ENTHUSIASTIC
INTERNATIONAL JOURNAL OF STATISTICS AND DATA SCIENCE
Volume 2, Issue 2, October 2022, pp. 82-96

AI-Based Personalized Virtual Therapist for Alcohol Relapse

Myat Noe Win[a,1], Lim Whei Han [a,2], Eric Kamalesh Kumar Samson Chandresh Kumar [a,3], Tan Yong Keat [a,4], and Sri Devi Ravana [a,5*]

[a] Faculty of Computer Science and Information Technology, University Malaya, Jln Profesor Diraja Ungku Aziz, 50603 Kuala Lumpur, Wilayah Persekutuan Kuala Lumpur, Malaysia
[1] myatnoewin786@gmail.com, [2] han9.2000@gmail.com, [3] erickk2805@gmail.com, [4] wie180044@siswa.um.edu.my, [5] sdevi@um.edu.my*
* corresponding author

ARTICLE INFO

Article history
Received August 10, 2022
Revised September 11, 2022
Accepted October 2, 2022

Keywords
Artificial intelligence
Alcohol addiction
Relapse
Prediction models
Virtual therapist

ABSTRACT

Binge drinking is one type of harmful alcohol use that has a variety of negative health impacts in both the drinker and others, either globally or in Malaysia. According to previous research, one in two current drinkers in Malaysia who are 13 years and older reported having engaged in binge drinking. Therefore, increased attention should be given to understand the drinking pattern of an individual and propose a solution that can help with addiction relapse. Thus, this study identified interventions that could assist alcohol relapse recovery and proposed a new generation of relapse prevention solution based on artificial intelligence (AI). By using a deep learning approach and machine learning based recommendation technique, it predicts the relapse rate of users, providing recovery consultation based on the user's data and clinical data through a chatbot. This study involved helpful data collection, advanced data modeling, prediction analysis to support the alcohol relapse recovery journey. Hence, the proposed AI solution acted as a personalized virtual therapist to help the addicts stay sober. The objective is to present the design and realization of the AI based solution for sober journey. The proposed solution was tested with pilot study and significant benefits of virtual therapists for alcohol addiction relapse is reported in this paper.

1. Introduction

According to WHO statistic, there are 3 million deaths from harmful use of alcohol globally every year, with 11.8% of Malaysia suffering binge drinking [1]. Malaysians have gradually increased their alcohol intake throughout the years. According to the public health agency, nearly half of Malaysians binge drink, which is roughly double the rate in countries like the United Kingdom (UK) and Indonesia [2]. Binge drinking is defined as consuming six or more drinks in a single sitting. Malaysians' drinking habits appear to be highly concerning when compared to other countries, especially given other countries have lower binge drinking patterns. Consequently, heavy consumption of alcohol in the form of drinks, underage drinking, and driving under the influence of alcohol are the most common forms of alcohol abuse in Malaysia [3]. Currently, alcohol addiction cases have risen, especially among the younger generation. It is due to the readily available alcohol and its low cost so that many people can afford.

Although the government have made efforts to curb the illegal sale of alcohol, those that are already addicted needs help to cure their addiction. However, before moving forward with the type of treatment

 p-ISSN 2798-253X
e-ISSN 2798-3153

needed to be given to the alcohol addict patients, it is required to identify the type of risk level for the patient because different risk level has different types of treatment to break free from alcohol addiction. Therefore, it is important to understand the behavioral and clinical data of the patient to predict the reliable risk level for a proper treatment.

As of today, the process of alcohol intervention includes the patient visiting the healthcare providers, taking the alcohol use disorder screening tests (such as AUDIT-C, AUDIT-10), having discussion with the healthcare personnel, and receiving the advice for alcohol use disorder recovery for individual. On the other hand, the current process of alcohol use disorder recovery support is solely dependent of manual process and face to face interactions. However, with the impact of global pandemic (such as COVID-19), there is a need to consider the virtual process of alcohol abuse recovery support. Since alcohol addiction solely is not considered as critical health issue requiring urgent attention (such as transmittable diseases, cancer, tumors, and any other life-threatening cases which speak of death and live situations). Therefore, the attention to the alcohol abuse recovery seems to drop during the pandemic, given the limitation that there is only face-to-face manual process for consultation and recovery assistance.

Similarly, the healthcare sector is closely associated with human interaction. Hospital administrators are spending their day in appointment scheduling and answering routine questions of patients. Continuing or repeating the same actions and words is neither necessary nor productive. Henceforth, although Malaysia is the world's tenth largest consumer of alcohol, few research have been conducted to determine the exact scope of effective relapse prevention and support for people suffering from alcohol use disorder [4]. Currently, the existing alcohol relapse support process consists of four major steps as shown in Fig. 1. These steps are labor intensive and repetitive tasks because the same process is repeated by the healthcare personnel for an individual patient.

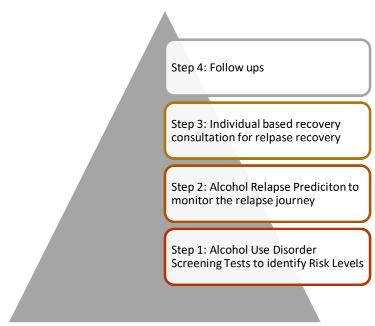

Fig. 1. Generic process of alcohol use disorder recovery support in clinical setting.

Additionally, with the increasing number of alcohol addiction and growing number of patients, the workload of healthcare personnel may also be increased. There can potentially be delay in response and in service which directly or indirectly impact the recovery progress of the patient with the given capacity limitation barriers. Furthermore, by following the standard practice, looking at the standard clinical data and imposing the fixed schedules/workload, healthcare personal will not have luxury to explore the insights or patterns from the existing data of patients to innovate further recommendations or generate new ideas which can support the recovery journey.

Hence, some of the aforementioned standard processes can be replaced with artificial intelligence (AI) solutions to hasten the process and improve the recovery support. For example, by gathering the relevant clinical and behavioral data, an artificial intelligence model can be presented to predict the alcohol relapse rate of an individual and provide 24/7 availability of the online recovery consultation for the recovery support.

