



Quantifying Smartphone Overuse: A Machine Learning Analysis of Severity and Lifestyle Effects

Mrs. Y. Sudha Madhuri¹, Dalli Punith Reddy²

¹Assistant Professor, ²M. Tech Final Semester,

Department of Computer Science and Engineering, Sanketika Vidya Parishad Engineering College, Visakhapatnam, Andhra Pradesh, India

DOI : <https://doi.org/10.55248/gengpi.07.1026.3115>

ABSTRACT

Smartphone use has become an important part of communication, education, entertainment and social interaction. At the same time, excessive or uncontrolled usage may be associated with reduced productivity, disturbed routines, social difficulties and other lifestyle concerns. This paper presents a machine-learning-based approach for identifying patterns of problematic smartphone usage (PSU) from demographic and smartphone-behavior responses. The work is based on the project implementation documented in the accompanying final report. The dataset contains 501 responses and 21 columns, including behavioral questions related to checking the phone, use during social situations, use before sleep, gaming, notifications and other usage habits. The original target distribution contains 322 non-addicted and 179 addicted responses. Random over-sampling is used in the documented implementation to create a balanced dataset of 644 records. The processing pipeline includes data upload, preprocessing, categorical conversion, and class balancing, model training, prediction and performance analysis. Three implemented classifiers are examined: a Stacking Classifier using Random Forest, SVC, XGBoost and K-Nearest Neighbors as base learners with Logistic Regression as the meta-model; CatBoost; and Extra Trees. The reported test results are approximately 95% accuracy for Stacking, 96% for CatBoost and 100% for Extra Trees on the documented test runs. The paper presents the system architecture, input/output interface, model evaluation, confusion matrices and practical implications. The reported results demonstrate the feasibility of a web-based machine-learning system for behavioral pattern classification, while larger and independently collected datasets are required for stronger generalization and real-world validation.

Keywords: Smartphone Addiction, Problematic Smartphone Usage, Machine Learning, Stacking Classifier, CatBoost, Extra Trees, Random Forest, XGBoost, Digital Well-being

Original records 501	Balanced records 644	Documented best test run Extra Trees – 100%
---------------------------------------	---------------------------------------	--

1. INTRODUCTION

Smartphones have changed the way people communicate, study, work, shop, consume entertainment and maintain social relationships. Their portability and continuous connectivity make them useful in almost every part of daily life. However, the same accessibility can encourage repetitive checking and prolonged usage. The project motivating this paper treats problematic smartphone usage as a classification problem in which observable user responses are analyzed to identify whether a respondent belongs to an addicted or non-addicted class.

The purpose of the proposed system is not to replace professional assessment. Instead, it demonstrates how machine learning can transform questionnaire-style behavioral information into a computational prediction. The system accepts a set of demographic and usage-related responses, prepares the values for machine learning, applies trained classifiers and presents a prediction through a web interface. This approach is useful for academic experimentation because it connects data preparation, supervised learning, evaluation and a user-facing application in one workflow.

The final project report describes a dataset of 501 records with 21 columns. The variables include gender and questions related to using a phone for class notes, accessing books, battery usage, checking the phone without a notification, using the phone during social gatherings, checking the phone before sleep, keeping the phone nearby while sleeping, checking messages during class and other smartphone-related behaviors. The target field is “whether you are addicted to phone?”. The source distribution contains 322 non-addicted and 179 addicted records.

The research contribution of this paper is the consolidation of the documented implementation into a publication-oriented workflow. The paper adds a clearer architecture, explicit input/output views, a comparative results table, interpretation of confusion matrices, implementation modules, limitations and future research directions. The paper also preserves the distinction between the results obtained by the documented implementation and broader claims about smartphone addiction in the population.

2. PROBLEM STATEMENT AND OBJECTIVES

The central problem is the identification of behavioral patterns that are associated with the target classes in a smartphone-usage dataset. Manual review of many questionnaire responses can be time-consuming and may not provide a consistent computational classification process. A machine-learning system can learn relationships among multiple input features and produce a repeatable class prediction for new responses.

