Impact of Food on Adoloscent Behavior

Divya Baskaran

Department of Computer Science

New York University

New York, NY 10012

db3619@nyu.edu

Abstract

Bullying is seen to be a persistent problem among the youth worldwide, especially in US. The project studies the general behavior of children and their food intake in 2009-2010 across US. It analyzes to find an association between the nutrient intake and bullying behavior based on the HBSC survey response and develop an application to monitor the exposure level of the student to anti-social acts.

Keywords—analytics, bully, nutrient intake, youth behavior, logistic regression, Spark, MLlib, HDFS, Scala

# **Introduction**

Humans have evolved and so has our body structure. Food plays a major role in building the constituents of this structure. Certain food that taste good need not necessarily be good for health. Such food causes the individual to behave unusually thus leading them to demonstrate anti-social acts.

Bullying and juvenile cases have been reported across various US states. This poses a serious concern on the cause for such anti-social behavior.

This project looks into the food habits of young people in different US states and their behavior at school towards other children. The objective of this project is to gain an insight on the impact of food in one’s behavior. There have been numerous reports on various food items and its effects on health. A clear study of those ingredients and its effects is required to analyze the behavior of an individual.

The project involves analysis of Health Behavior in School Aged Children (HBSC) survey and search for an association between the food intake and the behavior of the child. Such analysis is then used to train a logistic regression model to digest the strong association between dietary pattern and bullying behavior. The project then displays a form for the user of the application to fill their dietary pattern. The model then depicts the exposure level of the user to the bullying behavior. The model also helps to extract bullying statistics among students at large and depict the information via pie charts.

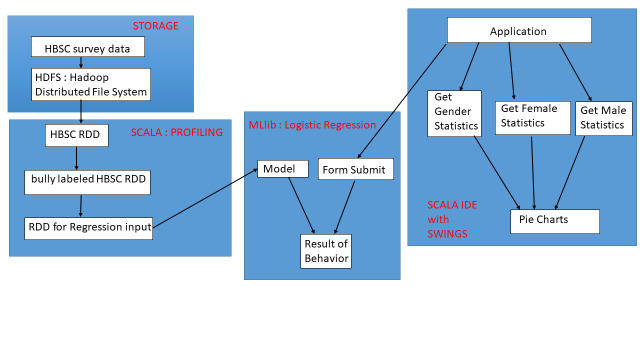
# **Motivation**

Bullying has proven to be dangerous to the health of the victims as well as to the bullies. It is a crime that is reported across various states in US often causing the victims to go into depression and other socio-psychological problems. To stop bullying, the cause for such crime needs to be analyzed in order to reduce the effect of it. Thus there is a need to find an association between dietary pattern and behavior of students. This paper researches on an association between the nutrition level and bullying behavior in adolescents thus creating awareness among the users of the application by displaying their behavior based on their dietary pattern and displaying statistics depicting strong correlation between the two.

# **Related Work**

Bullying has often been a topic of research. As discussed in prior research[1] poor nutrition often causes even non deviant youth to exhibit anti-social acts. Logistic regression techniques were used in [1] to determine a relationship between nutrient intake and bullying behavior across US. Regression techniques were further used to determine the impact of food on bullying. Scala for large datasets is often recommended for analysis as stated in [2]. MLlib is also recommended in [2] to run machine learning algorithms on large datasets. The Youth Risk Behavior Surveillance website reports statistics for each state on various components of the youth lifestyle. Such statistics serve a great purpose for analytics on each state. The website also inspires to create a map exhibiting the states where the association derived in [1] is demonstrated.

# **Design**



HBSC dataset contains survey response from children across US. The dataset includes questions on bullying and food intake by each child along with many other questions on their lifestyle. This dataset is profiled and cleaned to extract responses regarding bullying and nutrition. Responses with respect to bullying is extracted and given a value of 1 or 0 based on the response from each student. This value is defined as the label for the logistic regression model. Once every record is assigned a 0 or a 1, responses regarding nutrition are considered as features and analyzed using logistic regression techniques to find an association between Nutrition and bullying. Once an association is established, it is tested by taking input from the form response submitted via the Scala IDE through the application. The response is associated with the dietary pattern of the user. The trained model is then used to classify the user as a bully or not. The response regarding the behavior is displayed to user. The application also displays various options that correspond to the statistics detected by the regression model on the trained dataset. These results are depicted via pie charts.