In this case, the robotic automation under the umbrella of artificial intelligence which can be introduced to resolve the current limitations of alcohol recovery support process is chatbot applications. These applications are well known for automatic conversational agents that run on computer programming or a kind of artificial intelligence (AI) interaction between the users and machines with the intervention of natural language processing (NLP). Chatbots are potentially referred to as the most

p-ISSN 2798-253X
e-ISSN 2798-3153

promising and advanced form of human-machine interaction [5]. Therefore, these applications are highly relevant in the exact scope of effective relapse prevention and support for people suffering from alcohol use disorder, as well as assisting the clinical process of healthcare providers.

However, despite proven that chatbots are helpful for healthcare related purposes, the adoption of chatbot in healthcare domain is still slow compared to other domains such as businesses [6]. The query responses are mostly relying on human effort which is resource intensive and expensive. Chatbots can provide fast and instant responses to patients' healthcare-related questions regardless of the traffic and availability of healthcare personnel. This is especially challenging with the outbreak of COVID-19 pandemic. It is daunting for healthcare staffs or hospital to manage the flow of the high number of cases and cater to user enquiries at the same time. Especially in remote areas, it is becoming more difficult to consult a medical specialist regarding health and clinical matters [7].

Besides, training and deploying chatbot is also not a one-off effort. It is impossible to envision all the important questions that the users will ask from the beginning of the deployment. Chatbots in general need continuous maintenance and adaptation to perform their task appropriately [8]. The potential unpredictable variation in user input and new needs represent substantial challenges for the usefulness of the chatbot over time. To develop chatbots that are adaptive to the needs of users and conversational contexts, features such as flexibility and ease of retraining need to be considered. Developing a chatbot is not a one-solution-fits-all technology, but rather, developers and designers should work to understand new user needs and motivations. The shift of needs after deployment of chatbot should be studied but not only relying on common assumptions of user needs previously.

Henceforth, the major aim of this study is to predict the relapse rate and provide online recovery consultation of an individual using supervised machine learning in the backend process and interact with users using the multilingual chatbot in the front end to provide AI based personalized virtual therapy for alcohol relapse recovery. Hence, the major goal is to introduce a hybrid approach by integrating the predictive modelling and intent detection under the scope of machine learning. The major objectives of this study are explained as in below steps.

1. To predict the alcohol relapse rate of an individual using deep learning approach.
2. To provide automated online recovery consultation for an individual using collaborative filtering algorithms.
3. To develop a chatbot system using intent detection technique, which can capture useful customer interaction and is flexible to adapt to new intents discovered with ease.

On the other hand, achieving the above objectives, this study will significantly contribute to the healthcare domain as two out of four processes in standard alcohol abuse recovery support (as in Fig. 1) can be supported with proposed solution to eliminate the labor-intensive tasks and provide an opportunity to further explore insights from data for streamline recovery support process. Hence, it can help the current and future processes in this domain. Additionally, this research is essential as it will assist in resolving one of the major issues in public health, namely the risks and impacts of binge drinking. Therefore, effective support with alcohol recovery can lower binge drinking, which correspondingly reduces the risk and impact of binge drinking in the community.

Additionally, from the perspective of contribution to research work, this research introduces a new generation of relapse prevention modules which is not existed yet today in Malaysia healthcare services, especially in the scope of the alcohol recovery process. Therefore, it will open a new research avenue in machine learning for prediction modeling and the natural language processing (NLP) area to explore a different approach to creating a virtual intelligence chatbot. All in all, the major contribution of this research is the design and realization of AI-based personalized virtual therapy for alcohol relapse recovery that can assist with queries in real-time and minimize labor work.

From the perspective of scope of the study, this research focuses only on the relapse prediction in alcohol addiction domain in Malaysia. This study delivers a chatbot as part of an alcohol rehabilitation application for end/public users who are the patients or potential customers of an alcohol rehabilitation center. However, this study will not cover any integration with the existing systems or workflows.

2. Related Works/Literature Review

Since alcohol relapse recovery is growing interest in many research avenues, there are recent related studies in the area of chatbot implementation. However, the current related work focuses on the pure chatbot system as a question and answer and has not been integrating the idea of prediction modeling in the backend to fully utilize the power of artificial intelligence to resolve the current limitations of the traditional alcohol recovery process. In this section, this study reviewed current practices of technology-based interventions for alcohol recovery therapy, related work on chatbot and prediction modeling for alcohol recovery support, current trends, and implementation of chatbots in Malaysia for alcohol recovery support, and state-of-art approaches/algorithms for related work in alcohol relapse prediction modeling and chatbot implementation.

3.1. Current Practices of Technology-Based Interventions for Alcohol Recovery Therapy

As of today, there are several alcohol monitoring or substance use control e-health systems present in the literatures, including the DayBreak Program [9], Ria Treatment Platform [10], and another system by Scott, C.K., Dennis, M.L., and Gustafson, D.H. [11]. The literatures mostly revolve around information dissemination, motivation enhancement, and self-monitoring. Some systems provide users with self-assessment test, and then part of them will also suggest self-help intervention from the results. The system by Mitchell, M.M. *et al.* also uses external device to continuously monitor their users, such as monitoring their alcohol intake using a USB device periodically, to prevent them from drinking again. Some systems also have peer support where users undergoing the rehabilitation can form a social network and gain motivation from each other in there.

Lastly, when users have questions, they can also seek help from physicians, and physicians can answer them or provide live coaching. A system without live coaching will still require users to go to the rehabilitation center physically. Meanwhile, having a live coach to provide information or answer questions is common in the systems; there is no automated approach like a chatbot used to alleviate the task. In the literature review section, some alcohol rehabilitation center websites focusing on information dissemination will be reviewed to determine the kind of information that the end user landing page with the chatbot interface needs to provide.

There are many commercial and open-source options available for the development and management of a chatbot, and the number of chatbot-related technologies or NLU platforms is already overwhelming and is still climbing each day [12]. The NLU platforms allow developers to create applications capable of understanding natural languages. The leading ones identified are Google's DialogFlow, Facebook's wit.ai, Microsoft LUIS, IBM Watson Conversation, Amazon Lex, and SAP Conversation AI. All of them are supported by machine learning technologies. While these platforms share some standard functionalities, including a convenient interface for intent management, there are still some distinguishable functionalities, such as training on user endpoint utterance. The machine learning algorithms behind these chatbot platforms are not disclosed.

