Research Project Final Report

A Machine Learning-Based Tool for Predicting Depression in College Students: An Example of Interdisciplinary Integration for Future Skills Development

เครื่องมือทำนายภาวะซึมเศร้าในนักศึกษาโดยการใช้การเรียนรู้ของเครื่อง: ตัวอย่างการบูรณาการแบบสหวิทยาการเพื่อพัฒนาทักษะแห่งอนาคต

By

Asst. Prof. Dr. Pattaramon Vuttipittayamongkol

Dr. Kemachart Kemavuthanon

Dr. Pimrat Boonyapuk

This research is funded by Mae Fah Luang University

In fiscal year 2022

Research Project Final Report

A Machine Learning-Based Tool for Predicting Depression in College Students: An Example of Interdisciplinary Integration for Future Skills Development

เครื่องมือทำนายภาวะซึมเศร้าในนักศึกษาโดยการใช้การเรียนรู้ของเครื่อง: ตัวอย่างการบูรณาการแบบสหวิทยาการเพื่อพัฒนาทักษะแห่งอนาคต

By

Asst. Prof. Dr. Pattaramon Vuttipittayamongkol

Dr. Kemachart Kemavuthanon

Dr. Pimrat Boonyapuk

This research is funded by Mae Fah Luang University

In fiscal year 2022

Executive Summary

The underlying objective of the research project "A Machine Learning-Based Tool for Predicting Depression in College Students: An Example of Interdisciplinary Integration for Future Skills Development" is to build an example that could encourage students to gain knowledge in diverse fields and apply such knowledge in real-life problems or future work. This project serves as a development of an innovative approach that combines two disciplines, namely information technology and healthcare, with an emphasis on setting an example for teaching and learning in the School of Information Technology and Nursing at Mae Fah Luang University.

According to the report on "The Situation and Educational Productivity System of Undergraduate Level at Mae Fah Luang University" in 2017, based on surveys and tests conducted on a total of 4,034 students, it was found that students had a high prevalence of mental health issues, with 46% experiencing psychological symptoms and 22% of this group exhibiting symptoms of depression. These numbers are concerning and indicate the need for close monitoring and assistance. As of today, the School of Nursing at Mae Fah Luang University revealed the number of students experiencing mental health, especially depressive disorder, has arisen yearly.

The increasing number of individuals with mental health problems has overwhelmed the availability of medical personnel or relevant experts. The initial screening and categorization process for psychiatric patients is time-consuming, given the various types of mental health conditions, such as depression, anxiety, and panic disorders. Additionally, the students may not be aware of their own mental health status. The use of the Thai Mental Health Questionnaire (TMHQ), which consists of numerous questions (70 items), may discourage them from completing it. This further highlights the need of leveraging technology to assist in screening students with depression to alleviate the workload of professionals, reduce costs, and allow prompt interventions to provide assistance to students suffering from depression.

Over the past few decades, there has been a widespread utilization of data science and artificial intelligence in addressing diverse challenges. Among these applications is the realm of mental health. A recent research work involves creating of a depression detection system that employs machine learning methods on agricultural data from rural regions. The research findings indicate that by considering personal attributes and livelihood details, including age, gender, marital status, number of children, education, assets, income, and expenses, it is possible to develop an accurate predictive model for identifying depression among farmers, surpassing a 70% accuracy rate.

With this in mind, the research team recognizes the potential of leveraging artificial intelligence technology to assist in identifying the risk of depression among university students. This is achieved by utilizing a case study of students from Mae Fah Luang University, based on data collected in the research project on the undergraduate production system in the year 2017. The aim is to create an innovative approach that serves as an interdisciplinary example. Specifically, addressing mental health issues, particularly the problem of

depression, which closely affects students, will help them better understand the context of integrative learning. Furthermore, Mae Fah Luang University can utilize this developed innovation to benefit student depression screening and utilize the obtained results to establish policies for effectively supporting and assisting students' mental health.

The experiments in this research projects consisted of two main parts. The initial part focused on model building using machine learning. It involved data preprocessing steps such as removing irrelevant items, missing values, and outliers, grouping schools into categories, transforming categorical data, and splitting the dataset into training and testing sets. Various learning algorithms were used for model training. Resampling methods were applied to alleviate data imbalance, and the effects of adding selective questionnaire items were investigated. Evaluation metrics such as recall, specificity, and G-mean were used to assess classification results. The second part of the project focused on evaluating user satisfaction. Nursing students were selected as the study population, and an online questionnaire based on a Likert scale was used to assess user satisfaction. The collected data was analyzed to evaluate the effectiveness of the predictive model as an example of interdisciplinary integration. These approaches collectively provided valuable insights into machine-learning model building to predict depression in students and user satisfaction evaluation within the context of interdisciplinary integration.

In the first part of the experiments, the evaluation of the predictive model revealed challenges in detecting depression, with low recall rates but high specificity, possibly due to imbalanced sample classes. Resampling methods, especially RUS, showed promising improvements in recall and g-mean values. Furthermore, the inclusion of one to two non-disturbing TMHQ question items resulted in a significant improvement in the performance of predictive models. In evaluating the model as an interdisciplinary integration example for teaching and learning, the survey involving nursing students demonstrated the positive impact of the interdisciplinary integration model, increasing their understanding, recognition, and satisfaction with interdisciplinary learning. The research successfully addressed the challenges of depression detection, showcased the effectiveness of resampling methods and specific questionnaire items, and highlighted the positive impact of interdisciplinary integration on user comprehension and satisfaction.

In conclusion, this project showcases the successful application of interdisciplinary learning in depression detection among college students, utilizing a machine learning-based approach. By incorporating data resampling methods and various learning algorithms, the project effectively addresses the limitations of traditional mental health questionnaires and tackles the problem of under-representation. The achieved competitive classification accuracy underscores the efficacy of this interdisciplinary approach. Moreover, the high satisfaction rating towards the learning tool exemplifies its success in cross-disciplinary integration. This project not only provides valuable insights into depression detection but also demonstrates the practical application of interdisciplinary knowledge. The project emphasizes the importance of interdisciplinary collaboration in developing innovative solutions to complex societal issues and encourages further exploration of interdisciplinary approaches in diverse domains, fostering future innovations to address real-world challenges.

Abstract

This project presents a machine learning-based approach for depression detection in college students, aiming to create an innovation that serves as an example of interdisciplinary integration. By addressing the limitations of traditional methods such as visiting mental health experts or completing standard questionnaires, the proposed approach offers a more comfortable and efficient screening process. The model is trained on demographic information, physical health problems, relationships, and university life aspects, excluding direct mental health questions. The results demonstrate over 90% prediction accuracy, highlighting the potential of machine learning-based approaches for depression screening. The success of the predictive model can be attributed to the integration of knowledge from the fields of information technology and health science. User satisfaction survey on evaluating the predictive model as a learning tool for interdisciplinary integration was also conducted. The study population included a sample of fifty nursing students enrolled at Mae Fah Luang University. High user satisfaction ratings further support the success of this project.

บทคัดย่อ

โครงการวิจัยนี้นำเสนอโมเดลการตรวจจับภาวะซึมเศร้าในนักศึกษามหาวิทยาลัยด้วยเทคนิคการเรียนรู้ของเครื่อง โดยมีวัตถุประสงค์หลักเพื่อสร้างนวัตกรรมที่เกิดจากการบูรณาการความรู้ต่างศาสตร์สำหรับแก้ปัญหาในชีวิตจริง นวัตกรรมนี้เป็นตัวอย่างในการแก้ไขข้อด้อยและข้อจำกัดของวิธีการคัดกรองผู้ป่วยซึมเศร้าแบบทั่วไป เช่น การพบผู้เชี่ยวชาญด้านสุขภาพจิตที่ต้องใช้เวลาค่าใช้จ่ายและบุคลากรทางการแพทย์จำนวนมาก หรือการทำแบบทดสอบสุขภาพจิตที่มีข้อคำถามบางข้อที่อาจกระทบต่อความรู้สึกของผู้ทำแบบทดสอบ โมเดลทำนายภาวะซึมเศร้าที่ถูกพัฒนาขึ้นภายใต้โครงการวิจัยนี้ใช้ข้อมูลทั่วไปของนักศึกษาและข้อมูลปัญหาด้านต่างๆ เช่น สำนักวิชาที่เรียน อายุ เพศ สุขภาพกาย ความสัมพันธ์กับเพื่อน ความสัมพันธ์กับอาจารย์ ความรู้สึกต่อสิ่งแวดล้อมในมหาวิทยาลัย และความรู้สึกต่อสิ่งสนับสนุนการเรียนรู้ เป็นต้น ผลการทดสอบแสดงให้เห็นถึงความแม่นยำในการคัดกรองภาวะซึมเศร้าในนักศึกษาสูงกว่าร้อยละ 90 ความสำเร็จของโมเดลทำนายนี้เกิดขึ้นจากการรวมความรู้จากสาขาวิชาเทคโนโลยีสารสนเทศและวิทยาศาสตร์สุขภาพเข้าด้วยกัน ส่วนในด้านความพึงพอใจต่อโมเดลในเชิงของการเป็นตัวอย่างนวัตกรรมสำหรับการเรียนการสอนที่มุ่งเน้นให้ผู้เรียนเห็นถึงความสำคัญของการบูรณาการข้ามศาสตร์นั้นอยู่ในเกณฑ์ดีมาก โดยมีผู้ประเมิน คือ อาสาสมัครนักศึกษาสำนักวิชาพยาบาลศาสตร์ มหาวิทยาลัยแม่ฟ้าหลวง จำนวน 50 คน เป็นหลักฐานที่เน้นย้ำความสำเร็จของนวัตกรรมและโครงการนี้อีกด้วย

Contents

CHAPTER 1 Introduction **1**
1.1 Research Background 1
1.2 Objectives 2
1.3 Hypothesis and Scope of the Study 2
1.4 Definition of Terms 2

CHAPTER 2 Related Works **4**

CHAPTER 3 Methodology **7**
3.1 Predictive model building 7
3.1.1 Dataset and data selection 7
3.1.2 Data preprocessing 8
3.1.3 Learning algorithms 9
3.1.4 Parameter settings 10
3.1.5 Experimental framework 13
3.1.6 Evaluation measures 14
3.2 User satisfaction evaluation 14
3.2.1 Study population 15
3.2.2 Data collection tool 15
3.2.3 Data collection process 15
3.2.4 Data Analysis 16

