Bangladesh Army University of Science and Technology, Saidpur

Department of Computer Science and Engineering

Predictive Analysis of Student Smoking, Alcohol, and Psychological Wellness

Course Code: CSE 4140
Course Title: Machine Learning Sessional

Prepared By:
Md. Mominul Islam (220201033)

Prepared for:
Engr. Rohul Amin
Lecturer, CSE, BAUST
Nadim Rez
Lecturer, CSE, BAUST

TABLE OF CONTENTS

Introduction

- Psychological wellness is a critical factor influencing student academic performance, social interactions, and overall quality of life.

- University students face unique stressors:
  - Academic pressure
  - Adjustment to new environments
  - Social challenges

- Smoking and alcohol habits can increase stress and reduce wellness.

Goal: Predict student wellness using behavioral and demographic data to identify those needing support.

Objectives

- Examine the impact of student behaviors, demographics, and coping strategies on psychological wellness.
- Build and evaluate machine learning models for accurate wellness prediction.
- Detect at-risk students to inform targeted support and interventions.

Related Work

In [1], S. Smith, J. Doe, and R. Lee, 2020, proposed a “Machine Learning-Based Prediction of Mental Well- Being Using Health Behavior Data (ASEAN University Students).”

Shortcomings:

- Focused mainly on physical and academic features, missing detailed smoking and alcohol behaviors.
- Cross-sectional design, limiting causal inference.
- Generalizability limited to Southeast Asian students.

Related Work

In [2], A. Tan, K. Lim, and M. Chua, 2020, proposed "The Clusters of Health- Risk Behaviours and Mental Wellbeing among ASEAN University Students."

Shortcomings:

- Only clustering & no predictive model built.
- Limited behavioral measures, psychological scales underused.
- Findings limited to ASEAN university students.

Related Work

In [3], R. Petrauskiene, E. Vaitkeviciene, and G. Kriaucioniene, 2019, proposed "Lifestyle Factors and Psychological Well- Being: 10-Year Follow-Up Study (Lithuania)."

Shortcomings:

- Focused on middle-aged adults, not students.
- Did not include student-specific behaviors like stress-coping.
- Some causal relationships unclear despite longitudinal design.

Methodology

Dataset

- **Number of Students:** 1,174
- **Features:** 13 features including:
  - **Demographics:** Age, Year of Study, Field of Study
  - **Behavioral:** Smoking frequency, age of starting, alcohol consumption frequency, reasons
  - **Psychological:** Coping strategies, help-seeking behavior, interest to quit unhealthy habits, psychological wellness level

- **Data Source:** Aggregated from multiple public datasets.

