AN ENHANCED META-CLASSIFIER APPROACH FOR ALCOHOL ADDICTION PREDICTION

Myat Noe Win, *Sri Devi Ravana, and Liyana Shuib

Department of Information Systems, Faculty of Computer Science & Information Technology
Universiti Malaya
50603 Kuala Lumpur, Malaysia

Emails: 17219922@siswa.um.edu.my, *sdevi@um.edu.my, liyanashuib@um.edu.my

ABSTRACT

Chronic alcohol consumption poses significant public health challenges globally. In underserved regions, the lack of AI-based interventions for alcohol addiction highlights a critical gap in the healthcare system, particularly regarding the early detection of alcohol abuse. Henceforth, this research aims to raise awareness of alcohol use disorder and proposes a novel AI-powered solution designed with an improved classification algorithm to address this deficiency, with a primary focus on a cutting-edge prediction model. This research shifts the current reactive approach in alcohol addiction intervention to proactive approach by employing an enhanced meta-classification algorithm (EMC) that focuses on improving the interpretability, efficiency, and accuracy of predictions. The proposed EMC ultimately provides a robust tool for healthcare professionals and patients which fosters more effective and personalized intervention strategies for alcohol addiction recovery. The results demonstrate a remarkable 10.13% improvement in balanced accuracy and a 9.72% enhancement in the area under the curve compared to traditional ensemble and state-of-the-art methods. Thus, findings from this study will assist medical practitioners and policymakers in developing evidence-based strategies to combat alcoholism and enhance public health outcomes. By deriving insights from real-world case study, the outcome of this research represents a pioneering effort to betterment of healthcare in underserved regions, offering a low-cost, scalable solution for early detection, and has the potential to significantly improve outcomes in marginalized communities.

Keywords: Alcohol Addiction; Medical Informatics; AI-based Intervention; Meta Classification; Public Health.

1.0 INTRODUCTION

Alcohol addiction, also known as alcohol use disorder (AUD), is a significant public health concern, affecting millions of people globally. According to the World Health Organization (WHO), alcohol consumption is responsible for over 3 million deaths annually, which accounts for nearly 5.3% of all deaths worldwide [1]. This alarming statistic demonstrates the far-reaching consequences of alcohol addiction. From liver cirrhosis to cardiovascular diseases and mental health disorders, the physical and psychological toll of alcohol misuse is immense [2]. Economically, alcohol addiction leads to loss of productivity, increased healthcare costs, and expenses related to law enforcement. It also hampers national growth as working-age individuals struggling with addiction become less productive, contributing to the overall weakening of the labor force. The societal costs of alcohol addiction are equally devastating, leading to broken families, domestic violence, and social stigma [3].

With alcohol becoming increasingly accessible in many parts of the world due to globalization, the challenges associated with addiction continue to grow. As per latest statistics, the issue is even more pronounced in underdeveloped countries and among ethnic minorities, who often face systemic barriers to healthcare and education [4]. In many cases, alcohol addiction is exacerbated by poverty, lack of access to mental health services, and limited opportunities for early intervention. In these settings, alcohol misuse is often a coping mechanism for stress, trauma, or socio-economic hardships, which makes it harder to address. Ethnic minorities may also face cultural or language barriers that prevent them from accessing effective treatment.

Similarly, alcoholism has become all too frequent among indigenous community in Malaysia [5]. Studies show that some Indigenous groups in East Malaysia, such as the Orang Asli, have a high prevalence of risky drinking behaviors, with 37% of drinkers in these regions scoring 8 or more on the Alcohol Use Disorder Identification Test (AUDIT) scale. This is significantly higher compared to other regions in Malaysia. In Sabah, for example, it is estimated that 18.4% of the population consumes alcohol regularly, with Indigenous communities often engaging in binge drinking, which exacerbates health and social issues [6]. The lack of adequate healthcare services, particularly in remote areas where these communities reside, makes it harder for individuals to receive proper treatment and intervention for alcohol addiction. In 2019, head of the Kensiu Orang Asli tribe claimed that a negative side effect of modernity is the influx of drugs and cheap alcohol into the Orang Asli community.

Although police and Jabatan Kebajikan Orang Asli (JAKOA) jointly put into action to eliminate this impact, the problem with addiction still exists. Recently, Non-Governmental Organization (NGO) together with Sahabat Jariah Malaysia (SJM) investigated into alcohol abuse problem in Orang Asli and concluded a new implantable solution as exploring a more psychological approach to solve the problem [7]. However, the indigenous community as in their culture are not comfortable to visit hospitals, consult on addiction and receive intervention treatments. Consequently, individuals often do not receive timely interventions, exacerbating the problem. Only a tiny percentage of patients receive effective, evidence-based therapies for alcoholism, which is undoubtedly the most serious area of unmet medical needs in psychiatry [8]. Therefore, the ongoing concern about alcohol addiction stems from the failure to identify and intervene early, allowing addiction to progress to more severe stages. This lack of early identification is often due to insufficient public health infrastructure, lack of awareness, and cultural stigmas that discourage people from seeking help.

On the other hand, traditional methods of identifying alcohol addiction often rely on observable behaviors, social cues, and personal admissions. Such methods depend on the human judgements and bare eyes, where family and community members typically look for signs such as frequent intoxication, neglect of responsibilities, withdrawal symptoms, or a noticeable impact on relationships and employment. In many cultures, including Indigenous communities, elders or spiritual leaders might step in to counsel individuals when they notice these patterns. However, these methods depend heavily on subjective observations and the willingness of the person to admit their problem, which can lead to underreporting and delayed interventions. Traditional approaches often miss early-stage alcohol abuse, especially since people may hide their behavior, making it hard to detect without more objective tools like screening questionnaires or clinical assessments.

During the COVID-19 pandemic, healthcare systems worldwide were overwhelmed with addressing immediate health crises, leaving chronic conditions such as alcohol addiction under-prioritized. The pandemic environment itself, characterized by isolation, economic stress, and disrupted routines, increased alcohol consumption globally. There is no direct medication for alcohol abuse treatment if there are no other health complications, and access to rehabilitation services was often limited [9]. This led to a surge in addiction rates, highlighting the inadequacies of traditional methods in dealing with a complex, chronic issue like alcohol addiction. Objective assessment such as screening questionnaires or clinical assessments are also laborious and prone to human errors. Hence, traditional methods are not effective approach for raising awareness and preventing alcohol usage [10].

In order to address the limitations with traditional methods, technology-based interventions are introduced to assist in alcohol addiction prevention and treatment. Digital tools, such as mobile apps that track drinking habits, virtual counseling platforms offer a way to intervene before addiction becomes severe [11]. These tools can provide individuals with immediate access to support and self-assessment tools, enabling early detection of problematic drinking behaviors. The major issue remains the lack of widespread adoption of these technologies, especially in underdeveloped regions and marginalized groups [4]. Without predictive tools that can assess risk factors and alert users to potential addiction, alcohol abuse often remains undiagnosed until it causes severe health or social issues. Therefore, technology-driven solutions, including machine learning models and remote health monitoring, are essential to bridge this gap and provide timely, personalized interventions.

In brief, from a macro perspective, alcohol addiction poses a widespread public health issue that affects entire communities and economies. Limited healthcare access, underfunded addiction services, and a lack of focus on preventive measures have led to increased addiction rates globally. At a micro level, individuals suffering from alcohol addiction often avoid seeking help due to the stigma, denial, or lack of resources. Traditional methods of identifying addiction rely heavily on subjective assessments by family or community members, which are often insufficient to detect early-stage problems. Moreover, objective medical assessments are labor-intensive and tend to occur only after the addiction has progressed, delaying necessary interventions. However, although alcoholism has a negative impact on people's lives, early detection of binge drinking, and effective intervention support can lead to more successful therapies in the future, especially for the youngesters.

To overcome current challenge and filled the gap, this paper proposes the use of machine learning for predictive analysis in alcohol use disorder prediction by leverging the traditional predictive model to meta-model with improved interpretability, efficiency and accuracy. The goal is to overcome the constraints of individual classifiers and conventional stacking models by introducing a new method called metaclassifier on heterogeneous bootstrap-aggregated models. Proposed model aimed at addressing challenge of missing early-stage alcohol abuse with more accurate prediction for early detection of addiction. By utilizing improved prediction model, it will improve upon subjective assessments and provide more reliable results than conventional prediction tools. This automatic prediction system will reduce the laborious effort required for healthcare practitioners to assess patients, streamlining the process. Furthermore, the solution will address the under-prioritized criticality of addiction by offering an efficient, interpretable system that is easy to implement in healthcare settings. All in all, the proposed approach aims to enhance the accuracy and reliability of alcohol use disorder prediction models, contributing both practically and theoretically to the field. As for the novelty, this would be the first solution in Malaysia to provide AI-powered alcohol addiction intervention for the community. In brief, the overall contribution and experimental methods used for this study are summarized as below.

- Developed an enhanced meta-classification framework for alcohol addiction prediction, leveraging a diverse ensemble of kernel-based, instance-based, and deep learning classifiers.
- Introduced a novel integration of bootstrapped heterogeneous models combined with a linear model meta-learner to achieve a balance of accuracy, interpretability, and efficiency.
- Conducted rigorous testing and evaluation of the meta-classifier against traditional models to validate performance improvements on real-world datasets.
- Demonstrated the model’s effectiveness in early-stage alcohol addiction detection, addressing the critical need for proactive intervention in health monitoring systems.

Lastly, this paper is organized as Section 2 discusses state-of-art approaches and gap analysis; Section 3 explains the methodology followed throughout the research; Section 4 presents the results and discusses findings from results, recommends application to real-word scenario; and Section 5 with conclusions, contributions and future works.

2.0 LITERATURE REVIEW

A synthetic literature review is conducted from 38 papers to summarize existing alcohol brief interventions (ABIs), early identification tools, and technology-based interventions. The aim is to perform a critical review of current methodologies and thereof identify gaps that could be addressed by implementing a meta-classifier machine learning model. Therefore, this review collectively presents the features highlighted in heterogeneous studies and will serve as a baseline to draw out limitations from existing practices which eventually leads to suggest an AI-driven prediction model for early detection of alcohol addiction.

2.1 Classic Alcohol Interventions

ABIs are widely used for alcohol misuse prevention and typically involve short, structured conversations that encourage individuals to reduce their drinking. Classic alcohol addiction interventions include various approaches like Brief Interventions (BI), Cognitive Behavioral Therapy (CBT), Motivational Interviewing (MI), 12-Step Programs, and Family Therapy [12]. BI are short, structured conversations in primary care aimed at raising awareness and encouraging reduced alcohol consumption, typically using screening tools like the AUDIT. CBT focuses on altering thought patterns and behaviors linked to alcohol use, helping individuals manage triggers and cravings through structured therapy sessions. MI encourages individuals to find personal reasons for change, addressing ambivalence and motivating behavioral shifts. Peer support programs, such as Alcoholics Anonymous, follow a structured 12-step recovery process emphasizing mutual accountability and spiritual growth. Lastly, Family Therapy involves family members in the recovery process, addressing systemic issues contributing to addiction and enhancing communication to support the individual's sobriety.

Classic alcohol addiction interventions have both significant advantages and notable limitations. On the positive side, interventions like Cognitive Behavioral Therapy (CBT) and Motivational Interviewing (MI) offer personalized support by providing one-on-one sessions with trained professionals, allowing for tailored recovery strategies specific to each individual’s needs. Additionally, community-based programs like Alcoholics Anonymous (AA) offer strong peer support, which fosters a sense of belonging and accountability, enhancing long-term recovery for some individuals. Structured frameworks, such as the 12-step programs, provide a clear, step-by-step path that gives participants a sense of direction and progression in their journey toward sobriety [13].

Despite their benefits, these interventions often lack scalability, accessibility, and early detection capabilities, prompting a shift toward more efficient technology-based interventions. One major limitation is their restricted reach and scalability, as most of them rely on face-to-face interactions, making them resource-intensive and difficult to implement on a large scale [14]. For example, Brief Interventions (BI) require healthcare professionals to administer them, which can be time-consuming and limit the number of individuals who can be helped. Another disadvantage is the subjectivity and variability of outcomes, as the success of programs like MI largely depends on the counselor's skill and the participant’s readiness for change. Moreover, these interventions are often reactive, addressing alcohol addiction only after serious issues arise, thus missing opportunities for early intervention. Lastly, access barriers such as geographic, financial, and social constraints make it difficult for many individuals—especially those in underserved areas—to engage with these programs, further limiting their effectiveness [15].

In brief, although effective to some extent, these interventions have notable limitations, particularly in their generalization and scalability. Many studies show that ABIs often rely on self-reported data, which can be inaccurate due to the stigma surrounding alcohol addiction. Furthermore, ABIs are reactive rather than proactive, meaning they are often implemented after addiction symptoms are already present. Thus, classic interventions, while beneficial for some, are often inefficient due to their dependence on subjective, labor-intensive processes and in-person interactions. This has paved the way for more accessible, technology-driven solutions that can provide early detection, ongoing monitoring, and scalable support.

2.2 Technology-based interventions

In recent years, technology-based interventions, such as mobile health apps and digital counseling platforms, have grown in popularity due to their accessibility and flexibility. These interventions range from mobile health apps to online therapy platforms and telemedicine services. Mobile applications, such as "Drinkaware" and "Sober Grid," help users track their alcohol consumption, set goals, and monitor progress over time [16]. Online therapy platforms provide remote counseling sessions with licensed therapists, often through video calls, making it easier for individuals to receive professional help from the comfort of their homes. Telemedicine services offer virtual consultations and continuous health monitoring for those in recovery. These types of technology-based interventions leverage digital tools to provide individuals with ongoing support and education, ensuring that help is accessible at any time.