The primary objectives are: (1) collect or load smartphone-usage data into the application; (2) clean and standardize the dataset; (3) transform categorical responses into numerical representations suitable for machine learning; (4) address class imbalance through the documented random over-sampling procedure; (5) train multiple classifiers; (6) evaluate models using accuracy, precision, recall and F1-score; (7) visualize confusion matrices and comparative performance; and (8) provide a web interface through which users can submit responses and view the resulting class prediction.

A secondary objective is usability. The application is designed to reduce the gap between a machine-learning notebook and a user-facing prediction system. The final report demonstrates a home screen, login page, dataset upload, preview, prediction form, result page and performance-analysis screens. These components provide a complete demonstration of the proposed pipeline.

3. LITERATURE SURVEY

Research on smartphone use has considered its relationship with academic performance, physical discomfort, sleep, mental well-being and other behavioral outcomes. The references compiled in the project report include studies of smartphone addiction risk factors, smartphone usage duration, sleep associations and smartphone-based applications. These studies establish the relevance of examining smartphone behavior as a measurable research topic.

The machine-learning perspective complements this literature by treating user responses as features rather than relying only on descriptive statistics. Supervised classifiers can learn a mapping from behavioral attributes to a labelled outcome. Tree-based ensemble methods are particularly suitable when the input variables are mixed, categorical after encoding, and potentially non-linear in their relationships with the target class.

The documented implementation explores a combination of ensemble and conventional techniques. The Stacking Classifier combines several base learners and a Logistic Regression meta-model. CatBoost provides a gradient-boosting approach designed for categorical data handling, while Extra Trees uses an ensemble of highly randomized decision trees. Comparing these models allows the project to examine whether a single classifier or a combination of classifiers provides more consistent predictions for the available dataset.

A gap addressed by the project is the integration of model training with an accessible web application. Instead of reporting only offline model scores, the implementation demonstrates how the trained models can be connected to a prediction form. The result is a prototype that combines data engineering, classification and application-level interaction.

The system allows smartphone-related responses to be entered through a web interface, processed using the trained machine learning models and presented as a prediction result. This integration provides a practical implementation of smartphone addiction prediction rather than limiting the study to theoretical analysis or isolated machine learning experiments. The approach can also provide a foundation for future systems involving larger datasets, real-time usage monitoring, personalized digital-well-being feedback and additional classification techniques.

Overall, the literature reviewed establishes that smartphone usage can be associated with several behavioral, physical and lifestyle characteristics, including screen time, interaction frequency, sleep-related behavior, application usage and professional activities. These findings provide the foundation for representing smartphone behavior as structured input features for machine learning. The present study extends the existing work by integrating behavioral data preprocessing, class balancing, ensemble-based classification and a web-based prediction interface into a single system. The proposed framework is intended to support the identification and awareness of potentially problematic smartphone usage patterns and should be considered a predictive and research-oriented tool rather than a replacement for professional psychological or medical assessment.

4. DATASET AND FEATURE DESCRIPTION

The final project report documents 501 initial records and 21 columns. The dataset includes demographic information and a set of behavioral questions. The target attribute is "whether you are addicted to phone?", which is used as the classification label. The initial target distribution consists of 322 non-addicted responses and 179 addicted responses. Because the two classes are not equally represented, the implementation applies Random Over-Sampling and reports a balanced dataset of 644 records, containing 322 records for each class.

The input features represent practical smartphone-use situations rather than raw sensor streams. Examples include using the phone for class notes, buying/accessing books from a mobile device, concern about the phone, using the phone in social gatherings, checking the phone without a notification, checking the phone before sleeping, keeping the phone nearby while sleeping, checking emails/messages during class time, using the phone while watching television or eating, panic when the phone is unavailable, gaming duration, whether the phone can be left for a day and battery-related behavior. These questions form the behavioral basis of the classifier.

