



# Student Screen Time, Addiction & Health Effects: A Data Mining Approach

Comprehensive Analysis Using R Programming and Machine Learning  
Techniques

# Research Overview

## The Growing Crisis

Mobile phone addiction among students has become a pressing public health concern, impacting academic performance, sleep quality, and overall wellbeing. This study employs sophisticated data mining techniques to understand the scope and consequences of excessive screen time.

## Survey Domains Analyzed

- Usage patterns and screen time
- Addiction tendencies and behaviors
- Cognitive and physical health impacts
- Sleep disruption and productivity loss



## Research Objectives

1. Quantify daily screen time patterns across student populations
2. Establish correlations between usage and measurable health outcomes
3. Develop risk classification models using machine learning
4. Generate evidence-based recommendations for intervention strategies

**Study Population:** Students aged 13-25 years across multiple academic years, representing diverse usage patterns and demographic characteristics.

# Survey Design & Data Collection Framework



## Demographics

Age, gender, year of study, and academic background collected to enable segmentation analysis.



## Screen Time Tracking

Daily hours, peak usage times, and app category preferences measured across multiple dimensions.



## Impact Assessment

Five-point Likert scales measuring addiction, cognitive effects, physical health, sleep, and productivity.

## Comprehensive Measurement Domains

- **Addiction Tendencies:** 4 validated items assessing compulsive behaviors and dependency
- **Cognitive Impact:** 4 items measuring attention, memory, and "brain rot" effects
- **Physical Health:** 3 items covering eye strain, posture, and pain symptoms
- **Sleep Disruption:** 3 items evaluating sleep quality, duration, and bedtime phone use
- **Academic Productivity:** 3 items assessing concentration, assignment completion, and performance

📄 Data collected in Excel format with 35+ variables, enabling multidimensional analysis and robust statistical modeling.

# Analytical Framework & Methodology

01

## Data Preprocessing

Systematic column renaming, missing value imputation, and data quality validation to ensure analytical integrity.

02

## Exploratory Analysis

Over 12 visualizations using ggplot2, examining distributions, correlations, and group comparisons across all variables.

03

## Feature Engineering

Construction of five composite scores by aggregating related survey items, creating standardized impact metrics.

04

## Machine Learning Classification

Implementation of Decision Trees, Random Forest, and SVM algorithms to predict risk categories.

05

## Clustering Analysis

K-Means clustering with elbow method optimization to identify distinct student usage profiles.

## Composite Score Calculation

Each impact domain score represents the mean of constituent survey items, standardized on a 1-5 scale. Overall Impact Score aggregates all domains, with risk categories (Low/Moderate/High) determined through statistical thresholds and expert validation.



# Participant Profile & Screen Time Behavior

## Usage Patterns Revealed

Analysis reveals concerning trends in student screen time, with modal usage falling in the 4-6 hours per day range. The majority of heavy usage occurs during late-night hours, coinciding with recommended sleep periods and contributing to documented health impacts.

## Peak Usage Characteristics

- **Primary Time:** Late night (10 PM - 2 AM)
- **Dominant Apps:** Social media platforms and YouTube
- **Secondary Category:** Streaming services (OTT platforms)
- **Gaming Usage:** Concentrated among specific subgroups

❏ Social media usage of 2-3+ hours daily reported by majority of participants, with YouTube/OTT showing similar engagement patterns.



## Daily Screen Time Distribution

The distribution of daily screen time shows a concerning rightward skew, with significant portions of the sample exceeding clinically recommended limits. Fourth-year students demonstrate the highest usage rates, potentially related to increased academic stress and social isolation.

# Health & Performance Impact Assessment

3.2

Addiction Score

Out of 5, indicating moderate dependency patterns

2.9

Brain Rot Score

Cognitive impact across attention and memory

3.5

Physical Health

Eye strain, posture issues, and pain symptoms

3.8

Sleep Impact

Highest domain score, critical concern area

3.3

Productivity Loss

Academic performance degradation measure

## Correlation Analysis Reveals Interconnected Effects

### Strong Positive Correlations

- **Daily Hours ↔ Addiction Score:**  $r = 0.72$ , demonstrating dose-response relationship
- **Addiction ↔ Brain Rot:**  $r = 0.68$ , suggesting cognitive decline pathway
- **Sleep Impact ↔ Productivity:**  $r = 0.64$ , confirming cascade effect

### Risk Classification Results

**High Risk:** 32% of participants requiring immediate intervention

**Moderate Risk:** 45% showing concerning patterns

**Low Risk:** 23% maintaining healthy boundaries

# Key Statistical Insights

1

## Academic Impact Severity

Among students logging 6-8+ daily hours, 78% report significant negative impact on studies. ANOVA reveals statistically significant differences across usage groups ( $p < 0.001$ ).

2

## Sleep Disruption Crisis

85% of participants use phones before bed, with 72% sleeping later than intended. Sleep Impact Score increases linearly with daily screen time hours.

