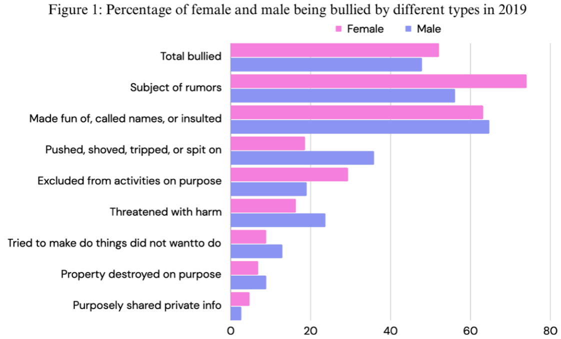


Figure 2: Percentage of female and male victims being bullied in different locations

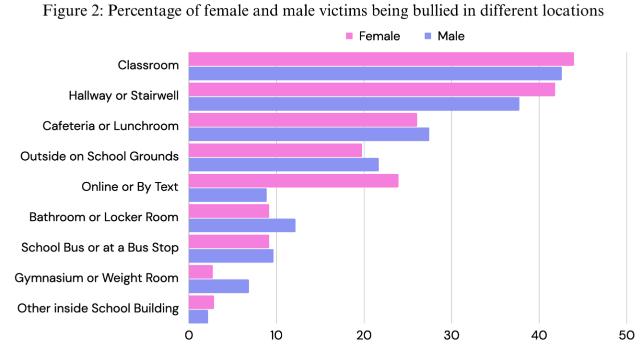
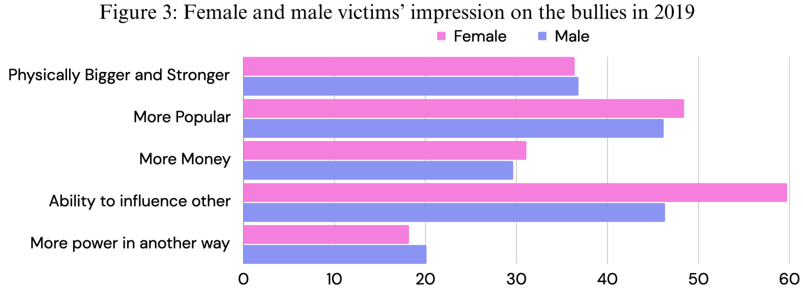


Figure 3: Female and male victims' impression on the bullies in 2019



(21.7%), and the lowest was reported in the 12th grade (15.8%). Public schools (22.7%) had slightly more students being bullied than private schools (21.7%). Schools located in rural areas had the highest percentage of students being bullied (27.7%), followed by city (22.4%), town (21.7%), and suburban (20.5%).

As shown in ***Figure 1**,* 52.1% of the students being bullied are female and 47.9% are male. Most of the victims, especially female victims, were the subject of rumors, or have been made fun of, called names, or insulted in a hurtful way. Male victims are more likely to be bullied and physically aggressive, like being pushed, shoved, tripped, or spit on, being threatened with harm, being forced to do something, and having property destroyed on purpose. Whereas female victims are more likely to be excluded from activities, social media, or other communications on purpose, and suffer from privacy breaches. Bullying locations also vary. As shown in ***Figure 2***, both female and male victims are mostly being bullied inside the classroom, or in the hallway or stairwell. Cafeteria or Lunchroom and outside on school grounds are also abuse-prone areas, especially for males. Female students are far more likely to be bullied online or by text, and male students are more likely to be bullied in the bathroom, locker room, gymnasium, or weight room. School Bus or bus stop is also a common bullying place.

Figure 4: Female and male speculation on reasons being bullied in 2019

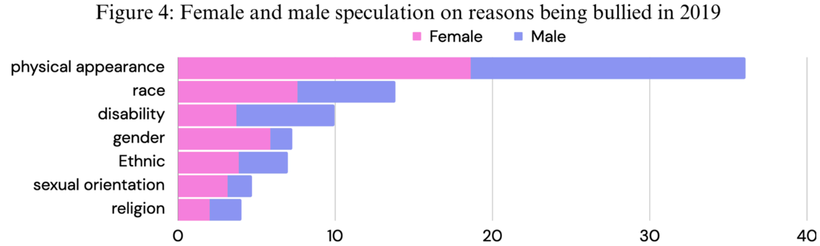
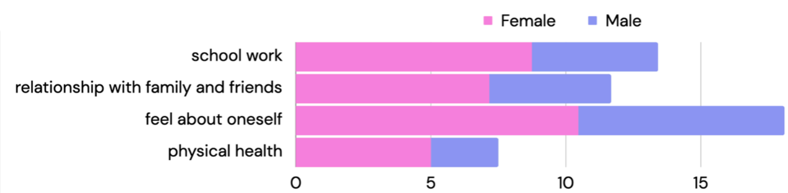


Figure 5: Negative effects on female and male victims in 2019



The survey also asked the victims about their impression of the bullies (***Figure 3***). The main impression for both male and female of the bullies are they're more popular and can influence how others think of them, especially for female victims. A higher percentage of female victims think the bullies are richer, and a higher percentage of male victims think the bullies are physically bigger or stronger. ***Figure 4*** illustrates how female and male victims speculate on the reason they were bullied. The main reason apparently for both female and male victims is their physical appearance. Race is also an important factor for both genders. A higher percentage of male victims believe it is because of their physical, mental, or developmental disability, whereas a higher percentage of female victims assume their gender and sexual orientation are key causes. Gender difference is not obvious in religion, and religion might be the least important reason the victims perceive. Bullying is hurtful and harmful, especially in how the victims feel about themselves (***Figure 5***). It also negatively affects one's academic achievement and relationships with family and friends. Some victims report their physical health is also affected.

Figure 3 and *Figure 4* provided the victim's perception of the bullies and themselves. Since in this research I'm going to list the key features for bullying victimization prediction, we can compare the result with the victim's perception and see if there are any overlaps factors or other factors.

All related codes of questions in the survey are listed in Appendix A.1.

Method

Data Preparation

Since the NCVS and SCS database contains information regarding different kinds of crime, I selected questions related to bullying victimization to serve my research, which are all listed in Appendix A.1. According to Bronfenbrenner's (1979) research, the risk factors of bullying victimization can be divided into four domains (individual, family, peer, and school). The individual's information includes age, original gender, sexual orientation, race, citizenship status, whether there are any physical disabilities (deaf, blind, cognitive, or physical limits), extracurricular activities, grades, and attendance. There is only one question regarding the family circumstance of the subject, which is the household income. As for the peer factor, I also include adults at schools. The questions focus on whether there are any adults or students at school who care about or listen to the target. The school factor includes its region, type, religion, locale, enrollment size, if there are any guards or assigned police officers, any adults supervising the hallway, locker checks, and code of student conduct, if the punishment is the same for everyone, if the school roles are strictly enforced, whether teachers treat students with respect, and if students can get alcoholic beverages or marijuana while at school. The target is whether or not the student is bullied. As we've seen in *Figure 1*, the survey asked 8 kinds of bullying-related actions. There is one more general question asking "Do you consider the actions that student did to you to be bullying?" (SCS278). I created the target variable 'bullied' whose value is 1 as long as there is at least one confirmed answer to the 8 actions and SCS278.

Since some of the questions focus on the same aspects that do not need to be examined separately (e.g. VS0029 through VS0035 asked if students attend different kinds of clubs/ activities that can be combined as 'extracurricular'), I combined them into single variables. I divided race into 5 domains (1 = Hispanic; 2 = non-Hispanic white-only; 3 = non-Hispanic black-only; 4 = non-Hispanic Asians; 5=non-Hispanic other races), and household income into 4 domains (1 = less than \$25,000; 2 = \$25,000-\$49,999; 3 = \$50,000-\$99,999; 4 = \$100,000 or greater).

There are a lot of missing values in various columns. To prevent data loss, I replace the missing value of 'sexual_orientation' and 'citizen', which are the features having most of the missing values, with the mode of each value. I drop the rows with missing in other categories, including 'transportation', 'grade', 'student_care', 'adult_care',' region', 'school_type', 'locale', 'school_size', 'security', 'supervision', 'school_code', and 'respect'. The final data set contains 5469 samples with 24 features. All 24 features are listed in Appendix A.2.

Since the dataset lost a lot of data after I dropped all the missing values, to improve the resolution of the dataset, I also used the SMOTE algorithms to oversample the data, which the algorithm increased the number of minorities in the training set. I divided the over-sampled dataset with a ratio of 7:3 for training and testing set for the full, female, and male datasets.

