Social Media Addiction Test: Web-Based and Machine Learning Platform

Abdelrahman Abdelazim Abdelrahman, Abdalla Hamid Mohamed, Dr. Ananta Ojha
School of CS & IT, Jain (Deemed-to-be-University)

***Abstract*—This study presents a web-based platform designed to assess social media addiction by integrating modern web technologies with advanced machine learning. The system combines an interactive questionnaire and a predictive model—a Random Forest classifier fine-tuned on a dataset of over 3,000 user responses and 400,000 social media posts. Data were collected via a secure web portal and preprocessing using label encoding and outlier removal (via isolation forest). The model achieved an overall accuracy of 88.6% with an AUROC of 0.65 on a balanced test set. Key predictors include average daily time on social media, distraction scores, and specific behavioural markers. These findings suggest that our hybrid system provides robust early detection of social media addiction risk and offers actionable insights for timely intervention.**

***Keywords*—Social media addiction, web application, machine learning, Random Forest, digital phenotyping, predictive analytics**

I. INTRODUCTION

The rapid expansion of social media usage has raised significant concerns regarding addictive behaviours and their negative impacts on mental health, productivity, and interpersonal relationships. With nearly 5 billion people actively engaging on social platforms, studies indicate that excessive use is correlated with increased anxiety, depression, and attention deficits. This project aims to develop an accessible web-based system that screens for social media addiction using both self-reported behavioural data and automated analysis of social media content. The primary objectives are to:

- Develop a user-friendly web interface (using Node.js and Express) for administering an addiction test.
- Train a Random Forest model on a large, multi-modal dataset—including demographic information and over 400,000 social media posts—to predict risk levels.
- Integrate both components to deliver real-time predictions and detailed user profiles to support early intervention strategies.

This hybrid approach builds on previous research in machine learning for behavioural health and digital phenotyping while addressing limitations in data representativeness and model interpretability.

II. RELATED WORK

Numerous studies have employed machine learning techniques to predict addiction-related behaviours by analysing social media activity and digital footprints. For example, deep learning models such as convolutional neural networks and transformer-based approaches have been used to detect depression from Twitter and Instagram posts. Other projects have applied Random Forest classifiers and logistic regression to analyse factors like online time, engagement patterns, and affective indicators, achieving AUROCs ranging from 0.60 to 0.72. Our work extends these methodologies by integrating additional behavioural metrics, psychometric scores, and detailed usage statistics into a comprehensive risk assessment model. With a dataset comprising over 3,000 user responses and 400,000 posts, our study enhances statistical reliability and model generalizability.

III. METHODOLOGY

3.1 System Architecture and Design

The system comprises two interconnected modules:

- Web/Application:
  Developed using Node.js and Express, the web server manages user authentication, routing, and database interactions. The frontend—built with modern JavaScript frameworks—guides users through a structured questionnaire and secure data upload process. All interactions are encrypted via HTTPS and stored in a MySQL database on AWS RDS.

- Machine Learning Service: Implemented in Python using scikit-learn and Py Torch, this service handles data preprocessing, feature extraction, and model training. The dataset includes over 3,000 user responses with demographic and behavioural information along with approximately 400,000 social media posts. A Random Forest classifier (configured with 100 trees, no maximum depth, and a minimum sample split of 2) is trained on 80% of the data, with the remaining 20% reserved for validation and testing. Model evaluation is conducted using 10-fold cross-validation.

3.2 Data Collection and Preprocessing

Data were collected via a secure web portal where users completed a detailed survey (demographics, time spent online, subjective distraction levels) and uploaded their social media data (extracted using platform APIs).

- Dataset Composition:
  - Over 3,000 user responses.
  - Approximately 400,000 posts spanning text, images, and comments.
  - Demographic variables (age, gender, socioeconomic status) and psychometric scores (e.g., addiction scales).

- Preprocessing Steps:
  - Data cleaning: Removal of missing or inconsistent entries.
  - Encoding: Categorical variables encoded using pre-trained label encoders.
  - Outlier removal: Detected via isolation forest.
  - Normalization: Data standardized using Robust Scaler and Standard Scaler from scikit-learn.
  - Segmentation: Posts exceeding 600 characters were segmented into chunks to meet transformer model input limits, with segmentation carefully managed to minimize class imbalance.

3.3 Model Training and Hyperparameter Tuning
A Random Forest classifier was trained using an 80:10:10 split (training, validation, test) with stratification.

