Reinforcement Learning for Personalized Digital Well Being

Student Tejaswi Kondaveeti
Qualification Fourth Year Data Science Student
University IISER Thiruvananthapuram
Country India

Introduction

In the digital age, social media has an integral part of the lives of every individual and especially young adults and teenagers. While platforms like Instagram, Youtube, and Whatsapp offer unprecedented avenues for connection, expression, and information sharing. They also contribute to rising concerns about Mental health in teenagers and young adults. This is majorly caused due to the increasing screen time, social comparison culture, cyberbullying, and information overload. These factors are directly linked to anxiety, depression, and loneliness in adolescents.

This study aims to explore the impact of social media usage patterns on the mental Health of young individuals through a data-driven approach. By analyzing survey responses and behavioral patterns, we assess users based on their Mental Health Score (MHS), and provide personalized suggestion to improve their well-being. To enhance the personalization and effectiveness of the recommendations, we applied Reinforcement Learning (RL) techniques and we have implemented Deep Q-Learning (DQN) and Advantage Actor Critic (A2C) algorithms. These models were trained to suggest mental health interventions based on users' unique characteristics and behavioral patterns, simulating the process of intelligent decision-making.

The objective was to understand the psychological effects of social media but also to develop a metric that uses AI to assist in promoting better mental health. The action suggestions fed to the model are habits that are renowned for creating a positive mindset and improving mental health. This work stands at the intersection of psychology, data science, and artificial intelligence, offering a novel approach to addressing a pressing issue affecting today's youth. The dataset that is currently available comprises 128 instances and very limited actions; this is the future scope of the study to increase the state space by collecting more data from individuals and improving model performance in providing personalized suggestions.

Data Collection

The foundation of this study is in the systematic data collection from the users. The survey covers various dimensions such as demographic information, social media usage patterns, mental health and emotional well-being, and a few custom questions about their hobbies. The survey carefully structured the individual's online behaviour and psychological state. The questionnaire included both objective and subjective questions, the following describes the major sections:

1. Demographic Information
    - Age group
    - Gender
    - Educational Background
2. Social Media Usage Patterns
    - Daily screen time on different platforms like Instagram, Whatsapp, Youtube, Snapchat, Twitter, etc.
    - Frequency of checking social media.
    - Type of content consumed such as educational, entertainment, lifestyle, etc.
    - Participation in online interactions like posting, commenting and liking.

3. Mental Health and Emotional Wellbeing
    - Self-reported levels of anxiety, stress, or low mood.
    - Sleep quality and duration
    - Impact of online interactions on self-esteem
    - level of social comparison experienced.
    - Frequency of feeling isolated despite online connectivity.

The survey was distributed digitally via Social Media across university student community. The responses were collected for a span of two weeks. The toal of 128 participants responded to the survey. A question seeking the consent was included in the survey. To maintain the privacy of the individuals, no personal information, including email, was collected, and participation in the survey was entirely voluntary.

Mental Health Scoring (MHS)

To quantify the mental health status of individuals based on their social media usage, a custom Mental Health Score (MHS) system was developed. This score represents an approximation of the user's mental health state, based on emotional feedback, platform behavior, and usage patterns.

Emotion Based Weighting

Users were asked to reflect on emotional states experience due to or during social media use. The following emotions were considered, with assigned weights based on their potential impact on mental well-being:

- **Happiness (10):** Positively contributes to mental health.
- **Inspiration (8):** Positive reinforcement and motivation.
- **Loneliness (5):** Mild negative influence.
- **Jealousy (4):** Moderate negative emotional response.
- **Anxiety (3):** Strongly associated with negative mental health effects.

The Higher the scores, the more positive the associated emotions. Positive emotions are defined as an increase in the MHS. The emotional dimension was a core component in assessing mental balance.

Let $r[f]$ be the value (response) for feature f in the user's data, and $w[f]$ be the weight assigned to feature f.

$$\text{Emotional Score} = \sum_{f \in \text{Emotions}} r[f] \cdot w[f]$$

Time spent on Social Media

' The time spent on social media was factored into the MHS with a weight of 2. Excessive usage correlates with a negative effect on mental health, as it is the second most likely score. It promotes the reduction of the MHS.

$$\text{Time of Use Score} = \sum_{f \in \text{TimeOfUse}} r[f] \cdot w[f]$$

Usage Weighting

Different platforms impact users differently based on their design and utility. Each majorly used platform was assigned with a weight reflecting its influence. The weights of the platform are as follows:

- **Instagram**: 4
- **Whatsapp**: 6
- **Youtube**: 9

- **Twitter**: 5
- **Facebook**: 5
- **LinkedIn**: 10
- **Reddit**: 7
- **Snapchat**: 4
- **Discord**: 7

$$\text{Platform Score} = \sum_{f \in \text{Platforms}} r[f] \cdot w[f]$$

The weights were assigned according to the usage pattern, which includes the purpose of the social media platform that is being used. Watching education videos has the highest weight of 9, and scrolling has the lowest weight of 2.

$$\text{Usage Pattern Score} = \sum_{f \in \text{UsagePatterns}} r[f] \cdot w[f]$$

Similarly, weights were assigned based on the time of day. Morning has the least score of 2, as starting the day with phone and social media is an unhealthy habit and the Evening having a higher score as it is the time for relaxing and chilling.

