Institute of Information Technology
Jahangirnagar University(JU)

Project proposal on-

"Utilizing Machine Learning for Effects of Alcohol on the Study"

Group no - 09

Name	Roll
Bushra Akter	1965
Zannat Hossain Tamim	1970
Sabina Yeasmim	1984

ABSTRACT

Alcohol consumption is a major public health problem that can have a significant impact on academic performance. College students are particularly vulnerable to the negative effects of alcohol, as they are still developing their brains and bodies. These problems can make it difficult for students to succeed in their studies. This project explores the effect of alcohol on academic outcomes. By examining existing literature and collecting and analyzing relevant data, including alcohol consumption patterns, study habits, and grades, this study intends to employ correlational analysis and regression modeling techniques to uncover potential links between these variables. This project aims to uncover potential correlations between alcohol intake and study efficacy. The findings could provide insights into the complex relationship between alcohol consumption and study performance. contribute to informed discussions on student well-being, and suggest strategies for promoting healthier study habits. The results of the project will be used to develop interventions to help students who are struggling with alcohol-related academic problems.

CHAPTER I

INTRODUCTION

In Today's fiercely competitive academic landscape, educational institutions must remain attuned to the multifaceted factors influencing student achievement. One such factor with implications for academic outcomes is alcohol consumption.

In contemporary society, student alcohol consumption remains a prevalent and intricate concern. Students engage in drinking for various motives, ranging from social interactions and gatherings to seeking relaxation. This behavior is intricately intertwined with factors like peer pressure, evolving cultural norms, and the delicate balance between academic stress and leisure pursuits. Yet, this practice is not without challenges, including potential harm to academic performance, health-related risks, and the potential to foster unhealthy coping mechanisms. In the midst of these complexities, comprehending the current landscape of student alcohol consumption assumes paramount importance.

To address this concern, we propose an investigation into the "Effects of Alcohol on Study." This project will harness the power of machine learning to delve into the intricate interplay between alcohol consumption and academic achievement.

In this problem, we use machine learning to understand alcohol's impact on studying for several distinct benefits. Machine learning excels at detecting intricate patterns in these datasets, revealing subtle correlations missed by traditional methods. It efficiently incorporates multiple variables, capturing the issue's complexity. Machine learning predicts outcomes, aiding scenario assessment and projecting how alcohol habits influence academics, a challenge for other approaches. It ensures unbiased analysis, another reason for selecting a machine learning approach.

The project aims to comprehensively explore the relationship between alcohol consumption and academic performance using machine learning. The primary goal is to

leverage advanced data analysis to uncover insights into how alcohol habits influence studying. Objectives include data collection, model development, and prediction of alcohol's impact on academic outcomes. By analyzing patterns and trends in the data, we aim to gain a better understanding of how alcohol consumption affects students' academic performance. This information could be used to develop interventions to help students who are struggling with alcohol abuse and to prevent academic problems from occurring in the first place.

CHAPTER 2

LITERATURE REVIEW

Understanding the effects of alcohol consumption on academic performance is a subject of growing importance. Prior research has highlighted the negative impact of alcohol on cognitive functions, memory retention, and overall academic achievement. Several studies have employed traditional statistical analyses to examine the relationship between alcohol consumption and academic outcomes. These studies often rely on self-reported surveys and questionnaires, which might suffer from response bias and inaccuracies in reporting. Moreover, while these studies offer valuable insights, they often fail to capture the complex and non-linear interactions between various factors that contribute to the effects of alcohol on studying.

There is a growing body of research on the use of machine learning to predict the effects of alcohol. A study by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) found that alcohol use can impair cognitive function, including memory, attention, and learning. The study also found that alcohol use can lead to academic problems, such as lower grades and increased dropout rates [1]. The studies did not always measure cognitive function in the same way. The studies did not always adjust for other factors that could have influenced the results, such as underlying mental health conditions.

Another study, published in the journal "Alcoholism: Clinical and Experimental Research," found that alcohol use can disrupt sleep patterns, which can also impair cognitive function. The study also found that alcohol use can lead to mood problems, such as anxiety and depression, which can also affect academic performance [2]. The study was limited to a predominantly white college student sample in the Midwest. The findings may not generalize to other populations due to the small sample size and the missing data, which was not completely random.

The project aims to address the limitations of current approaches by developing a machine learning model that can be trained on a large, high-quality dataset of data on alcohol use and its effects on learning. The model will be developed using a variety of machine-learning techniques, including supervised learning.