There are two approaches to developing a chatbot on intent detection depending on the techniques and algorithms used: pattern matching and machine learning. [12] Rule-based chatbots match the user input to a rule pattern and select a predefined answer from a set of responses with the implementation of pattern-matching algorithms and knowledge bases such as Artificial Intelligence Markup Language (AIML) or ChatScript. While the implementation can help with simple applications, it is not scalable to large-scale usage as the developers must write a pattern for every possible input of users. Maintaining the growing knowledge base periodically is also highly inefficient, and the learning cannot be automated.

Chatbots that adopt machine learning approaches instead of pattern matching extract the content and context from users' input using natural language processing (NLP), natural language understanding (NLU), or artificial neural network (ANN) techniques [13]. They have learning capabilities and typically need a decent dataset to perform well. Oftentimes, ANNs are used for the implementation of these chatbots. Machine learning models are further broken down into two categories, namely retrieval-based model and generative model. Retrieval-based models use a neural network to assign scores and select the most likely response from a set of responses. On the other hand, generative models synthesize the reply, usually using deep learning techniques, which could alter the information in an unwanted way.

p-ISSN 2798-253X
e-ISSN 2798-3153

3.2. Current Design of Alcohol Rehabilitation Center Web Applications in Malaysia

Current technology-based interventions for several well-known Malaysia alcohol rehabilitation centers are studied to observe the current trends according to Luxury Rehabs, namely Solace, The Wave Clinic, and Serene Retreat. All the websites provide some sort of information about their rehabilitation center, treatment program, and their contact to help the visitors to understand them better. Information about rehabilitation center ranges from the goals of the center, their location, the professional team to help with the treatment, the facilities they provide, to the awards they reaped. The treatment information includes approaches of the treatment, various scientific facts about the addictions, and helpful advice for treatment.

The information aims to convince the potential customer on the effectiveness and importance of treatment. The contact information includes their phone number and email, which will help the potential customers in admission and queries. Solace and The Wave Clinic also provide live chat services in case the content on the website is not enough to cater to the questions of the visitors. However, the live chat services are not available 24 hours every day. Moreover, the line could be busy sometimes. The same goes to the contact number. There is no automated approach in those websites like chatbot to alleviate the tasks. This could be resource intensive and time consuming for both of the parties. Table 1 below shows the comparison and summary of the features.

Table 1. Feature Comparison Between Alcohol Rehabilitation Center Web Applications

System	Rehab Centre Information	Program Information	Contact Information	Live chat	Chatbot
Solace	/	/	/	/	×
The Wave Clinic	/	/	/	/	×
Serene Retreat	/	/	/	×	×

3.3. Current Trends and Implementation of Chatbot in Chatbot Management Platforms

The different features of the well-known natural language understanding platforms for chatbot are reviewed, such as IBM Watson, Google Dialogflow, and Microsoft LUIS [14] and another chatbot platform also by a technology giant, which is Facebook Wit.ai. [15]. Generally, all of the mentioned platforms provide a user interface for chatbot intent management, support retraining, and have multilingual capabilities. Intent management are convenient with user interface to edit responses, add or delete intents, manage utterances, and adding entities.

Besides, the chatbots have fallback intent to show to the users when the confidence of the prediction is lower than certain threshold. In such case, several intents will also be offered to the user to select. The entity adding feature is only useful for task oriented chatbot such as hotel booking but has no direct effect on question answering. Having the utterance managing feature also means that the chatbot can be retrained and improved over time. An analytics module is present to help monitor the performance of the chatbot and the traffic.

Training on endpoint utterance is a feature where conversation history between the chatbot and end users can be traced, which assist in observing of whether there are any wrong predictions by the chatbot or new important intents the system does not currently cover, then retrain the chatbot on those messages. Only Dialogflow and LUIS support this feature. The feature is important to help the chatbot to solve real issues rather than only based on the developers’ assumptions on important questions.

Additionally, having a unified dataset for all languages will be a very convenient feature to have for the administrators, especially when their jobs are not specifically for the chatbot management. Besides, the chatbot does not automatically detect the users’ language in those systems and the user will have to manually select the language before the conversation. IBM Watson is an exception which uses translation service from the same ecosystem. However, extra programming work must be done for that feature. With the external translation service, it is possible for IBM Watson to only train one model for multiple languages. The feature comparison of the well-known systems is demonstrated as in Table 2.

p-ISSN 2798-253X
e-ISSN 2798-3153

Table 2. Feature Comparison Between Chatbot Building and Management Platforms

Chatbot	Intent	Fallback	Analytics	Training on	Single	Automatically
Dialogflow	/	/	/	/	×	×
Wit.ai	/	×	/	×	×	×
LUIS	/	/	/	/	×	×
IBM Watson	/	×	/	×	×	External service

3.4. State-of-Art Approaches in Related Work for Chatbot Algorithm

Algorithms as the working core of the chatbot system are reviewed to understand the knowledge gap in related work and propose relevant algorithm for this research. [16] proposed AidMe, an adaptive and sample efficient NLU module using semantic similarity-based model rather than classification model. While a wide diversity of machine learning models has been used to create intent detection classifiers, and these models performed extremely well, the motivation of [16] was to build a digital assistant which is flexible and can adapt to new intents with ease. A classification model can only predict classes on which it has been trained. When it discovers a new intent, a system based on a classification model needs therefore to be retrained. This can be an issue depending on the training time as the system would be unavailable for a while.

But above all, the main challenge would be to maintain a relative balance in the dataset between classes, otherwise some class could be “ignored” by the model. High performance classification methods like neural networks also overfits easily on small dataset, which makes them unfeasible for most conversational agents where the sample size is extremely small. Moreover, when an intent is discovered for the first time, few data would be available to train on and the performance on the new class would need time to become satisfactory.

The “half-shot” model by [16] works by comparing the user input with sentences of known intent using a semantic similarity model. The semantic similarity model used is trained on big semantic similarity dataset once and for all, thus is sample efficient and does not need to be retrained when new intents are discovered. Through users’ interaction, the digital assistant can learn more pattern and thus improve. However, the authors also pointed out that the sentences needed to be compared grow quickly and thus the efficiency will suffer as the time goes on.

