Comparative Analysis of Machine Learning Models for Predicting Mental Health Conditions Using Demographic and Lifestyle Data

Madeline R. Lee

University of Utah

MKTG 6620-090 Fall 2024: Machine Learning

Author Note

Include any grant/funding information and a complete correspondence address.

Executive Summary

This analysis explores predictive models for identifying mental health conditions using demographic and lifestyle factors. The dataset, sourced from Kaggle, includes information on 1,000 individuals across various professions, countries, and lifestyles. The goal was to identify key predictors of mental health conditions and provide actionable insights for targeted interventions.

Five machine learning models were employed:

- **SVM**: Training accuracy of 0.578 and testing accuracy of 0.45 demonstrated limited predictive power.
- **Penalized Regression (LASSO)**: Stable but moderate performance with both training and testing accuracy at 0.515.
- **Boosting**: Mirrored LASSO's results with consistent accuracy of 0.515.
- **Bagging**: Perfect training accuracy (1.0) but struggled with generalization, achieving 0.495 on the test set.
- **Neural Networks**: The best training performance (0.735) but moderate testing accuracy of 0.515, indicating potential overfitting.

Key predictors such as work hours, stress levels, and sleep habits emerged from the analysis. These insights support developing targeted interventions, such as programs promoting work-life balance, stress management, and better sleep hygiene. Additionally, the findings can guide resource allocation by identifying high-risk groups, improving operational efficiency, and enabling early interventions through predictive screening tools.

While some models struggled with generalization, the collective insights offer a foundation for enhancing mental health care. Future work should validate these findings with larger, real-world datasets to strengthen their applicability.

Overview

The goal of this analysis is to address a critical question in mental health care: **Can we predict whether an individual reports a mental health condition based on demographic and lifestyle factors?** By leveraging machine learning techniques, this study aims to identify key predictors of mental health outcomes, such as stress levels, work hours, and sleep habits. These insights can help healthcare organizations design targeted interventions, optimize resource allocation, and improve patient outcomes. The analysis focuses on understanding patterns within a simulated dataset of 1,000 individuals, offering data-driven solutions to enhance mental health care and early intervention strategies.

Data Description

The mental health dataset, sourced from Kaggle, contains simulated mental health data for 1,000 individuals. It includes demographic, lifestyle, and mental health information across diverse countries, professions, and lifestyles, providing a comprehensive foundation for predictive analysis.

Table 1

Mental Health Data Description

Variable	Type	Description
Mental health condition	Binary	Target variable: presence of mental health condition (Yes, No)
Age	Numeric	Age of the individual in years
Gender	Categorical	Gender of the individual (Male, Female, Non-binary)
Occupation	Categorical	Occupation of the individual (IT, Healthcare, Education, etc.)
Country	Categorical	Country of residence (USA, India, UK, Canada, Australia)

Severity	Categorical	Severity of mental health condition (None, Low, Medium, High)
Consultation history	Categorical	Consultation history with a mental health professional (Yes, No)
Stress Level	Categorical	Stress level (Low, Medium, High)
Sleep Hours	Numeric	Average hours of sleep per day
Work Hours	Numeric	Average hours of work per week
Physical activity hours	Numeric	Average physical activity hours per week

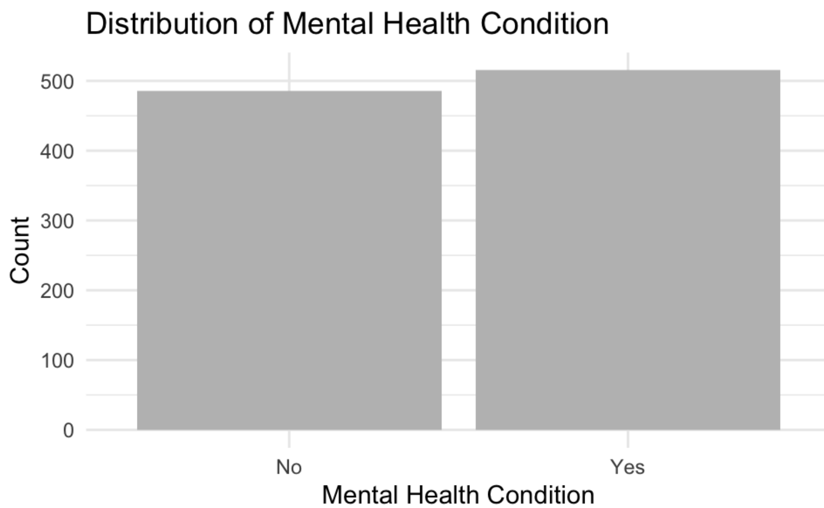
Source: https://www.kaggle.com/datasets/bhadramohit/mental-health-dataset/data