While technology-based interventions offer numerous advantages, their effectiveness is limited by the absence of accurate, efficient, and interpretable prediction models for alcohol addiction [17]. One significant issue is that current interventions lack the ability to proactively identify individuals at risk of developing alcohol dependency, which means they tend to respond only after addiction becomes a severe problem. Without reliable predictive models, early identification remains a challenge, leading to missed opportunities for timely intervention. Additionally, many of the algorithms used in existing systems are either too simplistic or too complex, resulting in inaccurate predictions or models that healthcare professionals struggle to interpret [18]. This makes it difficult to create tailored intervention plans that address specific needs, particularly in diverse populations. Furthermore, the absence of interpretable models leads to a lack of trust among healthcare providers, who may hesitate to rely on black-box machine learning solutions for clinical decisions. Hence, there is a pressing need for more accurate, interpretable, and efficient classification models to improve the early detection of alcohol addiction and enhance the efficacy of technology-based interventions.

In brief, these interventions offer convenience and accessibility, but they often lack predictive capability and depend heavily on user engagement, which may vary. There is a lack of tools that can predict alcohol addiction at an early stage, which is crucial for timely intervention. Current solutions are also limited in their ability to integrate large datasets for comprehensive analysis, and they rarely account for complex relationships between various risk factors. Therefore, the shift has moved towards AI-powerd machine learning models to fill the gaps and enhance the efficacy of technology-based interventions. Machine learning models can analyze vast amounts of data, identifying patterns and risk factors that traditional methods may overlook. By predicting potential addiction risks early, these models can help healthcare providers intervene before the addiction becomes severe, improving treatment outcomes.

2.2.1 Technology-Driven Alcohol Addiction Interventions: A Case Study on Malaysia's Minority Ethnic Communities

A case study has been conducted with a pilot study focusing on the Orang Asli community in Malaysia to understand their specific needs and assess how existing technology-based solutions can be adapted to suit this minority ethnic group. Main purpose is to position this research as a foundation for developing machine learning-based prediction models that cater to diverse real-world contexts and facilitate early identification of alcohol addiction.

Despite the fact that there are several web-based therapies for alcohol consumption management available online around the world, many of them were designed to meet their individual needs. These websites primarily use English as their primary language, which may not be appropriate for the majority of illiterate Orang Asli [19]. Similarly, the authors of [20] stated unequivocally that the indigenous community's failure to receive successful health promotion is due to a lack of culturally specific health promotional materials, a lack of community-based programs, and inefficiency in indigenous health data collection. A study on the prevalence, knowledge, attitude, and practices of noncommunicable diseases (NCDs) among adult Orang Asli and Malay ethnicity in Negeri Sembilan, Malaysia, found that NCDs (including alcohol use) are on the rise among Orang Asli, owing to low rates of healthy lifestyle practices and the need for immediate attention. For disease screening and prevention, adequate NCD education and promotion are essential [21].

According to clinical professionals from University of Malaya Centre for Addiction Sciences (UMCAS) and Orang Asli Gombak Hospital (HOAG), there are few limitations in current alcohol brief interventions. Firstly, there is no centralized data storage and data collection is performed manually in flat files which reflect to data missing in record and prone to human-errors. Secondly, score assessments (such as AUDIT-C and AUDIT-10) is conducted manually which is labor intensiveness and time-consuming. There is no existing workflow for alcohol relapse monitoring and tracking for individual patients to support recovery journey. There is no available dataset to utilize existing advanced technological interventions such as prediction models for relapse rate to better support in recovery curve. All in all, current interventions are lacking supportive features for alcohol abuse recovery such as online community-based programs, socio-cultural factors consideration (user-friendliness) for the minor community and relevant data collection and manipulation for the advantage of new generation solution in technology-based alcohol brief interventions.

According to a study on the digital inclusion of Peninsular Malaysia's Orang Asli, half of those polled believe that ICT will help to improve their health (53%), and 61.9% believe that ICT will bring about changes in their community [22]. Similarly, according to the Department of Orang Asli Development (JAKO), no previous effort had been made in terms of ICT-based intervention for Orang Asli, and signal coverage in Gombak is adequate; additionally, the majority of Orang Asli in the Gombak area use smartphones. This suggests that there is a potential to establish a computerized clinical decision support system for suggesting brief alcohol interventions with treatment-seeking persons from Malaysia's Orang Asli community by leveraging today's ICT readiness and advanced technology.

In brief, the case study observations highlight significant gaps in the current interventions for Malaysia’s minority ethnic groups, such as the Orang Asli. Classic interventions rely heavily on paper-based assessments like AUDIT-C and AUDIT-10, which are not flexible, scalable, or accessible to these communities. Due to their minority status and limited healthcare access, the interventions remain largely reactive, responding to addiction only after it has become a serious issue rather than proactively identifying at-risk individuals. This underlines the urgent need for technology-based solutions that can facilitate early identification of alcohol addiction.

2.3 AI-Powered Predictive Models

AI-powered models have been explored in healthcare for various applications, including addiction prediction. Machine learning classification models are instrumental in classifying health data into predefined categories, which can assist healthcare professionals in making informed decisions [23-24]. The literature reveals a variety of machine learning techniques, such as supervised learning, unsupervised learning, and reinforcement learning, which have been adapted for use in healthcare applications [25-26]. In the context of alcohol addiction prediction, deep learning classifiers have shown high accuracy in identifying alcoholics and classifying drinking behaviors from health records [27-28]. However, their black-box nature makes clinical interpretation challenging [29].

On the other hand, classification models such as random forest and machine learning methods like qualitative interaction trees and group LASSO interaction nets are employed for predicting alcohol addiction and treatment response [30]. The limitations of these models include challenges with data randomness, imbalanced datasets, and the need for more comprehensive integration with clinical practices to enhance their applicability and effectiveness in real-world settings [31]. Study conducted by [27] found that single dependent learners exhibited greater dominance compared to deep learning algorithms in the context of AUD research. Furthermore, [32] observed single dependent learners had a restricted ability to learn and perceive information. Likewise, [33] reported kernel-based methods guarantee structural risk minimization and global optimal solutions, but they may struggle with large-scale datasets due to computational complexity.

Similarly, [28] claimed that instance-based classifiers are often simple and interpretable but can suffer from high storage requirements and sensitivity to irrelevant features. Whereas deep learning methods can automatically learn complex features and have achieved state-of-the-art performance in many domains, but they require large amounts of labeled data and are less interpretable [34]. Hence, each single dependent learner has its own strengths and weaknesses [27][28][33][34] while deep learning classifiers offer high predictive accuracy for alcohol addiction prediction, their lack of interpretability is a significant limitation in a clinical context. Kernel-based classifiers offer better interpretability and theoretical guarantees but may face scalability issues. Instance-based classifiers are simple and interpretable but may not handle complex data as effectively.

Likewise, while DL and ensemble models offer advanced capabilities for predicting complex phenomena such as alcohol addiction, their inherent complexity can impede interpretability. This presents a challenge in clinical settings where explainability is essential for model acceptance and ethical decision-making. Efforts to enhance the interpretability of alcohol addiction classifiers without sacrificing performance are therefore of paramount importance in the field of alcohol addiction prediction where understanding model reasoning is critical [35-36].

In brief, most models focus on specific types of addiction, such as opioid use, or general health conditions, rather than alcohol addiction specifically. Additionally, these models often use single-classifier algorithms, which may not perform well across heterogeneous datasets. Therefore, AI-powered solution which can improve accuracy and generalizability, is required to capture the complex and multifactorial nature of alcohol addiction better than traditional AI models.

2.4 State-of-art Methods in AUD predictions

In conducting a synthetic literature review, a comprehensive collection of similar papers was gathered and categorized based on various themes related to alcohol addiction prediction. Each paper was critically reviewed and analyzed to identify prevailing methodologies and findings, providing a cohesive understanding of the current landscape in this research area.

Support Vector Machines (SVMs) have been employed for alcohol addiction prediction without substantial feature engineering efforts. While studies such as those by [37] demonstrate SVM's effectiveness in classifying alcohol use disorder, they often overlook the necessity of feature selection to enhance model interpretability and generalization. The reliance on SVMs without feature engineering may lead to suboptimal performance and poor interpretability which may hinder their practical application in clinical settings. Therefore, while SVMs can achieve reasonable accuracy, their usability in real-world scenarios is compromised, highlighting the need for enhanced methodologies.

The K-Nearest Neighbors (KNN) algorithm has also been applied to alcohol addiction prediction, with studies indicating its simplicity and ease of implementation. However, research by [38] shows that KNN lacks robustness and efficiency when applied without feature engineering, underscoring the importance of preprocessing in predictive modeling. The KNN algorithm alone suffers from inefficiencies in scalability, particularly with large datasets [39], that limits its utility for healthcare practitioners.

Multilayer Perceptrons (MLPs) are recognized for their potential in predicting alcohol addiction. Nevertheless, studies such as those by [27] that without proper feature engineering, MLPs may not fully utilize their capabilities, resulting in lower accuracy and slower convergence. MLP showcase their ability to model non-linear relationships; yet, they often yield models that are difficult to interpret [40]. The inherent complexity of MLPs can also lead to overfitting, impacting their accuracy when applied to new data [41]. As a result, this highlights a significant gap in interpretability and generalization that limits their applicability in clinical contexts.

Logistic Regression has been widely used for alcohol addiction prediction; however, recent studies indicate that the absence of feature engineering can limit the model's effectiveness. Research by [42] highlights how logistic regression may yield biased predictions due to unaccounted interactions among features. LR is commonly utilized for binary classification tasks like predicting alcohol use disorder; however, its simplicity may result in oversimplified models that fail to capture the intricacies of addiction [43]. Moreover, while LR is interpretable, its accuracy can be insufficient when handling complex datasets, leading to poor predictive performance. Consequently, this points to a gap in balancing interpretability and accuracy within LR approaches.

Decision trees (DTs) are frequently applied for their interpretability and ease of use in predicting alcohol addiction; however, they are prone to overfitting, particularly when dealing with imbalanced datasets [40]. While they provide clear visualizations of decision-making processes, the accuracy of predictions can be compromised, limiting their effectiveness in real-world scenarios. This illustrates a gap in accuracy that detracts from the potential benefits of DTs. Moreover, DTs are frequently employed in predicting alcohol addiction, but studies show that their performance can degrade without feature engineering.

Random forests (RFs) have been leveraged for alcohol addiction prediction, exhibiting high accuracy through ensemble learning; yet, their interpretability is often criticized due to the complexity of multiple decision trees [27]. Additionally, while RFs manage the issue of overfitting better than single decision trees, they still face challenges in computational efficiency and can produce models that are not easily interpretable for practitioners. Research by [44] indicates that while Random Forest can manage high dimensionality, lacking appropriate feature selection leads to inefficiencies and reduced accuracy.

On the other hand, there are improved single classifiers to achive higher perofmrnace of the prediction models. Among latest studies, state-of-art approaches are regarded as the use of Lasso with SVM and Random Forest Regularization on SVM for alcohol addiction prediction. The integration of Lasso regression with SVM has emerged as a promising approach in predicting alcohol addiction. Studies such as those by [45] indicate that Lasso can enhance the SVM model's performance by promoting sparsity and addressing multicollinearity among features. Utilizing Lasso with SVM can improve model performance, yet the reliance on a single model still raises concerns regarding interpretability and generalizability. Incorporating Random Forest regularization with SVM has shown to improve prediction accuracy in alcohol addiction cases. Research by [46] highlights that this combined approach can enhance the robustness of predictions by effectively managing feature interactions. Although SVM with Random Forest regularization enhances accuracy, it may still lack the interpretability required for practical applications, revealing an existing gap in developing interpretable predictive models.

Besides, Stacking models have been employed in the prediction of alcohol addiction, where multiple base classifiers are combined to improve predictive performance. While stacking has shown improvements in accuracy over single classifiers, the complexity of model aggregation can hinder interpretability, making it challenging for practitioners to understand the contribution of each base model to the final prediction [47]. This illustrates a gap in interpretability, highlighting the need for an improved stacking approach that can effectively balance accuracy with comprehensibility, thereby enhancing the applicability of stacking models in clinical settings.

Likewsie, Bagging techniques, particularly the random forest algorithm, have been applied to alcohol addiction prediction, achieving high accuracy by reducing variance through ensemble learning. However, the interpretability of bagging models remains an issue, as the aggregation of numerous decision trees often results in a “black box” effect that obscures the decision-making process [48]. Consequently, this presents an opportunity for enhanced classification algorithm to enhance the interpretability of bagging models while maintaining or improving accuracy, ultimately making them more useful for practitioners.

Similarly, Boosting models have demonstrated superior predictive capabilities in the context of alcohol addiction by sequentially focusing on misclassified instances. Despite their high accuracy, boosting algorithms, such as AdaBoost and XGBoost, can suffer from reduced interpretability due to their complex nature and the influence of numerous weak learners in the ensemble [49]. Therefore, integrating a relevant meta models could improve both interpretability and efficiency, enabling better decision-making and actionable insights for healthcare providers working with alcohol addiction cases.

2.5 Gap analysis

While several studies have explored machine learning for addiction prediction, very few focus specifically on alcohol addiction. Additionally, existing models do not fully address the problem of early-stage detection, and many are limited by the quality and scope of the training data. Similar works often lack interpretability, making them difficult for healthcare professionals to adopt in real-world settings.

Machine learning models for predicting alcohol addiction often either sacrifice interpretability due to their sophistication or lack necessary complexity because they are overly simplistic. This imbalance can hinder trust in real-world applications, as healthcare professionals may struggle to explain complex predictions to patients, ultimately limiting the effectiveness of interventions. Conversely, overly simplistic models can lead to significant inaccuracies, including high false-negative rates, which misclassify at-risk individuals as safe, particularly in underserved communities where timely support is crucial.