To support users of various demography in Malaysia, multilingual chat capabilities is crucial to give seamless experience. Often, multiple chatbots are developed to cater to this, increasing the complexity of maintenance having to keep several versions of data in different languages. [17] shows that aligning multiple languages embedding and model trained simultaneously on multilingual dataset can outperform model trained separately on each language. Their work opens the possibility to maintain only a single training set for multilingual chatbots, the collection of data from user interaction for online training is also easier and the extended sample size from multiple languages jointly improve the performance of chatbot. Users also do not have to explicitly specify their language, further improving their experience.

Another work by [15] also proposed semantic similarity-based model using hybrid method of cosine similarity and tf-idf which is much simpler, but it is shown that their work outperforms the proprietary chatbot wit.ai. In this simple approach, using matrix multiplication is much more scalable than ensemble machine-learning based models in inference stage. However, due to the properties of bag-of-word embedding, the authors must do synonym expansion before encoding the user input to classify utterance with out of vocab words properly. The comparison of different features from each related work is summarized as in Table 3.

Table 3. Feature Comparison Between Intent Detection Algorithms

System	Multilingual	Adaptability	Scalability	Contextual Embedding
Lair et al., 2020	×	/	×	/
Abbet et al., 2018	/	×	/	/
Chu & Huang, 2020	×	/	/	×

3.5. State-of-Art Approaches in Related Work for Prediction Modelling in Alcohol Addiction Domain

For the prediction models, [18] investigate the long-term effectiveness of online communities and social media in preventing relapse. Authors used Kaplan-Meier method and Cox regression to examine the relationship between the factors influencing the relapse and relapse rate for alcohol addiction. The findings concluded that there is direct positive impact with increased social engagement and alcohol relapse rate. However, authors claimed that there are limitations in the study as Falsified reporting posed a problem for the dataset. Individual inferences cannot be made based on the likelihood values of a relapse occurrence derived from survival analysis. Quantifying the extent to which involvement in an online support community complements efforts toward addiction cessation will be a focus of future research.

Similarly, [19] presented the structured prediction approach considering the linguistic and psycholinguistic features from user' and friends' tweets from social media. Hinge-loss Markov Random fields recovery prediction models and Latent Dirichlet Allocation (LDA) is used to perform the experiment of this study. The findings of the paper indicated that the proposed solution can encodes dependencies among them to model and understand recovery and relapse from/into AUD. The result is compared with Logistic Regression and observed the proposed solution outperformed the traditional Logistic Regression (LR).

All in all, based on the outcome of the literature review, there are four key limitations to be addressed as firstly, the current query responses in Malaysia rehab centers still relies fully on human response which is labor intensive [20]. Repeatedly responding to common queries by patients can jeopardize the productivity of medical staffs and time wasting for both parties [7]. The mundane task can be automated using chatbot technology to assuage the burden of medical staffs especially in this critical time in pandemic and improve service timeliness for patient.

Secondly, there will be evolving needs over time. Chatbot feedback loops and case-supported adaptability is important to provide great experience [21]. Chatbot cannot cater to all the user needs straight after rollout and the service could also spark new needs and interactions from users. Thus, the chatbot system shall allow the system admin to investigate the user endpoint messages to understand their real needs which may not be covered by current system. The chatbot should be flexible and adaptive to new needs easily. The chatbot performance on existing intents can also be improved through user feedback and more endpoint messages.

Thirdly, rich demography of users in Malaysia makes monolingual approach unfeasible. Current approach by chatbot platforms where one model is built for one language is inconvenient to maintain [22]. Current chatbot platforms either use translation or separated models to assemble multilingual chatbot. The models can only be trained using corpus from single language, making it hard to maintain especially when we are also trying to be adaptive to new customer messages which come in different languages. Ideally, the chatbot should be able to understand multiple languages using only one model so that the training samples do not need to be translated repeatedly. This will help to save time.

Fourthly, there is no effective integration as a solution for the prediction and interaction as online service for the alcohol relapse recovery support using machine learning model [23]. Most importantly, the current practices seem to be complex to be integrated into the hybrid solution and required simplified predictive modelling process.

In brief, the findings from the above literature reviews supports the justification of problem statement in this study as the major addressable issue as inefficient and ineffective alcohol recovery support solely depending on the human labor, and the need for artificial intelligence solution to improve the support process.

3. Material & Methodology

4.1. Data

For the research design, this research adopts quantitative research method. The target population is the healthcare providers and patients from Selangor, Malaysia. Therefore, for the data collection, survey

p-ISSN 2798-253X
e-ISSN 2798-3153

distribution is conducted to the sample group i.e., the healthcare providers and patients in Selangor. Both survey data and existing electronic clinically data is collected from the sample user group to further proceed with analysis. For data collection process, Online Google Form is used to create an online questionnaire to gather requirements and data from the stakeholders. The data collection period is from May 2022 to June 2022, with no limitation in gender, age, and occupation for the respondents.

Questionnaires are a list of questions either open-ended or close-ended for which the respondents can provide answers. This method allows the respondents to answer the questions at their own pace and is useful when it is difficult to arrive at a suitable appointment time for verbal interview. The aim of distributed survey is to understand the general workflow and design expectation for the proposed solution. There are three dedicated surveys to distribute to different dedicated users in target sampling group and collected the data related to the frequently asked questions / queries in alcohol addiction treatment departments, standard recovery consultation procedure, standard consultation queries and answers, user's interest data, clinical data related to addiction, and other relevant data provided by the sample group.

Once the data is collected from three different survey, it is consolidated, cleaned, and transformed using RapidMiner software to first filter the valid subset of dataset for further analysis and processing. Both descriptive analysis and correlation analysis is performed on the data to validate the reliability and validity of the dataset collected for further processing.

3.6. Method

Based on the outcome from literature review section, a few retrieval-based machine learning algorithms are reviewed, and requirements are collected to choose and implement suitable algorithm for proposed solution, to meet the objective of the study. There are three major components in the proposed solution to complete the objective of the study such as predictive modelling for relapse rate, prediction of recovery consultation responds for automate consultation service and real-time chatbot as virtual therapist.