# **Experiments**

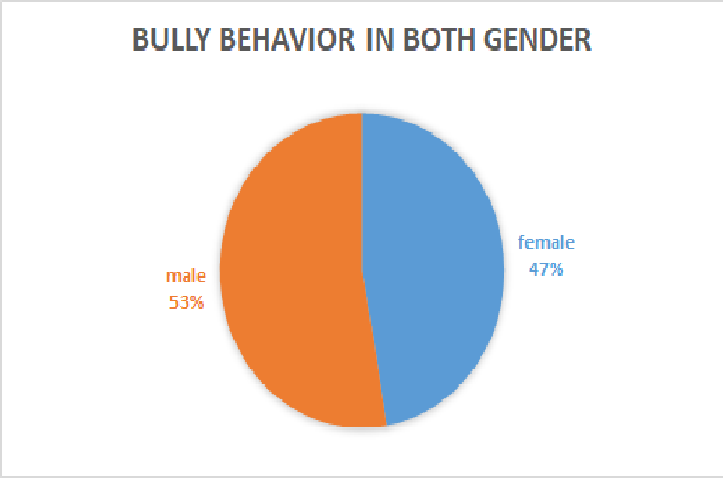
The application is built using Scala IDE with swings. The Health Behavior in school Aged Children ( HBSC ) survey data is stored in the Hadoop Distributed File System ( HDFS ). The dataset is profiled and cleaned to extract responses regarding bullying behavior. There were 12 questions regarding bullying that included responses ( 1 – Never bullied , 2 – once/twice , 3 – 2/3 times a month, 4 – about once a week , 5 – Several times a week ). Two approaches were designed to deal with these 12 responses that vary between 1–5. The first approach gave better results than the second approach. The first approach was to check if all the 12 questions had the response 1 or greater than 1. If the 12 responses were 1 or -9 ( missing ), the record was labeled 0 otherwise 1. The second approach involved Principal Component Analysis. PCA was used to reduce the 12 responses to a single value. Standard deviation was calculated and values among the 12 data points that lay in the first level of confidence were labeled 0 otherwise 1. Once each recorded was labeled 0 or 1 where 0 indicated that the student was not a bully and 1 indicated that he/she was a bully, features were extracted from the records.

The food habits were considered features for the logistic regression model. 5 questions involved the type of food such as fruits, vegetables, sweets, soft drinks and fast food while 4 questions involved the place and time of food such as breakfast during weekdays, weekends, midday meals and going to bed/school hungry. The responses were mapped between values ranging from 0 – 3. These constituted the features.

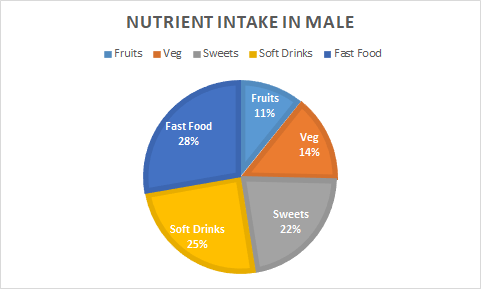
The MLlib library was used to construct the logistic regression model. The model was trained over the HBSC dataset that was labeled as 0 or 1.

Analysis were drawn from the model-trained data. The application displays these analysis in the form of pie charts. The following experiments and corresponding results were drawn:

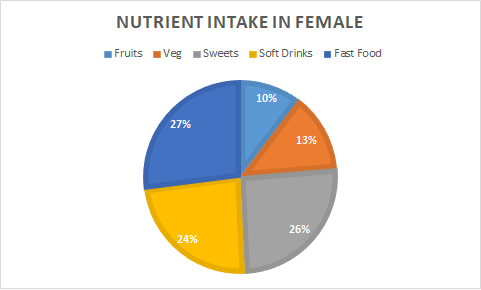
1. 53% of the male students were bullies while the remaining 47% were female bullies.



1. Among male bullies, fast food constituted 28%, soft drinks 25% and sweets 22% of the nutrient intake while vegetables and fruits only constituted 14% and 11% respectively.