Additionally, the datasets used in training these models are often hard to find and can be imbalanced, reflecting the rarity of positive cases compared to negative ones. Thus, it is essential to strike a balance that maintains both interpretability and accuracy, as this would improve the efficiency of interventions in real-world scenarios. Hernce, the existing literature emphasizes the need for a meta-classifier approach that can enhance interpretability, accuracy, and efficiency—essential elements that are currently not adequately addressed in existing single-classifier algorithms [50].

In brief, common gaps across related studies underscore the necessity for a improved prediction model, as existing solutions fall short in addressing the intertwined gaps of interpretability, accuracy, and efficiency and usability in real-world context. This prompts the consideration of meta classification algorithm to markup the deficiency of each model and bring the best combination of classifiers for improved prediction performance.

By offering a more proactive approach, this solution will be the first of its kind in Malaysia, with the potential to be adopted across other underserved regions in Southeast Asia, such as Myanmar, ultimately contributing to the betterment of healthcare systems in these areas. to improve public health domain in Malaysia and examine the benefits of intervention for patients and healthcare providers towards future advanced in health science. In short, through this synthetic review, the gaps in current approaches are systematically identified, establishing the need for a more advanced and reliable solution using an AI-powered meta-classifier.

3.0 METHODOLOGY

This methodology section outlines the research framework for developing an improved classification algorithm using an enhanced meta-classifier, which integrates aggregated bootstrapped heterogeneous models, including kernel-based, instance-based, and deep learning techniques. The section will detail the data sources utilized in the study, the development of the algorithm, the validation metrics employed to assess its performance, and the rationale for selecting these specific methods. The overview of machine learning experiment for the proposed AI-powered intervention design is discussed in below sections.

3.1 Overview of the proposed model

The proposed solution for alcohol addiction prediction leverages a robust methodology that begins with the collection of real-world data from both underage individuals (under 18) and adults (above 18). This dataset encompasses various dimensions, including demographic information, behavioral patterns, clinical data, and specific factors related to addiction. The diverse nature of the data ensures that the prediction model is not only comprehensive but also tailored to address the unique needs of different populations. The training of the prediction model aims to bridge the knowledge gaps identified in the literature review, focusing on enhancing accuracy, interpretability, and efficiency—critical factors for successful implementation in real-world scenarios.

To harness the advantages of heterogeneous models, the methodology employs a combination of base models from various categories, such as kernel methods, instance-based models, and deep learning algorithms. Each of these models undergoes a process called bootstrap aggregation (bagging) to improve stability and reduce variance, ultimately enhancing prediction performance. The conducted experiment involved comparing single classifiers, traditional stacking classifiers, and the proposed metaclassifier on kernel-based bagging, instance-based bagging and deep-learning based bagging. By aggregating predictions from multiple models, the approach aims to balance their respective benefits and drawbacks, enabling the meta-model to utilize the full potential of these varied algorithms, thus overcoming the limitations typically associated with single classifiers.

Unlike traditional methods that may lack interpretability or accuracy, enhanced meta-classifier balances the two ratios and delivers performance metrics comparable to existing state-of-the-art approaches in alcohol addiction prediction [27]. Compared to classic interventions, which often rely on static assessments and may overlook the nuanced changes in a patient's condition, the proposed AI-powered solution provides a dynamic, responsive system that adapts to individual needs and circumstances. Moreover, the data collected through user interactions with the web application can be utilized for retraining the model, enabling it to adapt and improve continuously based on real-world feedback and outcomes.

Once trained on preprocessed data, the prediction model is embedded within a web application that is accessible to users. This platform will enable users to self-assess their risk for alcohol addiction and receive tailored recovery recommendations. Notifications will be sent to assigned healthcare providers when users are identified as being at risk, allowing doctors to review cases and recommend appropriate interventions. The visibility of benefits will manifest in improved patient outcomes, increased engagement in recovery programs, and more effective use of healthcare resources. This AI-powered approach significantly improves the traditional intervention processes by providing real-time feedback and personalized care plans, ultimately enhancing the likelihood of successful recovery.

In brief, this research covers end to end solution with focus on improvement in classification algorithm to deliver AI-powered alcohol addiction intervention solution, firstly: the advancement in predictive model and secondly: the development of web-based intervention. The end goal is to deliver AI-powered predictive web-based intervention that performs three major tasks as learn, predict, and improve i.e., to learn the factors related to addiction and intervention, to predict better relapse prevention by artificial intelligence and to improve the AUD recovery process. The steps and objectives involved in the proposed system design are illustrated as in Fig. 1.

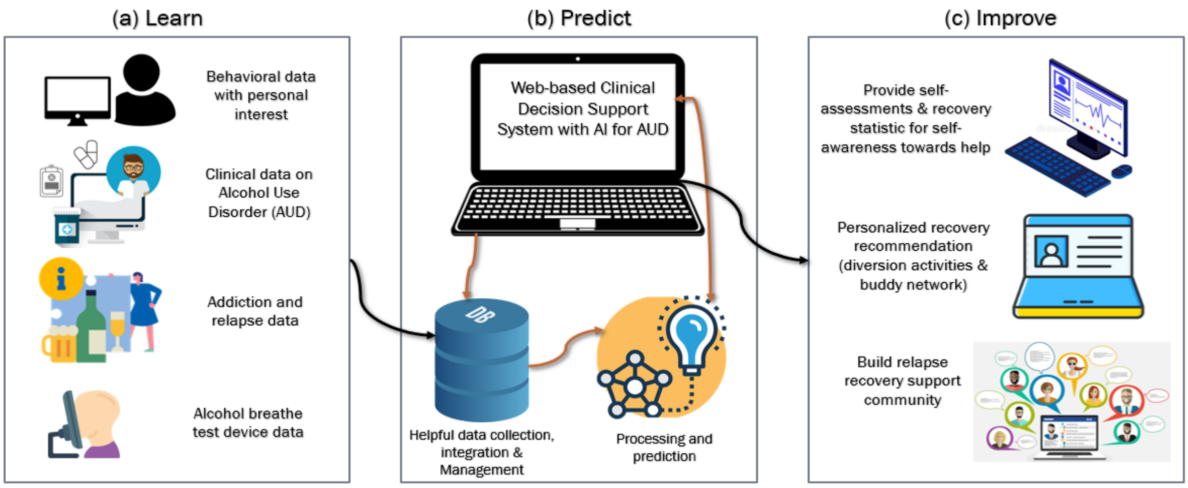

Fig. 1. Workflow and objective of proposed solution

3.2. Datasets

There are two datasets used in this paper i.e., Student alcohol addiction dataset and Adult Drug addiction dataset. Both datasets are widely used and cited in related studies, validated for the applicability for predictive classification experiments [51-54)]. The dataset collected is primary data with random sampling from publicly avaiable repository, Kaggle. Both datasets have an imbalanced class distribution, where records indicating addiction are significantly fewer than those for non-addiction. This imbalance is common in addiction-related datasets and presents challenges for model performance. To address this, stratified sampling is employed during the data split, ensuring that each fold in the training and testing sets maintains a proportionate representation of both classes. This approach helps the model generalize better, preventing bias towards the majority class.

3.2.1. Dataset 1

For this research, the primary dataset is student alcohol dataset. For student alcohol dataset, there are 33 columns with 1024 instances [51-53]. The data was obtained from two Portuguese secondary schools during the academic year 2005-2006, comprising 395 records from mathematics class and 649 entries from Portuguese language class. The dataset (Cortez & Silva, 2008) was compiled by authors who devised a survey with questions pertaining to demographics, social interactions, emotional well-being, and school-related matters. The study focused on two crucial inquiries: alcohol usage during workdays and weekends.

3.2.1. Dataset 2

Although the primary dataset used in this study is a student dataset, an additional alcohol dataset was utilized to repeat a subset of experiments to ensure the generalizability of the model's performance across different populations. Hence, another dataset named Adult Drug addiction dataset is collected, which consists of 32 columns with 1885 instances [54].

3.3. Data Pre-processing

After the data collection, a few steps are performed to prepare for feature engineering, such as data cleansing, data transformation, data pre-processing, labelling and splitting the data into seen and unseen for model validation as 80:20. For data pre-processing, the student dataset is filtered to include only records where the age is less than 18 to predict underage alcohol addiction, while for the adult dataset, individuals aged 18 and above are retained to focus on adult addiction prediction.

Both datasets undergo similar cleansing steps, including handling missing values, correcting inconsistent data entries, and normalizing features like demographic and behavioral data to ensure consistency. Additionally, any outliers are removed to maintain data quality. The student dataset lacks the target variable necessary for predicting alcohol addiction. Hence, a new variable called "addiction" was created from preexisting attributes utilizing the method described by [55], as in equation (1).

$$addiction = \frac{[(Walc \times 2)+(Dalc \times 5)]}{7} \quad (1)$$

The target variable, addiction, was subsequently transformed into a binary outcome, as demonstrated in equation (2) using a method suggested by [56]. This is a scale used to assess alcohol addiction, with a value of 1 indicating the presence of alcohol addiction and a value of 0 indicating the absence of alcohol addiction.

$$addiction \geq 3 = 1$$
$$addiction \ < 3 = 0 \qquad (2)$$

On the other hand, the target variable for the adult dataset is derived from the alcohol addiction score, which is calculated based on predefined thresholds in the dataset. This score categorizes individuals as either addicted or not addicted, serving as the key label for prediction in the machine learning model. For both datasets, the data is split into 80% training and 20% testing. Stratified K-fold cross-validation is applied to ensure that each fold has a balanced representation of both classes (addiction vs. non-addiction), mitigating the impact of imbalance during training. These pre-processing steps ensure that both datasets are ready for training the meta-classifier, with the adult addiction prediction tailored to real-world applications.

3.4. AI-powered AUD prediction

3.4.1. Kernel-based Classifier

Kernelized learners in machine learning provide extensible nonlinear hypothesis spaces over functions for learning latent functions from datasets [57]. They sometimes outperform autoencoders and limited Boltzmann machines [58]. However, the selection of an appropriate kernel can be a challenge, and noisy or outlier-laden data can impair kernelized graph-based approaches [59]. The concept of kernelization, which involves transforming input features using a kernel function, is applicable to various classifiers, such as Kernelized Logistic Regression, Kernelized Ridge Regression, Kernelized Naive Bayes, Kernelized Decision Trees, Kernelized Random Forests, Kernelized Neural Networks and Support Vector Machines (SVM).

With the popularity and being the state-of-art model for AUD, SVM is chosen for experimentation in this paper. A kernelized classifier, SVM may translate input data into higher-dimensional spaces using multiple kernel functions. SVM can capture complex decision boundaries and produce effective predictions for non-linearly separable datasets using the RBF kernel to evaluate data point similarity in converted space. SVMs using a Radial Basis Function (RBF) kernel, often known as the Gaussian kernel, use the following mathematical formula as in equation 3, where $\boldsymbol{f(x)}$ is the decision function to predict the class of new instance x, $\propto_{\boldsymbol{i}}$ as coefficients during the training phase, as class label of i, $||\boldsymbol{x} - \boldsymbol{x_i}||^{\boldsymbol{2}}$ as Euclidean distance between instances, as width parameter σ of the RBF kernel and b as the bias term.

$$f(x) = \sum_{i=1}^{n} \propto_i y_i \, exp(-\frac{||x-x_i||^2}{2\sigma^2}) + b \qquad (3)$$

3.4.2. Instance-based Classifer

Instance-based classifiers predict new instances based on their similarity to known instances in the feature space, as shown in Figure 3.5. Instance-based classifiers can handle complex, non-linear decision boundaries without training [60]. They demand large memory and are sensitive to noise, however improved instance selection can alleviate these problems [61]. Enhancing the effectiveness of instance-based classifiers is achieved by accurately identifying the most optimal instances from the dataset. However, the efficacy of instance selection strategies is contingent upon the classifier [62].

KNN is a widely used classifier in the prediction domain due to its robustness, incrementality, and low implementation complexity [63]. KNN is a classic instance-based classifier that assigns new data points to a class based on the majority class of their k-nearest neighbors in the feature space. Equation 4 represents the mathematical formula of the KNN algorithm, where x is the new instance and k is the neighbors considered.

$$f(x) = majority\ class\ of\ k\ nearest\ neighours\ of\ x \qquad (4)$$

3.4.3. Deeep-learning Classifier

Deep learning classifiers are machine learning models composed of multiple layers of interconnected neurons, designed to learn complex patterns directly from raw data. They are widely used in the prediction domain due to their ability to automatically discover intricate patterns in data, enabling high accuracy in tasks such as image recognition, natural language processing, and speech recognition [64]. Their hierarchical representation and feature learning capabilities make them particularly effective for handling large, high-dimensional datasets with complex structures, where traditional machine learning algorithms may struggle.

Out of all deep learning models, MLP is most applied model in health care domain due to its versatility and flexibility to handle various types of data, including textual, numerical, and image data, making them adaptable to diverse sources of information commonly encountered in public health research, such as medical records, social media posts, and imaging scans [65]. MLPs can be scaled up to accommodate large datasets and complex architectures, allowing researchers to leverage increasingly abundant and diverse data sources in alcohol addiction and public health research. Generic equation for MLP is described in equation (5), whereas X is the input data matrix, Y is the output of MLP, $W^{(1)}$ and $b^{(1)}$ are the weight matrix and bias vector for the $l^{(th)}$ layer respectively, σ(.) represents the activation function applied elementwise to the weighted sum, softmax(.) is the activation function for the output layer, often used for classification tasks.

$$Y = softmax\left(W^{(L)}.\sigma\left(W^{(L-1)}.\sigma(\ldots\sigma\left(W^{(1)}.X + b^{(1)}\right)\ldots) + b^{(L-1)}\right) + b^{(L)}\right) \quad (5)$$

3.1.6. Enhanced Meta classifer (EMC)

In the context of this research, meta-classifier is a classifier that is constructed with Base classifiers (Level-0 Models) and Meta-classifier (Level-1 Model), where Level 0 models are trained on the training data and make individual predictions and Level 1 model combines the predictions from the base models to make the final prediction. Enhanced meta classifier is an improved base classifiers with bootstrap aggregated and applied linear meta model for final prediction. Technically, Level 0 is constructed with bootstrap aggregated models from different classification families such as kernel-based (SVM), instance-based (KNN) and deep learning based (NN MLP). Level 1 is equipped with Logistic regression (LR).