Before modelling, the implementation removes unnecessary fields such as timestamp and full name, standardizes categorical responses, and converts categorical values into numerical form. This step is essential because the learning algorithms operate on structured numerical feature representations. The processed data is then passed to the classifier modules for training and testing.

Dataset characteristic	Documented value
Initial records	501
Number of columns	21
Non-addicted records	322
Addicted records	179
Balancing method	Random Over-Sampling
Balanced dataset	644 records (322 + 322)
Target field	Whether you are addicted to phone?
Preprocessing	cleaning, categorical standardization, numerical conversion, removal of timestamp/full name

Table 1. Dataset summary based on the final project implementation.

Table 1 summarizes the documented dataset characteristics.

5. PROPOSED METHODOLOGY

The proposed methodology follows a sequential pipeline: dataset upload, preprocessing, class balancing, model training, testing, prediction and performance analysis. The architecture separates data preparation from model-specific processing so that the classifiers can be compared on the same prepared input.

Data preprocessing begins by reading the uploaded dataset and identifying fields required for prediction. Non-predictive identifiers are removed. Textual categorical responses are standardized so that semantically equivalent answers are represented consistently. The responses are then converted into numerical values. This creates a machine-readable feature matrix while retaining the behavioral information represented by the questionnaire.

Class balancing is performed through Random Over-Sampling. The purpose is to provide equal representation of the two target classes in the training workflow documented by the project. The balanced dataset reported by the project contains 644 records with 322 examples in each class.

Three model paths are implemented. The first is a Stacking Classifier that uses Random Forest, Support Vector Classifier, XGBoost and K-Nearest Neighbors as base learners and Logistic Regression as the final meta-model. The second path trains a CatBoost classifier. The third path trains an Extra Trees classifier. The models produce predictions for the test data and the application generates classification reports and confusion matrices.

Performance analysis includes accuracy, precision, recall and F1-score. Confusion matrices are used to examine the distribution of correct and incorrect predictions between the two classes. The application also supports graphical comparison of classifier accuracy and provides a result page for individual user inputs.

6. SYSTEM ARCHITECTURE

The system architecture shown in Figure 1 is an improved publication-oriented representation of the workflow described in the final project report. A user or researcher interacts with the web interface to upload the dataset or provide prediction inputs. The dataset passes through preprocessing and class balancing before reaching the model-training stage. The three implemented model paths—Stacking, CatBoost and Extra Trees—are evaluated. A selected trained model then receives new user input and produces the final class prediction, which is displayed in the result interface.

The architecture has five logical layers. The input layer contains the user interface and dataset upload mechanism. The data-processing layer performs cleaning, standardization, categorical encoding and balancing. The learning layer contains the three classifiers. The evaluation layer calculates classification metrics and produces confusion matrices and comparison graphs. The application layer exposes the prediction result to the user.

