Title

Teen Smartphone Addiction Prediction Using Behavioral & Lifestyle Data

Authors

Denis Mulabegovic

Mostafa Zamaniturk

Wael Alhathal

Institution

Applied Artificial Intelligence 501

Shiley-Marcos School of Engineering

University of San Diego

Date

July 2025

Draft Proposal

Project Summary:

In today's world, it's common to see teens glued to their phones—whether scrolling through TikTok late at night or texting through class. Many parents and teachers worry that this constant screen time could be harming teens' well-being, yet lack the tools to objectively assess the risk. This project aims to bridge that gap by building a model to predict whether a teen has high or low smartphone addiction. The dataset we selected includes 3,000 synthetic records of teen behavior, mental health, academic performance, and phone usage patterns. This project will explore how these factors relate to addiction levels.

The data was generated using statistical modeling grounded in academic literature. Each record reflects plausible behavioral patterns, where addiction level is determined by a weighted combination of screen time, app diversity, and sleep disruption. Variables follow realistic distributions and correlations, with addiction labels derived from a blend of daily usage hours, number of phone checks, time spent on social media, and self-reported sleep hours. The dataset includes fields such as Age, Gender, School Grade, Anxiety and Depression Levels, Time on Gaming and Education, Sleep and Exercise Hours, and more. The diversity of attributes supports a multidimensional view of behavioral health and digital engagement.

This makes the dataset ideal for supervised learning applications, behavioral classification, and exploratory clustering experiments aimed at uncovering latent patterns in teen digital behavior.

Problem Statement and Methodology:

The main objective is to build a predictive model that classifies a teen's smartphone addiction level as either High or Low. Our approach begins with cleaning and preparing the dataset, including grouping records by addiction level and addressing missing or inconsistent data. We will then explore the data using visualization techniques and statistical tests to uncover significant patterns and relationships.

For the modeling phase, we will apply and compare several machine learning algorithms, including Logistic Regression, Random Forest, Gradient Boosting, XGBoost, and K-Nearest Neighbors. This gives us a diverse set of model types to ensure we obtain the best results possible. These models will be evaluated using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. To enhance performance and interpretability, we will also consider feature selection and dimensionality reduction techniques.

Our initial analysis shows that high smartphone addiction is strongly linked to frequent phone use, lack of sleep, and higher anxiety levels. As we build and test our models, we will explore these patterns further. Our approach will include supervised learning and classification, ensemble

methods, model evaluation using ROC analysis, feature selection, dimensionality reduction, and clustering to better understand behavior patterns.

Our system will classify teens into High or Low risk of phone addiction based on their behavioral profile. The goal is not only to build a predictive tool but to empower educators, parents, and healthcare providers with interpretable insights. The system is intended for use in digital wellness platforms or early intervention screening tools for adolescent populations.

Challenges and Priorities:

- Effective preprocessing for mixed data types
- Balancing model accuracy and interpretability
- Mitigating the risk of overfitting
- Identifying ethical implications in behavioral prediction
- Creating compelling and accessible visualizations

Conclusion:

Teen smartphone addiction has become a rising concern in today's digital society. By leveraging AI, this project aims to create a practical solution for the early identification of risk factors using accessible, interpretable data. Through rigorous modeling and precise analysis, we hope to offer both a technical contribution and a meaningful social impact.

References and Data Source:

Kaggle Dataset:
https://www.kaggle.com/datasets/khushikyad001/teen-phone-addiction-and-lifestyle-survey

List of contributions

Denis	Mostafa	Wael
- Model prep (splitting, scaling, etc.) - Model creation - Model tuning - Write up the model prep & selection sections for the report and presentation	- Provide result metrics for models - Provide result comparison graphs (ROC) - Write up the model analysis and evaluation for the report and presentation - Prepare the appendix	- Performed exploratory data analysis (EDA) - Performed data cleaning - Drafted the project proposal - Preparing the Introduction and data cleaning/EDA sections of the report and presentation