For the base models, different types of classifiers from different classification families are chosen as kernel-based, instance-based and deep-learning based, to prevent the highly correlated base models in stacking, which will lead to overfitting. Moreover, bagging is selected over boosting due to its nature of robustness against outliers and reducing of variance. On the other hand, boosting also reduces variance, however, it is not robust against noisy data. Hence, there are four main phases involved as Base model selection, Bootstrap aggregating, Meta-feature creation and meta-learner construction as shown in Fig 2.

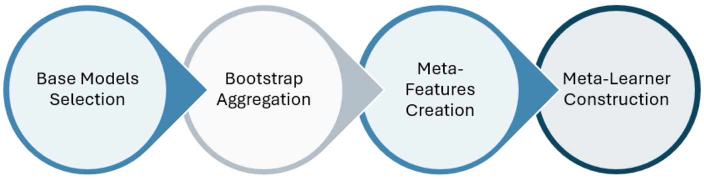

Fig. 2. High Level steps for EMC

On the other hand, below decision matrix is weighted to identify the current gap and generate high level conceptual idea for proposed solution. The chosen criteria of interpretability, accuracy, efficiency, scalability, and complexity are prioritized as they directly impact the practical application and effectiveness of machine learning models in real-world scenarios, especially in sensitive areas like alcohol addiction prediction. Weight represents the relative importance of each criterion in the context of comparing different machine learning models for alcohol addiction prediction decision making process. The weights are assigned based on a combination of expert opinions, relevant literature, and practical considerations in healthcare applications, emphasizing the necessity for interpretability and accuracy in models used for sensitive issues like addiction prediction [27, 40, 50]. Scores indicates a positive or negative impact associated with that model for the respective criterion (1: very low to 5: very high).

Criteria	Single Classifier	Traditional Stacking	Traditional Bagging	Meta Classifier
Intepretability	⇨3	⬇1	⬇2	⇨3
Accuracy	⬇2	⬆4	⇨3	⬆5
Efficiency	⇨3	⬇2	⬇2	⇨3
Simplicity	⬆4	⬇1	⇨3	⇨3
Scalability	⬇1	⬆4	⬆4	⬆4
Total:	⬇13	⬇12	⇨14	⬆18

High
Netural
Low

Fig. 3. Decision Matrix based on gap analysis outcome for determination of conceptual design

Fig. 3. explained rational behind single classifier such as SVM and bagging method such as DT or RF is highly recognized as the state-of-art approach in the alcohol addiction prediction domain, compared to traditional stacking. However, full potential of advtange from each approach can be utilized with meta-classifier. Hence, this research aims to address these gaps by enabling early identification through objective data analysis and prediction, facilitating more targeted and timely interventions. The expected outcome is effective predictive model which can be easily integrated into existing healthcare systems for early identification and intervention in alcohol addiction. This model can overcome the limitations of both traditional and technology-based interventions, providing a robust tool for early prediction and treatment of alcohol addiction. The purpose is to enhance the performance of existing inferior classification algorithms, enabling them to reach the same level of efficacy as superior classification models while maintaining the necessary interpretability for practical application.

3.1.6.1. Base Model Selection

For kernel-based, SVM with Radial Basis Function (RBF) kernel operates on the pre-processed data by projecting the input features (demographic, clinical, addiction-related variables) into a higher-dimensional space using a kernel function (RBF kernel). This transformation allows SVM to handle complex, non-linear relationships by solving a quadratic optimization problem to find an optimal hyperplane that separates different classes (e.g., alcohol addiction or no addiction). For instance-based, KNN processes the data by calculating the distance (Euclidean) between a test sample and the training samples. It then assigns the class label based on the majority vote from its closest K neighbors. KNN excels in identifying local structures in the data, which is particularly useful in detecting small clusters of similar addiction behaviors in the dataset.

For deep learning, MLP as a type of neural network, takes the pre-processed data and passes it through multiple layers of neurons, where each layer applies a non-linear transformation followed by a linear transformation (via learned weights). The training of MLP involves backpropagation, where the model learns by minimizing a loss function (e.g., cross-entropy) using gradient descent. This process adjusts the model weights to improve the prediction accuracy iteratively, to model complex patterns in the data, especially those involving non-linear relationships between variables like demographics and addiction behaviors.

3.1.6.2. Bootstrap aggregation

Bootstrap Aggregation is applied to each base model (SVM, KNN, and MLP). Bootstrapping involves generating multiple subsets of the training data by sampling with replacement. Each model (SVM, KNN, MLP) is trained on these bootstrapped samples, which helps in reducing overfitting by improving model robustness. Bagging creates diversity among the base models and averages their predictions.

Each base model is trained on a different bootstrap sample—a randomly selected subset of the data with replacement. This ensures that the model isn't overly sensitive to small fluctuations in the training data. The predictions from the multiple versions of each base model are then aggregated, using soft voting. Soft voting provides a probabilistic interpretation, allowing the meta-classifier to capture nuanced prediction probabilities instead of just class labels.

3.1.6.3. Meta-Features Creation

After training each base model (SVM, KNN, MLP) on the bootstrapped datasets, they produce probabilistic outputs (soft predictions). These predictions (e.g., probability of alcohol addiction) from each model are collected as meta-features. These meta-features encapsulate the "opinion" of each base model on the given instance. For instance, if SVM predicts a probability of 0.8 for class A, KNN predicts 0.6 for class A, and MLP predicts 0.9 for class A, these probabilities become features for the meta-classifier to analyze. For each test sample, the prediction probabilities from SVM, KNN, and MLP serve as new features (meta-features).

3.1.6.4. Meta-Learner Construction

The final step involves training a Logistic Regression (LR) model on the meta-features. Logistic Regression (LR) is then applied to the meta-features generated by the base models. The role of LR is to learn the best weights to assign to each base model's prediction based on their reliability. Unlike traditional stacking approaches where more complex models (e.g., decision trees) might be used, the use of LR enhances the interpretability of the meta-classifier while maintaining efficiency. It offers insight generation by assigning weights to each base model's prediction based on their performance, effectively learning which model to trust more in certain contexts.

The LR model learns from the meta-features and identifies the best combination of base model outputs to make the final prediction. This layer also adds interpretability, as logistic regression coefficients can be examined to understand how each base model's predictions contribute to the final outcome. LR solves a convex optimization problem, where it learns coefficients that maximize the likelihood of predicting the correct class given the base model outputs. Technically, LR ensures that the contribution of each base model to the final prediction is easily interpretable, making this approach more transparent compared to traditional stacking, which often uses more complex meta-models that reduce interpretability.

In brief, unlike traditional stacking, where each base model contributes directly to the meta-model's input features, here the bagged versions of the base models are treated as separate entities, and LR is trained on their predictions independently. In the context of using logistic regression (LR) as a meta-classifier on bootstrap aggregated base models, each bagged version of the base model (such as SVM, KNN, and NN) is considered as a separate entity or model in itself. Instead of directly combining the predictions of these bagged models into LR's input features, LR is trained independently on the predictions generated by each bagged base model. This means that LR learns to interpret and weigh the predictions from each bagged base model individually, without directly considering their original features or data. This approach allows LR to learn the optimal combination of predictions from the bagged base models, potentially capturing complex relationships between their predictions and improving the overall ensemble's performance. The algorithm for meta-classifier is described below.

Training phase:

For each base model M in {SVM, KNN, NN}: B_M = train_bagged_model (M, D)

For each base model M: $\alpha_M = predict\ (B_M, D)$

$F_{LR} = [\alpha_{SVM}, \alpha_{KNN}, \alpha_{NN}]$

$MC_{LR} = train\_logistic\_regression\ (F_{LR}\ , Y)$

Prediction phase:

For each base model M: $\alpha_{M,new} = predict\ (B_M, D_{new})$

$F_{LR,new} = [\alpha_{SVM,new}, \alpha_{KNN,new}, \alpha_{NN,new}]$

$MC_{LR\ final} = MC_{LR}.predict\ (F_{LR,new})$

Where:

- B_M represents the bagged version of the base model M trained on dataset D.
- α_M represents the predictions generated by the bagged base model M trained on dataset D.
- F_{LR} represents the features used to train the logistic regression model, which consist of the predictions from the bagged base models.
- $\alpha_{M,new}$ represents the predictions generated by the bagged base model M for new data D_{new}
- $F_{LR,new}$ represents the features used to make predictions for new data, consisting of the predictions from the bagged base models for the new data.
- $MC_{LR\ final}$ represents the final prediction made using logistic regression for the new data.

In alcohol addiction classification, it is preferrable to use simple and more interpretable models than complex models due to transparency and interpretability. For example, simple models, such as logistic regression (LR), provide straightforward interpretations of how input features relate to the predicted outcome. This transparency is essential in medical applications like addiction classification, where understanding the reasoning behind predictions is crucial for clinical decision-making.

Simple meta-learner on heterogeneous bagged models combines the advantages of bagging, such as diversity and reduced variance, with the predictive power of stacking, resulting in an ensemble that is often more robust, stable, and interpretable compared to traditional stacking approaches. Meta-classifiers provide a regularization effect, helping to generalize better to unseen data. LR imposes constraints on the model coefficients, which can prevent overfitting and lead to more robust performance on new data. High level illustartion as shown in Fig 4.

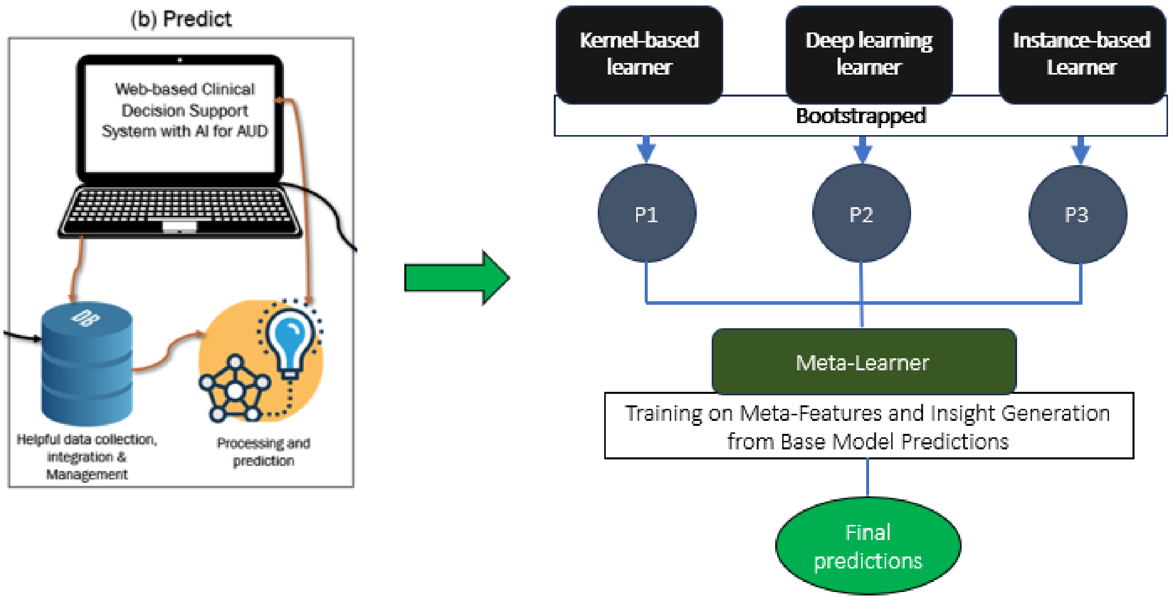

Fig. 4. Workflow of meta classifer approach

In brief, Logistic Regression (LR) as a meta-classifier on SVM, KNN, and NN bagging models is superior to traditional bagging, boosting, or stacking ensemble methods due to its simplicity, interpretability, and efficiency in handling diverse outputs from different base learners. LR excels in combining predictions, mitigating individual biases, and promoting model interpretability, making it a well-suited choice for integrating the diverse strengths of SVM, KNN, and NN bagging models. Training a meta-classifier on the predictions of bagged base models is computationally less intensive compared to training multiple complex base models.

3.5. Evaluation

After predictions are made by the meta-classifier, they are compared with the true labels in the test set. Different metrics are used to assess the performance, such as F1-score, and Area Under the Curve (AUC). Balanced accuracy, a key metric in this case, adjusts for any imbalance in the test data, ensuring that the performance is evaluated fairly. The AUC metric specifically evaluates the model's ability to distinguish between classes, which is critical in cases like alcohol addiction prediction where false negatives (missed cases) must be minimized.

The performance of the meta-classifier is compared with traditional stacking [47] and state-of-art methods [45-46]. This comparison is essential to demonstrate the improvement in accuracy, interpretability, and efficiency. The testing process not only validates the model's predictive capability but also highlights how the meta-classifier outperforms traditional methods in terms of both accuracy and interpretability, thus providing a technically superior solution for real-world alcohol addiction prediction. The end to end process of deployment and evaluation before releasing for practical utilization is depicted as in Fig 5.