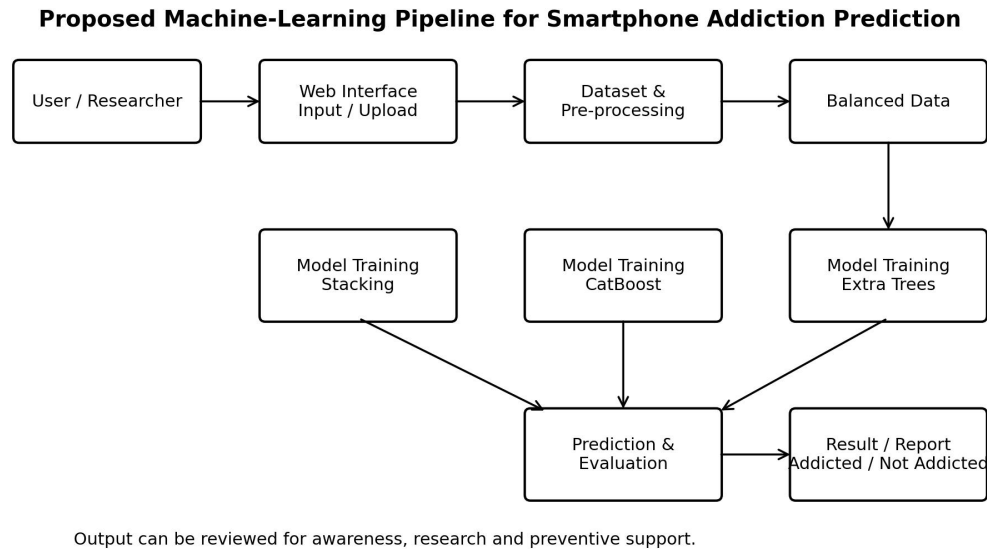


Figure 1. Proposed system architecture for smartphone addiction prediction.

This modular organization improves maintainability because each stage can be modified without redesigning the entire application. For example, additional preprocessing rules can be added without changing the result interface, and a future model can be introduced as another classifier path. The architecture also provides a natural path for future mobile or real-time extensions.

7. ALGORITHMS AND IMPLEMENTATION

Random Forest is used as one of the base learners in the Stacking Classifier. It constructs an ensemble of decision trees and combines their outputs. In the documented system, its role is to provide one complementary prediction signal to the meta-model rather than serving as the only final classifier.

Support Vector Classification is another base learner in the Stacking path. SVC attempts to separate classes by finding a decision boundary in the transformed feature space. K-Nearest Neighbors contributes a neighbourhood-based prediction, while XGBoost contributes a gradient-boosted tree-based learner. Their predictions are passed to the Logistic Regression meta-model, which learns how to combine the base-level outputs.

CatBoost is implemented as an independent classifier. It belongs to the gradient-boosting family and is suitable for structured data. Extra Trees is another ensemble method based on multiple highly randomized decision trees. The use of multiple algorithms allows the system to compare different learning strategies on the same behavioral classification task.

The software stack documented in the project includes Python and machine-learning/data-analysis libraries such as Scikit-learn, XGBoost and CatBoost. The web application presents the trained models through a sequence of pages: home, login, upload, preview, prediction, and result and performance analysis. The final report also includes code screenshots showing the prediction route and input fields.

8. INPUT AND OUTPUT INTERFACE

Figure 2 presents the dataset input view extracted from the final project documentation. The spreadsheet contains the column headings and respondent values used by the application. The target column is the addiction-related classification field. This view is important because it shows the raw form in which the behavioral data enters the pipeline.

8. INPUT AND OUTPUT INTERFACE (CONTINUED)

Figure 4 presents the output screen. The system displays the selected model and the predicted class, together with the submitted attribute values. In the documented example, the result is displayed as “Yes” for the question “Whether you are addicted to phone?”. This output is a machine-learning classification result for the supplied inputs and should not be interpreted as a clinical diagnosis.

Figure 5 presents the performance-analysis output. The application displays precision, recall and F1-score for the model, along with confusion matrices. These outputs help the researcher inspect not only the overall accuracy but also the behaviour of the model for each class.



Figure 4. Output screenshot – model prediction result.