CHAPTER 3

PROPOSED METHODS

3.1 Data Collection and Overview:

In the current situation in Bangladesh, alcohol consumption is a very common and growing threat [3]. Our main objective is to organize the machine learning concepts and try to make predictions about alcohol's impact on academic outcomes. So, primarily, we have collected data (Student Alcohol Consumption) from Kaggle [5].

3.2 Data Pre-processing:

We take a data set (Student Alcohol Consumption). If this data set contains any null values or messy data, then for our further processing, we will clean our data. Also, if there is any missing data or type mismatch features. Then, we will remove the features that lead to numerous percentages of missing values and impute other missing column values with their mean and frequency. Also, the data type of each feature will be checked and fixed. Finally, our data set will contain some samples with some features. We will use a label encoder to encode the data into a numeric format for further processing. We will perform a different statistical analysis to find important features among those features. Like the Chi-Square test. The Chi-Square test will give us better accuracy in the final score. So, after finding out the score of each feature, we will drop the less important column that will help the model become simpler and more powerful.

3.3 Classification:

Classification is a supervisory technique that categorizes the data into the desired number of classes [5]. The goal of this work is to find out the factors behind alcohol consumption and predict a person’s probability of being an alcoholic. In that manner, we can try to keep people, especially teenagers, away from this deadly situation. So, we have employed six classifiers: Decision Tree (DT), Random Forest (RF), Logistic

Regression (LR), Support Vector Machine (SVM), K-Nearest-Neighbour(KNN), and Linear regression. Finally, we will make a comparison of their performance based on different model evaluation metrics and find out the best-fitting algorithm for this piece of the problem.