Methodology

Dataset

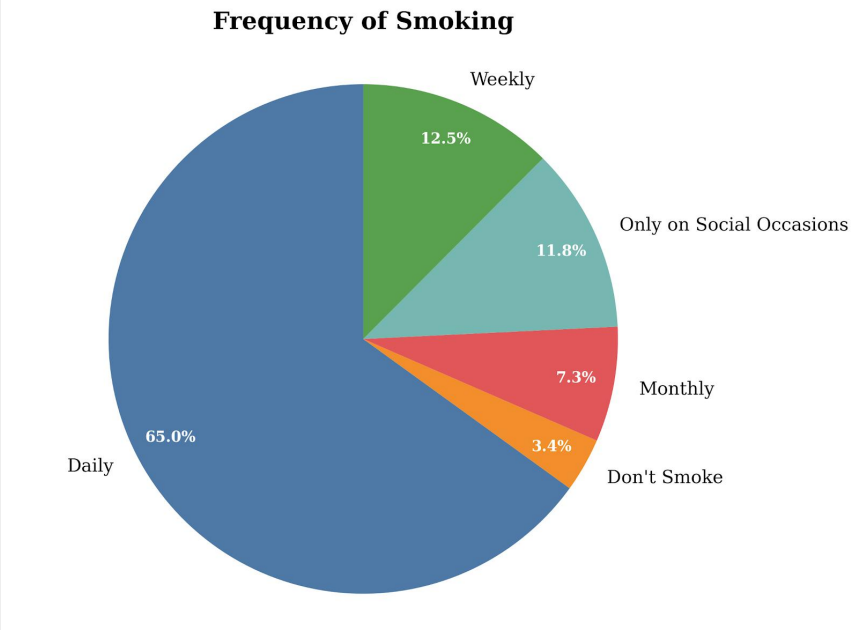


Fig-1: Frequency of Smoking

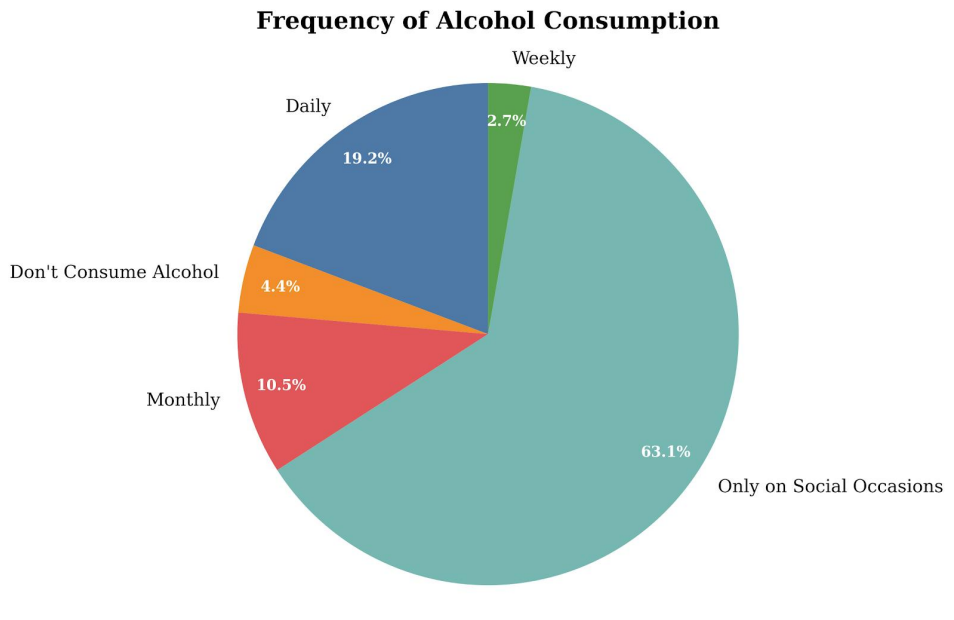


Fig-2: Frequency of Alcohol Consumption

Methodology

Dataset

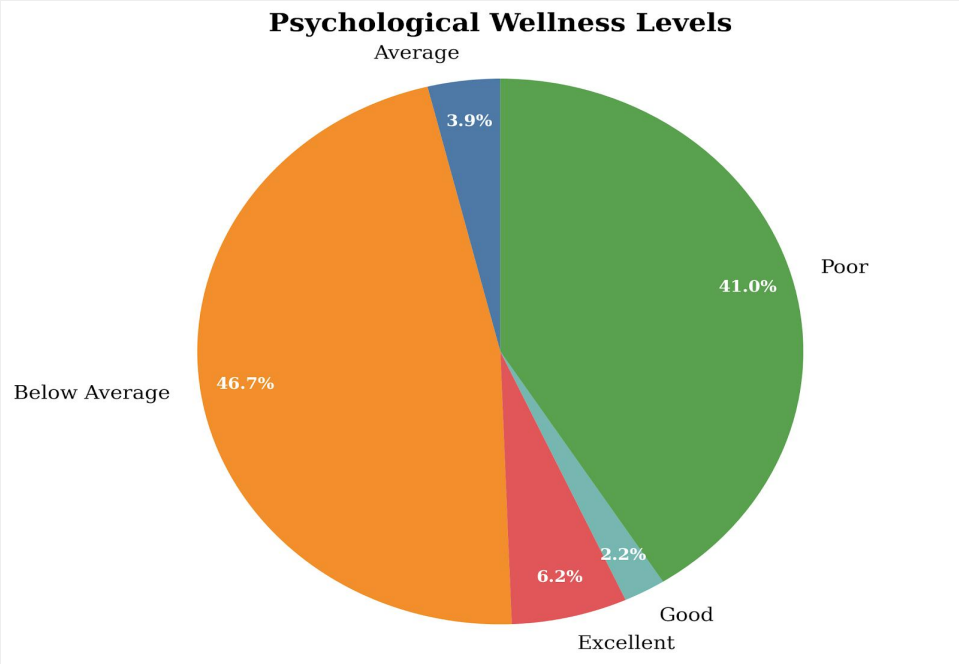


Fig-3: Psychological Wellness Levels