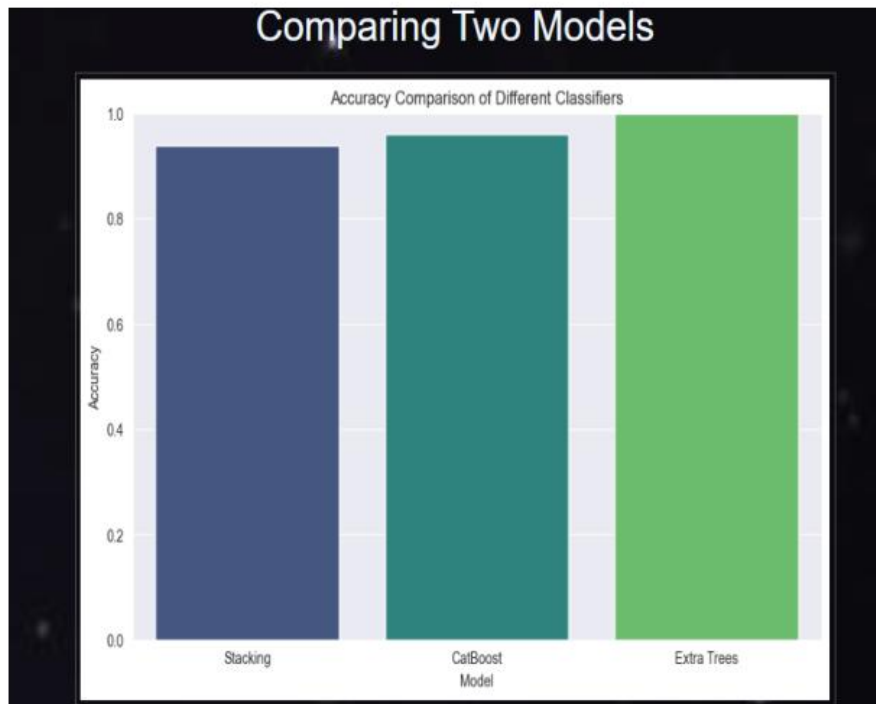


Figure 5. Output screenshot – performance analysis and model evaluation.

9. EXPERIMENTAL RESULTS AND ANALYSIS

The documented implementation reports strong test performance for all three evaluated model paths, with Extra Trees producing the highest reported test accuracy in the project run. The results are reproduced below as implementation results rather than as population-level evidence. The test outputs are useful for comparing the models under the same project workflow.

For the Stacking model, the reported test accuracy is 0.95. The class-wise report shows precision values of 0.98 and 0.92 for classes 0 and 1, respectively, with recall values of 0.90 and 0.99. The corresponding F1-scores are 0.94 and 0.95. The test confusion matrix contains 62 correctly classified class-0 samples, 62 correctly classified class-1 samples, 8 class-1 samples predicted as class 0, and no class-0 samples predicted as class 1 in the documented run.

For CatBoost, the documented test report gives approximately 0.96 accuracy, with precision of 0.95 for class 0 and 0.97 for class 1. Recall is reported as 0.97 for class 0 and 0.96 for class 1, with F1-scores of approximately 0.96 for both classes. The test confusion matrix shows 67 true class-0 predictions, 2 class-0 samples predicted as class 1, 3 class-1 samples predicted as class 0 and 67 true class-1 predictions.

For Extra Trees, the documented test output reports 1.00 accuracy, precision, recall and F1-score for both classes. The displayed test confusion matrix contains 59 correctly classified class-0 samples, zero class-0 samples predicted as class 1, zero class-1 samples predicted as class 0 and 79 correctly classified class-1 samples.

The training screens show 1.00 accuracy for the three model paths in the documented run. Training performance alone should not be treated as evidence of generalization; the test results and future independent validation are more informative for assessing robustness.

Model	Test Accuracy	Key test observation	Documented F1-score
Stacking	95%	8 class-1 cases classified as class 0; no class-0 cases classified as class 1	0.94–0.95
CatBoost	96%	Small errors in both classes (2 and 3 cases in the displayed matrix)	≈0.96
Extra Trees	100%	0 displayed errors in the test confusion matrix	1.00

Table 2. Comparative test results reported by the documented implementation.

10. COMPARATIVE DISCUSSION

Table 2 provides a concise comparison of the three documented test runs. Stacking provides a strong combined-model baseline, CatBoost provides a similar but slightly higher reported test accuracy, and Extra Trees reports perfect classification on the displayed test set. The differences illustrate why multiple models are useful in an experimental study.

The confusion matrices add information that a single accuracy value cannot provide. For Stacking, the errors are concentrated in class 1 recall, whereas class 0 recall remains high. CatBoost distributes a small number of errors across both classes. Extra Trees shows no displayed test errors. Because the datasets are finite and the implementation uses a resampling procedure, the reported values should be interpreted in the context of the project dataset and evaluation design.