CHAPTER 4 Results and Discussion **17**
4.1 Predictive models 17
4.1.1 Experiment I 17
4.1.2 Experiment II 17
4.1.3 Experiment III 19
4.2 User satisfaction 20

CHAPTER 5 Conclusion **25**

Appendices **29**
Appendix A. Ethical Approval 30
Appendix B. User satisfaction questionnaire 35

Appendix C. Conference proceeding 42

List of Figures

Figure 3.1 The experimental framework 8

Figure 4.1 The gender distribution of the respondents 20
Figure 4.2 The age distribution of the respondents 21
Figure 4.3 The breakdown of the academic disciplines studied by the respondents outside their own fields 21

List of Tables

Table 4.1 Classification performance of different learning algorithms 17
Table 4.2 Recall and specificity across different resampling methods 18
Table 4.3 G-mean across different resampling methods 18
Table 4.4 Classification results after adding TMHQ question items 19
Table 4.5 Survey results on understanding interdisciplinary integration before and after the model demonstration 23
Table 4.6 Satisfaction results on the predictive model as an example of interdisciplinary integration 23
Table 4.7 Satisfaction results regarding the utility of the model 24

Chapter 1

Introduction

1.1 Research Background

Major depressive disorder (MDD), also known as clinical depression, is one of the most common mental disorders among college students (Rudenstine et al., 2021). Depression in students can be caused by socioeconomic status (Ebert et al., 2019) and declining academic performance (Rudenstine et al., 2021). Entering a new society as an incoming freshman as well as high expectations from the parent elevate the level of stress and anxiety in college students, which could eventually lead to depression and college dropout (Lun et al., 2018). Physical impacts on depressed students include risky sexual behavior, smoking, alcohol and drug dependency, self-harm behaviors, and in the worst scenario, the risk of suicide.

The research project "The Situation and Educational Productivity System of Undergraduate Level at Mae Fah Luang University in 2017" reported that over 46% of students at Mae Fah Luang University who completed the survey had some indications of mental disorders while 22% of these students had symptoms of depression. These numbers tend to increase in these recent years partly due to the spread of Covid-19. During the breakout, teaching and learning was conducted mostly online preventing from social events and interactions with real people (AlAteeq et al., 2020). Also the fear of getting the disease created more stress, anxiety, and depression in them (Maia and Dias, 2020).

There are several ways to determine whether an individual is suffering from depression. Getting a direct consultation and examination by a psychiatrist is the most reliable practice; however, it is costly and takes up time. Various self-administered psychometric tests for assessing depression are also widely available (Bernaras et al., 2019). However, these tests are generally lengthy. A large number of questions discourages the taker to complete the test with precise answers. As a consequence, the assessment result would not be accurate. The main drawback of these traditional assessments is that the person realizes that he or she is being tested for depression. This may bring negative feelings to the subject and refusal of the assessment.

Data-driven and machine learning-based methods has been widely used to solve many real-world problems, especially in the public health domain (dos Santos et al., 2019). Khoory et al. (2022) studied the psychosocial impact of COVID-19 quarantine on par-

ents and their children and reported the relationship between emotional functioning and behavioral characteristics during the strict quarantine. Hassan et al. (2021) developed a depression detection system using machine learning algorithms on a dataset of farmers in rural areas. Results showed that demographic and livelihood information such as age, marital status, education, number of children, spending and earning, could be used for depression detection with the accuracy over 70%.

Our research team sees a promising possibility of using machine learning to aid the detection of depression in college students. The data collected in the research project "The Situation and Educational Productivity System of Undergraduate Level at Mae Fah Luang University in 2017" will be used for the study. Since the issue is highly related to the students, this could serve as a good example of interdisciplinary applications in class for both students in the School of Information and Technology and students in the School of Nursing. In addition, the university could use the predictive model to screen depression among students and handle the cases in a timely manner.

1.2 Objectives

1. To study the possibility of using machine learning to create a tool for early detection of depression in college students

2. To create an innovation of a predictive model for depression based on non-clinical information, which requires an understanding of both healthcare and data science domains, serving as an example of the integration of learning across disciplines

1.3 Hypothesis and Scope of the Study

We hypothesized that some non-clinical information of students can be associated with depression symptoms and machine learning algorithms could aid such a process. Our study focuses on university students and uses data collected from Mae Fah Luang University students in 2017. This includes 4,642 students, 811 of which were identified to have depression symptoms based on the Thai Mental Health Questionnaire (TMHQ) (Phattharayuttawat et al., 1999).

1.4 Definition of Terms

- Machine learning: the process of training computers to learn and improve from data without explicit programming

- Interdisciplinary integration: the collaboration and merging of knowledge, methods, and perspectives from multiple disciplines or fields of study

- Positive sample: A student sample who has a risk for depression based on the TMHQ measure test
- Negative sample: A student sample who is not likely to have a risk for depression based on the TMHQ measure test
- (Machine learning) Classification: the task of training a model to predict the class or category of a given input based on its features or attributes. It involves using a labeled dataset to teach the model to recognize patterns and make accurate predictions about unseen data
- General information (of samples): Unidentifiable personal information such as age, gender, financial status, health status, etc., which are non-clinical information

Chapter 2

Related Works

Machine learning is currently a popular field in data science that involves using a computer to automatically find patterns in data and using those patterns to make predictions or decisions (Géron, 2022). There are many learning algorithms available; however, the one that gives the best predictive model is unknown until many algorithms have been tried and tested. To build a model, an algorithm is first trained on a set of labeled data, where it learns to recognize patterns by adjusting its internal parameters. After training, the model's performance is evaluated on a separate testing dataset, and necessary improvements will be made using many techniques, such as tweaking hyperparameters, trying different algorithms or further processing the training data based on mathematical techniques. One of the most crucial steps in building machine learning models is data preparation, which includes data cleaning and data preprocessing (Kotsiantis et al., 2006). Data cleaning is the process of identifying and correcting or removing errors, inconsistencies, and inaccuracies in datasets before using them in machine learning models. It is an essential step in machine learning, as models can only be as good as the data they are trained on. Data cleaning involves tasks such as handling missing values, dealing with outliers, removing duplicates, correcting data types, and ensuring consistency across variables. These tasks can be performed using various techniques, including imputation, normalization, and scaling. Data cleaning is critical to ensure that the data is accurate, complete, and reliable, which can improve the performance and generalization of machine learning models. However, data cleaning can also be time-consuming and resource-intensive, and it requires both domain knowledge and expertise in data manipulation and programming.

Data preprocessing involves transforming raw data into a format that is suitable for training models. The purpose of data preprocessing is to clean and prepare the data, as well as to extract meaningful features that can be used by machine learning algorithms. Data preprocessing techniques include normalization, scaling, encoding categorical variables, handling missing values, and feature selection. These techniques can improve the performance and accuracy of machine learning models by reducing noise and making the data more relevant and informative. Data preprocessing requires knowledge of programming, statistics, and domain-specific knowledge to effectively handle the data and extract meaningful insights.

As mentioned above, cross-disciplinary knowledge is essential in data science projects because these projects require a combination of skills from different fields. Data science projects typically involve working with large and complex datasets, and this requires expertise in statistics, mathematics, and computer science. Additionally, to gain insights from data, you need domain-specific knowledge, such as understanding of economics, biology, or social sciences. Therefore, cross-disciplinary knowledge is needed in data science projects to understand the problem, identify the right data sources, select appropriate techniques for analysis, interpret the results, and communicate the findings to different audiences. Successful data science projects often involve collaboration among individuals with diverse expertise, including data scientists, subject-matter experts, software engineers, and business analysts. For example, in a project on building a model to predict depression, the data scientist will also need to have or work closely with a person who has the domain knowledge in psychology.

Emerging methods based on machine learning has been proposed to predict the likelihood of developing depression. Cho et al. (2021) used data consisting of sociological factors, physiological factors (e.g., the prevalence of chronic diseases), psychological factors, and laboratory tests(e.g.,blood cell counts) to create a classification model. LASSO, a regression method that also performs variable selection (Tibshirani, 1996), was used to learn the data. The predicting accuracy of over 80% in both classes was obtained with the 13 remaining variables. In this work, both clinical and non-clinical data were used. However, the requirement of laboratory results makes this method not readily accessible to average people.

Kumar et al. (2020) presented the use of several standard learning algorithms as well as neural network-based algorithms to predict the severity level of psychological disorders. The focused psychological problems were anxiety, depression, and stress. For each disorder, machine learning was employed to learn and predict the five severity levels based on 14 Likert-scaled items of a standard questionnaire used by psychiatrists that is the Depression, Anxiety, and Stress Scale questionnaire (DASS) (Parkitny and McAuley, 2010). Priya et al. (2020) later proposed a similar work. It should be pointed out that the use of machine learning would not be encouraged in this case. The severity levels were directly proportional to the sum of the answer scores; thus, using machine learning would not be a faster and more accurate way to classify the disorder level.

In (Zulfiker et al., 2021), the authors trained their predictive models to classify depressed and non-depressed persons based on socio-demographic and psychosocial information. Several feature selection techniques, namely, SelectKBest, mRMR, and Boruta, were applied resulting in 13 to 15 out of 30 remaining features for model learning. Results showed accuracy and recall of over 90% on most methods. However, some question items on the psychosocial information, for example, “Whether the participant has any suicidal thought recently or not”, could make the subject feel intimidated and uncomfortable to answer.

A recent review (Liu et al., 2022) summarizes 17 previous studies that have applied machine learning methods to text data from social media in order to detect depressive

symptoms of the post's owner. Despite not necessitating clinical test data, this social media text classification method still exhibits certain limitations. To collect the ground truth, the assessment of the user's mental status in some of these studies was based on either self-declaration or community membership. These approaches might not consistently provide correct identification leading to inaccurate prediction results. Moreover, the use of social media text data in the reviewed studies limited to English language, which could present an obstacle in research conducted in countries where English is not the primary language.