Methodology

Flow Chart

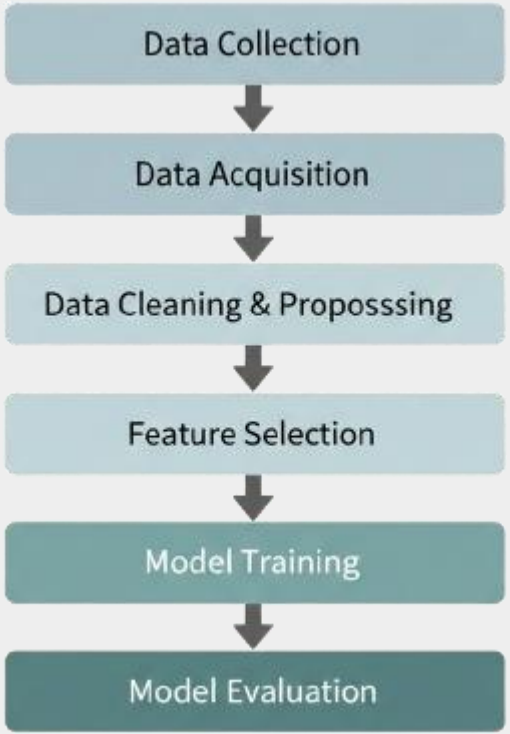


Fig-4: Flow chart of Methodology

Result

Model	Test Accuracy (%)
Random Forest	81.55
Logistic Regression	69.10
Decision Tree	78.54
K-Nearest Neighbors	72.10

Table-1: Accuracy Table

Result

Random Forest Model (81.55 %)