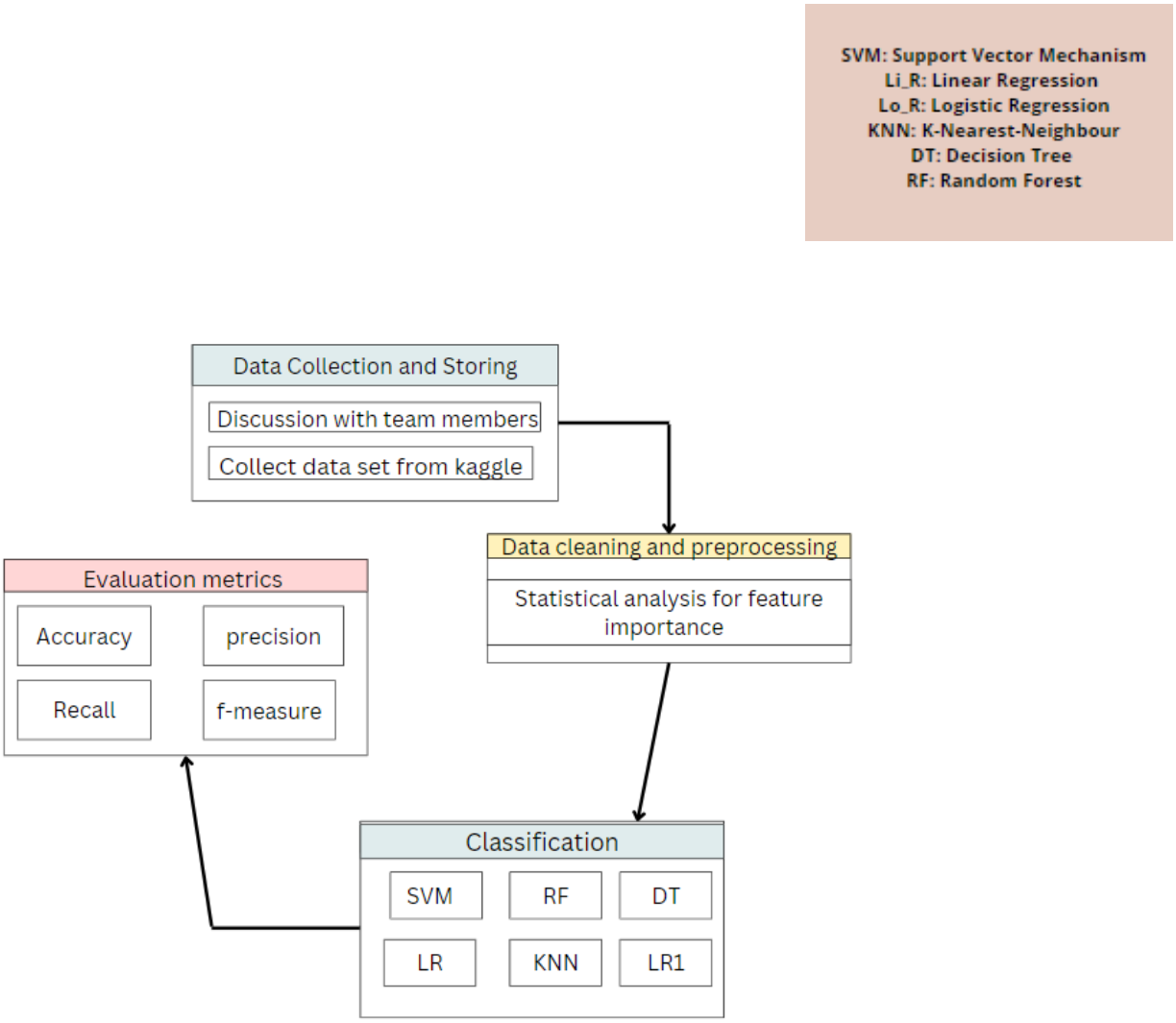

Figure 1. Methodology

CHAPTER 4

PROJECT PLAN

The project plan outlining the tasks, estimated timelines, milestones, and deliverables for the research project focused on understanding the effects of alcohol on academic performance:

Phase	milestone	Delibarable
Project Initiation and Preparation	Project objectives and scope defined	Project scope document outlining research goals and questions
Data Collection and Preprocessing	Data collected, cleaned, and ready for analysis.	Cleaned dataset and exploratory data analysis report.
Data Analysis and Model Development	Data analyzed, models developed, and initial performance assessment completed.	Statistical analysis results and trained machine learning models.
Results Interpretation and Insights	Research presentation and documentation completed.	Research presentation slides and documented research process.
Recommendations and Conclusion	Recommendations formulated and research summarized.	Recommendations report and research summary.
Presentation and Documentation	Research presentation and documentation completed.	Research presentation slides and documented research process.

Figure 2:Project Plan

4.1 Time distribution

TASK NO	TASK NAME	START TIME	FINISH TIME	DURATION
01	PROJECT INITIATION AND PREPARATION	01-05-23	31-07-2023	3 MONTH
02	DATA COLLECTION AND PRE-PROCESSING	01-08-2023	31-08-2023	1 MONTH
03	DATA ANALYSIS AND MODEL DEVELOPMENT	01-09-2023	10-09-2023	10 DAYS
04	RESULTS INTERPRETATION AND INSIGHTS	11-09-2023	20-09-2023	10 DAYS
05	RECOMMENDATIONS AND CONCLUSION	21-09-2023	25-09-2023	5 DAYS
06	PRESENTATION AND DOCUMENTATION	26-09-2023	30-09-2023	5 DAYS