The drawbacks of the aforementioned works can be listed as follows. Firstly, incorporating clinical or laboratory data into machine learning model development will not mitigate the limitations associated with traditional assessment methods. Secondly, employing question items from a professional diagnostic scale as features for model training would not yield depression identification results any more accurate than the traditional approach. Moreover, this approach fails to address the issue of individuals being required to answer some discomforting questions on the test. Lastly, individual language preferences on social media present challenges for data collection in countries where English is not the predominant language.

In this project, to tackle the above issues, we focus on the use of non-clinical information to predict depression in college students with machine learning methods. The main contributions of this work is that we investigated the possibility of using the general information related to college students in building machine learning models to detect students with depression. Unlike many existing works, we have dropped aggressive and intimidating questions that can be perceived by the subject as direct implications of depression from the required data for training a model, and we aimed to limit the use of remaining professional diagnostic scale's question items as much as possible. Our hypothesis was that general information of an individual can be associated with depression symptoms and machine learning algorithms could aid such a process.

Chapter 3

Methodology

In this chapter, the methodology of the proposed work is provided in detail. The work was split into two main parts: 1) Predictive model building and 2) Evaluating the model as an interdisciplinary integration example.

3.1 Predictive model building

This section provides methodology details on model building using machine learning. First, the dataset is introduced, and data preprocessing is discussed. Then, the flow of the experiment settings, which consists of three sets of experiments, is outlined. A detailed explanation of each experiment will be given. This includes the learning algorithms and the data resampling methods used. Finally, evaluation metrics are listed. The flow of the experimental framework is illustrated in Figure 3.1.

3.1.1 Dataset and data selection

The dataset used in this work was obtained from the research project "The Situation and Educational Productivity System of Undergraduate Level at Mae Fah Luang University in 2017". The data was collected in form of a questionnaire, which consists of six parts: general information, mental health measures, emotional intelligence assessment, learning style assessment, assessment of cognitive skills and assessment of student's skills according to the higher education qualifications framework. The questionnaire contains a total of 270 questions, 70 of which are the Thai Mental Health Questionnaire (TMHQ) (Phattharayuttawat et al., 1999). The TMHQ includes 20 indicators of depression, and most of them can be easily perceived by non-experts as mental health questions.

In this work, only the first part of the original questionnaire was used. As for the remaining sections, they comprise questions that are non-generic and either irrelevant or pertain to professional assessments. The selected part consists of 45 items covering demographic information, physical health problems, relationships with family and friends, existence of problems in different aspects, etc. The dataset has 4,642 instances representing the number of students participating in the questionnaire. Among these students, 811

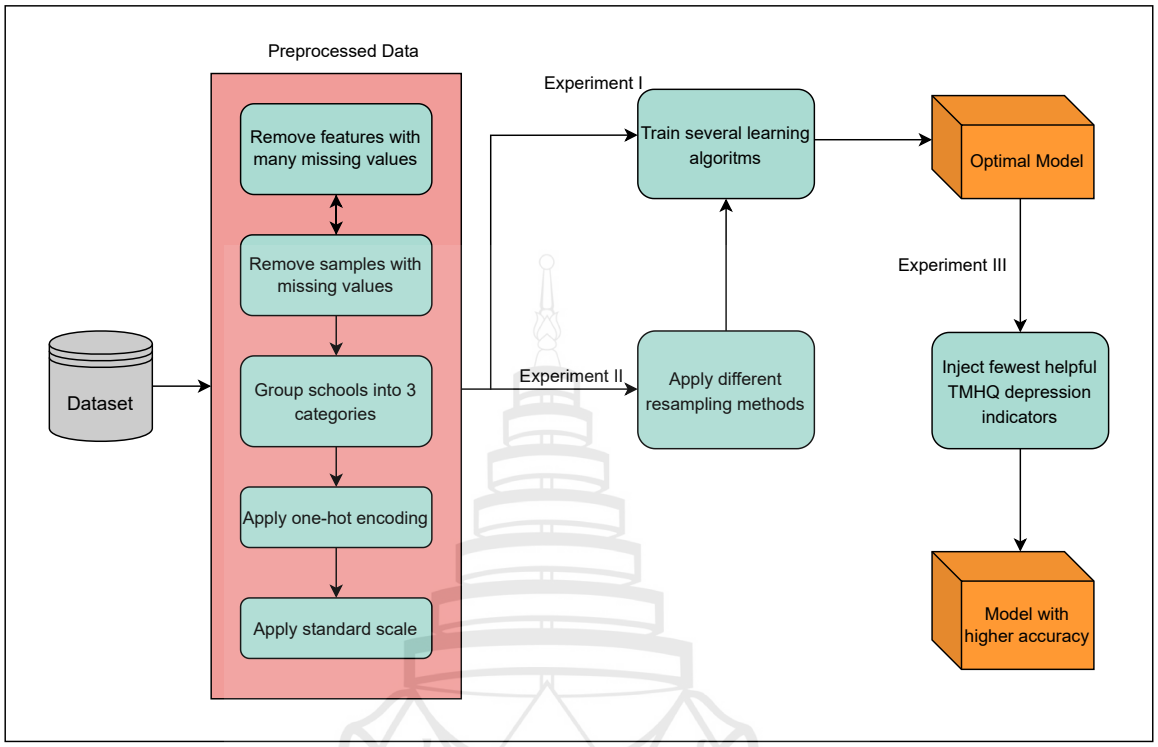

Figure 3.1: The experimental framework

of them had indications of depression-related symptoms. We identified them according to the 20 depression test scores available in the second part of the data based on the rubric of TMHQ. Thus, this results were not diagnoses made by a doctor. However, for the ease of discussion, we will simply refer to them as depression samples, and the others as non-depression samples. The difference in the number of samples in the two categories makes the dataset imbalanced with 3831:811 or 4.72 ratio of the negative (non-depression) to the positive samples (depression).

3.1.2 Data preprocessing

Firstly, the items with too much variation in responses were removed. This includes, for example, home address, relaxing activities, and health problems. Moreover, items containing a large number of missing values and excessive outliers were also excluded from the dataset. These were student income and family income, which contain 1,575 and 1,524 missing values, respectively. After the removals, there were 37 remaining attributes. The variables with few missing values that are GPAX and number of close friends, which have 11 and 6 missing values, respectively, were retained for the experiment. The samples with these missing values were instead removed from the dataset with a minimal impact on the dataset size.

The 13 schools were grouped into 3 main categories: 1) Science and Technology, 2)

Health and Biomedical Sciences, and 3) Social Sciences and Humanities based on the literature (Ebert et al., 2019). Furthermore, we employed one-hot encoding to convert categorical data to numerical data. To avoid bias brought by diverse ranges of feature values, the standard scale (z-score) was adopted.

For training and assessment, we split the dataset into 80:20 for training and testing, respectively. In the training phase, 5-fold cross-validation was employed for the purpose of model selection

3.1.3 Learning algorithms

- Decision tree
  A decision tree (DT) works by recursively splitting the dataset into subsets based on the features that best separate the data according to a chosen criterion, typically aiming to maximize information gain or minimize impurity at each step. This process continues until a stopping criterion is met, such as a predefined depth or a minimum number of samples in each leaf node.

- k-Nearest Neighbors
  k-Nearest Neighbors (kNN) makes predictions by finding the k-nearest data points (neighbors) in the training dataset to a given query point in the feature space. The prediction is then based on the majority class.

- Logistic Regression
  Logistic Regression (LR) is a statistical model used for binary classification tasks. Despite its name, logistic regression is a classification algorithm rather than a regression algorithm. It models the probability that a given input data point belongs to one of the two classes.

- Random Forest
  Random Forest (RF) is an ensemble learning algorithm. It is based on the idea of building multiple decision trees during training and combining their predictions to make more robust and accurate predictions.

- Bagging
  Bagging, also known as Bootstrap Aggregating, is an ensemble machine learning technique used to improve the performance and robustness of various base learning algorithms, such as decision trees, by combining their predictions. It works by creating multiple subsets of the training dataset through bootstrapping (random sampling with replacement) and training a separate base model on each subset. The final prediction is often determined by averaging (for regression) or voting (for classification) the predictions from these individual models.

- AdaBoost
  AdaBoost (Ada) comes from the name ‘Adaptive Boosting’. It is an ensemble ma-

chine learning technique used primarily for classification tasks. It is designed to improve the performance of weak learners (classifiers that perform slightly better than random guessing) by combining them into a strong ensemble model. AdaBoost assigns higher weights to instances that are difficult to classify and lower weights to correctly classified instances.

3.1.4 Parameter settings

The default parameters setting of the algorithms in the *sklearn* package in *python* were used, which are shown as follows.

- Decision tree
  Criterion: Default value is 'gini'. The criterion determines how the algorithm measures impurity or error at each split.

  Splitter: Default value is 'best', which means the algorithm will choose the best split point based on the selected criterion. An alternative option is 'random', where splits are chosen randomly.

  Max Depth: Default value is 'None', which means the tree will expand until all leaves are pure or contain fewer samples than the minimum samples required for a split (min samples split).

  Min Samples Split: Default value is 2, which is the minimum number of samples required to split an internal node.

  Min Samples Leaf: Default value is 1, which is the minimum number of samples required to be in a leaf node.

- k-Nearest Neighbors
  n Neighbors: Default value is 5. This parameter specifies the number of neighbors to consider when making predictions. By default, it is set to 5, meaning that the algorithm will look at the five nearest neighbors to the query point.

  Weights: Default value is 'uniform'. This parameter determines how the contributions of the nearest neighbors are weighted when making predictions. 'uniform' means that all neighbors have equal weight. Alternatively, you can choose 'distance,' which gives more weight to closer neighbors.

  Algorithm: Default value is 'auto'. This parameter specifies the algorithm used to compute the nearest neighbors. 'auto' will automatically choose the most appropriate algorithm based on the training dataset and other parameters. Other options include 'ball tree,' 'kd tree,' and 'brute.'

  p: Default value is 2. This parameter is the power parameter for the Minkowski distance metric. When p is set to 2, it uses the Euclidean distance metric, which is the most common choice. You can change this value to use other distance metrics, such as Manhattan distance (p=1).