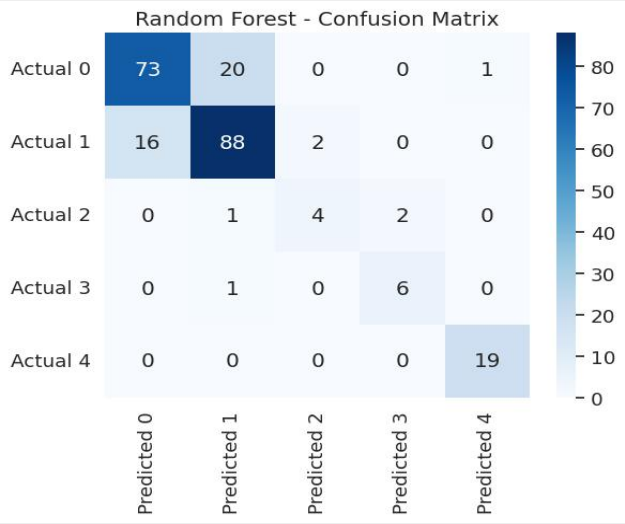


Fig-5: Confusion Matrix of Random Forest

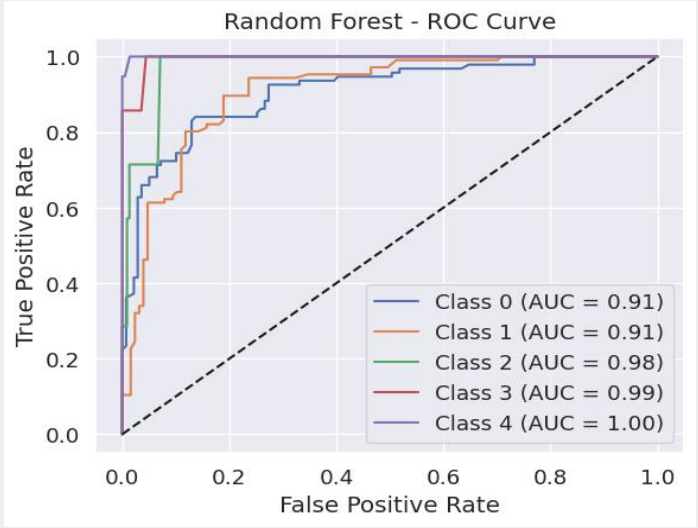


Fig-6: ROC Curve of Random Forest

Result

Logistic Regression Model (69.10%)

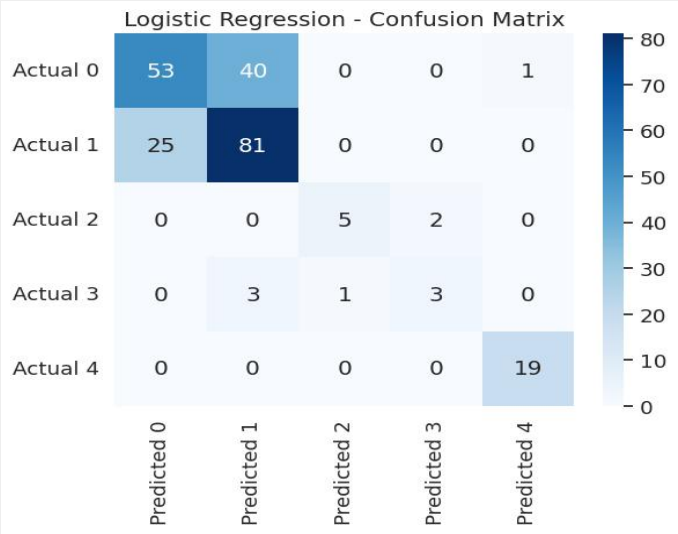


Fig-7: Confusion Matrix of Logistic Regression

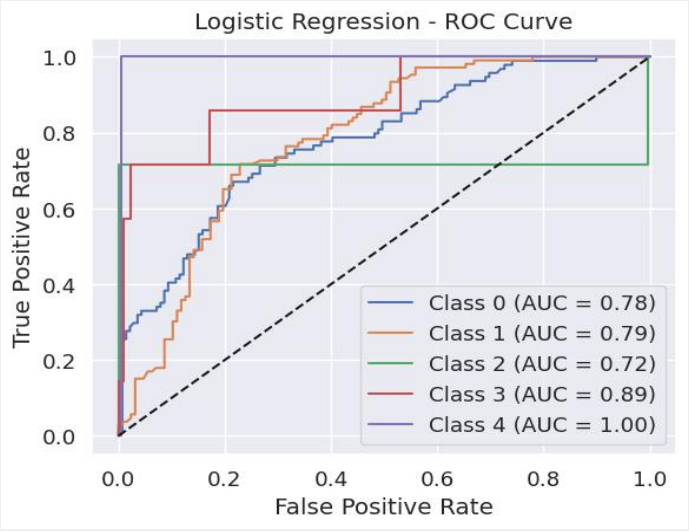


Fig-8: ROC Curve of Logistic Regression

Result

Decision Tree Model (78.54%)