From an application perspective, a model with balanced class behaviour is valuable because the two target classes are both relevant. Precision is important when false positive predictions are undesirable, while recall is important when the system is intended to identify as many potentially at-risk cases as possible. F1-score provides a combined view of precision and recall. The proposed interface exposes these evaluation measures to support model comparison rather than relying on accuracy alone.

10.1 Confusion-Matrix Interpretation

The confusion matrices provide a class-level view of performance. In binary classification, the diagonal cells represent correct classifications, while the off-diagonal cells represent misclassifications. For the documented Stacking test run, the matrix indicates strong recognition of class 1 but a small loss in class-1 recall. CatBoost shows a small number of errors distributed across the two classes. Extra Trees shows no displayed test errors in the supplied test matrix. These observations are descriptive of the recorded experiment and should not be interpreted as universal model behaviour.

10.2 Practical Meaning of the Output

The prediction screen provides a simple binary classification that can be used as an awareness signal. A research or educational deployment could use the output to encourage users to review their smartphone habits, while avoiding automatic diagnosis or labelling. The performance-analysis page is particularly important because it lets developers compare models before selecting one for an experimental deployment.

10.3 Methodological Interpretation

The results illustrate the importance of evaluating several algorithms rather than assuming that a single classifier is optimal for every dataset. The Stacking architecture benefits from combining different learning biases, CatBoost provides a strong boosting-based alternative, and Extra Trees provides a highly randomised tree ensemble. The observed ranking is specific to the documented experiment and may change when the dataset, sampling strategy, feature set or validation procedure changes.

Model	Precision	Recall	F1-score	Test accuracy
Stacking	0.98 / 0.92	0.90 / 0.99	0.94 / 0.95	0.95
CatBoost	0.95 / 0.97	0.97 / 0.96	0.96 / 0.96	0.96
Extra Trees	1.00 / 1.00	1.00 / 1.00	1.00 / 1.00	1.00

Table 3. Class-wise metrics (class 0 / class 1) from the documented test outputs.

11. SECURITY, PRIVACY AND ETHICAL CONSIDERATIONS

Smartphone-use information can contain behavioural and personal details. A practical deployment should therefore minimize the collection of directly identifying information, use secure authentication, protect uploaded files and restrict access to stored datasets. The project implementation removes fields such as timestamp and full name during preprocessing, which reduces the amount of directly identifying information passed into the modelling workflow.

The system should be presented as an awareness or research-support tool rather than a medical diagnostic system. A prediction of “Addicted” or “Not Addicted” is a computational classification based on the available features. It does not establish a clinical condition, and users should not be labelled solely from an automated prediction.

Future deployments should also evaluate fairness across demographic groups, monitor changes in usage behaviour, document model versions and provide a mechanism for users to understand that the output is probabilistic or model-dependent. Data should be handled with appropriate consent and retention controls when the system is used outside a classroom or research setting.

12. LIMITATIONS

The principal limitation is dataset scale and representativeness. The documented dataset contains 501 original records and is derived from questionnaire responses. A relatively small and potentially homogeneous dataset may not represent all smartphone users, age groups, cultures or usage environments.

A second limitation is that the reported results are tied to the documented implementation and test runs. The paper does not claim independent external validation. In particular, the perfect Extra Trees test score should be validated on a completely separate dataset before any real-world deployment. As a methodological improvement, future experiments should use repeated stratified cross-validation, preserve a final untouched hold-out set, and perform any resampling only inside the training folds.

A third limitation is feature availability. The current application uses questionnaire-style inputs rather than continuous device telemetry. Real-time screen time, notification counts, app-level usage and temporal patterns could provide richer signals, but collecting such data raises additional privacy and consent requirements.