- Hyperparameters:
  - *N estimators:* 100
  - *max depth:* None
  - *min samples split:* 2
  - *random state:* 42

- Training/Details:
  Training was conducted on a high-performance machine with an NVIDIA Titan XP GPU, Intel Xeon CPU, and 64 GB of RAM. The model was trained for 300 epochs with early stopping based on validation loss.

- Feature/Importance:
  Permutation importance (using ELI5) and embedded methods (Lasso CV, Elastic Net CV) identified key predictors such as average daily social media usage, distraction scores, and affective states (e.g., irritability).

3.4 Integration and Deployment

The machine learning service is deployed as a RESTful API. The web application sends user data in JSON format to the API, which returns real-time risk scores. This modular design supports independent updates and scalability for larger datasets.

IV. RESULTS

4.1 Model Performance and Evaluation
Our Random Forest model achieved:

- Accuracy: 88.6% on the test set.
- AUROC: 0.65 for risk classification.

- Precision/Recall/F1-Score:
  For the high-risk category, precision was 68.6%, recall was 76.6%, and the F1-score was 72.4%. Confusion matrix analysis showed that approximately 74% of positive cases were correctly classified, with balanced performance across low, medium, and high-risk groups.

4.2 Feature Analysis
Feature importance analysis revealed:

- Average Daily Social Media Usage: The strongest predictor.
- Distraction Score: A key indicator of risk.
- Engagement Metrics: Frequency of interactions with specific content types also significantly influenced predictions.

Additional data details:

- Data from 3,000+ users showed a mean social media usage time of 3.2 hours per day (SD = 1.8).

- High-risk users had a mean distraction score of 6.1 on a scale of 1–10 (SD = 2.4).

Permutation-based methods confirmed that the top five features accounted for over 60% of the model's decision power.

4.3 Comparative Analysis

- **Survey-Only Models:** Achieved an AUROC of 0.58.
- **Digital Trace Models:** Improved the AUROC to 0.65.
- **Combined Data Sources:** Yielded an AUROC of 0.73, demonstrating that integrating digital phenotypes with traditional survey measures significantly enhances predictive performance.

V. DISCUSSION

Our findings demonstrate that integrating traditional survey data with digital trace analysis from social media provides a robust method for predicting social media addiction risk. Key points include:

- **Digital Phenotyping Benefits:** The objective behavioural data (time usage, engagement patterns, affective signals) complement subjective self-report measures, leading to improved risk prediction.

- **Model/Generalizability:** With a large dataset from over 3,000 users and 400,000 posts, our model shows potential for broader application. However, since the sample is regionally concentrated, further validation across diverse populations is necessary.

- **Key/Predictors:** Time-related metrics and affective measures (e.g., irritability, distraction) emerged as significant predictors, supporting previous findings that both the quantity and quality of engagement are crucial to understanding addictive behaviours.

- **Limitations and Future Directions:** Limitations include potential bias due to class imbalance and challenges in capturing the full context of user behaviour. Future research should explore multi-modal fusion techniques, incorporate longitudinal tracking, and refine the interpretability of individual features.

VI. CONCLUSION

We developed and evaluated a web-based system using a hybrid approach—combining survey data with digital trace analysis—to predict social media addiction risk. Our system, built with modern web technologies and a Random Forest classifier, achieved an accuracy of 88.6% and an AUROC of 0.65. Combining traditional survey measures with automated digital phenotyping further increased performance to an AUROC of 0.73. These results validate the utility of integrating digital behavioural data into addiction risk assessments and open new avenues for targeted early interventions and further research into social media addiction.

ACKNOWLEDGEMENTS

We gratefully acknowledge our project guide, colleagues, and technical staff for their contributions during development, data collection, and analysis. We also thank [Funding Agency/Institution] for providing the necessary resources for this study.

REFERENCES

[1] Andreassen CS, Pallesen S, Griffiths MD. The relationship between addictive use of social media, narcissism, and self-esteem. *Addict Behave*. 2014.

[2] Lin CY, et al. Association between social media use and depression among young adults. *Depress Anxiety*. 2016.

[3] O'Keeffe GS, Clarke-Pearson K. The impact of social media on children, adolescents, and families. *Paediatrics*. 2011.

[4] SHENSA A, et al. Problematic social media use and depressive symptoms among U.S. young adults. *J Affect Discord*. 2017.

[5] Additional relevant literature on digital phenotyping and machine learning in behavioural health.