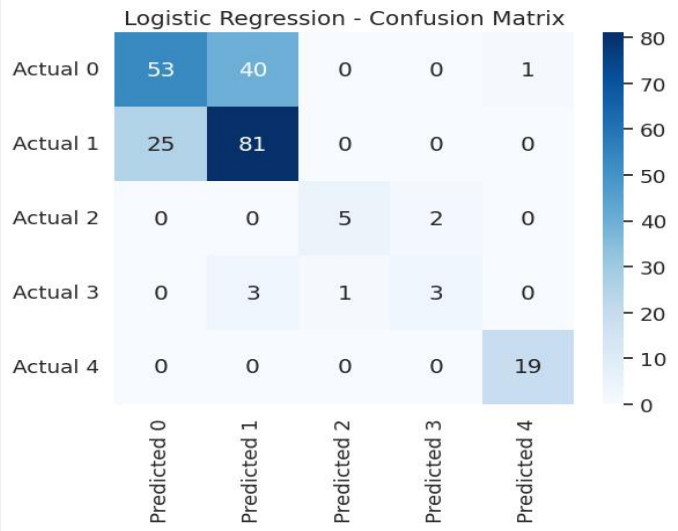


Fig-9: Confusion Matrix of Decision Tree

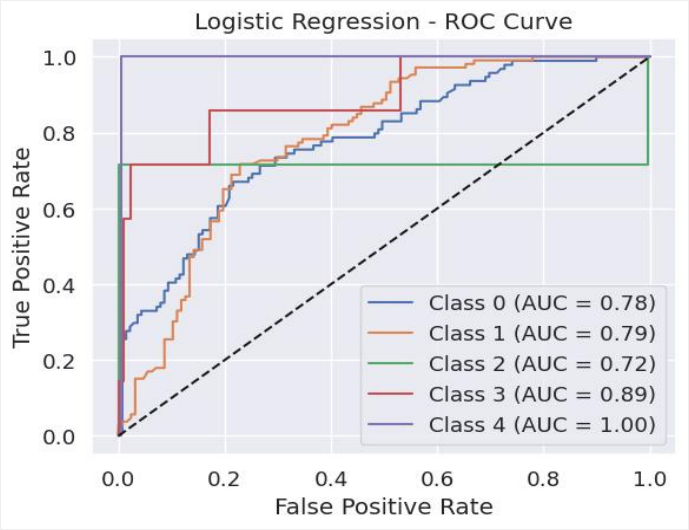
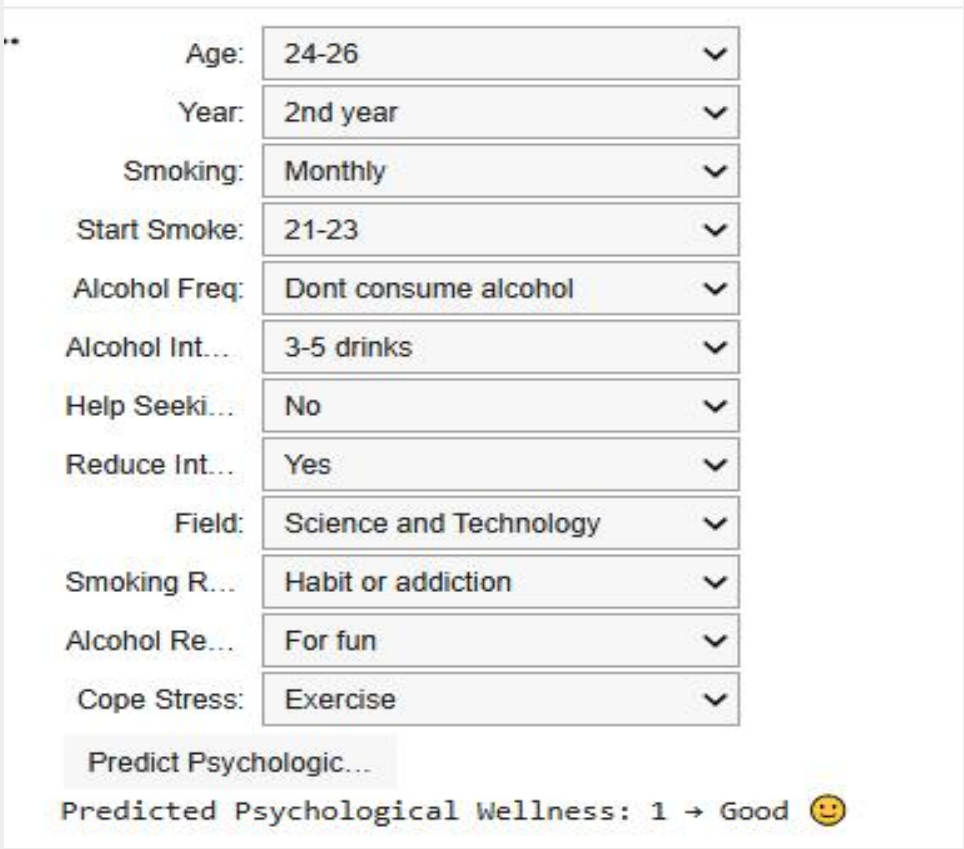