Metric: Default value is ‘minkowski’. This parameter determines the distance metric used when calculating distances between data points. ‘minkowski’ is a general metric that can be customized using the ‘p’ parameter, but other options include ‘manhattan,’ ‘euclidean,’ and more.

- Logistic Regression
  Penalty: Default value is ‘l2’. The ’penalty’ parameter determines the type of regularization applied to the model to prevent overfitting. ‘l2’ stands for L2 regularization (ridge regularization), which adds a penalty term based on the square of the coefficients to the loss function. You can also choose ‘l1’ for L1 regularization (lasso regularization) or ‘none’ for no regularization.

  C: Default value is 1.0. The ‘C’ parameter is the inverse of regularization strength. Smaller values of C indicate stronger regularization, which can help prevent overfitting. Larger values of C reduce the strength of regularization.

  Solver: Default value is ‘lbfgs’. This parameter specifies the algorithm used to optimize the model’s coefficients. The choice of solver depends on the size of the dataset and the type of regularization. Other options include ‘newton-cg’, ‘sag’, and ‘saga’ among others.

  Max Iter: Default value is 100. This parameter sets the maximum number of iterations for the solver to converge. If the solver doesn’t converge within this limit, it stops and returns the best result found so far.

  Multi Class: Default value is ‘auto’. For multi-class classification problems, this parameter specifies how the algorithm should handle multiple classes. ‘auto’ automatically selects ‘ovr’ (one-vs-rest) for binary classification and ‘multinomial’ for multi-class problems. Other options include ‘ovr’ and ’multinomial.’

  Random State: Default value is ‘None’. If set to an integer, it ensures reproducibility by seeding the random number generator for certain parts of the algorithm.

- Random Forest
  n Estimators: Default value is 100. This parameter determines the number of decision trees to build in the forest. Increasing the number of estimators generally leads to better model performance, up to a point.

  Criterion: Default value is ‘gini’. This parameter defines the criterion used to measure the quality of a split in each decision tree. ‘gini’ measures impurity for classification.

  Max Depth: Default value is ‘None’. This parameter controls the maximum depth of each individual decision tree in the forest. If set to ‘None,’ trees are expanded until they contain less than min samples split samples.

  Min Samples Split: Default value is 2. It specifies the minimum number of samples required to split an internal node. A higher value can lead to simpler trees.

Min Samples Leaf: Default value is 1. This parameter determines the minimum number of samples required to be in a leaf node. Increasing it can prevent overfitting.

Max Features: Default value is 'auto'. It represents the maximum number of features to consider for a split at each node. The choices include 'auto' (all features), 'sqrt' (square root of the number of features), 'log2' (log base 2 of the number of features), or an integer value.

Bootstrap: Default value is 'True'. It controls whether the Random Forest algorithm uses bootstrapped samples (sampling with replacement) during training.

Random State: Default value is 'None'. If set to an integer, it ensures reproducibility by seeding the random number generator.

- Bagging
  Base Estimator: Default is a decision tree classifier. This parameter specifies the base model that will be used for each subset of data.

  n Estimators: Default is 10. This parameter determines the number of base models (learners) to create, each trained on a different bootstrap sample.

  Max Samples: Default is 1.0 (which means all samples are used for bootstrapping). It specifies the proportion of the training dataset to use for creating each bootstrap sample. If set to a fraction less than 1, a random subset of the data for each base model is used.

  Max Features: Default is 1.0 (which means all features are used for each base model). It specifies the proportion of features to randomly select for training each base model. If set to a fraction less than 1, a random subset of features is used.

  Bootstrap: Default is 'True', which means bootstrapping is used to create the subsets of data. If set to 'False', the algorithm will use the entire training dataset for each base model.

  Bootstrap Features: Default is 'False', which means features are not bootstrapped (randomly sampled with replacement). If set to 'True', features are also randomly sampled with replacement when training each base model.

  OOB Score: Default is 'False'. If set to 'True', it calculates the out-of-bag (OOB) score, which is an estimate of the model's performance on unseen data based on the samples that were not included in each bootstrap sample.

  Random State: Default is 'None'. If set to an integer, it ensures reproducibility by seeding the random number generator for certain parts of the algorithm.

- AdaBoost
  Base Estimator: Default is a decision tree classifier (DecisionTreeClassifier(max depth=1)). This parameter specifies the base model or weak learner used for each iteration of AdaBoost.

n Estimators: Default is 50. This parameter determines the number of weak learners (classifiers) to train in the ensemble. Increasing the number of estimators often leads to better performance, but it also increases computation time.

Learning Rate: Default is 1.0. It controls the contribution of each weak learner to the final prediction. A smaller learning rate values (e.g., ¡ 1) will make the ensemble more robust and may require more estimators.

Algorithm: Default is 'SAMME.R'. This parameter specifies the AdaBoost algorithm to use. 'SAMME' stands for Stagewise Additive Modeling using a Multiclass Exponential loss, while 'SAMME.R' stands for the real-valued version of SAMME. 'SAMME.R' generally works better and is recommended, but 'SAMME' is suitable for discrete class problems.

Random State: Default is 'None'. If set to an integer, it ensures reproducibility by seeding the random number generator for certain parts of the algorithm.

3.1.5 Experimental framework

This section explains the experimental setup and evaluation metrics. Three sets of experiments were carried out as shown in Figure 3.1. In the first experiment, we investigated the performance of different learning algorithms to classify samples with depression and non-depression on the original dataset. Then, in the second experiment, several resampling methods were applied to improve the classification performance. Finally, in Experiment III, we will investigate the effects of adding selective TMHQ questions to train the optimal predictive model obtain in Experiment II. The detailed setup of each experiment and evaluation metrics used are as follows.

In Experiment I, the learning algorithms used were DT, kNN, LR, RF, Bag, and Ada. These algorithms were selected because they are non-deep learning methods that are commonly used for real-world applications with small-to-medium-sized data.

In Experiment II, different well-known data resampling methods that are well-known and easy to apply were used in an attempt to alleviate the data imbalance problem. The methods used were random undersampling (RUS), random oversampling (ROS), SMOTE (Chawla et al., 2002), SMOTE-NC (Mukherjee and Khushi, 2021), borderline-SMOTE (BL-SMOTE) (Han et al., 2005), and ADASYN (Tang and He, 2015). We followed the parameter settings for all of these methods as presented in their original works.

In Experiment III, question items from TMHQ for depression test were carefully selected by a psychiatric nurse and student mentors. Question items that tend to be disturbing to the users were filtered out. The model performance was tested by experimenting with conceivable combinations of those remaining question items. With the aim of minimizing the number of TMHQ question items used, we experimented with combinations of up to two questions. The baseline model for this experiment was the optimal model obtained from Experiment I and II.

3.1.6 Evaluation measures

To evaluate the classification results, common evaluation metrics for binary-class problems were adopted. The measures include recall (also known as sensitivity or the true positive rate), specificity (the true negative rate), and G-mean. The measures are formulated as shown in Equation 3.1 - 3.3, where TP is the true positive, P is the number positive samples, TN is the true negative, and N is the number negative samples.

$$recall = \frac{TP}{P} \tag{3.1}$$

The recall is a performance metric used in binary classification tasks. It measures the ability of a model to correctly identify all relevant instances of a particular class within a dataset. In other words, it calculates the proportion of true positive predictions (correctly identified instances of the class) out of all actual instances of that class, including both true positives and false negatives. A high recall indicates that the model is effective at capturing most of the relevant instances, while a low recall suggests that it is missing a significant portion of them.

$$specificity = \frac{TN}{N} \tag{3.2}$$

Specificity measures the ability of a model to correctly identify instances of the negative class within a dataset. In other words, it calculates the proportion of true negative predictions (correctly identified instances of the negative class) out of all actual instances of the negative class, including both true negatives and false positives. A high specificity indicates that the model is effective at distinguishing and correctly classifying negative instances, while a low specificity suggests that it is misclassifying.

$$G\text{-}mean = \sqrt{recall * specificity} \tag{3.3}$$

In the context of binary classification, G-mean is the geometric mean of the true positive rate (sensitivity) and true negative rate (specificity) to provide a single performance metric that considers both the ability to correctly identify positive cases and the ability to correctly identify negative cases. It provides a balanced measure of a model's performance in both sensitivity and specificity, making it a useful metric for situations where the accuracy rates of both classes need to be considered equally.

3.2 User satisfaction evaluation

This sections explains how the model was evaluated as an interdisciplinary integration example by users. The ethical approval certificate for the assessment process and user satisfaction tool, granted on May 1,2023, is presented in Appendix A.

3.2.1 Study population

The study population is some of the first-year nursing students of Mae Fah Luang University (academic year 2022). Nursing students were selected because their field of study is relevant to the research at hand, and the researchers find them more fitting compared to students from the School of Information Technology. This preference arises from the teaching methods and subjects in the School of Information Technology, which necessitate the integration of knowledge from other disciplines. Consequently, students across all programs in the School of Information Technology have gained experience in applying their knowledge across various fields from the very beginning of their studies in order to advance their work. The selection of first-year students was based on their likelihood of having the least experience in integrating their field of study with other disciplines, as this is often influenced by their relatively short duration of enrollment at the university.

We aimed for half of the first-year nursing students, which was 55 students out of 109, for the study population. The numbers of samples needed was calculated using Yamane's formula as shown in Equation 3.4, where y is the sample size, Y is the population of the study and e is the margin error in the calculation (Yamane, 1973). Using the common setting of e = 0.05, the numbers of samples needed to complete the questionnaire is 49 individuals.

$$y = \frac{Y}{1 + Y * e^2} \tag{3.4}$$

3.2.2 Data collection tool

In order to assess the model as an illustration of interdisciplinary integration, we gathered user satisfaction by means of an online questionnaire. The user experience satisfaction questions are based on a five-point Likert scale. The scoring system is 1–5, where 5 indicates strong agreement while 1 indicates strong disagreement. The questionnaire was self-developed based on the related literature (Harerimana and Mtshali, 2020) and the four steps of questionnaire development recommended by Davis (1996). The questionnaire is supplemented in Appendix B comprising three sections: Part 1 - General Information, Part 2 - Understanding of Interdisciplinary Learning, and Part 3 - Evaluation of the Presented Innovation.