After careful review and consideration on the methods relevant for the proposed solution, Recurrent Neural Network (RNN) is used to predict the possible relapse of alcohol addict patient. RNN is chosen as the less complex and effective machine learning model for time series-based prediction. The main advantage of RNN over other machine learning models in prediction of relapse rate is that it can model a collection of records (i.e., time collection) and presume that each pattern is reliant on the preceding one [24]. For all types of sequential data analysis, RNN is the best option. Because data varies over time in forecasting, and RNNs can learn changes in time domain, it may be a superior answer for prediction.

Secondly, collaborative filtering is used to provide relevant recovery consultation response based on user interest data. It is relevant for the scope of study because collaborative filtering is the process of gathering and analyzing data on user behavior, activities, and preferences to forecast what a user will like based on their similarities to other users [25]. It is a method of removing items that a user might enjoy based on the reactions of other users. It works by sifting through a huge group of people to locate a smaller group of users with similar likes to a specific user.

Thirdly, a BERT-based model with aligned multilingual embedding using distillation method is implemented for intent detection. It is chosen due to its state-of-the-art performance in semantic similarity tasks. It can also cope with current limitations in terms of multilingual capability, scalability, and adaptability.

System development will be done with python Flask as backend for AI framework integration, web languages as frontend, and MySQL as database. The overview of the proposed framework for the proposed solution is illustrated as in Fig. 2.

p-ISSN 2798-253X
e-ISSN 2798-3153

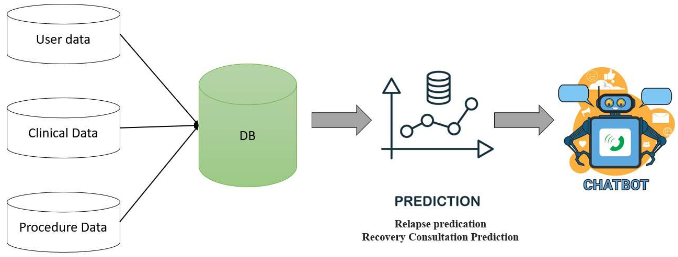

Fig. 2. Overall framework for the proposed solution

3.2.1. Alcohol Relapse Prediction Using RNN

RNN is implemented to forecast the addiction relapse of a patient. The proposed solution divides the time of the day into 4 quarters i.e., morning, afternoon, evening, and night. Then, the consumption amount within 4 quarters of the day is grouped. Afterward, the dataset is reduced into 4 quarters of the day data within a day for each user to correctly anticipate the period when the addict consumes the most alcohol in a day. The trained model forecasts the amount of alcohol drank every quarter by a person for the following day (i.e., the next 24 hours), as well as the time interval during which the maximum amount is consumed. With that information, the model detects the binge drinking period to provide the recovery consultation for the patient in that interval so that person can be diverted to another activity instead of involving in binge drinking. The overview of the aforementioned process is demonstrated in Fig. 3.

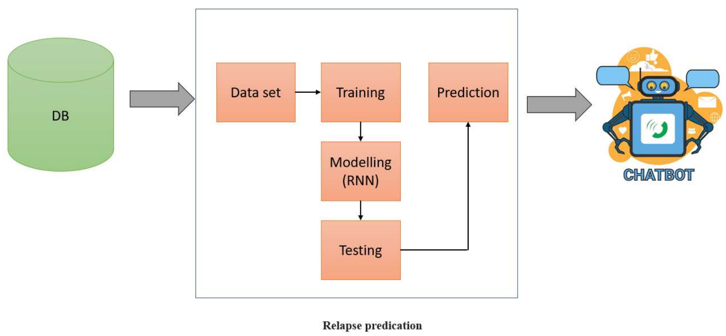

Fig. 3. Overview of steps in modelling with RNN for relapse prediction

3.2.2. Recovery Consultation Using Collaborative Filtering

For the online recovery consultation, collaborative filtering algorithm is used on the user characteristic dataset. Based on the user database, similarity matrix is calculated on the items in personal interest and search the relevant consultation respond related to it. Finally, the score prediction is calculated, and top-N is selected to respond the relevant consultation message to the patient to motivate abstain in binge drinking. Based on the similar traits, characteristic and activities of the patients, the relevant recovery consultation is automated as online recovery consultation service for individuals. The overview of recovery consultation model is illustrated as in Fig. 4 below.

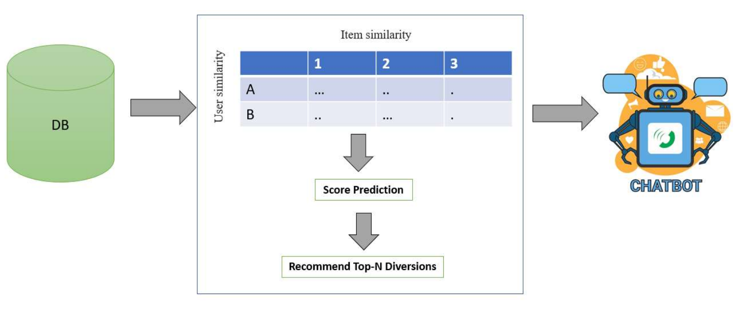

Fig. 4. Overview of recovery consultation model

p-ISSN 2798-253X
e-ISSN 2798-3153

3.2.3. Chatbot with Intent Detection

A chatbot management platform with user friendly interface is built for intent management and chatbot training. The queries elicited are annotated with their corresponding intents in the platform. Moreover, the question responses in different languages are also prepared so that the chatbot can respond in the language used by the users if the language is supported. The algorithm to be used is proposed and implemented by [26] in a python framework. The embedding space between languages is aligned using knowledge distillation method, so only one model is needed to be trained and maintained for multiple languages. The parallel texts for Malay, Tamil, Mandarin, and English are also available in the framework for the embedding space alignment. Comparison using cosine-similarity greatly reduces the inference time of sentence pair regression method originally used by Bidirectional Encoder Representations from Transformers (BERT) in semantic similarity task from a whooping 65 hours to 5 seconds with only very minimal loss in performance, yielding state-of-the-art embedding. The algorithm is adaptive to new intents, scalable, and multilingual.