Statistical Analysis

In this research, I'm going to build six machine learning models (Naive Bayes, K-Nearest Neighbors, Logistic Regression, Decision Tree, Random Forest, and Light Gradient Boosting Machine) for bullying victimization prediction by using the scikit-learn library, compare their performance by Accuracy, Precision, Specificity, Recall, and F1-score. Accuracy evaluates the percentage of true prediction.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Where TP represents True Positives that correctly predict the positive cases, TP represents True Negatives that correctly predict the negative cases, FP represents False Positives that incorrectly predict the positive cases, and FN represents False Negatives that incorrectly predict the negative cases.

Precision evaluates how precise the model is that out of the cases predicted as positive, how many of them are true.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

Specificity or True Negative Rate (TNR) examines the proportion of cases that were correctly predicted as negative out of all the true negative cases.

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{3}$$

Recall/Sensitivity or True Positive Rate (TPR) is the percentage of real positives out of all the cases that the model predicts as positive.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

F1 score is the harmonic mean of precision and recall, which balances both precision and recall.

$$\text{F1} = 2 \text{ x} \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

I would also perform the Receiver Operating Characteristics (ROC) curve and Area Under curve (AUC) to fully evaluate each machine learning algorithm. ROC is a probability curve of FPR against TPR, and AUC measures the separability. The AUC ROC score examines how well the model can distinguish between classes. The higher the score is, the better the model can separate students who might be bullied and students who might not, where 0.5 might be a sign of random guessing and 1 indicates perfect performance.

Then, I would like to examine the most important features of each gender on the best performance algorithm. Feature importance is calculated as the Gini importance for Random Forest, which is the total decrease in node impurity averaged over all trees of the ensemble. In LightGBM, we use the method of split feature importance. In this research, I would choose the top n features which combined could achieve 80% accuracy. Then, I would use Shapley Additive exPlanations (SHAP) to evaluate which features are the most important to predict bullying victimization for each gender. SHAP measures feature importance at the row level. It uses the idea of game theory which is calculated as the average of how a specific feature influences the performance of a model in a single row relative to other features. SHAP value shows how each feature affects the prediction of the model and the magnitude of the value indicates the significance of the feature compared to each other.

Results

Machine Learning Models Performance

From Table 2, we can see that the performance for the training set and testing set do not differ a lot for Naive Bayes and Logistic Regression for the full dataset and both genders. Decision Tree achieves a pretty high accuracy score for the training set for all three datasets, whereas the accuracy drops a lot in the testing set. The difference might be because the Decision Tree overfits the training set. The oversampling issue is largely improved by Random Forest, where the accuracy score is the same as Decision Tree but also maintains high accuracy in the testing set for all three datasets. Random Forest achieves the highest accuracy score and

LightGBM also performs very well for all three datasets. Generally, female achieves higher accuracy for both training and testing sets for all Machine Learning models than male

Table 2: Accuracy Score of Machine Learning algorithms for both gender

Model	Full Dataset		Female		Male	
	Training	Testing	Training	Testing	Training	Testing
Naive Bayes	0.6501	0.6339	0.6592	0.6737	0.6473	0.6573
KNN	0.8538	0.7571	0.8514	0.7647	0.8351	0.7386
Logistic Regression	0.6598	0.6461	0.6595	0.6710	0.6537	0.6652
Decision Tree	0.9958	0.7780	0.9977	0.7568	0.9951	0.7587
Random Forest	0.9958	0.8483	0.9977	0.8338	0.9951	0.8330
LightGBM	0.9110	0.8400	0.9384	0.8259	0.9299	0.8208

Table 3.1 and Table 3.2 list the Machine Learning Model performance for female and male datasets. To keep the completeness of the research, the performance table for the full dataset is in A.3. All performance scores are higher for females and males. Naive Bayes achieves a relatively high precision and accuracy score and lower recall for women and achieves a high accuracy score and a relatively lower score for precision, recall, and F1 for males. This suggests that Naive Bayes could generally give accurate positive predictions on whether a female student might be bullied, whereas the model also misjudges many positive cases as negative. Naive Bayes does not perform very well on male students (F1 = 0.5664), which achieves better than random guessing but does not capture specifically who might be bullied and who might not.

Logistic Regression made the lowest AUC score for both genders and the lowest accuracy, precision, recall, and f1 scores for females. Even though logistic Regression performs better on accuracy and precision for males than Naive Bayes, it is not very suitable for bullying prediction overall. All performance scores are higher in KNN for males and females, especially on recall, which suggests the model could capture more bullied cases accurately for both genders. The AUC scores for KNN for both genders are also relatively high (female = 0.7658, male = 0.7355), indicating that KNN might be a pretty good model for bullying victimization

Table 3.1: Machine Learning Model Performance for Female Dataset

Model	AUC	Accuracy	Precision	Recall	F1
Naive Bayes	0.6688	0.6737	0.6889	0.5709	0.6244
Logistic Regression	0.6664	0.6710	0.6835	0.5727	0.6232
KNN	0.7658	0.7647	0.7354	0.7882	0.7609
Decision Tree	0.7549	0.7550	0.7378	0.7514	0.7445
Random Forest	0.8354	0.8390	0.8827	0.7624	0.8182
LightGBM	0.8218	0.8259	0.8739	0.7403	0.8016

Table 3.2: Machine Learning Model Performance for Male Dataset

Model	AUC	Accuracy	Precision	Recall	F1
Naive Bayes	0.6406	0.6573	0.5805	0.5529	0.5664
Logistic Regression	0.6296	0.6652	0.6212	0.4428	0.5170
KNN	0.7355	0.7386	0.6633	0.7192	0.6901
Decision Tree	0.7461	0.7491	0.6760	0.7300	0.7020
Random Forest	0.8045	0.8270	0.8571	0.6868	0.7626
LightGBM	0.8007	0.8208	0.8342	0.6955	0.7585

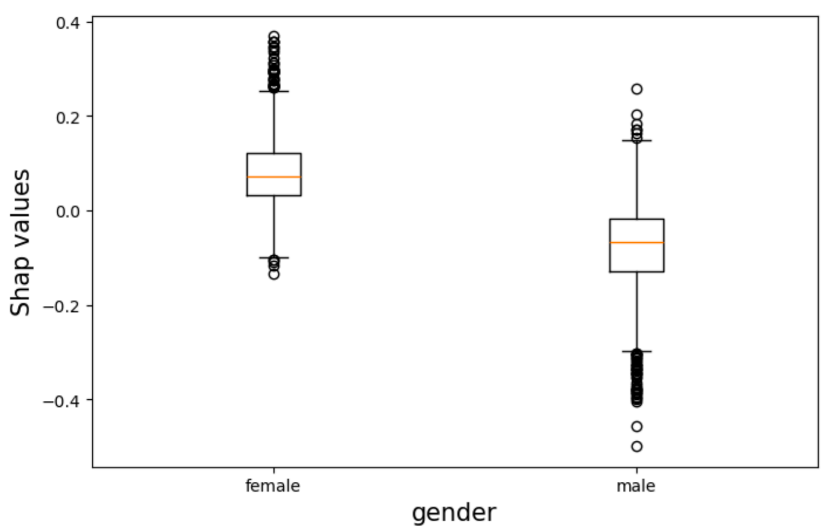
prediction. Decision Tree performs better on males but slightly worse than females than KNN, but overall achieves better performance in the overall data for all scores, suggesting that Decision Tree is also a very good choice for bullying victimization prediction. Random Forest of course further improves from the Decision Tree, which achieves the AUC highest score for both genders. LightGBM has the second-highest AUC score for both genders. All other scores of

Random Forest and Decision Tree are also the highest for females, whereas the recall for males is not one of the tops. This may suggest that even though Random Forest and LightGBM might now capture enough bullied cases for male victims, it generally is the best model for our prediction.

Feature Importance to Predicting Bullying Victimization

Firstly, I ran the feature importance of gender on the full dataset (Figure 6). Each dot in the graph represents a prediction. Positive SHAP values represent the feature that positively influences the prediction, and vice versa. The boundary is usually set at 0, which in our case a positive SHAP value indicates the category is more likely to be predicted as a bullied victim and a negative SHAP value represents less likely to be bullied. The result shows that generally female students are more likely to be bullied than male students.