Fig-10: ROC Curve of Decision Tree

Outcomes


Age: 24-26
Year: 2nd year
Smoking: Monthly
Start Smoke: 21-23
Alcohol Freq: Dont consume alcohol
Alcohol Int... 3-5 drinks
Help Seeki... No
Reduce Int... Yes
Field: Science and Technology
Smoking R... Habit or addiction
Alcohol Re... For fun
Cope Stress: Exercise
Predict Psychologic...
Predicted Psychological Wellness: 1 → Good

Comparison with Related Works

Reference	Technology Used	Test Accuracy	Limitations	Our Proposed Work
[1] Smith et al., 2020	Random Forest, Gradient Boosting	~75%	Only physical and academic features, detailed smoking/alcohol behaviors missing	Uses student-specific behavioral data (smoking frequency, reasons, age of start, coping behaviors) for prediction
[2] Tan et al., 2020	Clustering (Health-Risk Behaviors)	N/A	Cross-sectional clustering, no predictive modeling	Predictive ML models applied directly to behavioral variables to estimate psychological wellness
[3] Petrauskiene et al., 2019	Logistic Regression (longitudinal)	~70%	Focused on middle-aged adults, student-specific	Targets university students with comprehensive behavior and wellness

Table - 2: Comparison Table of all related work and this study

Future Research Direction

Current study focused on Predicting Psychological Wellness using ML.

In future we will:

- Integrate longitudinal data to track wellness changes over time.
- Develop real-time predictive tools for student support services.
- Give advice based on each student’s needs so they get the right help.

Conclusion

- We used student data and machine learning to predict psychological wellness. Random Forest gave the best results with 81.55% accuracy. Smoking, alcohol, and coping habits affect wellness. Cleaning and preparing the data carefully helped improve the model. In the future, more data and better models can make predictions even more accurate.

References

[1] S. Smith, J. Doe, and R. Lee, “Machine Learning-Based Prediction of Mental Well-Being Using Health Behavior Data (ASEAN University Students),” PLoS ONE, vol. 15, no. 8, pp. 1–15, Aug. 2020.

[2] A. Tan, K. Lim, and M. Chua, “The Clusters of Health-Risk Behaviours and Mental Wellbeing among ASEAN University Students,” BMC Public Health, vol. 20, no. 1, pp. 1–15, Jan. 2020.

[3] R. Petrauskiene, E. Vaitkeviciene, and G. Kriaucioniene, “Lifestyle Factors and Psychological Well-Being: 10-Year Follow-Up Study,” BMC Public Health, vol. 19, no. 1, pp. 1–10, Dec. 2019.

Thank You !

Any Question?