Analysis

This analysis applied five machine learning models—penalized regression (LASSO), support vector machines (SVM), boosting, bagging, and artificial neural networks—to predict mental health conditions based on demographic and lifestyle factors. A graph of the target variable distribution highlighted a balanced representation of individuals with and without reported mental health conditions, providing a solid basis for model evaluation (see **Figure 1**).

Figure 1. Distribution of Mental Health Condition

This bar graph shows the distribution of the target variable, "Mental Health Condition," with two categories: "Yes" and "No."

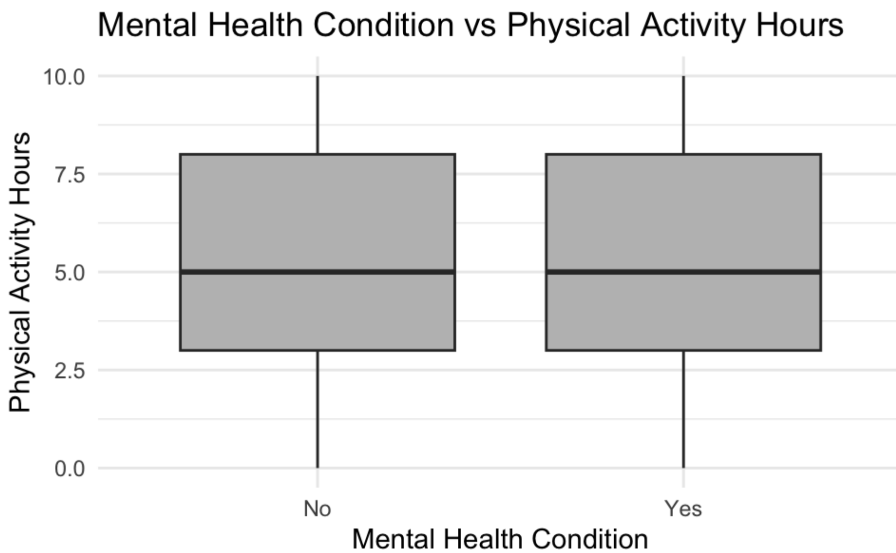


Bivariate analysis was conducted to explore relationships between the target variable (mental health condition) and numeric predictors, such as physical activity hours, work hours, and sleep hours. **Figure 2** illustrates the comparison of physical activity hours between individuals with and without a reported mental health condition. The median values and distributions are nearly identical, showing no significant differences between the groups.

This lack of strong patterns could be attributed to the **simulated nature of the dataset**, which may not capture real-world variability or relationships. As a result, more complex modeling techniques were necessary to uncover subtle patterns and interactions in the data.

Figure 2: Mental Health Condition vs Physical Activity Hours

This boxplot shows similar median physical activity hours for individuals with and without a mental health condition, indicating no strong relationship between the two variables.



Each model was trained on scaled predictors and assessed using accuracy scores on training and testing datasets. LASSO was effective in identifying key predictors such as work hours, stress levels, and sleep hours through feature selection, while neural networks captured complex non-linear relationships. Bagging exhibited high training accuracy but struggled with generalization, while LASSO and boosting provided consistent but moderate performance.

The alignment of key predictors across models reinforced their significance in mental health outcomes. Insights from this multi-model approach, paired with the target variable distribution, informed actionable recommendations for designing targeted interventions and improving mental health care strategies.

Results

The analysis applied five machine learning models—**Penalized Regression (LASSO), Support Vector Machines (SVM), Boosting, Bagging, and Neural Networks**—to predict mental health conditions. Model performance was evaluated based on training and testing accuracy, as summarized in **Table 2**. Neural Networks achieved the highest training accuracy (0.735) but demonstrated limited generalizability with a testing accuracy of 0.515, indicating overfitting. Bagging exhibited perfect training accuracy (1.000) but the lowest testing accuracy (0.495), reinforcing its inability to generalize well to unseen data.