Figure 6: SHAP value of gender for the Full Dataset



All 24 features are combined to run the algorithms to make predictions before. To check how many of them make the most effort to make predictions, I calculated the cumulative

Figure 8: Feature Importance for Female Dataset

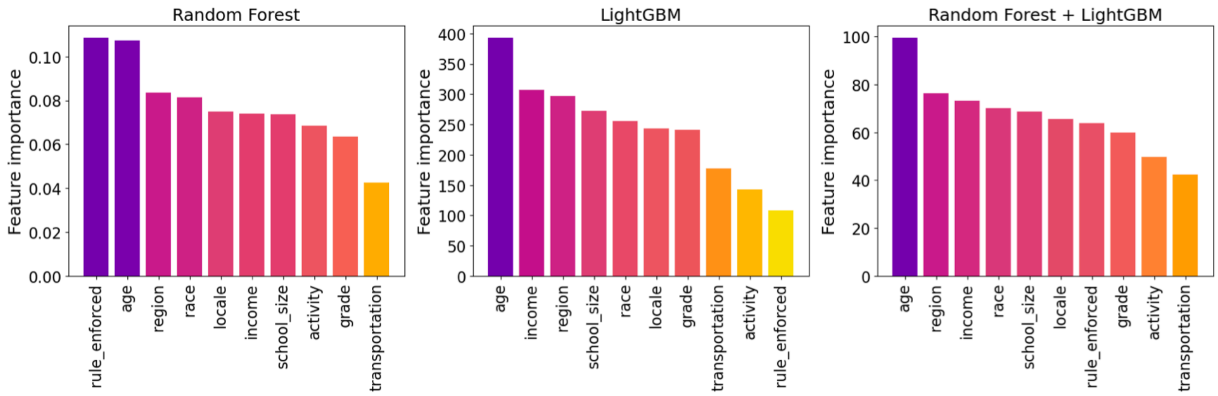
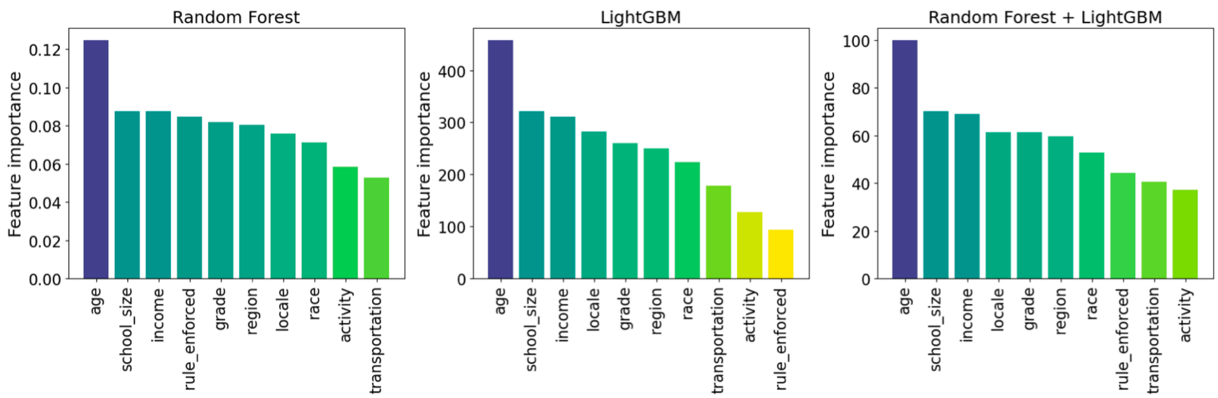


Figure 9: Feature Importance for Male Dataset



probability of the top-N feature importance for both female and male victims. From Figure 6, we can see that 10 features could make around 80% accuracy for both female and male victimization prediction, of which 80% is usually a good sign of high performance.

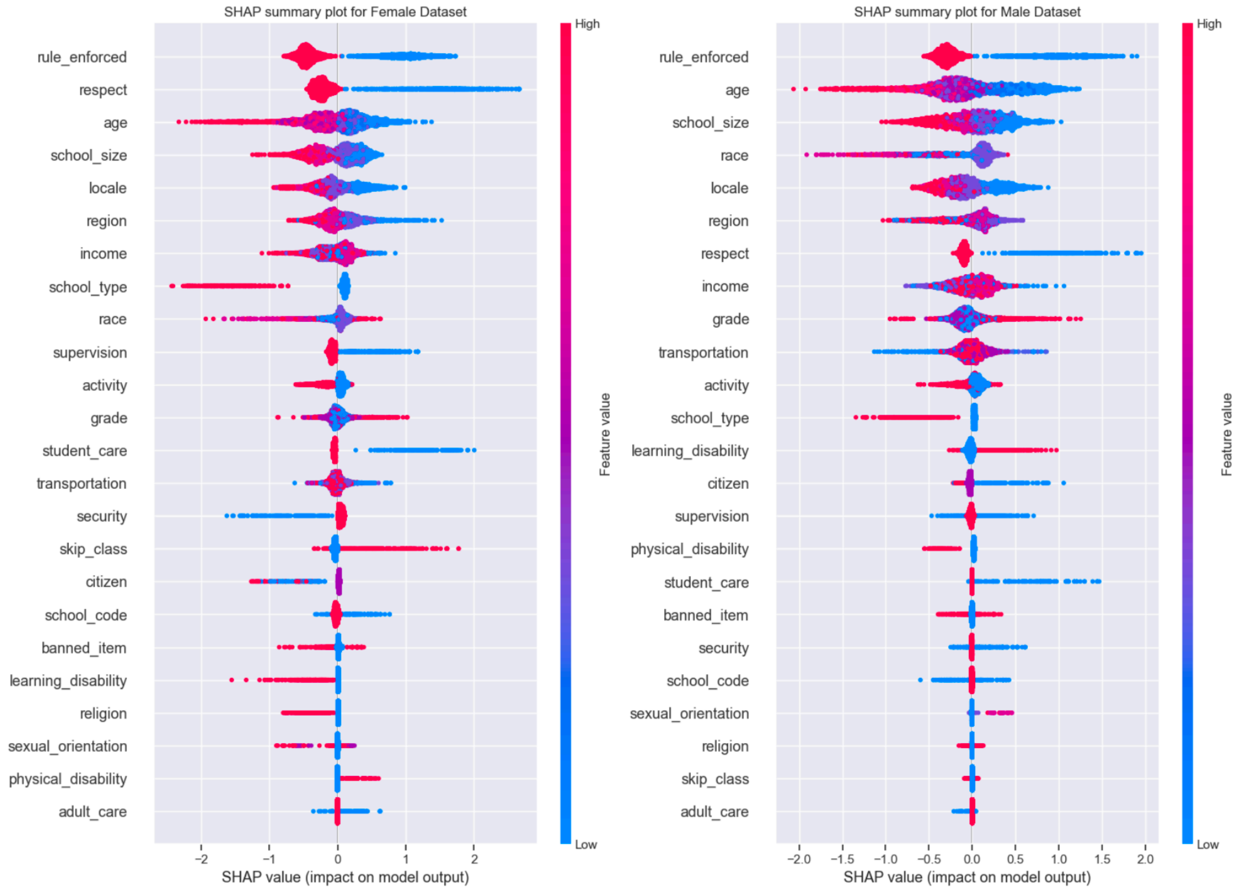
Then I display the top 10 feature importance of the best performance models from the last section — Random Forest, LightGBM, and the combined model for females (Figure 7) and males (Figure 8). The graphs showing all 24 features are shown in A.4. The combined model normalized each of the features' importance of Random Forest and LightGBM.

In Random Forest, whether the school enforces the rules properly (rule_enforced) and age are the most important factors for female students. School regions and students' race are also important factors for females. The locale of the school, the student's family income, and school size make similar contributions to the prediction for females. Whether the student attends any extracurricular activities (activity) and the student's grade are also determinants of whether a female student might be bullied or not. Transportation used for going to school is less important than the other factors. As for male students, age is the most important feature. School size, family income, and rule_enforced also contribute to bullying prediction. Student's grades and school regions make a similar contribution, locale, and race make fewer effects, and activity and transportation are much less important than the other features for male students. Random Forest picks the same top 10 features for both genders. Both genders consider age as the most important feature. The female dataset treated rule_enforced, region, race, and activity as more important features than the male dataset, whereas males consider schoool_size, income, and grade features more in the prediction.