Figure 3. Time distribution

4.2 Gantt chart

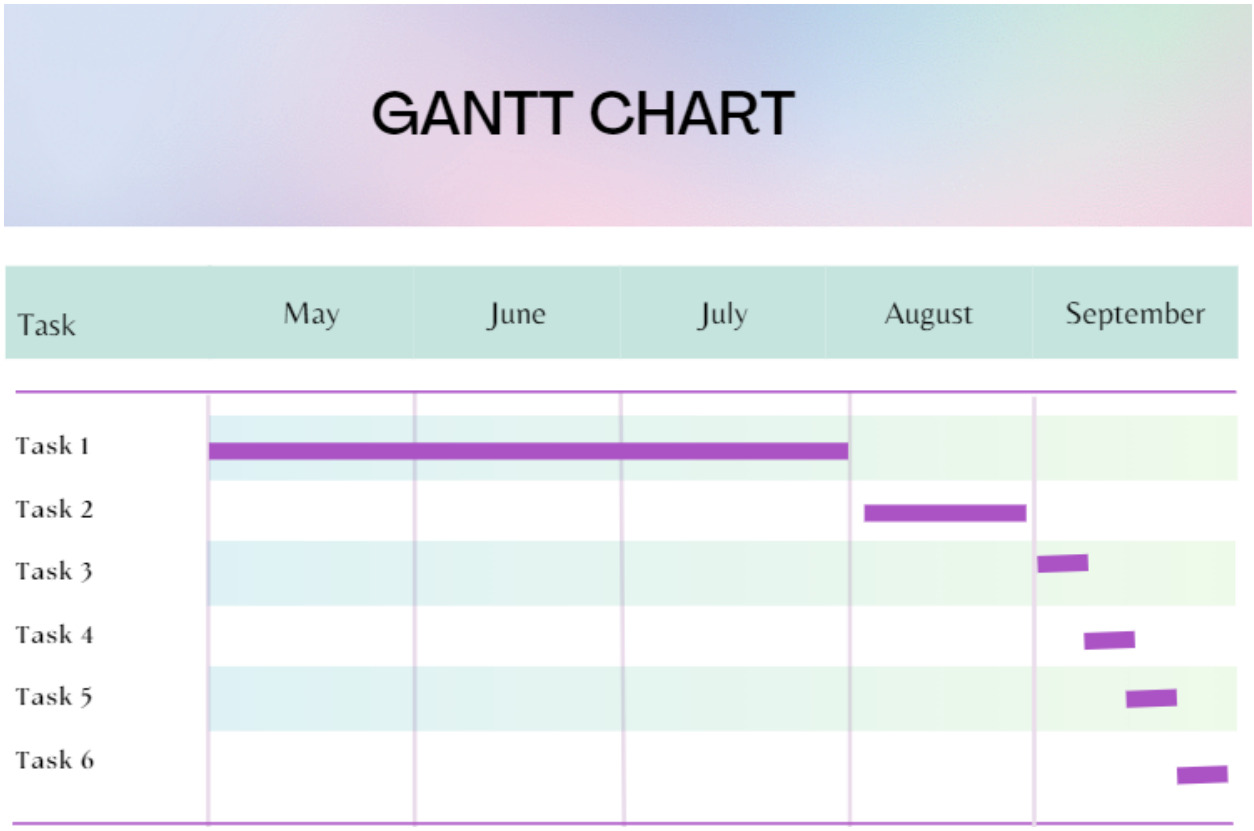

Figure 4. gantt chart

CHAPTER 5

EXPECTED RESULTS

5.1 Expected result

We expected that the project would endeavor to accomplish a multifaceted set of outcomes. These include the identification of meaningful correlations and significant factors that establish a clear connection between alcohol consumption and shifts in academic performance. The inverse relationship between alcohol consumption rates and academic grades becomes evident when considering the interplay of various key factors. Elements such as age, health, family status, frequency of going out, internet usage, and available free time significantly influence this dynamic.

5.2 Evaluation and performance metrics

For executing the different machine learning models and finding the best one, various evaluation metrics are available [7]. Different evaluation techniques are introduced based on confusion metrics such as accuracy, precision, recall, and f-measure and our model evaluation will be done based on these four evaluation criteria [6].The selected research methodology is well-structured to effectively achieve the project's goals of understanding the factors contributing to alcohol addiction and predicting the likelihood of an individual becoming drug addicted. The first phase, data collection, ensures a robust foundation by gathering pertinent data related to alcohol consumption factors. The second phase, data pre-processing, ensures data quality and reliability by addressing inconsistencies, missing values, and normalization. The third phase, classification, employs seven distinct classifiers to predict the likelihood of drug addiction based on identified factors. The methodology ensures the reliability and generalizability of the model's predictions by employing well-established classifiers. The methodology is a comprehensive and systematic approach to understanding the effect of Alcohol on the study. Our project aims to help students understand the potential future consequences of consuming high levels of alcohol, including how it can

negatively impact their studies. We believe that if students are aware of these risks, they are more likely to make informed decisions about their alcohol consumption.

REFERENCES:

[1] NIH, “, Is Alcohol Consumption Associated with Poor Academic Achievement in University Students?”, Jun. 27, 2021. Available:https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3843305/ [Accessed: Aug. 28, 2023]

[2] NIH, “, Is Alcohol Consumption Associated with Poor Academic Achievement in University Students?”, Aug. 27, 2021. Available:https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4850026/ [Accessed: Aug. 28, 2023]

[3] Rahman, Mohammad Mizanur, et al. ”Psycho-social factors associated with relapse to drug addiction in Bangladesh.” Journal of Substance Use [Accessed: Aug. 30, 2023]

[4]Kaggle,” Student Alcohol Consumption”,Jun. 27, 2024.Available:https://www.kaggle.com/datasets/uciml/student-alcohol-consumption [Accessed: Aug. 28, 2023]

[5] Kotsiantis, Sotiris B., I. Zaharakis, and P. Pintelas. ”Supervised machine learning: A review of classification techniques.” Emerging artificial intelligence applications in computer engineering.[Accessed Aug.30, 2023]

[6] Trivedi, Tanvi, and Devangi Kotak. ”Exploring Prediction Modeling of Students Alcohol and Drug Addiction Affecting Performance using Data Mining Approach.”[Accessed 31 Aug,2023]

[7] Handelman, Guy S., et al. ”Peering into the black box of artificial intelligence: evaluation metrics of machine learning methods.” American Journal of Roentgenology .[Acessed 31 Aug,2023]