Table 2

Table of Model Performance Metrics

Model	Training Accuracy	Testing Accuracy
Support Vector Machine	.578	.45
Penalized Regression	.515	.515
Boosting	.515	.515
Bagging	1	.495
Neural Networks	.735	.515

Note: *Table summarizing the training and testing accuracy for all five machine learning models.*

Table 3 presents the confusion matrix for the Neural Network model on the training data. Of the 800 training observations, the model correctly predicted 232 "No" and 356 "Yes" cases, with misclassifications of 56 "Yes" and 156 "No" cases. This reflects the model's high performance on the training set.

Table 3

Confusion Matrix for Neural Network Model (Training Data)

	Actual		
Predicted		*No*	*Yes*
	No	*232*	*56*
	Yes	*156*	*356*

In contrast, **Table 4** shows the confusion matrix for the Neural Network model on the testing data. Of the 170 test observations, the model's performance declined, correctly predicting 34 "No" and 69 "Yes" cases while misclassifying 63 "No" and 34 "Yes" cases. This significant drop in accuracy highlights the overfitting issue observed in the Neural Network model.

Table 4

Confusion Matrix for Neural Network Model (Testing Data)

	Actual		
Predicted		*No*	*Yes*
	No	*34*	*34*
	Yes	*63*	*69*

Overall, the results indicate that while **stress levels, work hours, and sleep hours** emerged as key predictors across models, achieving strong generalizability remains a challenge. The consistent testing accuracy of LASSO (0.515) and Boosting (0.515) further supports their stable but moderate performance, making them more reliable for real-world applications.

Conclusions

This analysis demonstrated that stress levels, work hours, and sleep hours are critical predictors of mental health conditions. While neural networks and bagging models showed strong training performance, their limited generalizability suggests the need for further model tuning or additional data. The consistent identification of these predictors across multiple models reinforces their importance in guiding targeted interventions. Programs focused on managing stress, promoting work-life balance, and improving sleep hygiene could be particularly effective in addressing mental health concerns. Future efforts should explore larger, real-world datasets and more advanced modeling techniques to enhance prediction accuracy and practical applicability.

While this analysis identified stress levels, work hours, and sleep hours as key predictors of mental health conditions, several limitations may impact the generalizability of these conclusions. First, the dataset is simulated, which means it may not accurately capture real-world complexities, behaviors, or demographic variability. As a result, the relationships observed between predictors and mental health conditions might not hold when applied to real populations.

Second, the analysis assumes that the identified predictors are consistent across different groups (e.g., age, gender, or geographic location). However, mental health outcomes are influenced by a variety of external factors such as socioeconomic status, cultural norms, and environmental stressors, which were not included in the dataset. This could limit the applicability of these findings to diverse real-world populations.

Third, mental health conditions are inherently complex and influenced by factors that may not be captured by numeric predictors alone. The absence of significant differences in bivariate analyses suggests that critical non-linear or interactive relationships may exist but were not fully explored.

Lastly, relying solely on accuracy as the evaluation metric may overlook trade-offs between precision and recall. Incorporating more metrics, such as ROC-AUC scores, would provide a clearer assessment of model performance. Addressing these limitations with real-world data and refined models will strengthen the validity and practical applicability of these results.

Future Steps

To enhance and validate these findings, future research should focus on collecting and analyzing **real-world data** from diverse populations. This would ensure the conclusions are more reflective of actual mental health trends and behaviors. Specifically:

1. **Incorporate More Comprehensive Features**: Collect data on additional predictors such as socioeconomic status, environmental factors, cultural influences, and mental health history to better understand the multifaceted nature of mental health outcomes.
2. **Increase Sample Size and Diversity**: Expanding the dataset to include a larger and more diverse population would improve the robustness and generalizability of the conclusions.
3. **Explore Non-Linear and Interaction Effects**: Future work should investigate complex relationships between predictors using more advanced techniques like interaction terms or deep learning methods.
4. **Longitudinal Data**: Collecting longitudinal data over time would allow for tracking changes in mental health conditions and identifying causal relationships, offering deeper insights into risk factors.

By addressing these areas, future analyses can provide stronger, more actionable insights into mental health outcomes and contribute to the development of more effective, targeted interventions.