In LightGBM, age is also the most important feature for both genders. Females take income and region seriously. School size is also an important feature. Race, local, and grade also have similar effects on female bullying prediction. Transportation makes fewer contributions, and comparatively activity and rule_enforced have the least effects among the top 10 features for female bullying prediction. As for males, school size and income are also very important, then is the locale. Grade, region, and race make similar contributions. Transportation, activity, and rule_enforced have similar few effects as they are on females on the male dataset. Generally, females consider region and race more important than males, and males take school size and locale more seriously than females.

Both Random Forest and LightGBM pick the same top 10 features; only the order is different. To evaluate the average of the two models, we look at the combined model. Indubidible

Figure 9: SHAP Summary for Female Victimization Prediction from LightGBM



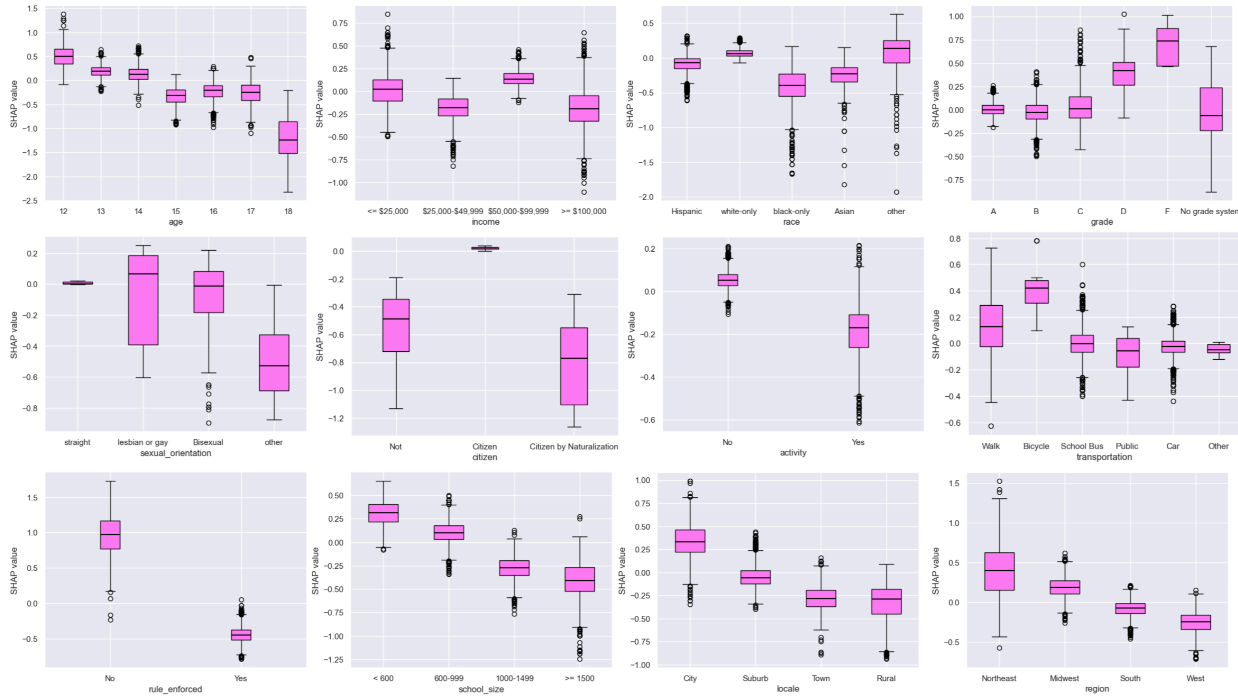
that age is the most important feature for both genders. Region, income, race, and school size are also very crucial for the prediction. Locale, rule_enforced make fewer effects, and activity and transportation are much less important. As for males, school size and income are also significant features for bullying prediction. Locale, grade, and region contribute similarly. Race is less important, and rules_enforced, transportation, and grade are the least important features for male bullying prediction.

Females take age and region as the most important features and transportation as the least important feature for all three models, and males take age, school size, and income as the top 3 important features and activity and transportation as the least important feature for bullying prediction. To see how each feature affects the prediction specifically, *Figure 9* shows in what direction each feature affects the prediction. Since there are some features containing more than two values that are hard to distinguish in the summary plot in Figure 9, I further plot how each value affects the prediction specifically for the top 10 features and multivariable features in *Figure 10* (female) and *Figure 11* (male).

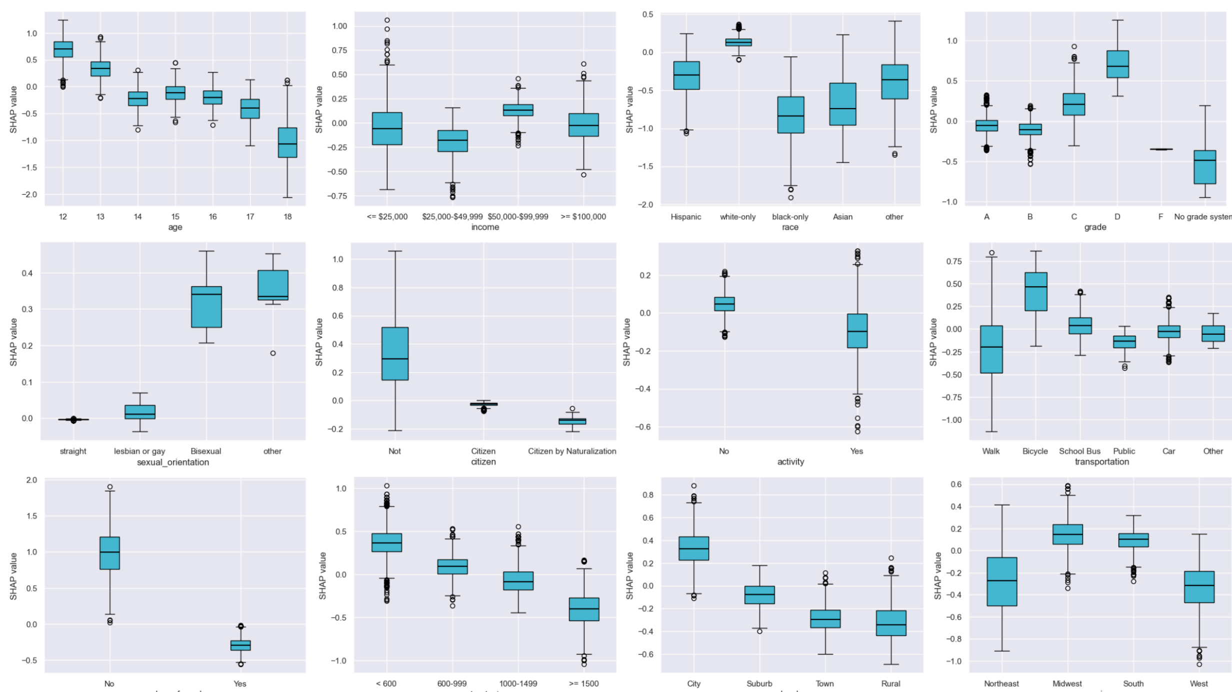
For both genders, the higher the student's age is, the more likely the student might be bullied. The SHAP value is positive when the age is under 15 for females and 14 for males, and reaches the lowest point at the age of 18 for both genders. Female students are more likely to be bullied if their family income is below $25,000 and $50,000 — $99,999, and male victims mainly focus on the range of $50,000 — $99,999. For female students, the SHAP value is higher when the race is white-only and in the 'other' category, and is lowest for black-only. Only white-only has a positive SHAP value. The value is lowest if the school doesn't have black-only is also

Figure 10: SHAP value of each value for Female Dataset
Figure 11: SHAP value of each value for Male Dataset

the lowest value. For both genders, the lower the grade, the more likely to be bullied. The value is lowest if the school does not have grades or does not give alphabetic equivalent grades. Lesbians are more likely to become victims, whereas all sexual orientation other than straight has positive SHAP value for males. Whether a student is a U.S. citizen does not explain a lot for female students, whereas noncitizen male students are more likely to be bullied. The SHAP value is positive for walking and bicycle for females, and bicycles and school buses for males. For both genders, the smaller the size of the school (population < 999), the more likely the student might suffer from bullying, especially under 600 students. Also, for both genders, the more metropolitan the locale is, the larger the SHAP value; the value is only positive in the city. For



females, the Northeast has the highest positive SHAP value, and next is the Midwest. For males,



the Midwest and South are prone areas of bullying.