Human in the loop approach is used to train and improve the chatbot over time. At first the bot will only have base knowledge on the frequently asked questions elicited from the stakeholders. The elicited frequently asked questions are annotated and stored in the intent table in the MySQL database. In the interaction with users, the users can give feedback if the responses from the chatbot are wrong, and the questions will be captured. After that, the admin from the rehabilitation center can investigate and annotate the captured questions to serve as new knowledge of the chatbot and the cycle goes on. The aforementioned learning process is illustrated in Fig. 5.

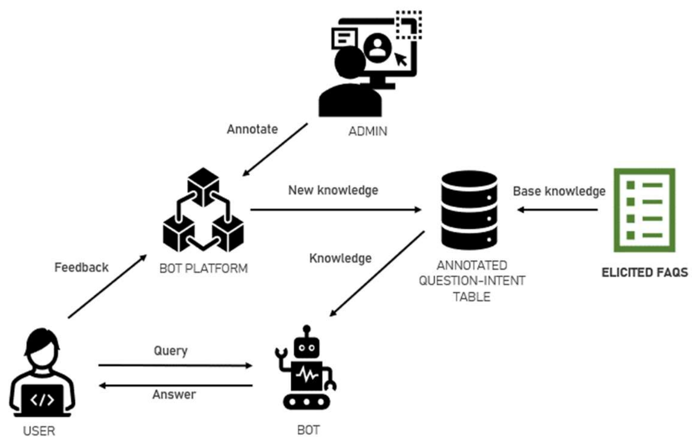

Fig. 5. Chatbot learning process

4. Results and Discussion

4.1. Result

In this section, the results of this study are presented. Based on the methodology proposed, AI based virtual therapist is developed as demonstrated in Fig. 6, where the landing page shows some information about the alcohol rehabilitation center. The developed AI solution includes chatbot module, user landing page, intent management module, chatbot retraining module, predictive module, and analytics module. The chatbot developed will have a mechanism for the users to give negative feedback on chatbot responses returned. The negative feedback rate will be used to monitor the performance of the chatbot, and the questions captured will be used for chatbot retraining to provide better and better experience for users over time. The overall solution is also designed to cater to multiple languages to give a better locale user experience. A monitoring dashboard is also designed as a part of analytic module to continuously monitor the performance of the chatbot, and user acceptance testing will be done to evaluate the effectiveness of chatbot system in terms of user friendliness and real-life use case applicability.

p-ISSN 2798-253X
e-ISSN 2798-3153

Fig. 6. Landing Page

Moreover, for the integration of prediction and interaction components, the user must first log-in to the system to access to the virtual therapist. The RNN prediction model in backend is connected to the chatbot and based on the user's input, the data from the trained model is fetched as an intent in user query and respond to the user's question. By validating the user ID, the chatbot recommends the recovery consultation respond in chatbot. User can follow the provided link in chatbot for further information with the given consultant advice relevant for the user. Fig. 7 demonstrated the sample recovery consultation for the user.

Fig. 7. The interface of the chatbot for online recovery consultation

Next, the activity flow for the interaction between the end users and chatbot is explained. First, the user sends a message, then the system detects the language of the message, so that it can automatically return the response in the corresponding language later. Next, the message is normalized, encoded and labelled with intent. The message returned will depend on the confidence level, if the confidence level is low, few selections will be given to improve the robustness, and users can give feedback if the response is wrong, and the message is captured. Then the user can continue the interaction or end the chat.

On the other hand, for the training or retraining of the chatbot, it can be achieved via the chatbot management interface. Initially, the "predicted" column will show the top response returned by the chatbot. To train the chatbot on specific question, for example, if it is from an existing intent, admin can select it from the dropdown list and click train. Otherwise, if it is a new intent, admin can select "new intent" and click train. The chatbot training on endpoint utterance is demonstrated as in Fig. 8.

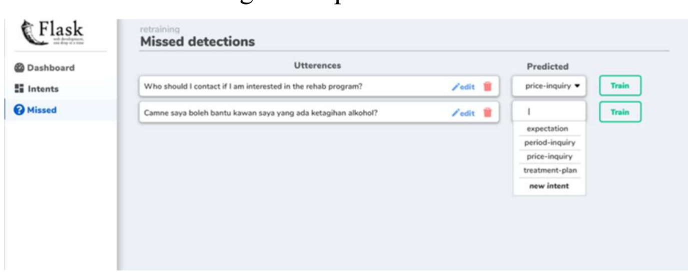

Fig. 8. Chatbot training on endpoint utterance steps

p-ISSN 2798-253X
e-ISSN 2798-3153

Additionally, if "new intent" is selected for training, the admin will be redirected to enter the intent name, the response for each language, and the short description for each language, which act as a summary that will be shown in the selection buttons when the confidence is low, as in Fig. 9. Additionally, admins can add, delete, or edit existing intents. They can also search for existing intents from the search bar. The responses, descriptions, and training samples are shown on the right-hand side. Admins can edit the responses, description, and training samples by clicking on the edit button. The bin button at the "Training samples" section allows them to delete samples while the add button below allows them to add more training samples.

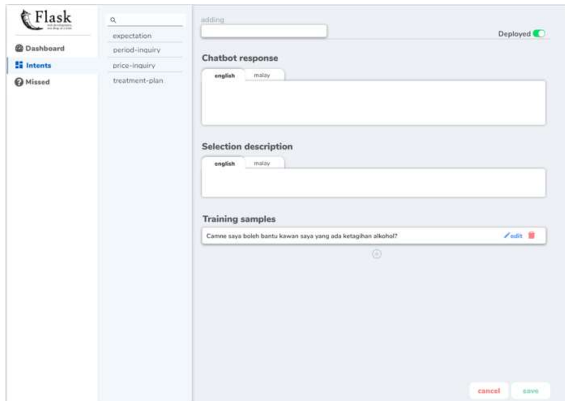

Fig. 9. Adding new intent to the chatbot training

Moreover, the overall training and chatbot performance can be evaluated with the designed dashboard. Admins can monitor the performance of the chatbot in this page by looking at the negative feedback/thumbs down rate. The thumbs down rate is also broken down into intent level so that the admin can see which intent is underperforming and need attention the most. They can also monitor the traffic using the chatbot and the bottom section. Sample dashboard interface is shown Fig. 10.

