Social Media, Exercise, and Happiness

Can Sports Balance the Effects of Screen Time?

Final Project: Programming for Data Science

Diego Avella • STU Miami, FL

Dataset Overview

Source & Observations

Mental Health & Social Media Balance Dataset (Kaggle)

1,000

Observations

Key Variables

- Daily screen time (hours)
- Exercise frequency (days/week)
- Happiness level (1-5 scale)
- Sleep quality, stress level
- Days without social media

Data Split

Training: 70%

Testing: 30%

Source: https://www.kaggle.com/datasets/prince7489/mental-health-and-social-media-balance-dataset

Research Questions

Question 1

Is unhappiness related to social media consumption?

Question 2

Can this unhappiness be balanced with sport regardless of social media consumption?

Key Findings from EDA

Screen Time Impact

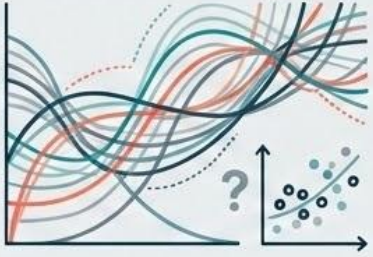
Higher daily screen time strongly associated with lower happiness levels across all exercise frequencies.

Exercise Benefit

Exercise shows slight positive correlation with happiness but does not fully offset high screen time effects.

Sleep & Stress

Better sleep quality and lower stress levels emerge as strong predictors of higher happiness.



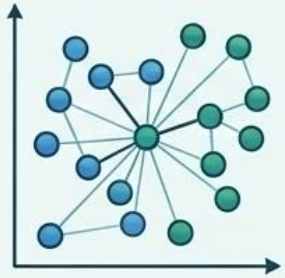
Non-linear Patterns

Relationships between variables show complex, non-linear patterns requiring advanced models.

Key Insight: Social media breaks (days without social media) have modest positive correlation with happiness.

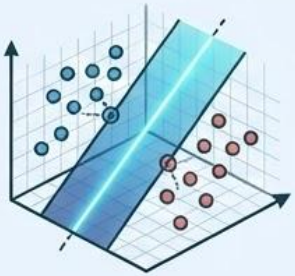
Methodology

Goal: Predict happiness level (1-5) using machine learning classification



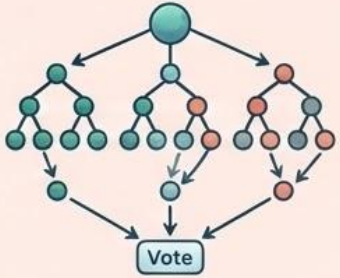
K-Nearest Neighbors

Instance-based learning algorithm that classifies based on similarity to K nearest data points



Support Vector Machine

Finds optimal decision boundaries to separate different happiness classes

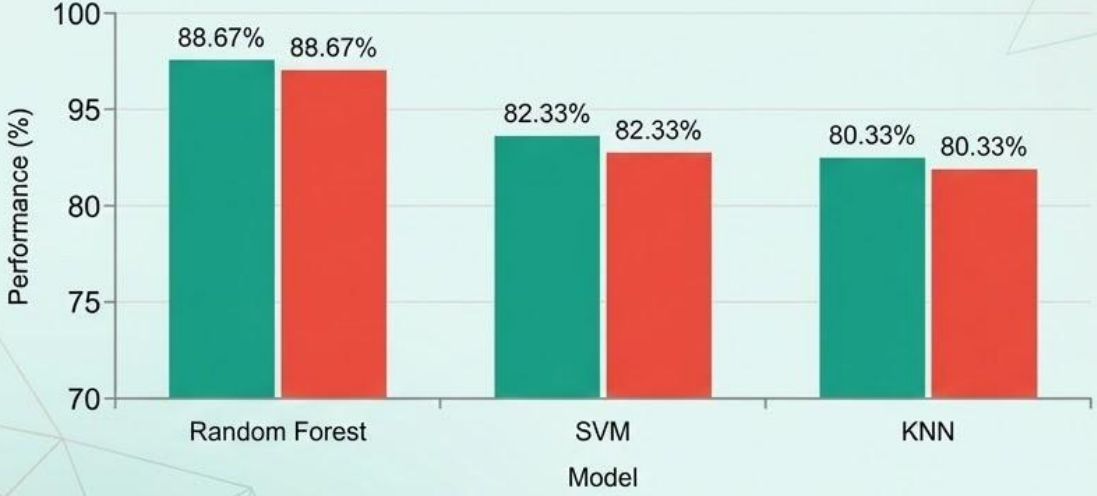


Random Forest

Ensemble method combining multiple decision trees for robust predictions

Features: Daily screen time, exercise frequency, sleep quality, stress level, days without social media

Model Performance



Best Performer

Random Forest

88.67% Accuracy

MODEL RANKINGS

1. Random Forest: 88.67%

2. SVM: 82.33%

3. KNN: 80.33%

Key: F1 scores showed similar pattern, confirming Random Forest's superiority in balanced prediction.

What Drives Happiness?

TOP PREDICTORS

1. Daily Screen Time

Most important factor - higher screen time strongly predicts lower happiness

2. Sleep Quality

Better sleep associated with higher happiness levels

3. Stress Level

Lower stress correlates with better happiness scores

Exercise and social media breaks show some importance but are less dominant than screen time, sleep, and stress.

Conclusions

Q1: Is unhappiness related to social media consumption?

Yes. Higher daily screen time is strongly associated with lower happiness. People spending many hours on screens report significantly lower happiness scores.

Q2: Can sport balance this unhappiness?

Partially. Exercise shows slight positive correlation but does *not fully compensate* for negative effects of very high screen time.

Bottom Line: Reducing screen time and promoting healthy sleep and stress management may be more effective than relying on exercise alone.

Limitations

Cross-Sectional Data

Data captured at single point in time. Cannot claim causation, only correlation between variables.

Self-Reported Measures

Happiness, stress, and sleep quality are subjective. May introduce bias and measurement error.

Sample Representativeness

Sample may not represent all populations (different countries, age groups, cultural contexts).

Social Media Detail

Limited information on how people use social media (content type, purpose, active vs passive use).

Important: These limitations do not invalidate findings but suggest caution in generalizing results and emphasize need for future longitudinal studies.

Future Work & Recommendations

Longitudinal Studies

Track individuals over time to better understand cause and effect relationships.

Detailed Usage Data

Collect information on content type, purpose, and time of day of social media use.

Advanced Modeling

Explore deep learning, hyperparameter tuning, and ensemble methods for improved predictions.

Key Takeaway

Holistic approach needed: reduce screen time, improve sleep, manage stress - exercise alone cannot offset digital overuse.

Practical Applications

- Digital wellness programs
- Screen time interventions
- Integrated health strategies
- Public health campaigns

Thank You

Questions & Discussion

Diego Avella

STU Miami, FL

Programming for Data Science