An Analysis of Teen Smartphone Usage and Its Impact

Predicting Teen Phone Addiction Levels Using Machine Learning Techniques

Made with GAMMA

Team Members

Salma Hassan Soliman

Basmala Adel Shalaby

Khaled Yasser Zedan

Khaled Adel EL-Sayed

Made with GAMMA

Project Overview

This project addresses the growing concern of problematic smartphone use among adolescents in the UK by leveraging machine learning to predict addiction risk.

- **Problem:** The increasing rates of problematic smartphone use among UK teenagers necessitate effective early identification methods.
- **Objective:** To develop a robust machine learning model capable of predicting individual teen phone addiction risk.
- **Impact:** Our work aims to enable early identification and facilitate targeted interventions for at-risk adolescents, promoting healthier digital habits.
- **Scope:** The analysis focuses on digital usage patterns and behavioural data collected from a specific UK adolescent cohort.

Made with GAMMA

Understanding the Challenge: Project Goals

1. **Predicting Addiction Levels**

   Our primary objective is to predict levels of phone addiction among teens by applying various machine learning models to behavioural and lifestyle data.

2. **Highlighting Importance**

   This project underscores the critical issue of smartphone addiction among teenagers and its significant impact on their health, academic performance, and social interactions.

3. **Data Foundation**

   The dataset utilised for this analysis is sourced from Kaggle, providing a rich collection of relevant information for our models.

Made with GAMMA

Dataset Insights: A Snapshot of Teen Behaviour

Our dataset comprises 25 features detailing various aspects of teen smartphone usage and its correlates, offering a comprehensive view for analysis.

Demographics	Teenager's age	15	Numerical	Age
Key Usage Metric	Average daily phone use	5.5	Numerical	Daily_Usage_Hours
Impact Metric	Daily sleep duration	6	Numerical	Sleep_Hours
Impact Metric	Self-reported stress	7	Numerical	Anxiety_Stress_Levels
Target Variable	Score binned into Low/Medium/High	2	Numerical (Target)	Addiction_Level

The dataset includes 25 features, such as **Daily_Usage_Hours** (avg. 4-6 hours) and **Addiction_Level**, collected from teens aged 13-18. It contains thousands of records, with 5808 training samples alone.

Made with GAMMA

Data Processing: From Raw Data to Features

Data Cleaning

We meticulously handled missing values and identified outliers to ensure data quality and integrity.

Encoding & Scaling

Categorical variables like Gender and Phone_Usage_Purpose were encoded using LabelEncoder, while numerical features underwent standardisation with StandardScaler.

Class Balancing

Due to severe class imbalance (e.g., only 6 samples for 'low addiction' vs. 482 for 'high'), SMOTE was crucial to create a balanced dataset for robust model training.

Feature Selection

SelectKBest with f_classif was employed to pinpoint the top 5 most impactful features, such as **Daily_Usage_Hours** and **Screen_Time_Before_Bed**.

We encoded categorical variables, scaled numerical features, and used SMOTE to balance the dataset, addressing the underrepresentation of low addiction levels.