The results for the binary answers are quite distinct. For both genders, the SHAP value is much higher when the school rules are not strictly enforced, teachers do not treat students with respect, no teachers or adults supervise the hallway or locker checks, no other students really care about you, and if the student skips a lot of classes. The SHAP value is only very low and negative for private schools. The SHAP value is zero when there are school codes, banned items are not able to be brought to school, and adults at school really care about you. Females are more likely to be bullied if they have a physical disability and less likely to be bullied if they have a learning disability, whereas males are the opposite.

Conclusion

In this research, we first run 6 machine-learning models on 24 features to predict whether an adolescent from age 12-18 is at risk of being bullied in the U.S. For full, female, and male datasets, Random Forest performs the best and LightGBM is also a comparable good choice. Then, I checked on the gender difference on the full dataset, and the result shows female students in general are more likely to be bullied than males. To further explore how bullying prediction differs between genders, I first calculate the cumulative probability on the Top-N features, and 10 features are enough to account for 80% prediction accuracy for both genders. Then, I compared the importance of the top 10 features of Random Forest, LightGBM, and the combined model of the two algorithms for both genders. The top 10 features from top to low for **females** in the combined model are **age, region, income, race, school size, locale, rule_enforced, grade, activity, and transportation**. For **males**, the sequence is **age, school_size, income, locale, grade, region, race, rule_enforced, transportation, activity**. Then, I draw the SHAP summary plot and SHAP value boxplot for each feature to examine further how each value of the feature influences the prediction. There are four domains of features for the prediction: individual, family, peer, and school. For students themselves, bullying most commonly happens during middle school (ages 12 to 14) for both genders. White-only, Hispanic, and mixed-race females are more likely to be bullied, and male victims are mainly white-only. Citizen is not a crucial indicator for females, whereas noncitizen male students are much more likely to be bullied than others. Sexual orientation seems to play an important part in bullying victimization. Nearly no students with straight orientation are bullied, whereas lesbians are most likely to become victims for females, and males with all other sexual orientations are potential targets. Females with physical disability and males with learning disability are more likely to be bullied. Going to school by bicycle seems to be related to bullying victimization for both genders. Students with better grades, skipping fewer classes, and attending extracurricular activities are much less likely to be bullied. As for the family domain, most of the victims for both genders are from upper-middle income families ($50,000 — $99,999). As for the peer domain, almost all students who have no other students who really care about them are at risk of bullying for both genders. Having a caring adult at school does not enhance or mitigate the chance of bullying, but having no such adults increases the chance of being bullied. In the school domain, female adolescents should pay more attention if they're living in the Northeast and Mideast, and male victims are mainly located in the Mideast and South of the U.S. The city is a prone area of bullying than town and rural areas. Bullying is less likely to happen in private schools in comparison to public schools. Bullying is most likely to take place in smaller schools (< 1000), especially having a population under 600. This might be because when the students are denser, they pay more attention to each other which might increase the possibility of bullying. Whether a student might be bullied is separated by whether the school strictly enforces the rules and if the teachers treat students with respect. Having security (guards, assigned police officers, or security cameras) at school, having school staff or other adults supervising the hallway and locker, and being unable to bring banned items including alcohol and marijuana have zero SHAP values, and might cause very high positive value if the answers are opposite.

Thus, the urgent need to address and mitigate bullying victimization among adolescents demands a comprehensive and multifaceted approach. Random Forest and LightGBM are great precautionary models to alert those who might be bullied. Adolescents themselves could attend more extracurricular activities, get better grades, skip fewer classes, and find someone (either other students or adults) who can truly listen to and care about their situations to reduce the chance of being bullied. As for bystanders, they could provide great help if they are willing to take more care of the minorities. Schools should always strictly enforce their rules, and teachers should always treat their students with respect. Certainly, the mechanisms behind bullying are complicated, and I believe misunderstanding might be one of the main reasons. Adults and schools should educate adolescents to stop hating each other based on gender, sexual orientation, race, citizenship, or disability of any form. Through the amalgamation of education, safety precautions, kind manners from peers, and personal development, we can create an environment that fosters empathy, understanding, and respect among adolescents.

Appendix

A.1) Questions selected from National Crime Victimization Survey: School Crime Supplement (ICPSR 37816) related to bullying victimization research:

WHETHER AND HOW IS BULLIED:
VS0073 - HAS ANOTHER STUDENT: MADE FUN OF YOU, CALLED YOU NAMES, OR INSULTED YOU, IN A HURFUL WAY? (BEING BULLIED)
SCS235 - USING TECHNOLOGIES: MADE FUN OF YOU, CALLED YOU NAMES, OR INSULTED YOU, IN A HURTFUL WAY?
VS0074 - SPREAD RUMORS
SCS236 - Q22V2B. DURING THIS SCHOOL YEAR, HAS ANOTHER STUDENT IN PERSON OR USING TECHNOLOGIES: SPREAD RUMORS ABOUT YOU OR TRIED TO MAKE OTHERS DISLIKE YOU?
VS0075 / SCS238 - THREATENED YOU WITH HARM
VS0076 / SCS239 - PUSHED YOU, SHOVED YOU, TRIPPED YOU, OR SPIT ON YOU
VS0077 / SCS240- TRIED TO MAKE YOU DO THINGS YOU DID NOT WANT TO DO
VS0078 / **SCS241**- EXCLUDED YOU FROM ACTIVITIES, SOCIAL MEDIA, OR OTHER COMMUNICATIONS ON PURPOSE
VS0079 / **SCS242**- DESTROYED YOUR PROPERTY ON PURPOSE
SCS237 - Q22V2C. DURING THIS SCHOOL YEAR, HAS ANOTHER STUDENT IN PERSON OR USING TECHNOLOGIES: PURPOSELY SHARED YOUR PRIVATE INFORMATION, PHOTOS, OR VIDEOS IN A HURTFUL WAY?
SCS278 - DO YOU CONSIDER [THAT THING/THOSE THINGS] THAT [ANOTHER STUDENT/OTHER STUDENTS] DID TO YOU TO BE BULLYING?
VS0090 / **SCS248**- DID YOU THINK THE BULLYING WOULD HAPPEN AGAIN
VS0091 / **SCS243**- DID MORE THAN ONE PERSON DO THIS TO YOU?
VS0092 / **SCS244**- DID THESE PEOPLE ACT ALONE, TOGETHER AS A TEAM, OR BOTH?

BULLY IMPRESSIONS:
VS0093 / SCS249 - WAS THIS PERSON OR ANY OF THESE PEOPLE PHYSICALLY BIGGER OR STRONGER THAN YOU?
VS0094 / SCS250 - WAS THIS PERSON OR ANY OF THESE PEOPLE MORE POPULAR THAN YOU?
VS0095 / SCS 251 - DID THIS PERSON OR ANY OF THESE PEOPLE HAVE MORE MONEY THAN YOU?
VS0096 / SCS252 - DID THIS PERSON OR ANY OF THESE PEOPLE HAVE THE ABILITY TO INFLUENCE WHAT OTHER STUDENTS THINK OF YOU?
VS0097 / SCS253 - DID THIS PERSON OR ANY OF THESE PEOPLE HAVE MORE POWER THAN YOU IN ANOTHER WAY?

REASONS:

SCS200 / SCS271 - RACE
SCS201 / SCS272 - RELIGION
SCS202 / SCS273 - ETHNIC BACKGROUND OR NATIONAL ORIGIN
SCS203 / SCS274 - DISABILITY (PHYSICAL, MENTAL, DEVELOPMENTAL)
SCS204 / SCS275 - GENDER
SCS205 / SCS276 - SEXUAL ORIENTATION
SCS206 / SCS277 - PHYSICAL APPEARANCE

WHERE:
VS0081 / SCS257 - CLASSROOM
VS0082 / SCS258 - HALLWAY OR STAIRWELL
VS0083 / SCS259 - BATHROOM OR LOCKER ROOM
SCS260 - GYMNASIUM OR WEIGHT ROOM
VS0087 / SCS261 - CAFETERIA OR LUNCHROOM
VS0084 / SCS265 - SOMEWHERE ELSE INSIDE THE SCHOOL BUILDING
VS0085 / SCS262 - OUTSIDE ON SCHOOL GROUNDS
VS0086 / SCS263 - SCHOOL BUS OR AT A BUS STOP
SCS211 / SCS264 - ONLINE OR BY TEXT

NEGATIVE EFFECTS:
SCS267 - SCHOOL WORK
SCS268 - YOUR RELATIONSHIPS WITH FRIENDS OR FAMILY
SCS269 - HOW YOU FEEL ABOUT YOURSELF
SCS270 - YOUR PHYSICAL HEALTH