3.2.3 Data collection process

The gathering of volunteers took place on May 21, 2023. The session commenced with a brief overview of key concepts such as integrative learning, machine learning, depression, depression screening, and associated tools. Subsequently, the predictive model was introduced and its functions and usage were explained to the volunteer nursing students. In order to evaluate user satisfaction, the volunteers were provided with a self-administered online

questionnaire (Google Form). The data collected from this questionnaire will be utilized to evaluate the effectiveness of the showcased cross-disciplinary learning innovation.

3.2.4 Data Analysis

The analysis method used in this part involved examining the results of a survey conducted to evaluate various aspects of the model's effectiveness in interdisciplinary integration. The survey responses were analyzed by calculating average scores for different statements related to comprehension, satisfaction, and other relevant factors. These scores were then discussed and interpreted to gain insights into participants' perceptions and experiences.

Chapter 4

Results and Discussion

4.1 Predictive models

4.1.1 Experiment I

Table 4.1 shows the performance of the learning algorithms across different measures. The highest value in each measure is shown in bold. It can be seen that all learning algorithms provided extremely low recall, where the lowest value was 0.01. Even though the best recall of 0.32 was achieve by decision tree, it is less than random probability indicating failure in detecting depression cases. In contrast, high specificity was achieved across the different algorithms. Specifically, all but DT gave specificity close to 1, where the highest value was 0.99. The major difference in recall and specificity obtained was potentially caused by the imbalance distribution of the sample classes. This is because typical learning algorithms are designed to maximize the overall accuracy and hence bias towards the majority class (Vuttipittayamongkol et al., 2021). DT offered the highest g-mean of 0.52 indicating the optimal trade-off between the accuracy of the positive class and the negative class; however, a higher g-mean would be preferable.

4.1.2 Experiment II

Table 4.2 and 4.3 present results after application of various data resampling techniques. The top four recall and g-mean values are highlighted in bold in Table 4.2 and 4.3, respectively. The highest specificity values are not highlighted here because it has been shown in

Table 4.1: Classification performance of different learning algorithms

	DT	kNN	LR	RF	Bag	Ada
Recall	**0.32**	0.1	0.09	0.01	0.09	0.12
Specificity	0.83	0.97	0.97	**0.99**	0.96	0.97
G-mean	**0.52**	0.31	0.3	0.1	0.29	0.34

Table 4.2: Recall and specificity across different resampling methods

	DT		kNN		LR		RF		Bag		Ada	
	Recall	Spec	Recall	Spec	Recall	Spec	Recall	Spec	Recall	Spec	Recall	Spec
ROS	0.61	0.73	0.46	0.69	0.59	0.68	0.08	0.98	0.13	0.94	0.61	0.73
RUS	0.57	0.62	**0.63**	0.65	0.59	0.65	**0.66**	0.65	0.47	0.68	**0.65**	0.67
SMOTE	0.29	0.84	**0.62**	0.59	0.59	0.65	0.1	0.98	0.13	0.95	0.3	0.91
SMOTE-NC	0.27	0.83	0.61	0.57	0.6	0.67	0.1	0.97	0.15	0.95	0.31	0.9
BL-SMOTE	0.27	0.83	0.61	0.57	0.6	0.67	0.09	0.97	0.15	0.95	0.31	0.9
ADASYN	0.38	0.84	0.49	0.68	0.47	0.8	0.1	0.97	0.2	0.95	0.3	0.92

Table 4.3: G-mean across different resampling methods

	DT	kNN	LR	RF	Bag	Ada
ROS	**0.67**	0.56	0.63	0.28	0.35	**0.67**
RUS	0.59	**0.66**	0.62	**0.65**	0.57	**0.66**
SMOTE	0.49	0.56	0.62	0.31	0.35	0.52
SMOTE-NC	0.47	0.59	0.63	0.31	0.38	0.53
BL-SMOTE	0.47	0.59	0.63	0.3	0.38	0.53
ADASYN	0.56	0.58	0.57	0.31	0.44	0.53

Section 4.1.1 that high specificity without a good balance with recall is not meaningful.

As can be seen in Table 4.2, RUS, on average, could provide better recall than other resampling methods and led to the three highest recall values of 0.66, 0.65, and 0.63 with RF, Adaboost, and kNN, respectively. In addition, RUS with these learning algorithms also resulted in top g-mean values of 0.66 and 0.65 as shown in Table 4.3. Although DT and Adaboost with ROS achieved higher g-mean of 0.67, their recall of 0.61 was lower than those of RF and Ada with RUS. Generally, higher recall with decent g-mean is the priority in the medical and health domain; thus, it can be said that RF and Ada with RUS are the two most preferred methods in this work.

The higher performance of RUS over the other resampling methods could be attributed to the removal of some majority class instances from the overlapping region, which enhances the visibility of minority class instances to the learning algorithm (Vuttipittayamongkol and Elyan, 2020). Similarly, ROS and SMOTE-family methods help increase minority class instances in the overlapping region, but all overlapped majority class instances remain. This results in higher recall, yet to some extent. Lastly, ADASYN has a similar objective to remove overlapped majority class instances; however, less removal is performed (Vuttipittayamongkol et al., 2021).

4.1.3 Experiment III

After filtering out disturbing question items by the expert and student mentors, the remaining items from the TMHQ test are as follows.

Q1: I am experiencing difficulties in maintaining focus on my regular daily activities.

Q2: I am experiencing a lack of enthusiasm in my hobbies.

Q3: I no longer derive the same enjoyment from activities as I used to.

Q4: Within a month, I have unintentionally lost 1-2 kilograms of weight without any underlying physical illness or deliberate effort.

Q5: I am experiencing irregular sleep patterns that are causing disruption and discomfort.

Q6: I am currently experiencing an extended period of time required for me to fall asleep at night.

Table 4.4: Classification results after adding TMHQ question items

Additional Variables	Recall	Specificty	G-mean
Q1	0.88	0.8	0.84
Q2	0.81	0.85	0.83
Q3	**0.92**	0.81	0.86
Q4	0.72	0.79	0.75
Q5	0.85	0.81	0.83
Q6	0.78	0.75	0.76
Q1+Q2	0.86	0.82	0.84
Q1+Q3	0.9	0.84	**0.87**
Q1+Q4	0.8	0.85	0.82
Q1+Q5	0.87	0.86	0.86
Q1+Q6	0.86	0.84	0.85
Q2+Q3	0.88	0.81	0.84
Q2+Q4	0.85	0.84	0.84
Q2+Q5	0.84	**0.87**	0.85
Q2+Q6	0.86	0.84	0.85
Q3+Q4	0.86	0.85	0.85
Q3+Q5	0.89	0.85	**0.87**
Q3+Q6	0.88	0.85	0.86
Q4+Q5	0.78	0.85	0.81
Q4+Q6	0.81	0.81	0.81
Q5+Q6	0.86	0.83	0.84

Each of the aforementioned question items and their corresponding two-item combinations was sequentially incorporated into the dataset utilized in Experiment I and II. This resulted in 21 attempts. The baseline method used in this experiment was Adaboost with RUS, which was among the model achieving the highest sensitivity and G-mean. Results are shown in Table 4.4, where the highest value in each measure is highlighted in bold.

The evident improvement in the performance of the predictive models across all measures indicates the significant impact of incorporating TMHQ question items. This enhancement can be attributed to the direct correlation between these items and the sample labels. As can be seen in Table 4.4, adding Q3 question item alone provided the highest recall of 0.92 and the second highest G-mean of 0.86. This level of recall achieved was not attainable through other options. Although the choice of Q1+Q3 gave comparable result, adding the fewest item(s) is preferred. Thus, according to our objective, the best model that we could achieve was based on 38 demographic and non-clinical questions items including Q3.

4.2 User satisfaction

For assessing the model's success as an example of interdisciplinary integration, we received 50 survey responses. Demographic information of the respondents is presented in Figure 4.1- 4.3.

Figure 4.1 and Figure 4.2 show that most volunteers were female in the age of 19. The results are consistent with the nature of the volunteers, which are first-year nursing students. According to Figure 3, the majority (84%) of volunteer nursing students had not experience interdisciplinary learning; meanwhile, a portion of the students had engaged in learning out-of-field courses, such as information technology, liberal arts, management, and law. The accumulated grade point average (GPAX) of the respondents was 3.25.