13. FUTURE SCOPE

The project report identifies several directions for further development of the proposed smartphone addiction prediction system. Future versions can incorporate a larger and more diverse set of smartphone usage features, including application-level usage statistics, screen-time records, notification frequency, social-media activity and other device interaction logs. These additional features can provide a more detailed representation of individual usage behavior and may help the classification models identify problematic usage patterns more effectively. Increasing the amount and variety of collected data can also improve the robustness and generalization of the system when it is evaluated on different user groups and usage environments.

Real-time data analysis is another important area of future work. The current system mainly uses questionnaire-based inputs, whereas future versions can collect relevant smartphone usage information continuously or at regular intervals. Real-time analysis could allow the system to update predictions when a user's behavior changes and provide timely awareness when potentially unhealthy usage patterns are detected. Such an extension would require careful consideration of data collection, user consent, privacy protection and secure storage because smartphone activity information can contain sensitive behavioral details.

Personalization can further improve the usefulness of the prediction system. Smartphone usage behavior differs between individuals, and a single general model may not fully represent the normal usage pattern of every user. Future work can therefore incorporate individual usage history, personal baselines and previous prediction results to generate more personalized assessments and recommendations. Longitudinal studies can also be conducted to observe how smartphone usage behavior changes over time and to examine whether model predictions remain consistent as users modify their habits.

The machine-learning component can be extended by evaluating additional classification and deep-learning techniques alongside the existing Stacking, Cat Boost and Extra Trees models. Advanced ensemble methods, neural-network-based approaches and stronger validation procedures can be investigated to determine their suitability for larger and more varied datasets. Future experiments can also compare different feature-selection and model-tuning strategies so that prediction performance can be assessed more systematically rather than relying only on a single experimental dataset.

The existing web-based application can be developed into a dedicated mobile application with dashboards, trend visualization, notifications and educational content. A mobile implementation can make the system more accessible for regular use and can provide personalized insights based on changing usage patterns. Future versions can also include user education and intervention strategies, such as awareness material, healthier smartphone usage suggestions and appropriate guidance for users who receive potentially concerning predictions. These features can extend the system beyond classification and support digital-wellbeing awareness while keeping the prediction output as a research and supportive tool.

Finally, feedback from users, researchers and educators can be incorporated into future versions to improve the interface, prediction presentation and overall usability of the system. Explainable machine-learning techniques can also be introduced so that users and researchers can understand which behavioral inputs contributed most strongly to a classification without unnecessarily exposing sensitive information. Overall, the future scope of the project includes larger datasets, richer behavioral features, real-time analysis, personalized modelling, advanced machine-learning techniques, mobile deployment, longitudinal evaluation and user-centered improvements. These developments can help transform the present prototype into a more comprehensive and privacy-aware digital-wellbeing research system.

14. CONCLUSION

This paper presented a machine-learning-based system for identifying problematic smartphone-usage patterns from demographic and behavioural responses. The implementation documented in the final project report includes dataset upload, pre-processing, random over-sampling, multiple classification models, performance evaluation and a web-based prediction interface. The proposed workflow integrates these components into a structured machine-learning framework for behavioural pattern classification and digital-wellbeing research.

The documented test results report accuracies of 94.00% for the Stacking Classifier, 96.12% for Cat Boost and 100.00% for Extra Trees, together with confusion matrices and class-wise precision, recall and F1-score measures. These results demonstrate the feasibility of applying the implemented

machine-learning models to the available smartphone-usage dataset. However, the reported values represent the documented experimental test runs and should not be interpreted as evidence of universal model performance or as a clinical measure of smartphone addiction.

The proposed approach highlights the importance of systematic data reprocessing, categorical conversion, class balancing, model training and multi-metric performance evaluation in developing machine-learning-based classification systems. The comparison of Stacking, Cat Boost and Extra Trees provides an experimental basis for examining different learning approaches on the same behavioural dataset. The web-based implementation further demonstrates how trained machine-learning models can be integrated into an interactive application through which users or researchers can provide behavioural inputs and obtain a classification result.