Risk factors from 4 domains:
Individual
V3014 - AGE
V3017 - SEX (ORIGINAL)
V3023A- RACE RECODE
V3024- HISPANIC ORIGIN
V3_V4526H3A- DEAF
V3_V4526H3B - BLIND
V3_V4526H5 - DIFFICULT LEARN, REMEMBER, CONCENTRATE
V3_V4526H4 - LIMITS PHYSICAL ACTIVITIES
V3083 - CITIZENSHIP STATUS
V3084 - SEXUAL ORIENTATION
V3085 - GENDER IDENTITY AT BIRTH
VS0024 - HOW DO YOU GET TO SCHOOL MOST OF THE TIME THIS SCHOOL YEAR?
VS0029 - ATHLETIC TEAMS
VS0030 - SPIRIT GROUPS
VS0031 - PERFORMING ARTS
VS0032 - ACADEMIC CLUBS

VS0033 - STUDENT GOVERNMENT
VS0034 - COMMUNITY SERVICE/VOLUNTEER CLUBS
VS0035 - OTHER SCHOOL CLUBS/ACTIVITIES
VS0138 - DURING THIS SCHOOL YEAR, ACROSS ALL SUBJECTS HAVE YOU GOTTEN MOSTLY - (GRADE)
VS0136 - DURING THE LAST 4 WEEK, DID YOU SKIP ANY CLASSES?

Family
SC214A - HOUSEHOLD INCOME

People around at school
VS0146 - THERE IS AN ADULT AT SCHOOL WHO REALLY CARES ABOUT YOU
VS0148 - THERE IS AN ADULT AT SCHOOL WHO LISTENS TO YOU
SCS186 - THERE IS A STUDENT AT SCHOOL WHO REALLY CARES ABOUT YOU
SCS187 - THERE IS A STUDENT AT SCHOOL WHO LISTENS TO YOU

School
SCS214 - SCHOOL REGION
VS0019- SCHOOL TYPE: PUBLIC OR PRIVATE
SCS215: SCHOOL TYPE: PUBLIC, PRIVATE-NO RELIGIOUS AFFILIATION DATA REPORTED, PRIVATE-ROMAN CATHOLIC, PRIVATE-OTHER RELIGIOUS, PRIVATE-NONSECTARIAN
VS0021 - RELIGION
SCS216 - SCHOOL LOCALE CODE: CITY, SUBURB, TOWN, RURAL
SCS218 - SCHOOL ENROLLMENT SIZE
VS0036 - GUARDS OR ASSIGNED POLICE OFFICES
VS0037 - OTHER SCHOOL STAFF OR OTHER ADULTS SUPERVISING THE HALLWAY
VS0041 - LOCKER CHECKS
VS0043 - SECURITY CAMERAS
VS0044 - CODE OF STUDENT CONDUCTVS0045 - DO YOU HAVE A WAY TO REPORT IT TO SOMEONE IN AUTHORITY WITHOUT GIVING YOUR NAME?
VS0050 - PUNISHMENT IS THE SAME TO NO MATTER WHO YOU ARE
VS0051 - THE SCHOOL RULES ARE STRICTLY ENFORCED
VS0053 - TEACHERS TREAT STUDENTS WITH RESPECT
SCS230 - GETTING ALCOHOLIC BEVERAGES WHILE AT SCHOOL
SCS231 - GETTING MARIJUANA WHILE AT SCHOOL

A.2) Features used for Bullying Victimization Prediction

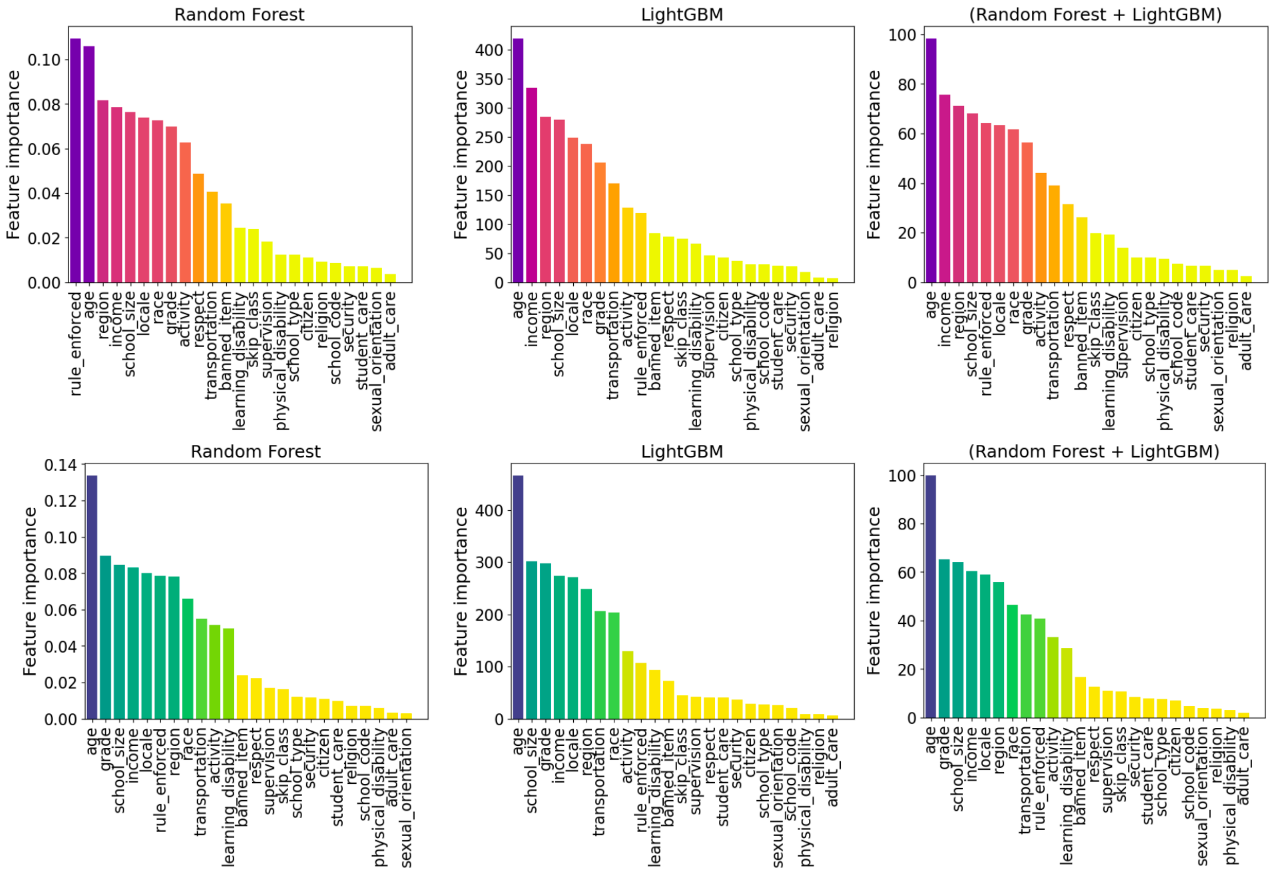
Domain	Feature	Description	Value
Individual	Age	Age	12, 13, 14, 15, 16, 17, 18
	Gender	Birth sex	0 = Female, 1 = Male
	Sexual_orientation	Sexual orientation	1 = Straight; 2 = Lesbian or gay; 3 = Bisexual; 4 = Something else
	Race	Race	1 = Hispanic; 2 = non-Hispanic white-only; 3 = non-Hispanic black-only; 4 = non-Hispanic Asian; 5 = non-Hispanic other races
	Citizen	U.S. Citizenship status	0 = not a citizen; 1 = citizen; 2 = citizen by naturalization
	Physical_disability	Deaf, blind, or limits physical activites	0 = no; 1 = yes
	learning_disability	Difficult learn, remember, concentrate	0 = no; 1 = yes
	Activity	Athletic teams, spirit groups, performing arts, academic clubs, student government, community service/volunteer clubs, or other school clubs/activities	0 = no; 1= yes
	Transportation	How to ge to school most of the time this chool year?	1 = Walk; 2=Bicycle; 3=School Bus; 4=Public bus, subway, train; 5=Car; 6=Some other way
	grade	What grade have you gotten mostly across all subejcts?	1 = A; 2=B; 3=C; 4=D; 5=F; 6=' School does not give grades/no alphabetic grade equivalent '
	skip_class	Skip class in the last semester	0 = no; 1= yes
Family	Income	Household income	1 = less than $25,000; 2 = $25,000-$49,999; 3 = $50,000-$99,999; 4 = $100,000 or greater
Peer	student_care	There is a student at School who really cares about you/listens to you.	0 = disagree; 1= agree
	adult_care	There is an adult at School who really cares about you/listens to you.	0 = disagree; 1= agree
School	Region	Where the school is located	1 = Northeast; 2=Midwest; 3=South; 4 = West