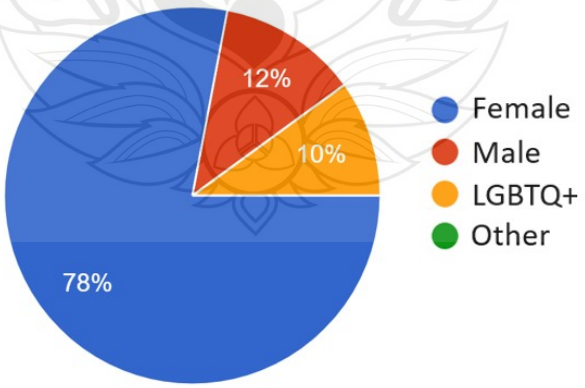

Figure 4.1: The gender distribution of the respondents

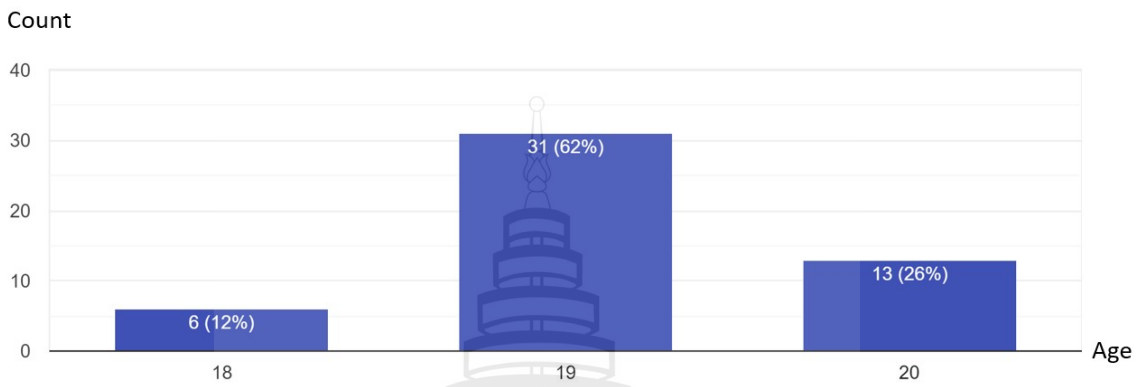

Figure 4.2: The age distribution of the respondents

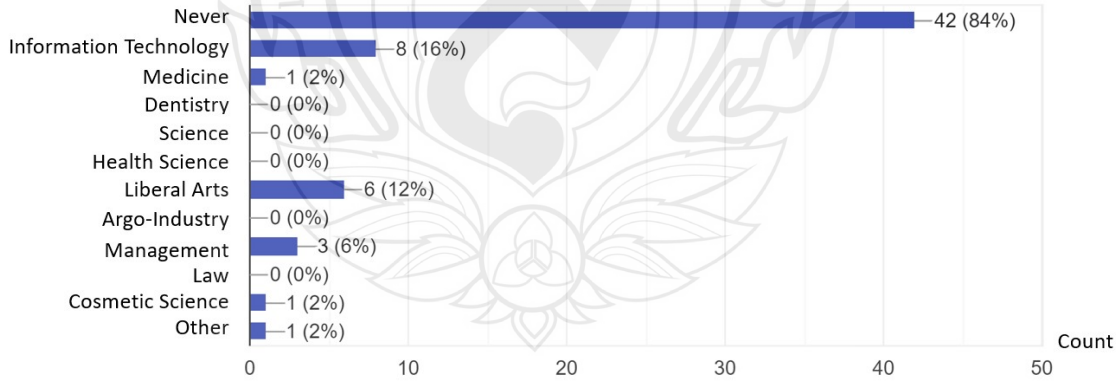

Figure 4.3: The breakdown of the academic disciplines studied by the respondents outside their own fields

Table 4.5 to 4.7 present the mean scores from the survey. Table 4.5 shows the scores regarding comprehension of interdisciplinary integration prior to and following the model demonstration. The increase in scores across all aspects following the demonstration indicates a greater understanding and recognition of the significance of interdisciplinary learning. Their comprehension scores increased from 3.76 to 3.98, showing a better grasp of the concept. Additionally, participants had a clearer picture of the practical application of interdisciplinary integration, with scores increasing from 3.88 to 4.20. They also recognized the importance of interdisciplinary knowledge for their future careers, as indicated by scores increasing from 3.86 to 4.36. Moreover, participants showed increased interest in acquiring knowledge from other fields, with scores rising from 3.88 to 4.10. Overall, it can be concluded from these results that the model demonstration positively influenced participants' comprehension, application, recognition, and interest in interdisciplinary integration.

The survey results in Table 4.6 indicates that the participants highly appreciated the innovation as an instructional medium for interdisciplinary integration, with an average score of 4.16. They found it to be a valuable example that helped them understand the practical application of interdisciplinary knowledge, with an average score of 4.26. The innovation was also recognized as relevant to a nearby and timely problem, receiving an average score of 4.28. Furthermore, the participants strongly agreed that the use of technology in conjunction with medical science could reduce the workload of medical personnel, with an average score of 4.34. The positive feedback reflects the participants' agreement that the predictive model can be serve as a good example of interdisciplinary integration for learners. By using a real-world problem that was closely relevance to the participants and their future career as an example application, the learners will be able to clearly see the importance of interdisciplinary integration.

Results in Table 4.7 indicate a high level of satisfaction with the utility of the model. Participants rated the number of questions in the depression assessment section as appropriate, with a score of 4.22. They also felt that the questions did not have a psychological impact on them, receiving a score of 4.04. Participants found that this set of questions elicited a better emotional response compared to direct mental health assessment questions, scoring it at 4.06. The language used in the model was deemed accurate and effective in conveying the intended message, receiving a score of 4.14. Participants also considered the model's operational principles to be reliable, giving it a score of 4.00. Overall, the satisfaction results indicate that participants found the model's utility and effectiveness to be favorable across various aspects, including question appropriateness, psychological impact, emotional response, language accuracy, and operational reliability. This demonstrates the achievement of our initial goal to investigate the feasibility of employing machine learning for the development of a tool to identify depression in college students at an early stage.

In conclusion, participants expressed an overall satisfaction rating of 4.12 towards this learning tool example for cross-disciplinary integration. While the first objective of this project, which is 'to study the possibility of using machine learning to create a tool for early detection of depression in college students', has been addressed by the result from Table 4.7, the second objective, which is 'to create an innovation of a predictive model

Table 4.5: Survey results on understanding interdisciplinary integration before and after the model demonstration

Question	Before	After
You understand about interdisciplinary integration.	3.76	3.98
You see a clear picture of the application of interdisciplinary integration in solving real problems.	3.88	4.20
You realize the importance of having knowledge in other fields for your future career.	3.86	4.36
You are interested in getting knowledge in other fields to integrate with your own field of study.	3.88	4.10

Table 4.6: Satisfaction results on the predictive model as an example of interdisciplinary integration

Question	Score
This innovation is an interesting instructional medium for interdisciplinary integration.	4.16
It is a good innovative example that helps learners understand the practical application of interdisciplinary knowledge.	4.26
It is an innovation example that relates to a nearby and timely problem.	4.28
The use of technology in conjunction with medical science can help reduce the workload of medical personnel.	4.34

for depression based on non-clinical information serving as an example of the integration of learning across disciplines', is demonstrated through the results from Table 4.5 and Table 4.6. Taking into account the aforementioned results, it can be inferred that this project has successfully achieved all of its objectives.

Table 4.7: Satisfaction results regarding the utility of the model

Question	Score
The number of questions in the section for assessing depression is appropriate	4.22
There is no question that have a psychological impact on the respondents.	4.04
This set of questions provides a better emotional response for the evaluators than direct mental health assessment questions.	4.06
The language used is accurate and conveys the message correctly.	4.14
This model's operational principles are reliable.	4.00

Chapter 5

Conclusion

This project successfully demonstrates the application of interdisciplinary learning through a machine learning-based approach for depression detection in college students. By experimenting six non-deep learning algorithms, the project proves the feasibility of using non-clinical test data together with machine learning techniques in predicting depression. This tackles the issue of the discomfort experienced by individuals when using conventional mental health questionnaires. The achieved competitive classification accuracy highlights the effectiveness of this interdisciplinary approach.

These findings have significant implications for both education and mental health fields. By offering a more comfortable and efficient screening process, this interdisciplinary approach can help identify potential students with depression without relying on lengthy or intimidating questionnaires. It promotes early intervention and support, contributing to improved student well-being and academic success.

The results also fulfill the objective of using the predictive model as an example of interdisciplinary learning. The overall satisfaction rating towards the learning tool example for cross-disciplinary integration further supports the success of this approach. By integrating knowledge from fields such as machine learning, mental health, and education, this project bridges the gap between different disciplines and showcases the practical application of interdisciplinary knowledge.

To encourage further study upon the results attained in this project, the following actions could be undertaken. Firstly, it would be interesting to explore whether this tool could similarly exemplify interdisciplinary learning when extended to students in different academic disciplines. Given that our experimental findings were exclusively drawn from a sample of nursing students, who were closely tied to one of the integrated domains in the model development, the participants might find relatively straightforward to understand the materials. Moreover, in this study, the survey participants gave satisfaction evaluations based solely on model demonstration. To expand upon this research, one could implement and assess a practical application, such as a web or mobile application. Finally, further exploration and utilization of interdisciplinary approaches in diverse domains should be encouraged to meet the needs of learners and address real-world challenges.

Bibliography

Sasha Rudenstine, Kat McNeal, Talia Schulder, Catherine K Ettman, Michelle Hernandez, Kseniia Gvozdieva, and Sandro Galea. Depression and anxiety during the covid-19 pandemic in an urban, low-income public university sample. *Journal of Traumatic Stress*, 34(1):12–22, 2021.

David D Ebert, Claudia Buntrock, Philippe Mortier, Randy Auerbach, Kiona K Weisel, Ronald C Kessler, Pim Cuijpers, Jennifer G Green, Glenn Kiekens, Matthew K Nock, et al. Prediction of major depressive disorder onset in college students. *Depression and anxiety*, 36(4):294–304, 2019.

K Wc Lun, CK Chan, P Ky Ip, S Yk Ma, WW Tsai, CS Wong, C Ht Wong, TW Wong, and D Yan. Depression and anxiety among university students in hong kong. *Hong Kong Med J*, 24(5):466–472, 2018.

Deemah A AlAteeq, Sumayah Aljhani, and Dalal AlEesa. Perceived stress among students in virtual classrooms during the covid-19 outbreak in ksa. *Journal of Taibah University Medical Sciences*, 15(5):398–403, 2020.

Berta Rodrigues Maia and Paulo César Dias. Anxiety, depression and stress in university students: the impact of covid-19. *Estudos de Psicologia (Campinas)*, 37, 2020.

Elena Bernaras, Joana Jaureguizar, and Maite Garaigordobil. Child and adolescent depression: A review of theories, evaluation instruments, prevention programs, and treatments. *Frontiers in psychology*, 10:543, 2019.

Bruno Samways dos Santos, Maria Teresinha Arns Steiner, Amanda Trojan Fenerich, and Rafael Henrique Palma Lima. Data mining and machine learning techniques applied to public health problems: A bibliometric analysis from 2009 to 2018. *Computers & Industrial Engineering*, 138:106120, 2019.

Bassem J Khoory, Maya W Keuning, Anne C Fledderus, Roberta Cicchelli, Vassilios Fanos, Janin Khoory, Deborah Nervi, Eyad Elyan, Pattaramon Vuttipittayamongkol, Matthijs WN Oomen, et al. Psychosocial impact of 8 weeks covid-19 quarantine on italian parents and their children. *Maternal and Child Health Journal*, pages 1–10, 2022.

Md Mehedi Hassan, Md Asif Rakib Khan, Khan Kamrul Islam, Md Mahedi Hassan, and MM Fazle Rabbi. Depression detection system with statistical analysis and data mining approaches. In *2021 International Conference on Science & Contemporary Technologies (ICSCT)*, pages 1–6. IEEE, 2021.