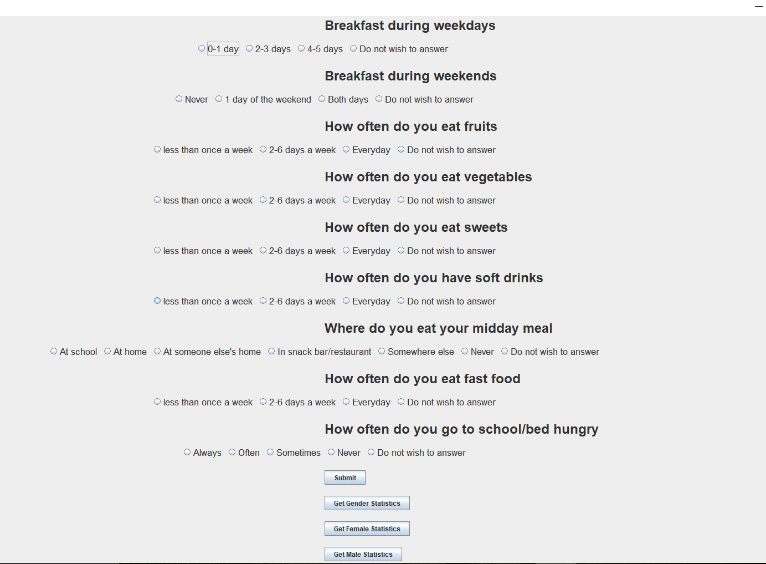
1. Among female bullies, fast food constituted 27%, soft drinks 24% and sweets 26% of the nutrient intake while vegetables and fruits only constituted 13% and 10% respectively.



The above statistics concluded that students with high intake of soft drinks, fast food and sweets were bullies.

The application provides a form for the user to fill data regarding their food habits. Questions were regarding the 9 features that constituted the features for the model. These questions were given options. When the user hits the submit button, the regression model tests the input and predicts/ classifies the user as a bully or not. Questions in the form were answered in various ways to view the classification results i.e to view if the user is a bully or not.

The application Home Page:



1. Results Based on the food type :
2. Association of non-healthy food with bullying behavior

Users with high intake of fast food, sweets and soft drinks were classified as bullies.

b. Association of low F&V with bullying

Users with low intake of fruits and vegetables and high intake of fast food, sweets and soft drinks were classified as bullies.

c. Association of high F&V with bullying

Users with high intake of fruits and vegetables and low intake of fast food, sweets and soft drinks were classified as good.

1. Results Based on the time and place of food :

a.

Users having no breakfast during the week, having no midday meal and always going to bed hungry were classified as bullies.

b.

Users having breakfast everyday, always having their midday meal at school and never going to bed hungry were classified as good.

