UNDERSTANDING AND MODELING HUMAN BEHAVIOR IN PREPARING FOR QUITTING SMOKING. (EFFECTIVENESS OF REINFORCEMENT LEARNING)

GHIYATH ALASWAD
CSE3000 RESEARCH PROJECT

OBJECTIVE

- Design a Reinforcement learning model to decide whether to involve a human expert in the conversation and for which user
- Analyze the effectiveness of the model on the long-term change in smoking/vaping behavior and identity.

RESEARCH QUESTIONS:

Research Question:

1. How effective is a reinforcement learning model that optimizes the change in smoking/vaping-related behavior and identity?
   a. Defining the reinforcement learning model (e.g., state space, action space, reward function, discount factor).
   b. Analyzing the (long-term) effectiveness of the model components based on the provided data

Sub-questions:

- How can the state space be reduced to improve the accuracy of the result, which states to consider as a part of the model, and how to choose these states?
- What is the optimal reward function, considering the cost of human feedback and the limitations of availability?
- How effective is the model, and how reliable are the predictions it proposes compared to other algorithms?

RL MODEL COMPONENTS:

state space: (G-algorithm is used to select the most relevant states(Q1))

1. state energy.
2. state human support.
3. state importance.

actions space:
Provide/Don't provide human feedback

reward function:
combined sparse reward function based on:

1. Smoking/vaping frequency (50%).
2. change in user identity score (50%).
3. cost factor as penality (0.1)

discount factor:
The discount factor is set at 0.85, reflecting a value close to one, which emphasizes a greater focus on long-term changes.

ANALYSIS

Reward Function Analysis (Q2):
We explored different ways to design the reward function:

- Behavioral Weights: We tried several weight distributions for different behaviors and found that giving equal weight (50/50) to each behavior worked best.
- Cost Factor Values: We experimented with various values for the cost factor to see how they affected the outcomes.
- Improvement Rewards: We tested two methods for rewarding improvements: one that gave rewards based on the percentage of improvement and another that gave a fixed reward for reaching a goal. In the end, we found no significant difference between these two methods.

To assess the accuracy of the model's predictions, one user was excluded from the training data. The model was then trained on the remaining data, and predictions were made for the excluded user. The predicted outcomes were compared to the actual values to calculate the reward and evaluate the model's performance(Q3)

CONCLUSION

Finding the perfect solution for complex problems is always challenging, and there's always room for improvement by trying different methods and combinations. For this project, given the time constraints, we couldn't achieve a highly accurate model. However, with more resources like a larger dataset and better implementations, we could build a stronger model.

Additionally, the AI field is evolving rapidly, and new algorithms are constantly being developed. These advancements could offer even better solutions to problems like this in the future.