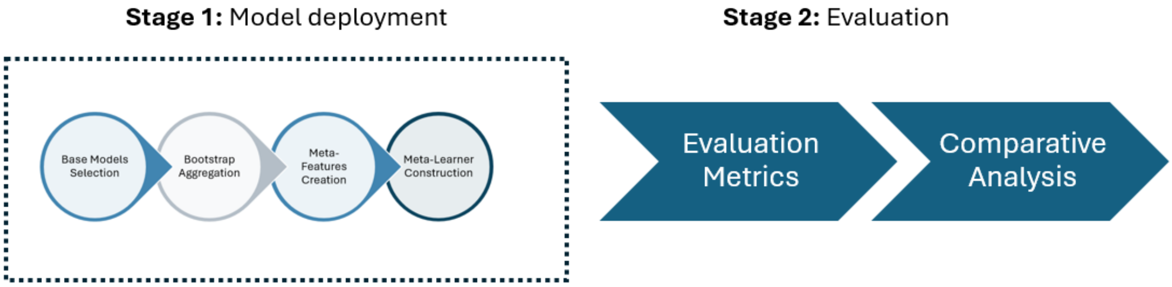

Fig. 5. Workflow for deployment and evaluation

3.5.1. Balanced Accuracy

The evaluation method for this research utilizes the F1 score as a measure of balanced accuracy, as in equation 6. The F1 score is selected over other metrics because it provides a harmonic mean of precision and recall, offering a single metric that accounts for both false positives and false negatives [27]. This is particularly important in the context of alcohol addiction prediction, where both types of errors can have significant consequences. Other methods, such as simple accuracy, may not be as effective in scenarios with imbalanced datasets, which is common in medical informatics. By using the F1 score, the model's performance can be more accurately assessed, ensuring it is both reliable and effective in real-world applications.

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision \times Recall} \tag{6}$$

3.5.2. Area Under the Curve

The Receiver Operating Curve (ROC) or AUC curve evaluates binary classification problems. The True Positive Rate (TPR) versus False Positive Rate (FPR) probability curve at various threshold settings separates the ‘signal’ from the ‘noise.’ Thus, it displays classification model performance at all thresholds [27]. The Area Under the Curve (AUC) summarizes the ROC curve and measures a binary classifier's class distinction ability, as in equation 7, where the Receiver Operation Characteristic curve, and the inverse of the decision threshold applied to the models’ predicted probabilities with integral calculation ranging rom 0 to 1. The model distinguishes positive and negative classes better with greater AUC [20].

$$AUC = \int_0^1 ROC\ (f^{-1}\ (t))\ dt \tag{7}$$

3.6. Methodological Selection and Justification

Traditional stacking typically combines base models without a clear strategy to maintain both interpretability and accuracy [47]. By using a diverse set of models (SVM, KNN, MLP) and incorporating bagging, this method maximizes the strengths of each model (SVM for decision boundaries, KNN for local patterns, MLP for complex non-linear relationships) while mitigating their weaknesses through bootstrapping. The decision matrix for selection of SVM, KNN and MLP is described as in Fig 7.

Criteria	SVM	KNN	MLP
Handling Nonlinear Data	3	2	3
Model Interpretability	2	3	1
Scalability	3	1	2
Tolerance to Noisy Data	2	1	3
Training Speed	1	1	1
Flexibility in High Dimensions	3	1	2
Simplicity	1	3	1
Total:	15	12	13

High
Netural
Low

Fig. 7. Decision Matrix for base models selection

The weighted decision matrix demonstrates that SVM, KNN, and MLP complement each other in critical area and therefore, these three models are ideal for constructing a meta-learner, leveraging the advantages of each to improve performance in alcohol addiction prediction. The novelty also lies in combining bootstrapping with meta-learning in a single unified process, which improves robustness, accuracy, and interpretability—an advantage that traditional stacking and bagging methods do not achieve simultaneously. Bagging not only stabilizes SVM and MLP by averaging predictions from multiple bootstrapped instances but also minimizes variance in KNN, ultimately leading to better generalization and convergence stability.

On the other hand, LR, has low computational overhead, making it suitable for running as meta-learner. LR is chosen to improve the interpretability of the combined predictions from the bagged models, addressing the lack of clarity in how individual models contribute to the final outcome and enhancing overall predictive accuracy. It achieves by transforming complex ensemble predictions into actionable insights, which addresses the interpretability challenge often faced with traditional models. This structured and technically sound approach is designed to handle real-world complexities, such as dataset imbalance, while ensuring the model remains interpretable and efficient for healthcare professionals working on alcohol addiction prediction.

For evalution, F-measure and AUC were used for this experiment due to their ability to capture different aspects of model performance. F-measure is ideal for binary classification applications because it balances precision with recall, which is important when false positives and negatives carry significant consequences. AUC, on the other hand, provides a complete assessment of a model's class discrimination capacity, making it suited for imbalanced datasets. In this study, AUC is prioritized as major metric for comparison because discriminative power is crucial in practice for alcohol addiction prediction in clinical setting. It measures a model's ability to correctly distinguish between different classes, which is essential for reliable decision-making between addicted and non-addicted.

4.0 RESULTS AND DISCUSSION

The experimental results derived from Student dataset, along with a subset of experiments conducted using Adult dataset is presented in this section. A thorough evaluation and comparative analysis of the outcomes is provided, followed by a critical discussion aimed at addressing real-world problems highlighted by this research.

Support Vector Machine (SVM) is recognized as state-of-art model for alcohol addiction [27] and many studies centered improvement around the kernel-based SVM for alcohol addiction prediction state-of-art approahces. [47] conducted traditional stacking which is limited with the evaluation only with recall and observed insiginifcant improvement compared to single Naive Bayes. [45] implemented Support Vector Machine (SVM) with LASSO and PCA, which claimed as benchmark prediction model for the alcohol addiction. [46] introduced SVM on regularized random forest as best performing model for alcohol addiction prediction. Thus, this research selected above studies as comparative studies to evaluate the performance of meta classifer.

4.1. Results from Dataset 1

4.1.1. Experiment 1 - Baseline Comparison: Performance of base classifiers versus meta-classifier

For modelling, the first set of experimentation is performed to compare meta modelling verusus single dependent learners. The experimentation result is reported as in table 1.

Table 1: Baseline Comparison: Performance of base classifiers versus meta-classifier (Student Alcohol Dataset)

Experiment I - Baseline Comparison:	**F-score**	**AUC**
Single kernel-based dependent learner: SVM on all features	0.83	0.64
Single instance-based dependent learner: KNN on all features	0.82	0.51
Single deep learning based dependent learner: NN on all features	0.85	0.68
Enhanced Meta-classifier	0.86	0.79

In this comparison, where Meta-classifier exhibits superior performance in terms of AUC compared to all single dependent learners [66–68], while F-measure results are comparable across all models, it can be inferred that proposed model demonstrates stronger discriminative power in distinguishing between positive and negative instances.

Nonetheless, improved F-measure of 0.86 suggests that the models achieve comparable balance between precision and recall, indicating consistent performance in correctly identifying true positives while minimizing false positives and false negatives. Finding is pragmatic as deep learning model such as NN provides better accuracy due to its nature of automatically discovering complex patterns in data, enabling high accuracy in tasks. Deep learning can handle large, high-dimensional datasets with complex structures, where traditional machine learning algorithms struggle. Therefore, achieving comparable performance with deep learning is a significant improvement for the interpretable meta classifier. Result of F-score across single learners compared to enhanced meta classifier is shown in Fig 8.

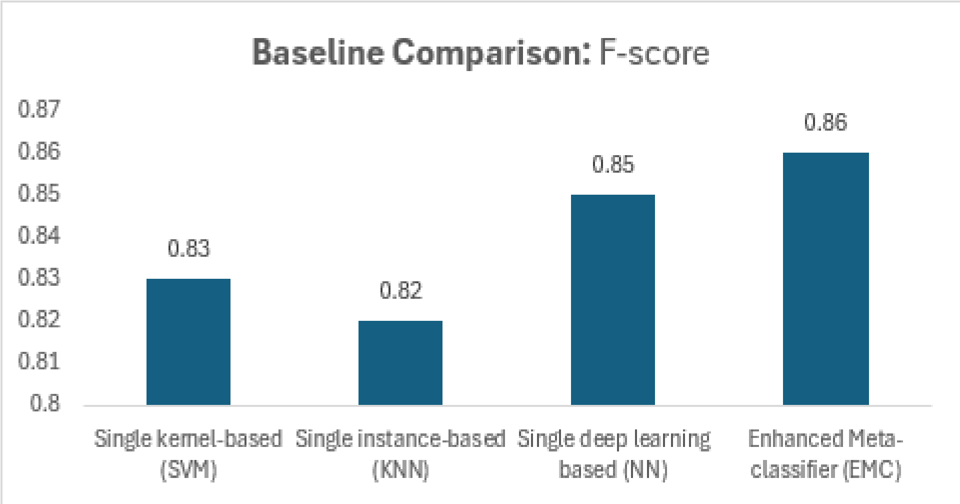

Fig. 8. Baseline performance comparison with single dependent learnerr in terms of balanced accuracy

On the other hand, the findings highlight the proposed model as a favorable choice for tasks prioritizing AUC as the primary evaluation metric, demonstrating its effectiveness in capturing class separation compared to other models. The result clearly demonstrated that AUC achieved 16.18% improvement compared to the closed competitors, which is Deep learning versus metaclassifier with AUC from 0.68 to 0.79, as shown in Fig 9.

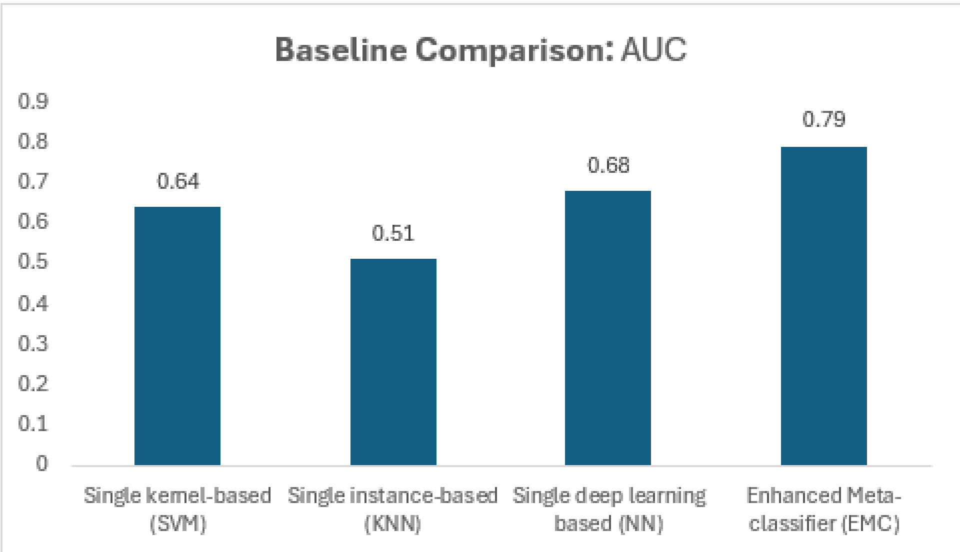

Fig. 9. Baseline performance comparison with single dependent learnerr in terms of discriminative power

Additionally, effective size is calculated using Cohen's d measure for each of the performances to determine the significant difference in performance between the single learners and proposed meta leaner as in Table 2. Effect size provides a standardized measure of the magnitude of the difference between compared models, irrespective of sample size. This helps in understanding the practical significance of the difference between models. Effect size offers more informative and reliable estimates compared to p-values alone. For example, a small difference in predictive accuracy may not be clinically meaningful, even if it is statistically significant. Effect size and confidence intervals are more robust and less sensitive to sample size variations compared to p-values from t-tests, especially in the context of large sample sizes where even small differences can become statistically significant.

Table 2: Effect size between base classifiers and meta-classifier

Effect Size	F-score	AUC
Effect size (NN versus Meta-classifier)	4.24	1.92
Effect size (SVM versus Meta-classifier)	1.01	1.42
Effect size (KNN versus Meta-classifier)	1.06	0.76

For NN versus meta classifier, the effect size is 4.24 for F-score, suggesting a large effect size. This indicates a substantial difference between the two models in terms of F1-score. For the AUC, the effect size is 1.92 indicating a moderate effect size. For SVM versus meta classifier, a Cohen's d value of 1.01 for F-score and 1.414 is considered large, suggesting that there is a substantial difference between the models in terms of both F1-score and AUC. For KNN versus meta classifier, the effect size is approximately 1.061, indicating a moderate effect size. For the AUC, the effect size is approximately 0.758, also indicating a moderate effect size. This suggests a notable difference between the two models in terms of F-score and AUC. All in all, from the aspect of effect size, meta-classifier performs better than single dependent classifiers, regardless of kernel-based, instance-based or deep learning based.

4.1.2. Benchmark Comparison: Performance of baseline papers versus meta-classifier

Subsequently, the experiment is conducted compared to benchmark approaches. For benchmark, traditional stacking model is chosen to set fair comparison as similar technique to proposed solution, followed by traditional bagging to determine performance on bagging compared to enhanced meta classifier. The result is presented in Table 3.

Table 3: Benchmark Comparison: Performance of traditional stacking and bagging classifiers versus meta-classifier (Student Alcohol Dataset)

Experiment II: Benchmark Comparison	F-score	AUC
Traditional Stacking (SVM + KNN + NN + LR)	0.83	0.72
Traditional Bagging	0.76	0.71
Enhanced Meta-classifier	0.86	0.79

As per the result shown in Table 3, the experiment results demonstrate that the Enhanced Meta-classifier outperforms both Traditional Stacking and Bagging techniques. The Enhanced Meta-classifier achieved the highest F-score (0.86) and AUC (0.79), indicating a better balance between precision and recall, as well as superior discrimination ability across class labels. The strength of the meta-classifier lies in its ability to aggregate and learn from diverse base models, leveraging their complementary strengths to address both variance and bias issues, which traditional stacking and bagging alone cannot fully resolve.