Made with GAMMA

Methodology: Machine Learning Approaches

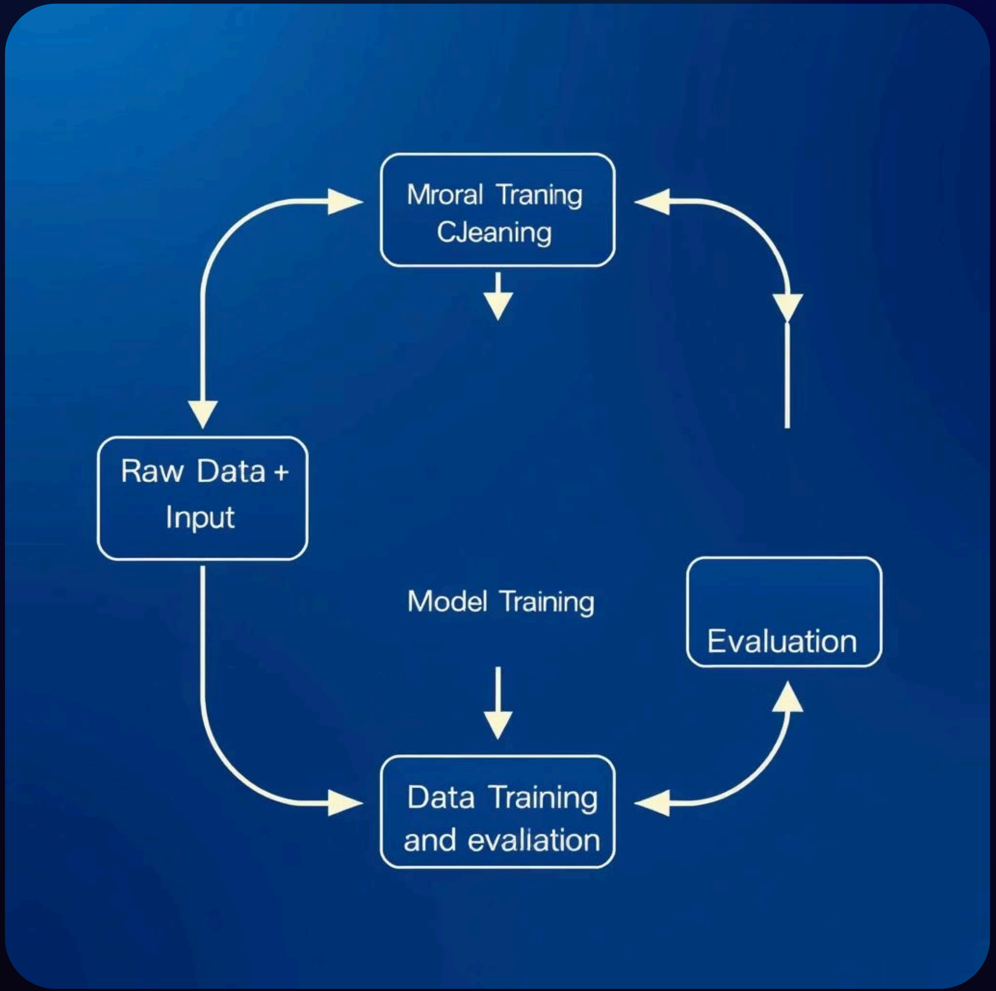
Rigorous Model Testing

- XGBoost
- Random Forest
- SVM
- Logistic Regression
- GaussianNB
- KNN

We trained 9 machine learning models to classify addiction levels, with extensive hyperparameter tuning to maximise performance.