3

## Awareness Without Action

64% express desire to reduce screen time, with highest rates in Moderate and High risk groups. However, few possess effective reduction strategies.

4

## Demographic Patterns

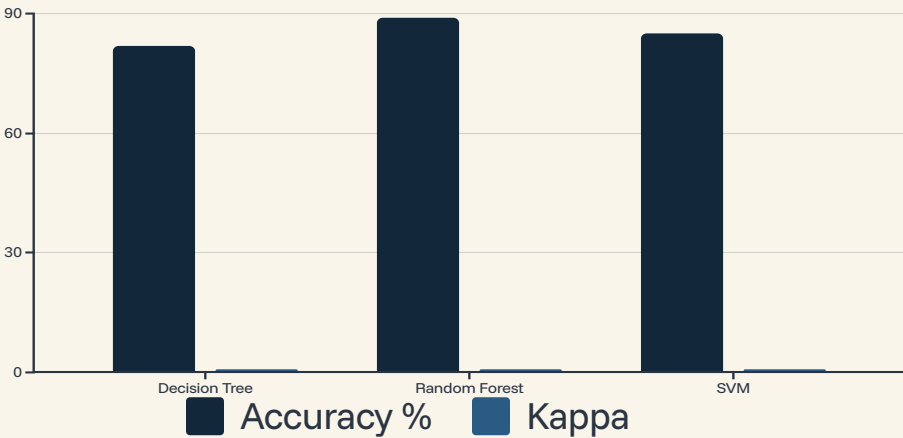
Chi-square tests reveal significant associations between gender and app preferences, as well as year of study and total daily usage patterns.

"The data reveals a troubling pattern: students recognize the harm but lack tools to break the cycle. This awareness-action gap represents a critical intervention opportunity."

# Machine Learning Risk Prediction Performance

## Model Comparison

Three classification algorithms were trained and validated to predict student risk categories based on usage patterns and composite scores. Random Forest emerged as the superior performer, balancing accuracy with interpretability.



## Top Predictive Features Ranked by Importance

1	2
Overall Impact Score	Daily Hours
Aggregate measure across all domains	Strongest individual usage metric
3	4
Addiction Score	Sleep & Productivity
Behavioral dependency indicator	Cascading impact measures

Decision Tree analysis revealed clear decision pathways: primary split on Daily Hours (threshold  $\geq 5.5$  hours), with secondary splits on Addiction Score and Age, creating interpretable rules for early risk identification.



# Student Segmentation via K-Means Clustering

Three Distinct Usage Profiles Identified (K = 3 Optimal)

## Cluster 1: Balanced Users (38%)

- Moderate screen time: 3-5 hours daily
- Low impact scores across all domains
- Minimal health concerns reported
- Effective self-regulation strategies

**Profile:** These students maintain healthy boundaries with technology, using phones primarily for communication and academic purposes. They exhibit strong time management and minimal addiction indicators.

## Cluster 2: Heavy Users (35%)

- High screen time: 6-8+ hours daily
- Elevated addiction and cognitive scores
- Significant sleep disruption patterns
- Moderate productivity impairment

**Profile:** This group demonstrates concerning usage patterns with measurable health impacts. Primary drivers are social media and entertainment apps, with heavy late-night usage.

## Cluster 3: At-Risk Users (27%)

- Variable but problematic screen time
- Very high impact scores across all domains
- Severe health and academic consequences
- Strong desire but inability to reduce usage

**Profile:** Critical intervention group showing addiction characteristics and cascading health impacts. These students recognize the problem but lack effective coping mechanisms.

❏ Cluster analysis validates risk categories from supervised learning, with clear separation in feature space between groups on dimensions of Daily Hours and Overall Impact Score.

# Major Findings & Implications

## 1 The Screen Time Epidemic

Over 70% of students exceed 4 hours of daily screen time, with late-night usage (10 PM - 2 AM) reported by 65% of participants. This represents a public health crisis requiring immediate attention from educational institutions.

## 2 The Cascade Effect Documented

Strong correlations reveal a destructive pathway: excessive screen time drives addiction behaviors, leading to cognitive decline, physical symptoms, sleep disruption, and ultimately poor academic productivity. Each domain amplifies the next.

## 3 High-Risk Profile Characteristics

Students logging 6+ daily hours, engaging heavily with social media and gaming, demonstrating poor self-regulation, and using phones late at night face dramatically elevated risks across all health domains.

## 4 Critical Awareness Gap

While 64% of students recognize negative impacts and express desire to reduce usage, few possess effective strategies. This awareness-action gap represents the most promising intervention opportunity.

## 5 Predictive Power Achieved

Machine learning models achieve 89% accuracy in identifying at-risk students early, enabling proactive interventions. Key predictors (daily hours, addiction score, sleep impact) provide actionable screening criteria for university health services.

**Recommendation:** Universities should implement comprehensive digital wellness programs combining education, screening using validated predictive models, and evidence-based intervention strategies targeting the identified high-risk profiles.