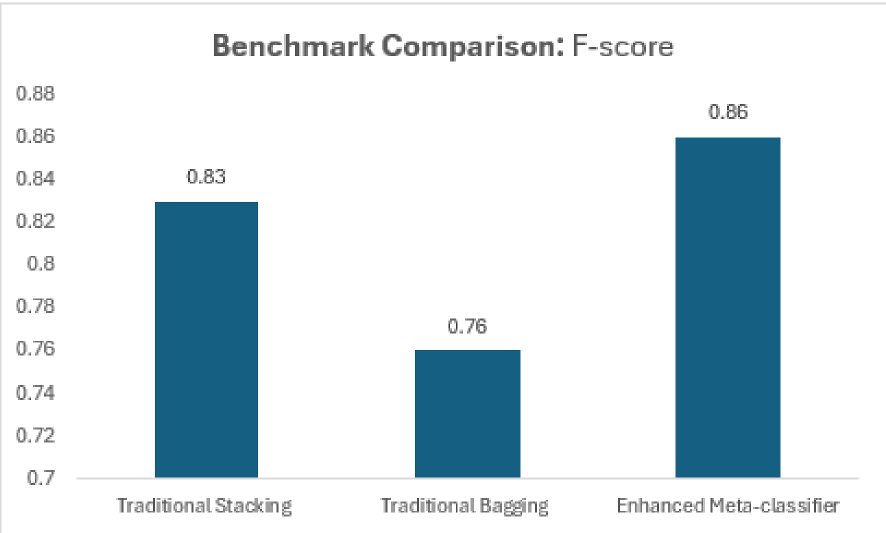

Fig. 10. Benchmark performance comparison with stacking and bagging learners in terms of balanced accuracy

As shown in Fig 10, the F-score measures the harmonic mean of precision and recall, highlighting the model's capability to correctly classify positive cases. The Enhanced Meta-classifier scored 0.86, surpassing traditional stacking (0.83) and bagging (0.76). The F-score for traditional stacking (0.83) is slightly lower than the meta-classifier (0.86) due to the limitations of basic model integration in traditional stacking, where predictions are simply aggregated without deeper insight extraction from base models. Likewise, bagging tends to reduce variance but may fail to correct bias, leading to a lower F-score. Improvement from EMC can be attributed to meta-classifier intelligent learning from base model outputs, leveraging a more interpretable approach to combine them, which enhances the precision and recall trade-off, leading to a higher F-score.

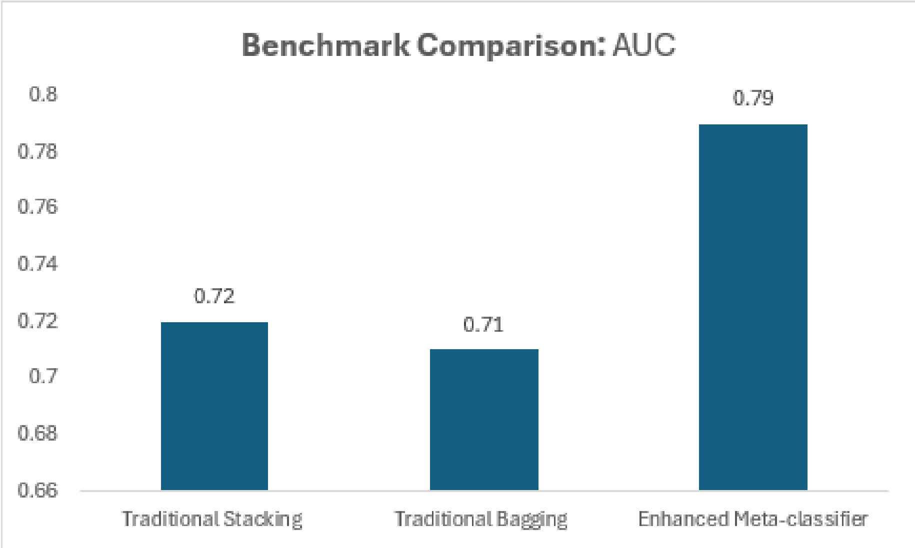

Fig. 11. Benchmark performance comparison with stacking and bagging learners in terms of discriminative power

As in Fig 11, the Enhanced Meta-classifier achieved an AUC of 0.79, higher than stacking (0.72) and bagging (0.71). This suggests that the meta-classifier excels in ranking predictions, benefiting from the fine-tuned ensemble of diverse models. Traditional stacking's AUC (0.72) is lower than the enhanced meta-classifier (0.79) because it treats the stacked models' predictions as static without leveraging their internal dynamics. The meta-classifier applies a more refined learning process using logistic regression on the meta-features, which enables improved ranking and classification, leading to better class separation and, consequently, a higher AUC. . Bagging's lower AUC indicates its lsimitation in addressing bias, whereas the meta-classifier benefits from a more sophisticated combination of base learners.

4.1.3. State-of-art comparison: Performance of state-of-art methods versus meta-classifier

The state-of-the-art approaches chosen for comparison include SVM on Lasso regularization and SVM with Random Forest regularization, which have demonstrated strong performance in predicting alcohol addiction, especially in imbalanced datasets. These methods were selected because they represent the latest advancements, with claims of achieving high classification accuracy in addiction prediction, making them relevant benchmarks for evaluating the proposed meta-classifier.

Table 4: State-of-art Comparison: Performance of state-of-art approaches versus meta-classifier (Student Alcohol Dataset)

Experiment III: State of art Comparison	**F-score**	**AUC**
State-of-art 1 (SVM on Lasso)	0.83	0.57
State-of-art 2 (SVM on Random Forest Regularization)	0.78	0.63
Enhanced Meta-classifier	0.86	0.79

The results from Table 4 show that the Enhanced Meta-classifier outperforms both state-of-the-art approaches—SVM on Lasso and SVM on Random Forest Regularization—in terms of both F-score and AUC. The Enhanced Meta-classifier achieved an F-score of 0.86 and AUC of 0.79, surpassing the best F-score (0.83) from the SVM on Lasso and the best AUC (0.63) from the SVM on Random Forest Regularization. This improvement demonstrates that combining heterogeneous models through meta-learning and bagging is highly effective in dealing with noisy and imbalanced datasets, especially for the complex problems with inbalanced data.

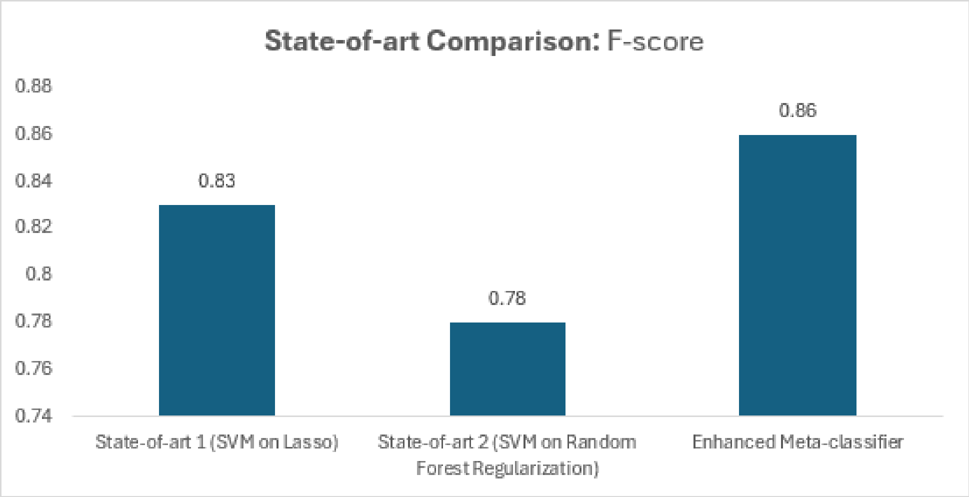

Fig. 12. State-of-art performance comparison with SVM on Lasso and RFR learners in terms of balanced accuracy

As in Fig 12, when focusing on the F-score, the Enhanced Meta-classifier shows clear improvement over the state-of-the-art. The SVM on Lasso and SVM on Random Forest Regularization methods show lower F-scores (0.83 and 0.78, respectively), mainly due to their inability to handle diverse data dimensions as effectively. The Enhanced Meta-classifier benefits from combining SVM, KNN, and MLP, which brings a robust balance of interpretability, scalability, and accuracy, thus improving the overall prediction quality.

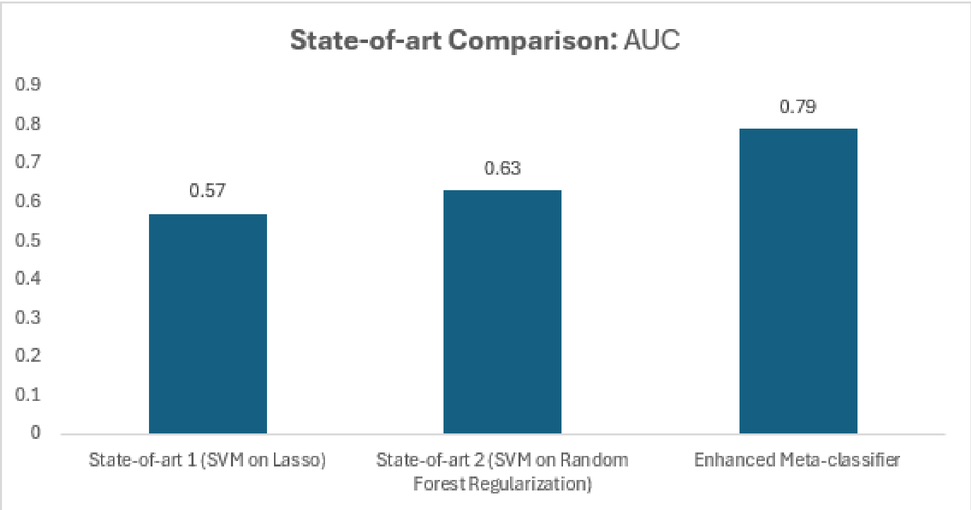

Fig. 13. State-of-art performance comparison with SVM on Lasso and RFR learners in terms of discriminative power

As in Fig 13, for the AUC metric, which assesses the classifier's ability to distinguish between classes, the Enhanced Meta-classifier again outperforms (AUC of 0.79). The SVM on Lasso (AUC 0.57) and SVM on Random Forest Regularization (AUC 0.63) show lower performance, which can be attributed to their susceptibility to noise and inability to capture complex non-linear relationships in the dataset. The ensemble nature of the Meta-classifier, supported by bagging, allows for better generalization and improved performance in distinguishing between addicted and non-addicted individuals.

However, one significant findings resultant here as in this experiment, SVM on Random Forest Regularization (RFR) shows a higher AUC (0.63) but a lower F-score (0.78) compared to other models. RFR's ability to generalize well comes from its ensemble nature, which helps in capturing the broader, global patterns in the data, resulting in better distinction between classes. This is why the AUC, which measures the model's capacity to separate addicted from non-addicted cases across all thresholds, is higher for SVM with RFR. The regularization helps control overfitting and improves the model's ability to handle imbalanced datasets, making it more robust in distinguishing classes.

The F-score is more sensitive to the balance between precision and recall, and the RFR's ensemble mechanism might prioritize recall over precision, causing it to perform worse in situations where a balanced trade-off between false positives and false negatives is critical. This leads to a lower F-score, as it may correctly classify more cases (hence the higher AUC) but struggle to optimize for the balance needed for a high F-score. Thus, SVM on RFR captures the data's global structure well, leading to a higher AUC but suffers in terms of precision-recall trade-offs, resulting in a lower F-score.

4.2. Results from Dataset 2

4.2.1. Baseline comparison: Performance of baseline papers versus meta-classifier

Experiments are selectively repeated with different datasets to validate model performance across varied data distributions and ensure the robustness of findings. Not all experiments are repeated due to resource constraints, time limitations, and the need to prioritize experiments based on their significance and potential impact. Selective repetition allows for efficient allocation of resources while still ensuring thorough validation of key findings and insights. The findings from the second dataset are summarized in Table 5.

Table 5: Baseline comparison: Performance of baseline papers versus meta-classifier (Adult Drug Dataset)

Experiment IV: Adult Drug Dataset	**F-score**	**AUC**
Traditional Stacking (SVM + KNN + NN + LR)	0.79	0.55
State-of-art 1 (SVM on Lasso)	0.81	0.54
Meta-classifier	0.89	0.68

Similar to the findings from Dataset 1 (Student dataset), the meta-classifier demonstrates the highest F-score of 0.8921, indicating superior precision and recall balance, while also achieving the highest AUC of 0.68, suggesting strong discrimination power. compared to traditional stacking, and SVM on Lasso, the meta-classifier. This finding (as in Fig 14) declares that meta-classifier combines the predictions of multiple base models, potentially capturing more complex relationships within the data compared to simpler individual models or traditional ensemble methods. It also aggregates predictions from multiple base models using a weighted scheme, allowing it to harness the strengths of each base model while mitigating their weaknesses, thereby improving overall performance.

Synthetically, the meta-classifier provides consistent result on both dataset 1 and dataset 2. In both experimental scenarios, a common observation is that the proposed model achieves higher AUC scores compared to other metrics, indicating superior discriminative power. This performance edge is especially important in alcohol addiction prediction, where reducing false negatives is critical. A model with high discriminative power minimizes missed cases of addiction, ensuring that individuals at risk are identified and addressed. False negatives in this context mean failing to detect early-stage or ongoing addiction, which can lead to delayed intervention, worsening health outcomes, and increased strain on healthcare resources. In medical diagnostics, where timely detection can prevent severe consequences, a low false-negative rate is essential. By reducing missed cases, the proposed model contributes directly to more effective interventions and improved public health outcomes, emphasizing the importance of prioritizing AUC as a key metric.