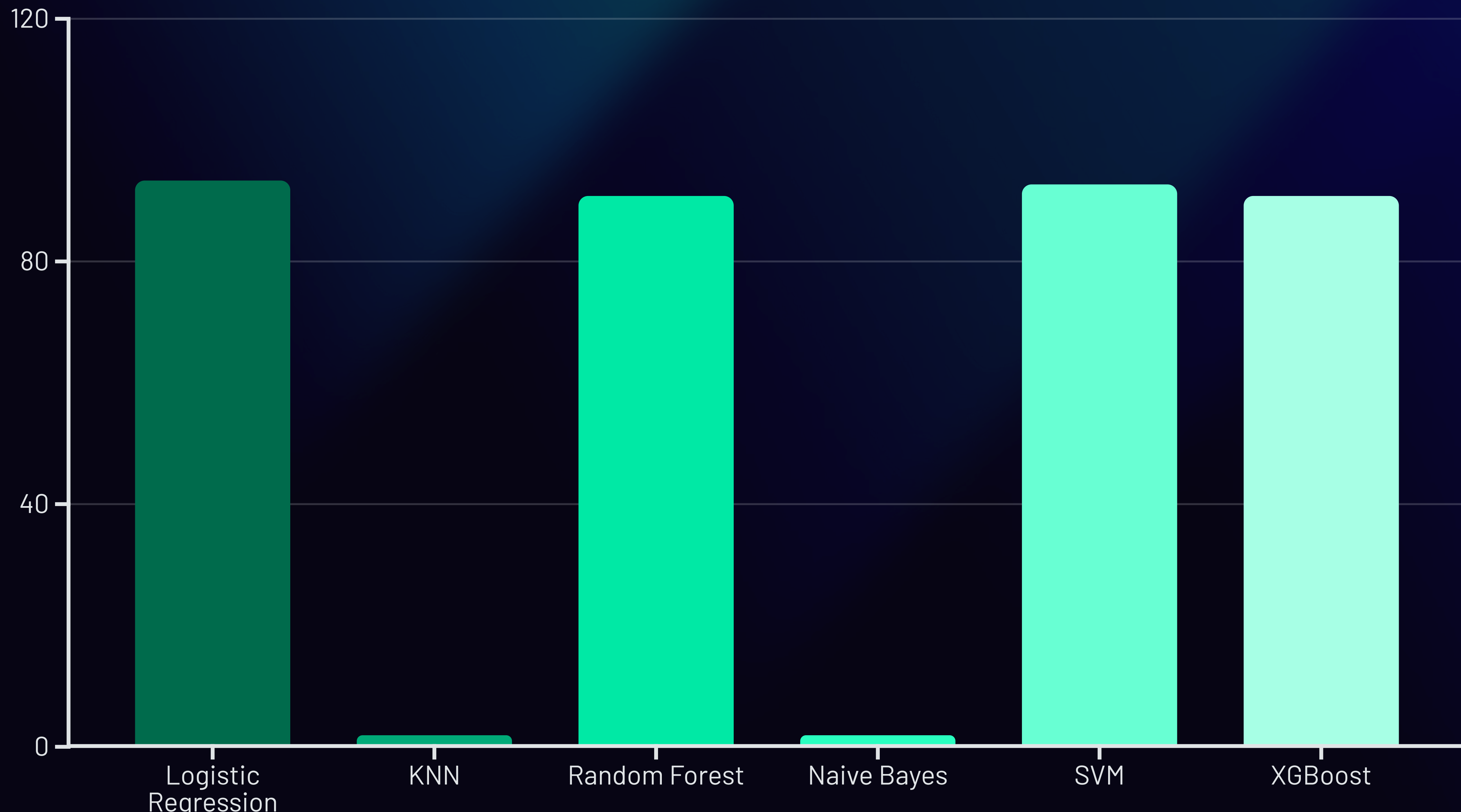
Optimisation & Validation

A 80% training (5808 samples) and 20% testing (600 samples) split was used. Hyperparameter tuning, including combinations like 'n_estimators' and 'max_depth' for Random Forest, was performed using RandomizedSearchCV for XGBoost, Random Forest, and MLPClassifier.



Made with GAMMA

Model Performance: Comparative Results



Logistic Regression achieved the highest accuracy at 93.33%, with SVM closely behind at **92.66%**. While overall performance was strong, Class 0 (Low addiction) exhibited lower precision due to the limited number of available samples.

Made with GAMMA

Model Deployment on the Web Steps for Deploying the "Social Media Addiction Level Prediction" Project:

1. Model Development an AI model was developed to predict social media and phone addiction levels based on daily usage data.
2. Using Streamlit The interactive user interface was built with Streamlit, allowing users to input their own data and instantly check their addiction level on the website.
3. The entire application was deployed on Hugging Face, using a decoder model to make the project accessible online for everyone
4. Easy Access and Interaction Anyone can try the tool through a simple, fast interface without the need to install any software.

Screenshot from the Deployed App: [Social Media Addiction Prediction Project](attached_image): Benefits of This Deployment:

Easy access to the project from anywhere online.

Interactive and intuitive user interface for all users.

Real-time experience of the model's predictions.

Made with GAMMA

Conclusion and Future Outlook

Key Insights:

- **Feature Importance:** Daily_Usage_Hours, Screen_Time_Before_Bed, and Phone_Checks_Per_Day emerged as the strongest predictors of addiction levels.
- **Class Imbalance:** The low sample size for Class 0 (Low addiction, only 6 samples) negatively impacted its precision, highlighting a data limitation.
- **Model Strength:** High accuracy for Class 2 (High addiction) demonstrates the models' ability to identify at-risk individuals effectively, despite challenges with other classes.

Daily phone usage and screen time before bed were consistently the strongest predictors of addiction levels in our analysis.

Made with GAMMA