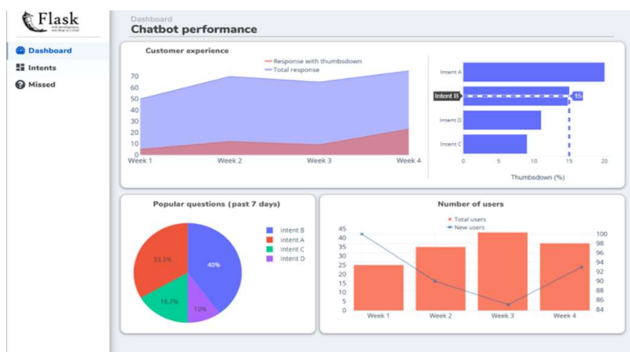

Fig. 10. Sample dashboard interface for chatbot performance monitoring

4.2. Discussion

All in all, this study achieved the mentioned three objectives. Firstly, the relapse rate of an individual is predicted using recurrence neural network to identify the high-risk binge drinking period of the day. Secondly, relevant recovery consultation respond is automated for an individual by predicting using collaborative filtering algorithms to provide respond based on the risk and period indicated by RNN model. Thirdly, a chatbot system is developed using intent detection technique and with the help of relapse predictive model and recovery consultation model, which can capture useful customer interaction and is flexible to adapt to new intents discovered with ease.

From the aspect of solution evaluation, two approaches are used. Firstly, a monitoring dashboard is developed to monitor the performance of the chatbot by looking at the percentage of negative feedback by end users. Secondly, user acceptance testing is carried out to evaluate the effectiveness of the chatbot

p-ISSN 2798-253X
e-ISSN 2798-3153

system in terms of real-life use case capability and user friendliness. As the outcome of user acceptance testing by stakeholders from the target sampling group, the proposed solution is acknowledged as usable in real-life scenario and use-case to support alcohol recovery process.

Moreover, case by case comparison is also conducted to highlight the contribution of this study in response to the findings of knowledge gaps from the literature review. Firstly, the proposed solution (i.e., AI based virtual therapist) can be integrated into the existing web applications of alcohol rehabilitation centers. The feature comparison between alcohol rehabilitation center web applications and proposed solution is explained in Table 4.

Table 4. Feature Comparison with Current Design of Alcohol Rehabilitation Center Web Applications in Malaysia

System	Rehab Centre Information	Program Information	Contact Information	Live Chat	Chatbot
Solace	/	/	/	/	×
The Wave Clinic	/	/	/	/	×
Serene Retreat	/	/	/	×	×
Proposed solution	/	/	/	/	/

Secondly, the proposed solution considered the existing limitations of current chatbot building platforms such as DialogFlow in building multilingual chatbots so that only one model will need to be built and managed to give the admin a more user-friendly experience when managing the chatbot. This feature will be helpful especially when we are retraining the chatbot on endpoint utterances from time to time as user messages come in different languages. It eliminated the need to translate the messages for each model multiple times with the presence of this feature. The feature comparison with current state of art approaches and proposed solution is explained in Table 5.

Table 5. Feature Comparison with Current State-of-Art Approaches and Proposed Solution

Chatbot	Intent	Fallback	Analytics	Training	Single	Automatically
Dialogflow	/	/	/	/	×	×
Wit.ai	/	×	/	×	×	×
LUIS	/	/	/	/	×	×
IBM Watson	/	×	/	×	×	External service
Proposed	/	/	/	/	/	/

Thirdly, the proposed solution covered the knowledge gap in the recent related work. Addition of the contextual embedding, this study included the multilingual aspect of training in the chatbot and consideration of BERT for further adaptability and scalability purposes. The highlight of proposed solution filling the gaps in the related studies is explained in Table 6.

Table 6. Fulfilment of Knowledge Gap in Previous Studies by The Proposed Solution

System	Multilingual	Adaptability	Scalability	Contextual
Lair, N. et al.	×	/	×	/
Abbet, C. et al.	/	×	/	/
Chu, E. T.-H. & Huang, Z.-Z.	×	/	/	×
Proposed solution	/	/	/	/

Therefore, the proposed solution is evaluated theoretically as well as practically to ensure the usability of the solution in real world healthcare domain. Overall, the proposed solution can be implemented in alcohol abuse recovery process for the foreseen benefits such as reduce of labor intensiveness and improve of streamline support process virtually.

p-ISSN 2798-253X
e-ISSN 2798-3153

5. Conclusion

In conclusion, a chatbot targeting a local alcohol rehabilitation center and its management system have been designed as AI based virtual therapist for alcohol recovery to reduce the labor work in queries responses on traditional consultant process for alcohol recovery support. The chatbot serves as an interaction with the user, however, there are backend processes such as predictive modelling and recovery consultation models are integrated to improve the usability of chatbot. The management system aims to make the chatbot case-supported and adaptive to future needs to address important questions by users and provide them with better experience.

Recurrent Neural Network (RNN) is used for predictive modelling on relapse rate to identify the period where binge drinking occurs. Collaborative filtering is used to provide online recovery consultation for the patient to abstain the patient from binge drinking and support relapse recovery. BERT-based model architecture as intent detection algorithm is used to overcome the limitations of current chatbot building platforms such as DialogFlow.

For the evaluation of system performance, user acceptance testing is conducted with target group to assess the effectiveness of the solution to ensure the project objectives are achieved. Additionally, a monitoring dashboard is designed to help continuously monitor the performance such as negative feedback rate of the chatbot. On the other hand, the exploration on the different prediction models (such as SVM, NBC, LR) can be a future study on this research to experiment different combination of prediction model to improve the performance and prediction accuracy of AI based virtual therapist.

Acknowledgement

This research was supported by the Impact-Oriented Interdisciplinary Research Grant Programme (IIRG) Universiti Malaya (IIRG005B-2020SAH).

References

[1] Abdalrazak, H., *et al.* (2021). Review on Substances Abuse in Malaysia-Way Forward and Challenging. Annals of the Romanian Society for Cell Biology, 4930-4943.

[2] Ng, E. (2020). Alcohol Consumption in Malaysia. Retrieved 18 June 2022, from https://andatech.com.my/blogs/news/drug-and-alcohol-consumption-in-malaysia.