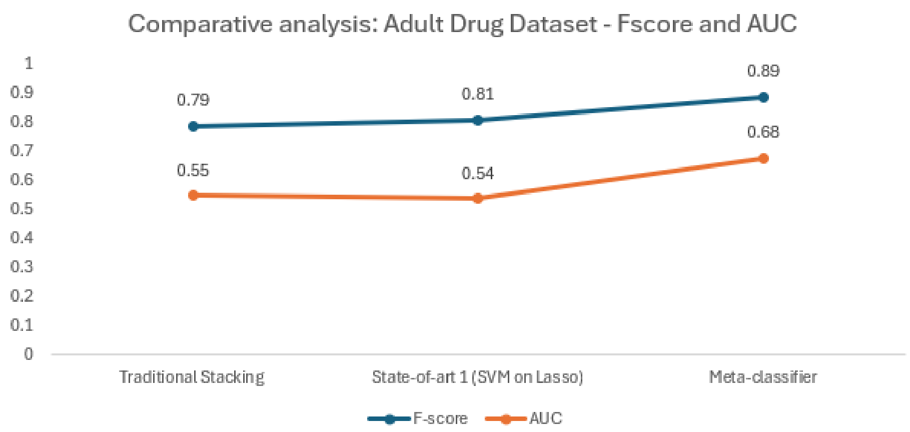

Fig. 14. Comparative analysis with Adult Drug dataset on benchmark and state-of-art learners

However, compared to dataset 1, all models perform relatively poorer on dataset 2, with lower AUC values and marginally reduced F-scores. This suggests potential dataset-specific challenges or differences in data characteristics between the two datasets, highlighting the importance of evaluating model performance across multiple datasets to ensure robustness and generalizability of findings.

4.3. Comparison with previous studies

The outcome of experiments on two datasets affirmed the problem statements mentioned in this research can be resolved with the proposed solution. According to the author's knowledge, there is no previous work on the enhanced meta classifier approach. Therefore, phase by phase experiment is performed on three different scenarios (i.e., baseline, benchmark and state-of-art) on two different datasets for evaluating the performance with proposed approach.. Hypothesis is preserved as enhanced meta classifier produces better discriminative power and balanced accuracy compared to traditional methods, while maintaining the accuracy, interpretability and efficiency. The comparison with previous studies are summarized as in below Table 6.

Table 6: Comparative analysis with previous studies compared to enhanced meta classifier

Papers	Approach	F-score	AUC	Efficiency	Interpretability
Park et al. (2021) [67]	Single kernel-based: SVM	0.83	0.64	1-5 mins (RBF kernel adds complexity, quadratic scaling)	Low (RBF kernel makes it harder to interpret)
Xiong et al. (2023) [66]	Single instance-based: KNN	0.82	0.51	<1 min ($O(n^2)$ scaling with dataset size)	High (Simple to interpret based on distance metrics)
Kumari et al. (2018) [68]	Single deep-learning: MLP	0.85	0.68	3-7 mins (high computational complexity)	Low (black-box, challenging to interpret)
Buniyamin (2022) [47]	Traditional Stacking	0.83	0.72	5-10 mins (multiple models combined, increases training time)	Low (difficult to interpret when many models are stacked)
Ebrahimi et al. (2021) [69]	Traditional Bagging	0.76	0.71	5-8 mins (depends on base models)	Medium (improves variance control, moderate interpretability)
Priya & Thilagamani (2022) [45]	SVM on Lasso	0.83	0.57	1-5 mins (Lasso regularization improves efficiency)	Medium-High (better interpretability due to feature selection)
Enhanced Meta Classifier	Bootstrap aggregated SVM, KNN, MLP with Linear meta learner	0.86	0.79	1 – 7 mins (parallelization possible)	Medium-High (meta-learner LR improves interpretability of ensemble)

As in table 6, improvements from enhanced meta classifer underscore the efficacy of the metaclassifier approach in enhancing model discrimination capabilities, surpassing the performance of alternative methodologies [45-47]. Interpretability concern is also supported by latest studies for deep learning models such as [70] highlighted it can be complex and require substantial computational resources, which might not be feasible in all healthcare settings. Likewise, [71] claimed that implementing and interpreting ensemble or bagging models like Random Forest can be challenging, especially in clinical settings without technical expertise. In contrast, with these compelling results, the proposed approach not only demonstrates a substantial improvement in accuracy but also advances the forefront of machine learning applications in the realm of public health and presents a promising new research avenue in the machine learning domain.

Henceforth, the proposed meta-classifier outperforms traditional stacking methods by 10% in terms of the F1 score due to its innovative approach to generating and leveraging meta-features. Unlike conventional stacking, proposed method involves training the LR model on meta-features derived from different subsets of the training data, which allows for a more nuanced understanding of the relationships between the base model predictions. This diverse subset approach enhances the meta-learner's ability to generalize better across various scenarios. Additionally, the LR model in proposed approach examines how each bootstraped base model (SVM, KNN, NN) performs on these subsets, providing a richer set of meta-features compared to traditional stacking methods.

Furthermore, by combining these predictions into meta-features, the LR model can identify patterns and interactions between the base models that would be overlooked in a traditional stacking setup. This comprehensive insight generation allows the meta-learner to correct biases and errors more effectively. In contrast, standard methods like SVM on Lasso or random forest regularization often rely on a single perspective of the data, which can lead to suboptimal performance in complex datasets. The proposed method's ability to dynamically adapt to the strengths and weaknesses of each base model on different data subsets results in a more balanced and accurate final prediction.

Moreover, traditional stacking techniques may suffer from overfitting due to their reliance on the entire training dataset for meta-features, while proposed method's subset approach mitigates this risk by promoting diversity in the training process. This robustness against overfitting contributes significantly to the improved F1 score. The enhanced interpretability of meta-classifier also allows for better model diagnostics and refinements, ensuring continuous performance improvements. Overall, the proposed method's strategic use of meta-features and comprehensive insight generation leads to superior predictive performance, as evidenced by the 10% increase in the F1 score.

In brief, the enhanced meta-learner embraces the strengths of each base model (SVM for handling complex decision boundaries, KNN for local instance-based learning, and MLP for capturing deep, non-linear relationships) and addresses their individual weaknesses (overfitting, noise sensitivity, complexity) through bagging. The final use of Logistic Regression as a meta-learner offers an additional layer of interpretability and stability compared to traditional stacking or bagging approaches, making it more robust and explainable for real-world applications. This balance of improving interpretability without sacrificing performance makes this approach superior.

4.4. Future work: Application to the real-world problems

The ongoing issues surrounding alcohol abuse in the Orang Asli community, Malaysia, such as missed early-stage detection, reliance on subjective assessments, laborious evaluation processes, and limited healthcare access, can be effectively addressed through the findings from this study. By employing machine learning (ML)-based predictions, early detection of alcohol abuse, and timely intervention for at-risk individuals can be promoted. Improved prediction model offers objective and standardized results, significantly enhancing accuracy over traditional methods. Additionally, automating the assessment process alleviates the burden on healthcare practitioners, enabling them to focus on patient care rather than administrative tasks. Finally, the proposed solution is designed to be efficient and interpretable, making it easy to implement within existing healthcare systems, thereby increasing accessibility for remote communities and prioritizing the urgent need for effective intervention.

The integration of our ML-based predictive model into a web-based solution can significantly enhance accessibility and usability for both healthcare practitioners and community members. A user-friendly interface can be designed to allow healthcare workers to input relevant patient data easily, with the system providing real-time predictions and risk assessments. Visual dashboards could display critical metrics, such as individual risk scores for alcohol abuse, trends in community health data, and alerts for high-risk cases requiring immediate attention. Such a system would not only streamline the assessment process but also facilitate better communication among healthcare providers, enabling them to collaborate effectively in addressing patient needs

In brief, the findings from this research have the potential to create a profound impact on the Orang Asli community by improving the detection and management of alcohol abuse. With the existing strong use-case from Hospital Oang Asli Gombak (HOAG), this research findings can be implemented as prototype to enable betterment in Alcohol addiction management among Orang Asli with the expected outcome to significantly reduce the instances of missed early-stage alcohol abuse, leading to more effective prevention strategies.

5.0 CONCLUSION AND FUTURE WORK

Alcohol addiction remains a pressing public health issue that affects countless individuals and communities, necessitating effective early detection and intervention strategies. Although technology-based interventions remains popular, existing machine learning models for alcohol addiction often struggle to find the right balance between accuracy, interpretability, and efficiency, limiting their practical application in real-world healthcare settings. Hence, in this enhanced meta-classifier approach, combining heterogeneous models (SVM, KNN, MLP) ensures a balance between robustness and diversity, while using logistic regression as the meta-learner provides a clear, interpretable model. By using bagging at the base model level and logistic regression at the meta-learning stage, the solution maintains interpretability, improves accuracy through diversity, and ensures efficiency by leveraging soft voting and meta-features, thus overcoming the limitations of traditional approaches.Therefore, the proposed meta-classifier represents a significant advancement in predictive modeling by introducing a novel approach that enhances model performance through improved meta-feature generation and insight extraction. This method outperforms state-of-art and traditional techniques by 10.13% in balanced accuracy and 9.72% in discriminative power.

The practical application of this research is realized through its integration into a web-based intervention, providing a user-friendly tool for healthcare professionals and patients. The theoretical contributions include a deeper understanding of meta-classification dynamics and its impact on model accuracy. The study's limitations include the reliance on a limited number of datasets, which may impact the generalizability of the model across diverse populations. Additionally, implementing the model in underserved areas presents challenges due to limited infrastructure and resource constraints, potentially affecting accessibility and adoption rate.

Future work will involve extending experimentation to diverse datasets and incorporating additional evaluation metrics such as Log Loss and confidence intervals to further validate the model's robustness. Additionally, a prototype of the AI-powered web-based intervention is aimed to develop specifically for Hospital Orang Asli in Malaysia, enhancing its relevance and applicability in real-world healthcare settings. This ongoing work will refine the model and intervention, ensuring both practical and theoretical advancements in the field of medical informatics.

ACKNOWLEDGEMENT

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

REFERENCES

[1] Heilig, M., Augier, E., Pfarr, S. et al. (2019). Developing neuroscience-based treatments for alcohol addiction: A matter of choice?. *Translational Psychiatry*, 9, 255.

[2] Buchanan, R., & Sinclair, J. (2020). Alcohol use disorder and the liver. *Addiction* 116, no. 5 (2021): 1270-*1278.*

[3] Nooreen, Z., Shukla, A., Shukla, A., & Verma, P. (2023). A sustainable approach towards prevention and treatment of hepatic and other disorders associated with alcohol consumption. *Current Drug Therapy* 19, no. 5 (2024): 511-528.

[4] Shield, K., Manthey, J., Rylett, M., Probst, C., Wettlaufer, A., Parry, C., & Rehm, J. (2020). National, regional, and global burdens of disease from 2000 to 2016 attributable to alcohol use: a comparative risk assessment study. *The Lancet. Public health*, 5 1, e51-e61 .

[5] Sholehah, A. R., WT, W. R., Ramle, A., Jamilah, M. S., & Rafizah, M. R. (2018). Hypertension genes among Orang Asli in Peninsular Malaysia and ethno-medicinal plant. Asian Journal of Medicine and Biomedicine, 2(1), 25-35.

[6] Joseph, A., Lasimbang, H., James, S., & Seok, C. (2020). Effects of Alcohol towards Quality of Life in the Indigenous Groups of the West Coast Division, Sabah, Malaysia. *Borneo Journal of Medical Sciences (BJMS) 14, no. 2 (2020): 19-19.*

[7] Aswad, M. (2019, November). Helping the Kensiu tribe fight problems of drugs and alcohol | New Straits Times. [online] NST Online. Available at: https://www.nst.com.my/news/nation/2019/11/539482/helping-kensiu-tribe-fight-problems-drugs-and-alcohol [Accessed 3 March 2022]

[8] Barnett, A. I., Hall, W., Fry, C. L., Dilkes-Frayne, E., & Carter, A. (2018). Drug and alcohol treatment providers' views about the disease model of addiction and its impact on clinical practice: A systematic

review. *Drug and alcohol review*, 37(6), 697-720.

[9] Kranzler, H.R. and Soyka, M. (2018). Diagnosis and pharmacotherapy of alcohol use disorder: a review. *Jama*, 320(8), pp.815-824.

[10] Jayachandra, V., Kesidi, R., Yang, Z., Zhang, C., Pan, Z., Sheng, V. and Jin, F. (2020, December). BeSober: Assisting relapse prevention in Alcohol Addiction using a novel mobile app-based intervention. In 2020 *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (ASONAM) (pp. 944-947). IEEE.

[11] Caponnetto, P., & Casu, M. (2022). Update on Cyber Health Psychology: Virtual Reality and Mobile Health Tools in Psychotherapy, Clinical Rehabilitation, and Addiction Treatment. *International Journal of Environmental Research and Public Health* 19, no. 6 (2022): 3516.

[12] Steele, D., Becker, S., Danko, K., Balk, E., Saldanha, I., Adam, G., Bagley, S., Friedman, C., Spirito, A., Scott, K., Ntzani, E., Saeed, I., Smith, B., Popp, J., & Trikalinos, T. (2020). Interventions for Substance Use Disorders in Adolescents: A Systematic Review. https://doi.org/10.23970/ahrqepccer225.

[13] Heteny, T., & Kotta, I. (2022). 12 Steps and Abstinence: How Do I Stay Sober?. *Studia Universitatis Babeș-Bolyai Psychologia-Paedagogia* (2022): 5-23.

[14] Bernard, C., Grigg, J., Volpe, I., Lubman, D., & Manning, V. (2020). Client Experiences of a Telephone-Delivered Intervention for Alcohol Use: a Qualitative Study. *International Journal of Mental Health and Addiction*, 20, 522-540.