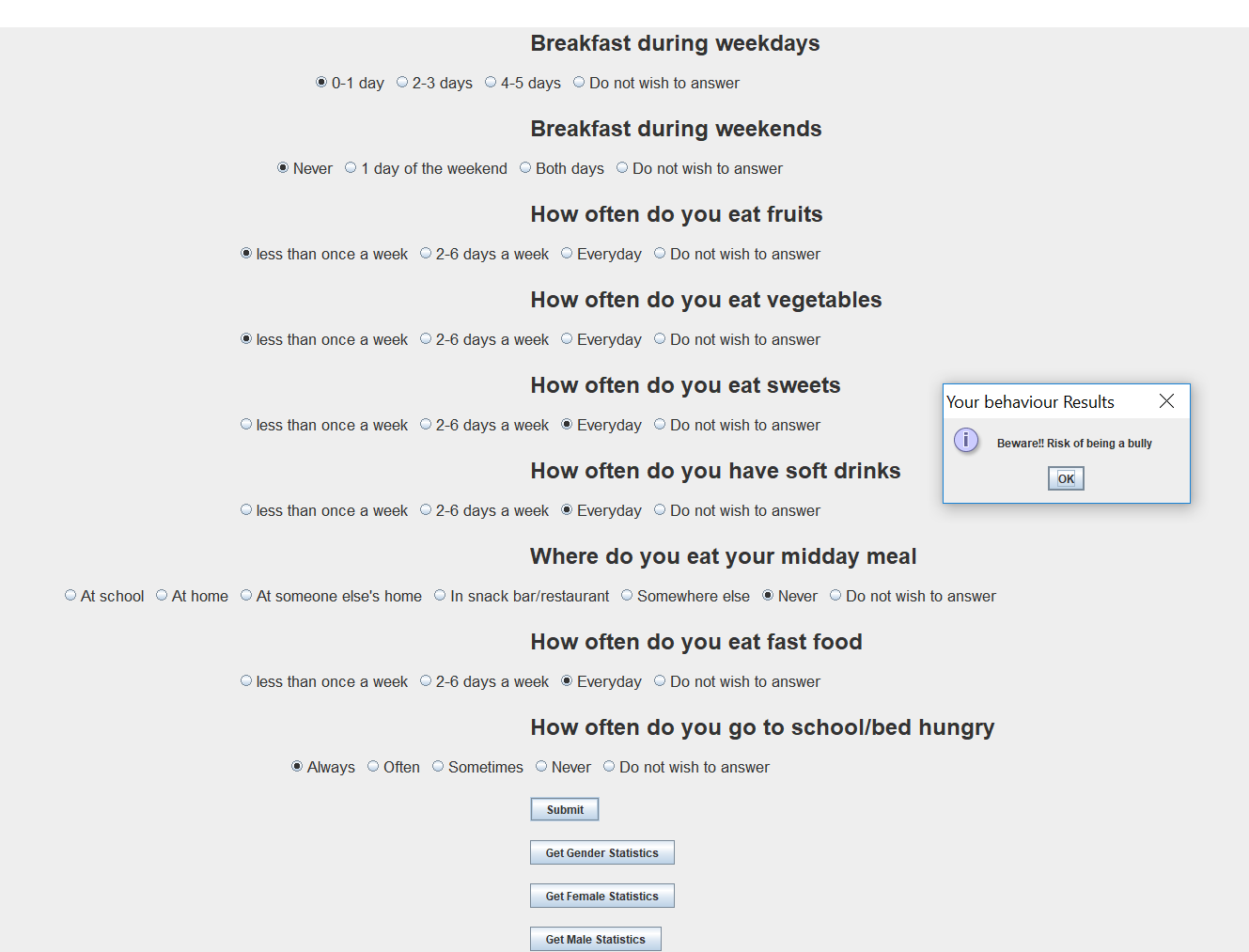
1. Combined results :

a.

Users having no breakfast during the week, having no midday meal and always going to bed hungry with low intake of fruits and vegetables and high intake of fast food, sweets and soft drinks were classified as bullies.

b.

Users having breakfast and midday meal everyday, never going to bed hungry with high intake of fruits and vegetables and low intake of fast food, sweets and soft drinks were classified as good.



# **Conclusion**

The project proved to be a detailed study of the association between food habits and the behavior in students. The application performed statistical analysis and concluded that there is a strong relationship between dietary pattern and bullying behavior. While establishing this strong bond, the application further proves the accuracy of this stated association by taking in user input and proving the correctness of the analysis by predicting the behavior of the user with the given input. The general statistics derived from the model- trained data further strengthens this analysis.

The application is a source for the user to be warned on their dietary pattern and control their nutrient intake to impact their actions. Bullying is not always associated with bad attitude. It is not an uncontrollable behavior. The food we eat often characterizes such behavior.So children should have the right amount of nutrients to maintain their body constituents in the right balance thus maintaining a positive attitude and behavior.

##### **Acknowledgment**

1. NYU HPC for providing access to dumbo ( Hadoop cluster )
2. WHO to provide the HBSC survey dataset
3. YRBS website for data regarding bullying and nutrient intake across states.

##### **References**

[1] Dylan B. Jackson, Michael G. Vaughn, Christopher P. Salas-Wright. Poor nutrition and bullying behaviors: A comparison of

deviant and non-deviant youth: Journal of Adolescence, March 2017

[2] T. White. Hadoop: The Definitive Guide. O’Reilly Media Inc., Sebastopol, CA, May 2012

[3] Ian Janssen, PhD; Wendy M. Craig, PhD; William F. Boyce, PhD and William Pickett, PhD : Associations between Overweight and Obesity with Bullying Behaviors in School-Aged Children