[3] Chhoa, K.H., Zakaria, H., and Abd Rahman, F.N. (2019). Problematic Alcohol Use and Depression in Secondary School Students Iin Miri, Malaysia. *Pediatrics International*, 61(3), 284-292.

[4] Hong, S.A., and Peltzer, K. (2019). Early Adolescent Patterns of Alcohol and Tobacco Use in Eight Association of South-East Asian Nations (ASEAN) Member States. *Substance Use & Misuse*, 54(2), 288-296.

[5] Murtarelli, G., Gregory, A., & Romenti, S. (2021). A Conversation-Based Perspective for Shaping Ethical Human–Machine Interactions: The Particular Challenge of Chatbots. *Journal of Business Research*, 129, 927-935.

[6] Laranjo, L., *et al.* (2018). Conversational Agents in Healthcare: A Systematic Review. *Journal of the American Medical Informatics Association: JAMIA*, 25(9), 1248–1258. https://doi.org/10.1093/jamia/ocy072.

[7] Battineni, G., Chintalapudi, N., and Amenta, F. (2020). AI Chatbot Design during an Epidemic Like the Novel Coronavirus. Healthcare (Basel, Switzerland), 8(2), 154. https://doi.org/10.3390/healthcare8020154.

[8] Brandtzaeg, P., and Folstad, A. (2018). Chatbots: Changing User Needs and Motivations. *Interactions*, 25(5), 38-43. https://doi.org/10.1145/3236669.

[9] Tait, R.J., Paz Castro, R., Kirkman, J., Moore, J. C., and Schaub, M. P. (2019). A Digital Intervention Addressing Alcohol Use Problems (the "Daybreak" Program): Quasi-Experimental Randomized Controlled Trial. *Journal of Medical Internet Research*, 21(9), e14967. https://doi.org/10.2196/14967.

[10] Mitchell, M.M., Mendelson, J., Gryczynski, J., Carswell, S.B., and Schwartz, R. P. (2020). A Novel Telehealth Platform for Alcohol Use Disorder Treatment: Preliminary Evidence of Reductions in Drinking. *The American Journal of Drug and Alcohol Abuse*, 46(3), 297–303. https://doi.org/10.1080/00952990.2019.1658197.

[11] Scott, C.K., Dennis, M.L., and Gustafson, D.H. (2017). Using Smartphones to Decrease Substance Use via Self-Monitoring and Recovery Support: Study Protocol for a Randomized Control Trial. *Trials*, 18(1), 374. https://doi.org/10.1186/s13063-017-2096-z.

[12] Adamopoulou, E., and Moussiades, L. (2020). Chatbots: History, Technology, and Applications. *Machine Learning with Applications*, 2. https://doi.org/10.1016/j.mlwa.2020.100006.

[13] Pandey, S., Sharma, S., and Wazir, S. (2022). Mental Healthcare Chatbot Based on Natural Language Processing and Deep Learning Approaches: Ted the Therapist. *International Journal of Information Technology*, 1-10.

[14] Abdellatif, A., Badran, K., Costa, D., and Shihab, E. (2021). A Comparison of Natural Language Understanding Platforms for Chatbots in Software Engineering. *IEEE Transactions on Software Engineering*, 1-1. doi: 10.1109/TSE.2021.3078384.

[15] Chu, E.T.-H., and Huang, Z.-Z. (2020). DBOS: A Dialog-Based Object Query System for Hospital Nurses. *Sensors*, 20(22), 6639. doi:10.3390/s20226639.

[16] Lair, N., Delgrange, C., Mugisha, D., Dussoux, J.-M., Oudeyer, P.-Y., and Dominey, P. F. (2020). User-in-the-Loop Adaptive Intent Detection for Instructable Digital Assistant. *Proceedings of the 25th International Conference on Intelligent User Interfaces*, Cagliari, Italy, March 2020. IUI. https://doi.org/10.1145/3377325.3377490.

[17] Abbet, C., *et al*. (2018). Churn Intent Detection in Multilingual Chatbot Conversations and Social Media. *Proceedings of the 22nd Conference on Computational Natural Language Learning*, Brussels, Belgium, October 2018. Association for Computational Linguistics.

[18] Tamersoy, A., Chau, D.H., and De Choudhury, M. (2017, July). Analysis of Smoking and Drinking Relapse in an Online Community. *In Proceedings of the 2017 International Conference on Digital Health* (pp. 33-42).

[19] Zhang, Y., Ramesh, A., Golbeck, J., Sridhar, D., and Getoor, L. (2018, April). A Structured Approach to Understanding Recovery and Relapse in AA. *In Proceedings of the 2018 World Wide Web Conference* (pp. 1205-1214).

[20] LuxuryRehab (n.d.) Find a luxury Rehab Centre in Malaysia. LuxuryRehab. https://luxuryrehabs.com/malaysia/.

[21] Brenier, J. (2018). Amplifying User Intelligence with Chatbot Feedback Loops. Chatbots Magazine. https://chatbotsmagazine.com/amplifying-user-intelligence-with-chatbot-feedback-loops-b8e6ded391ec.

[22] Louis, H. (2018). Chatbots Orchestration and Multilingual Challenges. IBM Cloud. https://www.ibm.com/blogs/cloud-archive/?p=163089&preview=true&preview_id=163089.

[23] Blanchard, B.E., Stevens, A.K., Sher, K.J., and Littlefield, A.K. (2020). Reexamining the Psychometric properties of the substance use risk profile scale. Assessment, 27(3), 454-471.

[24] Solares, J.R. *et al.* (2020). Deep Learning for Electronic Health Records: A Comparative Review of Multiple Deep Neural Architectures. *Journal of Biomedical Informatics*, 101, 103337.

[25] Najafabadi, M.K., Mahrin, M.N.R., Chuprat, S., and Sarkan, H. M. (2017). Improving the Accuracy of Collaborative Filtering Recommendations Using Clustering and Association Rules Mining on Implicit Data. *Computers in Human Behavior*, 67, 113-128.

[26] Reimers, N. and Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, Hong Kong, China, 3982–3992. doi:10.18653/v1/D19-1410.

p-ISSN 2798-253X
e-ISSN 2798-3153