[15] King, S., Richner, K., Tuliao, A., Kennedy, J., & McChargue, D. (2019). A Comparison between Telehealth and Face-to-Face Delivery of a Brief Alcohol Intervention for College Students. *Substance Abuse*, 41, 501 - 509.

[16] Humphreys, G., Evans, R., Makin, H., Cooke, R., & Jones, A. (2020). Identification of Behavior Change Techniques From Successful Web-Based Interventions Targeting Alcohol Consumption, Binge Eating, and Gambling: Systematic Review. *Journal of Medical Internet Research* 23, no. 2 (2021): e22694.

[17] Symons, M., Feeney, G., Gallagher, M., Young, R., & Connor, J. (2019). Machine learning vs addiction therapists: A pilot study predicting alcohol dependence treatment outcome from patient data in behavior therapy with adjunctive medication.. *Journal of substance abuse treatment*, 99, 156-162.

[18] Burkhardt, G., Adorjan, K., Kambeitz, J., Kambeitz-Ilankovic, L., Falkai, P., Eyer, F., Koller, G., Pogarell, O., Koutsouleris, N., & Dwyer, D. (2020). A machine learning approach to risk assessment for alcohol withdrawal syndrome. *European Neuropsychopharmacology*, 35, 61-70.

[19] Yee, A., Salina, M., Rusdi, A. R., Zahari, M. M., & Sharmilla, K. (2017). Psychometric Properties of the Malay Version of Alcohol, Smoking and Substance Involvement Screening Test (ASSIST-M). *Malaysian Journal of Psychiatry*, 25(2), 3-13.

[20] Aniza, I., & Norhayati, M. (2016). Barriers to health promotion for indigenous communities: Lessons for Malaysia. *Malaysian Journal of Public Health Medicine*, 16(1), 6-14.

[21] Ithnina, M., Julianab, N., Norc, N. A. U. M., Effendyd, N. M., & Ranic, M. D. M. (2020, December). Behavioural Risk Factors of Non-communicable Diseases (NCDs) Determinant in the Context of an Orang Asli (Indigenous peoples): A Qualitative Study. In *Proceedings of the 4th UUM International Qualitative Research Conference* (QRC 2020) (Vol. 1, p. 3).

[22] Bala, P., & Tan, C. (2021). Digital inclusion of the Orang Asli of Peninsular Malaysia: Remote virtual mechanism for usability of telecentres amongst indigenous peoples. *The Electronic Journal Of Information Systems In Developing Countries*. doi: 10.1002/isd2.12171.

[23] Shah, B., Rajdev, D., Salunkhe, R., Ramrakhiani, P., & Deshpande, H. (2021, August). Prognosis of Supervised Machine Learning Algorithms in Healthcare Sector. In 2021 *International Conference on Recent Trends on Electronics, Information, Communication & Technology* (RTEICT) (pp. 257-261). IEEE.

[24] An, G., Omodaka, K., Tsuda, S., Shiga, Y., Takada, N., Kikawa, T., ... & Akiba, M. (2018). Comparison of machine-learning classification models for glaucoma management. *Journal of healthcare engineering* 2018, no. 1 (2018): 6874765.

[25] Iyyanki, M., Jayanthi, P., & Manickam, V. (2022). Machine learning for health data analytics: a few case studies of application of regression. In *Research Anthology on Machine Learning Techniques, Methods, and Applications* (pp. 1038-1061). IGI Global.

[26] Joshi, A., Pant, I., & Dhiman, Y. (2022). Efficient Analysis in Healthcare Domain using Machine Learning. In Telemedicine: *The Computer Transformation of Healthcare* (pp. 125-134). Cham: Springer International Publishing.

[27] Ebrahimi, A., Wiil, U. K., Schmidt, T., Naemi, A., Nielsen, A. S., Shaikh, G. M., & Mansourvar, M. (2021). Predicting the risk of alcohol use disorder using machine learning: a systematic literature review. IE*EE*

Access, 9, 151697-151712.

[28] Rahman, S., Sharma, T., & Mahmud, M. (2020, September). Improving alcoholism diagnosis: comparing instance-based classifiers against neural networks for classifying EEG signal. In *International Conference on Brain Informatics* (pp. 239-250). Cham: Springer International Publishing.

[29] Ellis, C. A., Sattiraju, A., Miller, R. L., & Calhoun, V. D. (2023, June). Novel Approach Explains Spatio-Spectral Interactions in Raw Electroencephalogram Deep Learning Classifiers. In 2023 *IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops* (ICASSPW) (pp. 1-5). IEEE.

[30] Zhao, J., Wang, P., & Pan, Y. (2022). Predicting liver disorder based on machine learning models. *The Journal of Engineering*, 2022(10), 978-984.

[31] Grodin, E. N., Montoya, A. K., Cruz, A., Donato, S., Baskerville, W. A., & Ray, L. A. (2024). Identifying Treatment Responders to Varenicline for Alcohol Use Disorder Using Two Machine-Learning Approaches. *Clinical Psychological Science*, 12(1), 133-146.

[32] Xu, R., Lin, H., Lu, K., Cao, L., & Liu, Y. (2021). A forest fire detection system based on ensemble learning. *Forests*, 12(2), 217.

[33] Nikhitha, N. K., Afzal, A. L., & Asharaf, S. (2021). Deep Kernel machines: a survey. *Pattern Analysis and Applications*, 24, 537-556.

[34] Yang, T., Yu, X., Ma, N., Zhang, Y., & Li, H. (2022). Deep representation-based transfer learning for deep neural networks. *Knowledge-Based Systems*, 253, 109526.

[35] Kook, L., Götschi, A., Baumann, P. F., Hothorn, T., & Sick, B. (2022). Deep interpretable ensembles. arXiv preprint arXiv:2205.12729.

[36] Gao, Y., Miyata, S., & Akashi, Y. (2022). Interpretable deep learning models for hourly solar radiation prediction based on graph neural network and attention. *Applied Energy*, 321, 119288.

[37] Mak, K. K., Lee, K., & Park, C. (2019). Applications of machine learning in addiction studies: A systematic review. *Psychiatry research*, *275*, 53-60.

[38] Alam, F., Kashef, R., & Jaseemuddin, M. (2021). Enhancing The Performance of Network Traffic Classification Methods Using Efficient Feature Selection Models. *2021 IEEE International Systems Conference (SysCon)*, 1-6.

[39] Shokrzade, A., Ramezani, M., Tab, F., & Mohammad, M. (2021). A novel extreme learning machine based kNN classification method for dealing with big data. *Expert Syst. Appl.*, 183, 115293.

[40] Vandewiele, G., Lannoye, K., Janssens, O., Ongenae, F., Turck, F., & Hoecke, S. (2017). A Genetic Algorithm for Interpretable Model Extraction from Decision Tree Ensembles. In *Trends and Applications in Knowledge Discovery and Data Mining: PAKDD 2017 Workshops, MLSDA, BDM, DM-BPM Jeju, South Korea, May 23, 2017, Revised Selected Papers 21* (pp. 104-115). Springer International Publishing.

[41] Zhang, Y., Tiňo, P., Leonardis, A., & Tang, K. (2020). A Survey on Neural Network Interpretability. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 5, 726-742. https://doi.org/10.1109/TETCI.2021.3100641.

[42] Symons, M., Feeney, G., Gallagher, M., Young, R., & Connor, J. (2020). Predicting Alcohol Dependence Treatment Outcomes: A Prospective Comparative Study of Clinical Psychologists vs 'Trained' Machine Learning Models. *Addiction*, *115*(11), pp.2164-2175.

[43] Christodoulou, E., Ma, J., Collins, G., Steyerberg, E., Verbakel, J., & calster, B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models.. *Journal of clinical epidemiology*, 110, 12-22.

[44] Singh, A., Singh, V., Gourisaria, M. K., & Sharma, A. (2022, December). Alcohol Consumption Rate Prediction using Machine Learning Algorithms. In 2022 OITS *International Conference on Information Technology* (OCIT) (pp. 85-90). IEEE.

[45] Priya, A. M., & Thilagamani, S. (2022). Prediction of Arterial Stiffness Risk in Diabetes Patients through Deep Learning Techniques. *Information Technology and Control*, 51(4), 678-691.

[46] Lottering, R., Hans, R., & Lall, M. (2020). A Machine Learning Approach to Identifying Students at Risk of Dropout: A Case Study. *International Journal of Advanced Computer Science and Applications*, 11(10), 417-422.

[47] Buniyamin, N. (2022). Predicting engineering students' academic performance using ensemble classifiers-a preliminary finding. *Journal of Electrical and Electronic Systems Research* (JEESR), 20, 92-101.

[48] Ibarguren, I., Pérez, J. M., Muguerza, J., Arbelaitz, O., & Yera, A. (2022). PCTBagging: from inner ensembles to ensembles. A trade-off between discriminating capacity and interpretability. *Information Sciences*, *583*, 219-238.

[49] Arif, M. A. I., Sany, S. I., Sharmin, F., Rahman, M. S., & Habib, M. T. (2021). Prediction of addiction to drugs and alcohol using machine learning: A case study on Bangladeshi population. *International Journal of Electrical and Computer Engineering*, 11(5), 4471.

[50] Wickersham, M., Bartelo, N., Kulm, S., Liu, Y., Zhang, Y., & Elemento, O. (2023). Using Machine Learning Methods to Assess the Risk of Alcohol Misuse in Older Adults. *Research Square*.

[51] Pagnotta, F., & Amran, H. M. (2016). *Using data mining to predict secondary school student alcohol consumption*. Department of Computer Science, University of Camerino, 1-9.

[52] Palaniappan, S., Hameed, N. A., Mustapha, A., & Samsudin, N. A. (2017). Classification of alcohol consumption among secondary school students. *JOIV: International Journal on Informatics Visualization*, 1(4-2), 224-226.

[53] Sisodia, D. S., Agrawal, R., & Sisodia, D. (2019). A comparative performance of classification algorithms in predicting alcohol consumption among secondary school students. In *Machine Intelligence and Signal Analysis* (pp. 523-532). Springer, Singapore.

[54] Fehrman, E., Muhammad, A. K., Mirkes, E. M., Egan, V., & Gorban, A. N. (2017). The five factor model of personality and evaluation of drug consumption risk. In *Data science* (pp. 231-242). Springer, Cham.

[55] Shukla, A. K., Singh, P., & Vardhan, M. (2018). Predicting alcohol consumption behaviors of secondary level students. In *Proceedings of 3rd international conference on internet of things and connected technologies* (ICIoTCT) (pp. 26-27).

[56] Liu, Z., Japkowicz, N., Wang, R., Cai, Y., Tang, D., & Cai, X. (2020). A statistical pattern based feature extraction method on system call traces for anomaly detection. *Information and Software Technology*, 126, 106348.

[57] Samo, Y. (2017). *Advances in kernel methods: towards general-purpose and scalable models* (Doctoral dissertation, University of Oxford).

[58] Maggu, J., & Majumdar, A. (2017). Kernel transform learning. *Pattern Recognition Letters*, 98, 117-122.

[59] Manna, S., Khonglah, J. R., Mukherjee, A., & Saha, G. (2021). Robust kernelized graph-based learning. *Pattern Recognition*, 110, 107628.

[60] Gokcay, E. (2020). An information-theoretic instance-based classifier. *Information Sciences*, 536, 263-276.

[61] Cavalcanti, G. D., & Soares, R. J. (2020). Ranking-based instance selection for pattern classification. *Expert Systems with Applications*, 150, 113269.

[62] Blachnik, M., & Kordos, M. (2020). Comparison of instance selection and construction methods with various classifiers. *Applied Sciences*, 10(11), 3933.

[63] Sziburis, T., Nowak, M., & Brunelli, D. (2021). KNN Learning Techniques for Proportional Myocontrol in Prosthetics. ArXiv, abs/2109.08917.

[64] Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J. (2020). Deep Learning--based Text Classification. *ACM Computing Surveys* (CSUR), 54, 1 - 40.

[65] Bhatt, C., Kumar, I., Vijayakumar, V., Singh, K., & Kumar, A. (2020). The state of the art of deep learning models in medical science and their challenges. *Multimedia Systems*, 27, 599-613.

[66] Xiong, Y., Ding, Y., Li, J., Zhang, W., & Ni, S. (2023). Prediction of Final Academic Performance of Secondary School Students Based on Alcohol Consumption and Diverse Backgrounds. *LNEP*, 2, 567-574.

[67] Park, S. J., Lee, S. J., Kim, H., Kim, J. K., Chun, J. W., Lee, S. J., ... & Choi, I. Y. (2021). Machine learning prediction of dropping out of outpatients with alcohol use disorders. Plos one, 16(8), e0255626.

[68] Kumari, D., Kilam, S., Nath, P., & Swetapadma, A. (2018). Prediction of alcohol abused individuals using artificial neural network. *International Journal of Information Technology*, 10(2), 233-237.

[69] Ebrahimi, A., Wiil, U. K., Mansourvar, M., Naemi, A., Andersen, K., & Nielsen, A. S. (2021, May). Deep Neural Network to Identify Patients with Alcohol Use Disorder. In *MIE* (pp. 238-242).

[70] Mohamed, A. T., Santhoshkumar, S., & Varadarajan, V. (2022). Intelligent deep learning based predictive model for coronary heart disease and chronic kidney disease on people with diabetes mellitus. *Malaysian Journal of Computer Science*, 88-101.

[71] Ramesh, T. R., Lilhore, U. K., Poongodi, M., Simaiya, S., Kaur, A., & Hamdi, M. (2022). Predictive analysis of heart diseases with machine learning approaches. *Malaysian Journal of Computer Science*, 132-148.