Distractions and lifestyle habits

Predicting Academic Success in College Students.

Samirah Jeanty

School of Computing and Data Science

Wentworth Institute of Technology

Boston, MA 02115

jeantyn@wit.edu

Abstract

This project looks into the impact of students' lifestyle habits and its correlation to student success; those habits range from sleeping patterns, fast food and energy drink consumption, physical activity and breakfast frequency. The dataset that I will use is a self-reported survey from Florida Gulf Coast University where about 614 responses were collected. In this study, linear regression is applied to predict GPA from student's lifestyle habit. It will help provide insight into the correlation between healthy lifestyle patterns and academic performance in undergraduate students.

Keywords

Lifestyle habits; academic success; GPA; Regression analysis; undergraduate students

1.Introduction

College students should balance their lifestyle choices with their academic success. This project will look into those lifestyle choices and see if they can help predict academic success by using self-reported GPA. Being able to understand the relationship between those values can help with improving student's success, even in the smallest changes such as reducing energy drink or increasing sleeping habits. This study can help contribute to a broader, data-driven understanding on how routines influence academic outcomes.

2.Data

2.1. Source of dataset

- Title and source of data set:

Student Health Behavior and Academic Performance.
https://dataverse.fgcu.edu/dataset.xhtml?persistentId=doi:10.60863/SF/HWI1MX

- About the Dataset:

The dataset has survey responses from about 614 undergraduate students at Florida Gulf Coast University. It contains self-reported GPA, sleeping patterns, breakfast frequency, physical activity, energy drink or alcohol consumption and many other factors.

- Limitations and credibility:

Since this is a self-reported GPA dataset, there might be inaccuracies or biased in the data. These caveats will be presented as a limitation in the report.

2.2. Characters of the datasets

The dataset has about 614 rows.

Cleaning the data and new variables: removing missing values, looking for outliers and converting the appropriate numerical form. For visualization, might create a category section for lifestyle habits, for example low, medium or high.

Variable	Description	Unit / Type
GPA	Self-reported academic performance	Continuous
Sleep Hours	Average daily hours of sleep	Continuous
Breakfast Frequency	Days per week students eat breakfast	Integer
Fast-Food Frequency	Weekly fast-food consumption	Integer
Energy Drink Use	Number of energy drinks per week	Integer
Physical Activity	Times per week engaging in exercise	Integer

3.Methodology

This project will have two predictive modeling methods that will help us understand which lifestyle factors are more connected with GPA and/or GPA prediction.

3.1 Random Forest Regression

This will help compare with a nonlinear model that deals with variable and interactions relationships. This captures nonlinear relationships and provides another approach to understanding and validating regression results.

Metrics: R^2, RMSE, cross-validation accuracy.

3.2 Multiple Linear Regression

This will help predict GPA from independent variables like breakfast, fast food and energy drink consumption, and other factors. GPA is continuous and regression quantifies the effect size of each variable. Python modules will be pandas, sklearn and statsmodels.

4.Results

To be completed after analysis.

This section will have descriptive statistics, regression results, correlation analysis and random forest findings.

5.Discussion

In this section, I will discuss:

- Limitations of the data such as bias
- Strengths of the data
- Possibility of some variables not included such as different types of majors.
- Future research possibilities

6.Conclusion

This report will help understand the relationship between daily habits and academic success. This section will highlight how daily habits (the variables) have an impact on

students' academic success (GPA) and how this data could be used to support and encourage students to adopt healthier routines. The overarching impact will show whether simple lifestyle adjustments can translate into better academic outcomes.