Domain	Feature	Description	Value
	school_type	what type of school is it?	1 = Public; 2=Private
	Religion	Is it religious school	0=no; 1=yes
	locale	School locale code	1=City; 2=Suburb; 3=Town; 4=Rural
	school_size	School enrollment size	1= less than 600; 2=600 to 999; 3=1000 to 1499; 4=over 1500
	Security	Whether there is guards or assignd police offices/ security cameras	0=no; 1=yes
	Supervision	school staff or other adults supervising the hallway/ locker check	0=no; 1=yes
	school_code	there is a code of student conduct	0=no; 1=yes
	Rule_enforced	The school rules are strictly enforced and punishment is the same to no matter who you are.	0=no; 1=yes
	Respect	Teachers treat students with respect	0=no; 1=yes
	banned_item	Students are able to get alcoholic beverages or marijuana while at school.	0=no; 1=yes

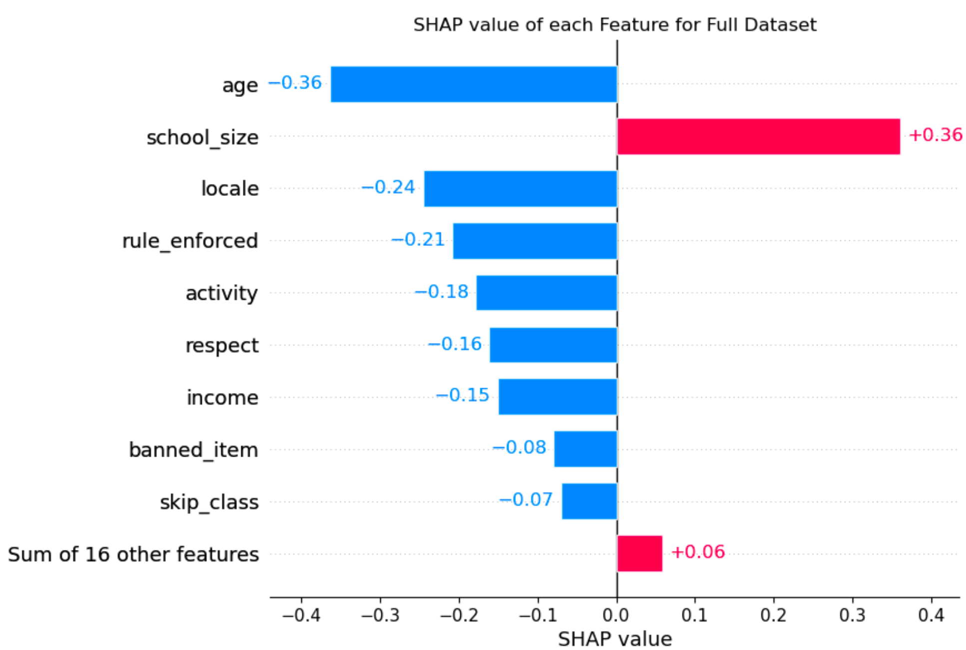
A.3) Machine Learning Model Performance for the full dataset

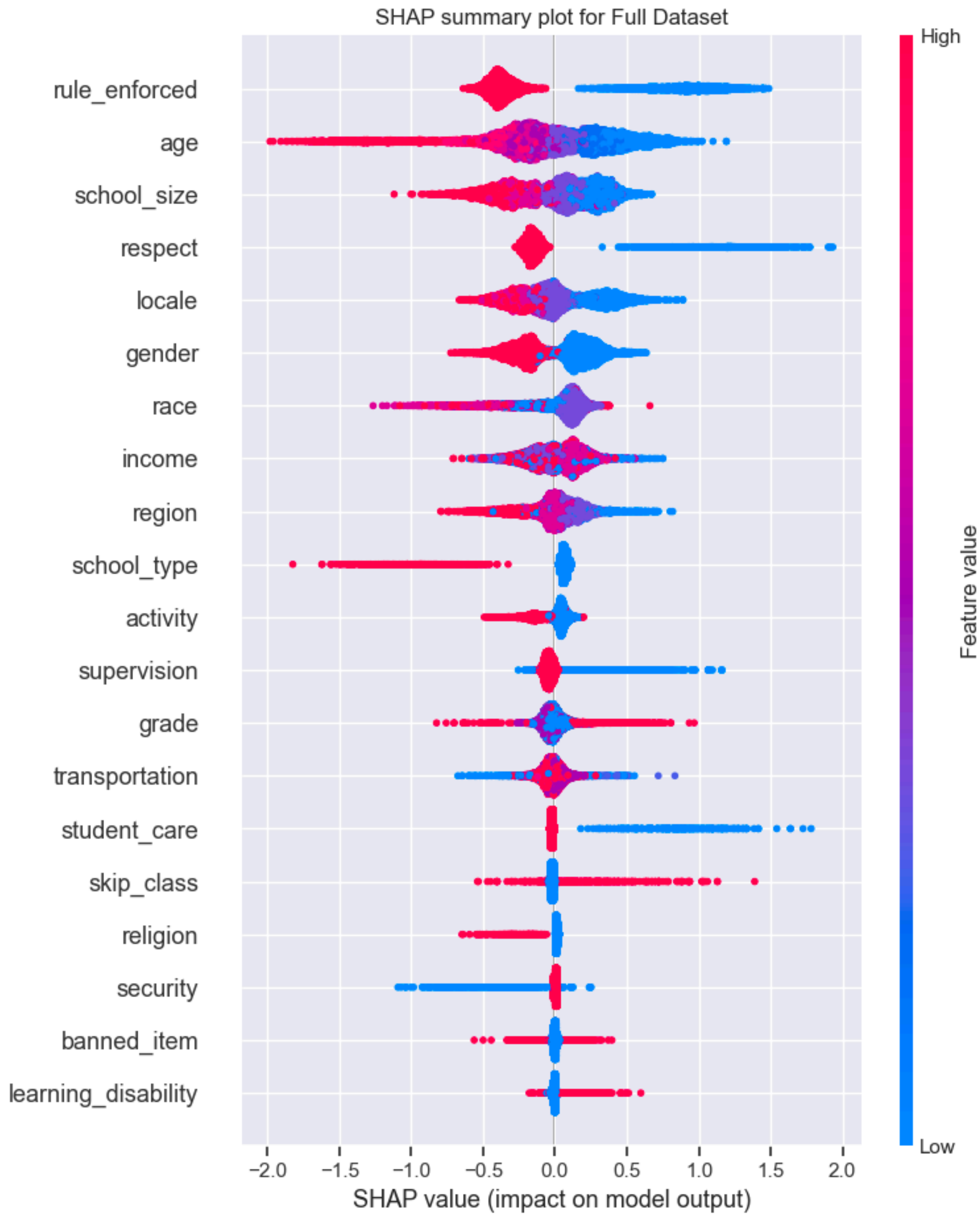
Model	AUC	Accuracy	Precision	Recall	F1
Naive Bayes	0.6338	0.6339	0.6441	0.5946	0.6183
Logistic Regression	0.6460	0.6461	0.6608	0.5970	0.6273
KNN	0.7572	0.7571	0.7298	0.8147	0.7699
Decision Tree	0.7839	0.7780	0.7843	0.7656	0.7748
Random Forest	0.8462	0.8483	0.9088	0.7736	0.8358
LightGBM	0.8398	0.8400	0.9165	0.7472	0.8234

A.4) Feature Importance of all 24 Features for top performance models and combined model



A.5) Feature Importance Analysis for the Full Dataset




SHAP summary plot for Full Dataset
rule_enforced
age
school_size
respect
locale
gender
race
income
region
school_type
activity
supervision
grade
transportation
student_care
skip_class
religion
security
banned_item
learning_disability
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
SHAP value (impact on model output)
High
Low
Feature value

References

Biswas, T., Scott, J. G., Munir, K., Thomas, H. J., Huda, M. M., Hasan, Md. M., David de Vries, T., Baxter, J., & Mamun, A. A. (2020). Global variation in the prevalence of bullying victimization amongst adolescents: Role of peer and parental supports. *EClinicalMedicine, 20*, 100276. https://doi.org/10.1016/j.eclinm.2020.100276

Burger C. School Bullying Is Not a Conflict: The Interplay between Conflict Management Styles, Bullying Victimization and Psychological School Adjustment. Int J Environ Res Public Health. 2022 Sep 19;19(18):11809. doi: 10.3390/ijerph191811809. PMID: 36142079; PMCID: PMC9517642.