Sucheera Phattharayuttawat, Thienchai Ngamthipwattana, and Kanokrat Sukhatungkha. The norm profile for 'the thai mental health questionnaire'. 15(20):212, 1999.

Aurélien Géron. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. " O'Reilly Media, Inc.", 2022.

Sotiris B Kotsiantis, Dimitris Kanellopoulos, and Panagiotis E Pintelas. Data preprocessing for supervised leaning. *International journal of computer science*, 1(2):111–117, 2006.

Seo-Eun Cho, Zong Woo Geem, and Kyoung-Sae Na. Predicting depression in community dwellers using a machine learning algorithm. *Diagnostics*, 11(8):1429, 2021.

Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

Prince Kumar, Shruti Garg, and Ashwani Garg. Assessment of anxiety, depression and stress using machine learning models. *Procedia Computer Science*, 171:1989–1998, 2020.

Luke Parkitny and James McAuley. The depression anxiety stress scale (dass). *Journal of physiotherapy*, 56(3):204, 2010.

Anu Priya, Shruti Garg, and Neha Prerna Tigga. Predicting anxiety, depression and stress in modern life using machine learning algorithms. *Procedia Computer Science*, 167: 1258–1267, 2020.

Md Sabab Zulfiker, Nasrin Kabir, Al Amin Biswas, Tahmina Nazneen, and Mohammad Shorif Uddin. An in-depth analysis of machine learning approaches to predict depression. *Current research in behavioral sciences*, 2:100044, 2021.

Danxia Liu, Xing Lin Feng, Farooq Ahmed, Muhammad Shahid, Jing Guo, et al. Detecting and measuring depression on social media using a machine learning approach: systematic review. *JMIR Mental Health*, 9(3):e27244, 2022.

Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.

Mimi Mukherjee and Matloob Khushi. Smote-enc: A novel smote-based method to generate synthetic data for nominal and continuous features. *Applied System Innovation*, 4 (1):18, 2021.

Hui Han, Wen-Yuan Wang, and Bing-Huan Mao. Borderline-smote: a new over-sampling method in imbalanced data sets learning. In *International conference on intelligent computing*, pages 878–887. Springer, 2005.

Bo Tang and Haibo He. Kerneladasyn: Kernel based adaptive synthetic data generation for imbalanced learning. In *2015 IEEE congress on evolutionary computation (CEC)*, pages 664–671. IEEE, 2015.

Taro Yamane. Statistics: an introductory analysis-3. 1973.

Alexis Harerimana and Ntombifikile Gloria Mtshali. Using exploratory and confirmatory factor analysis to understand the role of technology in nursing education. *Nurse Education Today*, 92:104490, 2020.

Alice E Davis. Instrument development: getting started. *Journal of Neuroscience Nursing*, 28(3):204–208, 1996.

Pattaramon Vuttipittayamongkol, Eyad Elyan, and Andrei Petrovski. On the class overlap problem in imbalanced data classification. *Knowledge-based systems*, 212:106631, 2021.

Pattaramon Vuttipittayamongkol and Eyad Elyan. Neighbourhood-based undersampling approach for handling imbalanced and overlapped data. *Information Sciences*, 509:47–70, 2020.

Appendices

Appendix A. Ethical Approval

The Mae Fah Luang University Ethics Committee on Human Research
333 Moo 1, Thasud, Muang, ChiangRai 57100
Tel: (053) 917-170 to 71, (053) 916-551 Fax: (053) 917-170 E-mail: rec.human@mfu.ac.th

หนังสือรับรองด้านจริยธรรมการวิจัย

COA: 101/2023 **รหัสโครงการวิจัย: EC 23038-13**

ชื่อโครงการวิจัย : เครื่องมือทำนายภาวะซึมเศร้าในนักศึกษาโดยการใช้การเรียนรู้ของเครื่อง: ตัวอย่างการบูรณาการแบบสหวิทยาการเพื่อพัฒนาทักษะแห่งอนาคต

ชื่อผู้วิจัยหลัก: ผู้ช่วยศาสตราจารย์ ดร.ภัทรมน วุฒิพิทยามงคล

สำนักวิชา: เทคโนโลยีสารสนเทศ

ผู้สนับสนุนทุนวิจัย: โครงการวิจัยเพื่อพัฒนาการเรียนการสอนในชั้นเรียน สถาบันนวัตกรรมการเรียนรู้ มหาวิทยาลัยแม่ฟ้าหลวง ประจำปีงบประมาณ พ.ศ. 2565

การรับรอง :

(1) โครงร่างการวิจัย ฉบับที่ 2 วันที่ 22 เมษายน 2566
(2) เอกสารข้อมูลและขอความยินยอมเข้าร่วมการวิจัย ฉบับที่ 2 วันที่ 22 เมษายน 2566
(3) แบบสอบถาม ฉบับที่ 2 วันที่ 22 เมษายน 2566
(4) ผู้วิจัย และผู้วิจัยร่วม
- ผู้ช่วยศาสตราจารย์ ดร.ภัทรมน วุฒิพิทยามงคล
- อาจารย์ ดร.พิมพ์รัตน์ บุณยะภักดิ์

ขอรับรองว่าโครงการดังกล่าวข้างต้นได้ผ่านการพิจารณารับรองจากคณะกรรมการจริยธรรมการวิจัยในมนุษย์ มหาวิทยาลัยแม่ฟ้าหลวง ว่าสอดคล้องกับแนวทางจริยธรรมสากล ได้แก่ ปฏิญญาเฮลซิงกิ (Declaration of Helsinki) รายงานเบลมองต์ (Belmont Report) แนวทางจริยธรรมสากลสำหรับการวิจัยในมนุษย์ของสภาองค์การสากลด้านวิทยาศาสตร์การแพทย์ (CIOMS) และแนวทางการปฏิบัติการวิจัยที่ดี (ICH GCP)

วันที่รับรองด้านจริยธรรมของโครงร่างการวิจัย: 1 พฤษภาคม 2566

วันสิ้นสุดการรับรอง: 30 เมษายน 2567

ความถี่ของการส่งรายงานความก้าวหน้าของการวิจัย: 1 ปี

ลงนาม .....

(รองศาสตราจารย์ พลตรีหญิง แพทย์หญิง แสงแข ชำนาญวนกิจ)

ประธานคณะกรรมการจริยธรรมการวิจัยในมนุษย์ มหาวิทยาลัยแม่ฟ้าหลวง

The Mae Fah Luang University Ethics Committee on Human Research
333 Moo 1, Thasud, Muang, ChiangRai 57100
Tel: (053) 917-170 to 71, (053) 916-551 Fax: (053) 917-170 E-mail: rec.human@mfu.ac.th

ผู้วิจัยที่โครงร่างการวิจัยผ่านการรับรองจากคณะกรรมการจริยธรรมการวิจัยในมนุษย์ มหาวิทยาลัยแม่ฟ้าหลวง ต้องปฏิบัติดังต่อไปนี้

(1) ดำเนินการวิจัยตามที่ระบุในโครงร่างการวิจัยที่ได้รับการรับรองอย่างเคร่งครัด

(2) ใช้เอกสารข้อมูลและขอความยินยอมเข้าร่วมการวิจัย/แบบสอบถาม/แบบบันทึกข้อมูล ที่มีตราประทับของคณะกรรมการจริยธรรมการวิจัยในมนุษย์ มหาวิทยาลัยแม่ฟ้าหลวง เท่านั้น

(3) ผู้วิจัยต้องส่งแบบรายงานความก้าวหน้าของการวิจัย และขอต่ออายุการรับรองโครงการวิจัย (AP 05/2022) ตามระยะเวลาที่กำหนดโดยคณะกรรมการ <u>ภายใน 30 วัน ก่อนหมดอายุการรับรองครั้งล่าสุด</u> ในกรณีที่การเก็บข้อมูลจากอาสาสมัครยังไม่แล้วเสร็จ ผู้วิจัยต้องส่งรายงานความก้าวหน้าของการวิจัย พร้อมทั้งขอต่ออายุการรับรอง

(4) เมื่อมีการแก้ไขเพิ่มเติมโครงร่างการวิจัย ผู้วิจัยต้องส่งแบบรายงานส่วนแก้ไขเพิ่มเติมโครงร่างการวิจัย (AP 06/2022) และโครงร่างการวิจัยที่มีการแก้ไขเพิ่มเติม เพื่อแจ้งให้คณะกรรมการพิจารณารับรองก่อนดำเนินการตามที่ได้แก้ไขเพิ่มเติม ยกเว้นในกรณีที่ การแก้ไขเพิ่มเติมนั้นกระทำเพื่อความปลอดภัยของอาสาสมัคร

(5) เมื่อมีเหตุการณ์ไม่พึงประสงค์ชนิดร้ายแรง เกิดขึ้นกับอาสาสมัครที่เข้าร่วมในการวิจัยที่ผู้วิจัยรับผิดชอบ ผู้วิจัยต้องส่งแบบรายงานเหตุการณ์ไม่พึงประสงค์ชนิดร้ายแรง ในสถาบัน (AP 07/2022) หรือแบบรายงานตามที่ผู้สนับสนุนการวิจัยกำหนด

(6) เมื่อมีการเบี่ยงเบนจากโครงการวิจัยที่ผ่านการรับรอง หรือไม่ปฏิบัติตามข้อกำหนด ผู้วิจัยต้องส่งแบบรายงานการเบี่ยงเบนจากโครงการวิจัยที่ผ่านการรับรอง หรือการไม่ปฏิบัติตามข้อกำหนด (AP 08/2022)

(7) เมื่อการวิจัยเสร็จสิ้น ผู้วิจัยต้องส่งแบบรายงานสรุปผลการวิจัย (AP 09/2022)

(8) เมื่อมีการยุติโครงการวิจัยก่อนกำหนด ผู้วิจัยต้องส่งแบบรายงานการยุติโครงการวิจัยก่อนกำหนด (AP 10/2022)

<u>หมายเหตุ</u> สามารถ Download แบบรายงานต่าง ๆ ได้ที่ https://ec.mfu.ac.th

ข้าพเจ้าในฐานะ ผู้วิจัย ยินยอมที่จะปฏิบัติตามข้อกำหนดดังกล่าว

.....