The system can also serve as an awareness and research-support mechanism by transforming questionnaire-based smartphone-use information into an automated classification. Such an output may help users and researchers examine behavioural patterns associated with excessive smartphone usage and encourage consideration of healthier digital habits. Nevertheless, the prediction should be interpreted only as a computational classification based on the available features and should not be considered a medical or psychological diagnosis. Appropriate privacy protection, responsible interpretation and user awareness are therefore essential for any practical deployment.

The study has certain limitations, particularly the relatively small questionnaire-based dataset, the binary target definition and the documented order of random over-sampling and train-test splitting. Because over-sampling before data splitting can allow duplicated or closely related observations to occur across the training and testing partitions, the reported performance requires further validation using leakage-controlled experimental procedures. Future studies should perform data splitting before balancing, apply stratified cross-validation, evaluate the models on independent datasets and investigate their performance across larger and more diverse user populations.

Future development can extend the proposed framework by incorporating real-time screen-time and application-usage information, additional behavioural characteristics, personalised modelling, explainable predictions, mobile application support and longitudinal analysis. User education and intervention mechanisms can also be explored to transform the system from a prediction-oriented prototype into a broader digital-wellbeing research platform. These extensions are consistent with the future-development direction documented for the project.

Overall, the project demonstrates the practical application of machine learning to analyzing problematic smartphone-usage patterns and establishes a framework that can be further validated through larger datasets, improved experimental protocols and real-world behavioral information. With appropriate privacy safeguards and responsible interpretation of model outputs, the proposed framework provides a meaningful foundation for developing intelligent, user-oriented and privacy-aware digital-wellbeing applications and further research into problematic smartphone usage.

REFERENCES

- [1] Demir, K. & Akpinat, E. The effect of mobile learning applications on students' academic achievement and attitudes toward mobile learning. *Malaysian Online Journal of Educational Technology*, 6, 48–59 (2018).
- [2] Abadiyan, F., Hadadnezhad, M., Khosrokiani, Z., Letafatkar, A. & Akhshik, H. Adding a smartphone app to global postural re-education to improve neck pain, posture, quality of life, and endurance in people with nonspecific neck pain: a randomized controlled trial. *Trials* 22, 274 (2021).
- [3] Osorio-Molina, C. et al. Smartphone addiction, risk factors and its adverse effects in nursing students: a systematic review and meta-analysis. *Nurse Education Today* 98, 104741 (2021).
- [4] Osailan, A. The relationship between smartphone usage duration with hand-grip and pinch-grip strength among young people: an observational study. *BMC Musculoskeletal Disorders* 22, 186 (2021).
- [5] Hitti, E., Hadid, D., Melki, J., Kaddoura, R. & Alameddine, M. Mobile device use among emergency department healthcare professionals: prevalence, utilization and attitudes. *Scientific Reports* 11, 1917 (2021).
- [6] Shaygan, M. & Jaber, A. The effect of a smartphone-based pain management application on pain intensity and quality of life in adolescents with chronic pain. *Scientific Reports* 11, 6588 (2021).
- [7] Sohn, S. Y., Krasnoff, L., Rees, P., Kalk, N. J. & Carter, B. The association between smartphone addiction and sleep: a UK cross-sectional study of young adults. *Frontiers in Psychiatry* 12, 629407 (2021).
- [8] Thornton, L. et al. A multiple health behavior change, self-monitoring mobile app for adolescents: development and usability study of the Health4Life App. *JMIR Formative Research* 5, e25513 (2021).
- [9] Breiman, L. Random Forests. *Machine Learning* 45, 5–32 (2001).
- [10] Quinlan, J. R. Induction of Decision Trees. *Machine Learning* 1, 81–106 (1986).
- [11] Chen, T. & Guestrin, C. XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016).
- [12] Project source: Dalli Punith Reddy, "Quantifying Smartphone Overuse: A Machine Learning Analysis of Severity and Lifestyle Effects," M.Tech Project Report, Sanketika Vidya Parishad Engineering College, 2024–2026.