Centers for Disease Control and Prevention. Youth risk behavior surveillance—United States, 2019. Morbidity and Mortality Weekly Report–Surveillance Summaries 2020; 69(SS1). Available from https://www.cdc.gov/healthyyouth/data/yrbs/pdf/2019/su6901-H.pdf

De Goede IH, Branje SJ, Delsing MJ, Meeus WH. Linkages over time between adolescents' relationships with parents and friends. J Youth Adolesc. 2009 Nov;38(10):1304-15. doi: 10.1007/s10964-009-9403-2. Epub 2009 Mar 11. PMID: 19779807; PMCID: PMC2758153.

Dirks M.A., Dunfield K.A., Recchia H.E. Prosocial Behavior with Peers: Intentions, Outcomes, and Interpersonal Adjustment. In: Bukowski W.M., Laursen B., Rubin K.H., editors. *Handbook of Peer Interactions, Relationships, and Groups.* The Guilford Press; New York, NY, USA: 2018. pp. 243–264. [Google Scholar]

Ericson, Nels (June 2001). "Addressing the Problem of Juvenile Bullying" (PDF). *OJJDP Fact Sheet #FS-200127*. **27**. Archived (PDF) from the original on 2015-06-26. Retrieved 2013-10-28.

Ferguson CJ, San Miguel C, Kilburn JC, et al. The effectiveness of school-based anti-bullying programs: a meta-analytic review. *Crim Justice Rev.* 2007;32(4):401–414. [Google Scholar] [Ref list]

Gladden, R.M., Vivolo-Kantor, A.M., Hamburger, M.E., & Lumpkin, C.D. Bullying Surveillance Among Youths: Uniform Definitions for Public Health and Recommended Data Elements, Version 1.0. Atlanta, GA; National Center for Injury Prevention and Control, Centers for Disease Control and Prevention and U.S. Department of Education; 2014.

Kaltiala-Heino, Riittakerttu; Fröjd, Sari; Marttunen, Mauri (2009-07-09). "Involvement in bullying and depression in a 2-year follow-up in middle adolescence". *European Child & Adolescent Psychiatry*. **19** (1): 45–55. doi:10.1007/s00787-009-0039-2. ISSN 1018-8827. PMID 19588185. S2CID 23077909.

Kim YS; Leventhal B (2008). "Bullying and suicide. A review". *International Journal of Adolescent Medicine and Health*. **20** (2): 133–54. doi:10.1515/IJAMH.2008.20.2.133. PMID 18714552. S2CID 41651943.

Kipling D. Williams; Joseph P. Forgas; William Von Hippel (13 May 2013). *The Social Outcast: Ostracism, Social Exclusion, Rejection, and Bullying*. Psychology Press. pp. 2–. ISBN 978-1-135-42338-4. Archived from the original on 3 January 2014. Retrieved 2013-10-29.

KLOMEK, ANAT; MARROCCO, FRANK; KLEINMAN, MARJORIE; SCHONFELD, IRVIN S.; GOULD, MADELYN S. (January 2007). "Bullying, Depression, and Suicidality in Adolescents". *Journal of the American Academy of Child & Adolescent Psychiatry*. **46** (1): 40–49. doi:10.1097/01.chi.0000242237.84925.18. ISSN 0890-8567. PMID 17195728.

Lamb J, Pepler DJ, Craig W. Approach to bullying and victimization. *Can Fam Physician.* 2009;55:356–60. [PMC free article] [PubMed] [Google Scholar] [Ref list]

Lansford, J. Adolescent Development. 2023. Duke University.

Legkauskas V., Magelinskaitė-Legkauskienė Š. Social Competence in the 1st Grade Predicts School Adjustment Two Years Later. *Early Child Dev. Care.* 2019;**191**:83–92. doi: 10.1080/03004430.2019.1603149. [CrossRef] [Google Scholar]

Merrell KW, Gueldner BA, Ross SW. How effective are school bullying intervention programs? A meta-analysis of intervention research. *School Psychol Q.* 2008;23(1):26–42. [Google Scholar] [Ref list]

Meyer, Doug (2016). "The Gentle Neoliberalism of Modern Anti-bullying Texts: Surveillance, Intervention, and Bystanders in Contemporary Bullying Discourse". *Sexuality Research and Social Policy*. **13** (4): 356–370. doi:10.1007/s13178-016-0238-9. S2CID 148471672.

Nelson H.J., Kendall G.E., Burns S.K., Schonert-Reichl K.A., Kane R.T. Measuring 8 to 12 Year Old Children's Self-Report of Power Imbalance in Relation to Bullying: Development of the Scale of Perceived Power Imbalance. *BMC Public Health.* 2019;**19**:1046. doi: 10.1186/s12889-019-7375-z. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

Olweus D. *Bullying at School: What We Know and What We Can Do.* Blackwell Publishing; Hoboken, NJ, USA: 1993. [Google Scholar]

Padgett, S., Notar, C., Bystanders are the Key to Stopping Bullying. Jacksonville State University. 2013.

Rahim M.A. Toward a Theory of Managing Organizational Conflict. *Int. J. Confl. Manag.* 2002;**13**:206–235. doi: 10.1108/eb022874. [CrossRef] [Google Scholar]

Salmivalli, C., (2014) "Participant Roles in Bullying: How Can Peer Bystanders Be Utilized in Interventions?," *Theory Into Practice,* 53:4, 286-292, DOI: 10.1080/00405841.2014.947222. To link to this article: https://doi.org/10.1080/00405841.2014.947222

Schoeler T., Duncan L., Cecil C.M., Ploubidis G.B., Pingault J.-B. Quasi-Experimental Evidence on Short- and Long-Term Consequences of Bullying Victimization: A Meta-Analysis. *Psychol. Bull.* 2018;**144**:1229–1246. doi: 10.1037/bul0000171. [PubMed] [CrossRef] [Google Scholar]

Shetgiri R. Bullying and victimization among children. Adv Pediatr. 2013;60(1):33-51. doi: 10.1016/j.yapd.2013.04.004. Epub 2013 Jul 12. PMID: 24007839; PMCID: PMC3766526.

Smokowski PR, Kopasz KH. Bullying in school: an overview of types, effects, family characteristics, and intervention strategies. *Child Sch.* 2005;37(2):101–110. [Google Scholar] [Ref list]

Thornberg, R., Tenenbaum, L., Varjas, K., Meyers, J., Jungert, T., & Vanegas, G. (2012). "Bystander Motivation in Bullying Incvalueents: To Intervene or Not to Intervene?" *Western Journal of Emergency Medicine,* 13(3), 247–252. http://doi.org/10.5811/westjem.2012.3.11792 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3415829/

United Nations Educational, Scientific and Cultural Organization. (2018). School violence and bullying: Global status and trends, drivers and consequences. Education 2030. Surveillance Summaries. 65(6):1-174.

Valente S., Afonso Lourenço A., Németh Z. *Interpersonal Relationships.* IntechOpen; London, UK: 2022. School Conflicts: Causes and Management Strategies in Classroom Relationships. [Google Scholar]

Valerie E. Besag (1989). *Bullies and victims in schools: a guide to understanding and management*. Open University Press. ISBN 978-0-335-09542-1. Archived from the original on 2014-07-07. Retrieved 2013-10-29.

Williams, Ray (3 May 2011). "The Silent Epidemic: Workplace Bullying". *Psychology Today*. Archived from the original on 2015-11-21. Retrieved 2016-11-13.

Zapf D., Gross C. Conflict Escalation and Coping with Workplace Bullying: A Replication and Extension. *Eur. J. Work Organ. Psychol.* 2001;**10**:497–522. doi: 10.1080/13594320143000834. [CrossRef] [Google Scholar]

Zygo M, Pawłowska B, Potembska E, Dreher P, Kapka-Skrzypczak L. Prevalence and selected risk factors of suicidal ideation, suicidal tendencies and suicide attempts in young people aged 13-19 years. Ann Agric Environ Med. 2019 Jun 17;26(2):329-336. doi: 10.26444/aaem/93817. Epub 2018 Aug 24. PMID: 31232067.