Tf-IDF Based Sentiment Analysis

Custom text responses were collected from the users, the answers to these questions needs to be given a neutral a less important weightage as the responses weren't well defined or standard. It was given a weight of 2.

$$\text{TF-IDF Score} = \left(\sum_{i=1}^{n} \text{TFIDF}_i \right) \cdot w_{\text{TFIDF}}, \quad \text{for } n \text{ open-ended features}$$

Final Scoring

The final MHS was computed as a weighted aggregation of the above mentioned individual weights, resulting in a numerical score. This score was used to assess and recommend personalized suggestions for the users by using Reinforcement learning models (DQN and A2C). The individual scores and the MHS were used to track patterns between the usage behavior and emotional health.

$$\text{MHS} = \sum_{g \in G} \sum_{f \in g} r[f] \cdot w[f] + \text{TF-IDF Score}$$

- Let G be the set of all feature groups:

$$G = \{\text{Emotions}, \text{TimeOfUse}, \text{Platforms}, \text{UsagePatterns}, \text{Impacts}, \text{AdditionalFeatures}\}$$

- Let $r[f]$ be the user's response value for a given feature $f \in G$
- Let $w[f]$ be the assigned weight for the feature f

Reinforcement Learning Environment

The next step is the environment setup and implementation of RL on the MHS data frame. This work presents a custom OpenAI Gym environment designed to simulate how different lifestyles and behavioural actions impact various dimensions of an individual's mental health. The environment serves as a framework for training RL agents to suggest optimal interventions for mental well-being. The dataset prepared in the previous steps contains user

Table 1: Mental Health Score (MHS) Breakdown

Emotional Score	Time Score	Platform Score	Usage Score	Impact Score	Additional Score	TF-IDF Score	MHS
84	13	20	18	27	38	0	200
56	13	10	2	0	35	0	116
97	23	19	18	61	34	0	252
114	13	19	11	54	63	0	274
87	13	19	18	25	10	0	172

behaviour analytics. Each row represents a user's mental health state in seven dimensions, detailed in Table 1. Each of these scores ranges from 0 to 300, providing a quantitative foundation for simulating user states.

The suggestions for the users were defined with weight modifications of scores they would offer if implemented by the user. There were seven available actions from which the model could select one for the user based on their current state. The actions available are:

- Engage in Outdoor Activities
- Reduce Social Media Usage
- Watch Educational Videos
- Increase Mindfulness Practices like Yoga
- Interact with positive communities
- Avoid Stress-Inducing content
- Follow a Structured Daily Routine

State Initialization: The initial state is the current state of the user, calculated based on the scores of the user. This state is reset after evaluation of one individual.

Action effects: Each action that was defined earlier has predefined effects on the individual scores which inturn contribute to the overall mental health score of the individual. Each action contributes to the effect on the seven dimensions present in the Mental Health Score dataframe. These actions are relevant to the real world and are expected that the user makes these changes in their life according to the suggestions. The inclusion of these habits in the users life contributes to the change of the MHS in the next round of monitoring the mental health. These effects are designed to simulate the real-world influence of these actions.

Reward Function: The reward function, in our case, is the increase in the MHS of the individuals. According to the current state of assigned weights and actions, the maximum MHS can go up to 300. The goal is to suggest an action suitable to the user, which increases the current state MHS to a better score.

Model Implementation and Evaluation

In the model implementation, two RL agents were set up: **Deep Q-network** and **Advantage Actor-Critic**.

Deep Q-Network

DQN is a value-based reinforcement learning algorithm. It estimates the Q-value for each action in a given state. It essentially learns how beneficial each action is for the user's mental well-being. The agent interacts with the RL environment and learns from the reward feedback, and updates its policy. DQN is particularly effective in discrete action spaces, like our case with 7 fixed action spaces.

Advantage Actor Critic

A2C is a hybrid algorithm that combines policy-based and value-based methods. It maintains an actor, that decides the action to be taken and a critic, that evaluates the value of the state. This dual nature helps in more stable and efficient learning. A2C is useful when we want smoother learning and faster convergence in complex environments.

Both the models are trained on the RL environment and saved for future use. Each action corresponds to a mental health intervention and contributes to the increase on MHS. The reward function in our environment likely measures the improvement in the user's mental health state due to an action. The trained models are run through each user, limiting to a single user per episode to simulate single suggestion per user. It collects key metrics such as *total rewards*, which indicate the effectiveness of the actions taken; *action counts*, showing how frequently each intervention is selected; and *user suggestions*, a log of personalized recommendations made for each user. These outputs not only help in comparing model performance but also in drawing meaningful, user-centric insights. For instance, *user suggestions* can reveal what specific strategies the model recommends to different individuals, while action counts highlight which actions are generally favoured or potentially more effective across the population. This evaluation framework effectively bridges AI-based decision-making with real-world therapeutic relevance, allowing for the development of personalized and data-driven mental health support strategies.

Conclusion

This project successfully demonstrates how reinforcement learning can be applied to the domain of mental health recommendation systems. By modeling the problem as a custom Gym environment, we simulate real-world scenarios where an agent learns to suggest beneficial interventions based on a user's mental health profile. Through training with algorithms like DQN and A2C, the models learn to optimize well-being by selecting contextually relevant actions. The evaluation results based on metrics such as total rewards, action frequency, and user-specific suggestions highlight the potential of these models to generate personalized mental health recommendations. Overall, this work presents a promising step toward AI-driven therapeutic support systems that are adaptive, interpretable, and impactful. Future work can extend this framework by incorporating multi-step decision-making, richer user profiles, and integration with real-time feedback to make the system even more robust and practical for real-world applications.