(ผู้ช่วยศาสตราจารย์ ดร.ภัทรมน วุฒิพิทยามงคล)

วันที่.....01/06/23.....

The Mae Fah Luang University Ethics Committee on Human Research
333 Moo 1, Thasud, Muang, ChiangRai 57100
Tel: (053) 917-170 to 71, (053) 916-551 Fax: (053) 917-170 E-mail: rec.human@mfu.ac.th

CERTIFICATE OF APPROVAL

COA: 101/2023 **Protocol No: EC 23038-13**

Title: A Machine Learning-Based Tool for Predicting Depression in College Students: An Example of Interdisciplinary Integration for Future Skills Development

Principal investigator: Asst. Prof.Pattaramon Vuttipittayamongkol, Ph.D.

School: Information Technology

Funding support: The research project for the development of teaching and learning in classrooms, under the Institute for Innovative Learning of Mae Fah Luang University, for the fiscal year 2022

Approval:

1) Research protocol Version 2 Date April 22, 2023
2) Information sheet and informed consent documents Version 2 Date April 22, 2023
3) Questionnaire Version 2 Date April 22, 2023
4) Principal investigator and Co-investigators
   - Asst. Prof.Pattaramon Vuttipittayamongkol, Ph.D.
   - Pimrat Boonyapuk, Ph.D.

The aforementioned documents have been reviewed and approved by the Mae Fah Luang University Ethics Committee on Human Research in compliance with international guidelines such as Declaration of Helsinki, the Belmont Report, CIOMS Guidelines and the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use - Good Clinical Practice (ICH GCP)

Date of Approval: May 1, 2023

Date of Expiration: April 30, 2024

Frequency of Continuing Review: 1 year

.....

(Assoc. Prof., Maj. Gen. Sangkae Chamnanvanakij, M.D.)

Chairperson of the Mae Fah Luang Ethics Committee on Human Research

The Mae Fah Luang University Ethics Committee on Human Research
333 Moo 1, Thasud, Muang, ChiangRai 57100
Tel: (053) 917-170 to 71, (053) 916-551 Fax: (053) 917-170 E-mail: rec.human@mfu.ac.th

For research protocol approved by the Mae Fah Luang University Ethics Committee on Human Research (MFU EC), the investigators must comply with the followings:

1. Strictly conduct the research as required by the protocol
2. Use only the information sheet, consent form, questionnaire and case record form bearing the MFU EC stamp of approval
3. Send a progress report form and renewal of previously approved protocol (AP 05/2022) for continuing review and for renewing the approval within 30 days before expiration date.
4. When there are changes of the protocol, the investigator must send a protocol amendment report (AP 06/2022) with amended protocol for MFU EC approval before implementing any changes in the research (Unless those changes are required urgently for the safety of the research subjects).
5. When there is any unanticipated problem or severe adverse event, the investigator must send a safety report or a serious adverse event (SAE) report (AP 07/2022).
6. When there is any deviation or non-compliance with the approved protocol, the investigator must send a protocol deviation/protocol violation/non-compliance report (AP 08/2022).
7. When the research finishes, the investigator must send a final report (AP 09/2022).
8. When the research stops before planned schedule, the investigator must send a premature study termination report (AP 10/2022).

Please go to https://ec.mfu.ac.th to download MFU EC forms for reporting.

I, as an investigator, agree to comply with the above obligation.

.....

(Asst. Prof.Pattaramon Vuttipittayamongkol, Ph.D.)

Date 01/06/23

Appendix B. User satisfaction questionnaire

แบบสอบถามความพึงพอใจในสื่อการเรียนการสอน เครื่องมือทำนายภาวะซึมเศร้าในนักศึกษาโดยการใช้การเรียนรู้ของเครื่อง: ตัวอย่างการบูรณาการแบบสหวิทยาการเพื่อพัฒนาทักษะแห่งอนาคต

แบบสอบถามนี้เป็นส่วนหนึ่งของโครงการวิจัย "เครื่องมือทำนายภาวะซึมเศร้าในนักศึกษาโดยการใช้การเรียนรู้ของเครื่อง: ตัวอย่างการบูรณาการแบบสหวิทยาการเพื่อพัฒนาทักษะแห่งอนาคต" ของมหาวิทยาลัยแม่ฟ้าหลวง เพื่อใช้ในการประเมินนวัตกรรมสำหรับการเรียนการสอบแบบบูรณาการข้ามศาสตร์ โดยข้อมูลที่ได้รับจะนำไปใช้เพื่อเป็นประโยชน์ทางการศึกษาและการวิจัยในการปรับปรุง พัฒนา หรือสร้างนวัตกรรมให้ตรงตามความต้องการของผู้เรียนในอนาคตต่อไป

ความคิดเห็นของท่านจึงเป็นส่วนสำคัญต่อ ความสำเร็จของงานวิจัยนี้และในอนาคต ข้อมูลของท่านจะถูกเก็บรักษาเป็นความลับ ผู้วิจัยจึงใคร่ขอความร่วมมือ จากท่านในการตอบแบบสอบถามอย่างครบถ้วน ขอขอบคุณมา ณ ที่นี้

แบบสอบถามชุดนี้ประกอบไปด้วย 3 ส่วน ดังนี้
ส่วนที่ 1 ข้อมูลทั่วไป
ส่วนที่ 2 ความรู้ความเข้าใจในการบูรณาการข้ามศาสตร์
ส่วนที่ 3 ประเมินตัวอย่างนวัตกรรม

* Indicates required question

ส่วนที่ 1 ข้อมูลทั่วไป

1. เพศ *

Mark only one oval.

- ( ) หญิง
- ( ) ชาย
- ( ) LGBTQ+
- ( ) อื่นๆ

2. อายุ *

_____

3. ผลการเรียนเฉลี่ยสะสม (ตัวอย่าง 3.20) *

_____

4. ในระหว่างที่เรียนที่มหาวิทยาลัยแม่ฟ้าหลวง ท่านเคยเรียนรู้ศาสตร์ในสาขาวิชาอื่นด้วยตนเองหรือลงเรียนในรายวิชาด้วยความสนใจหรือไม่ หากใช่ โปรดเลือกศาสตร์ทั้งหมด *

Check all that apply.

- [ ] ไม่เคย
- [ ] เทคโนโลยีสารสนเทศ
- [ ] แพทย์ศาสตร์
- [ ] ทันตแพทย์ศาสตร์
- [ ] วิทยาศาสตร์
- [ ] วิทยาศาสตร์สุขภาพ
- [ ] ศิลปศาสตร์
- [ ] อุตสาหกรรมเกษตร
- [ ] การจัดการ
- [ ] นิติศาสตร์
- [ ] Other: _____

ส่วนที่ 2 ความรู้ความเข้าใจในการบูรณาการข้ามศาสตร์

โปรดเลือกตัวเลขตรงกับความคิดเห็นของท่านมากที่สุด

1 = ไม่เห็นด้วยอย่างยิ่ง
2 = ไม่เห็นด้วย
3 = เฉยๆ
4 = เห็นด้วย
5 = เห็นด้วยอย่างยิ่ง

5. ก่อนการสาธิตในวันนี้ *

Mark only one oval per row.

	1	2	3	4	5
ท่านมีความเข้าใจเกี่ยวกับการบูรณาการข้ามศาสตร์	◯	◯	◯	◯	◯
ท่านเห็นภาพของการนำการบูรณาการข้ามศาสตร์ไปใช้แก้ปัญหาจริงอย่างชัดเจน	◯	◯	◯	◯	◯
ท่านตระหนักถึงความสำคัญในการมีความรู้ในสาขาวิชาอื่นในการประกอบอาชีพในภายภาคหน้า	◯	◯	◯	◯	◯
ท่านมีความสนใจที่จะเรียนรู้ศาสตร์ในสาขาวิชาอื่นเพื่อบูรณาการกับสาขาวิชาของท่าน	◯	◯	◯	◯	◯

6. หลังการสาธิตในวันนี้ *

Mark only one oval per row.

	1	2	3	4	5
ท่านมีความเข้าใจเกี่ยวกับการบูรณาการข้ามศาสตร์					
ท่านเห็นภาพของการนำการบูรณาการข้ามศาสตร์ไปใช้แก้ปัญหาจริงอย่างชัดเจน					
ท่านตระหนักถึงความสำคัญในการมีความรู้ในสาขาวิชาอื่นในการประกอบอาชีพในภายภาคหน้า					
ท่านมีความสนใจที่จะเรียนรู้ศาสตร์ในสาขาวิชาอื่นเพื่อบูรณาการกับสาขาวิชาของท่าน					

ส่วนที่ 3 ประเมินตัวอย่างนวัตกรรม

โปรดเลือกตัวเลขตรงกับความคิดเห็นของท่านมากที่สุด

1 = ไม่เห็นด้วยอย่างยิ่ง
2 = ไม่เห็นด้วย
3 = เฉยๆ
4 = เห็นด้วย
5 = เห็นด้วยอย่างยิ่ง

7. ภาพรวมของนวัตกรรม *

Mark only one oval per row.

	1	2	3	4	5
นวัตกรรมนี้เป็นสื่อการเรียนการสอนสำหรับการบูรณาการข้ามศาสตร์ที่น่าสนใจ	⬭	⬭	⬭	⬭	⬭
เป็นตัวอย่างนวัตกรรมที่ทำให้ผู้เรียนเข้าใจถึงการประยุกต์ใช้ความรู้ต่างศาสตร์ได้เป็นอย่างดี	⬭	⬭	⬭	⬭	⬭
เป็นตัวอย่างนวัตกรรมที่เกี่ยวข้องกับปัญหาใกล้ตัวและเป็นปัจจุบัน	⬭	⬭	⬭	⬭	⬭
นวัตกรรมนี้แสดงให้เห็นว่าการนำเทคโนโลยีมาใช้ร่วมกับศาสตร์ทางการแพทย์สามารถช่วยลดภาระงานของบุคลากรทางการแพทย์ได้	⬭	⬭	⬭	⬭	⬭

8. ส่วนของการเรียนรู้ของเครื่อง *

Mark only one oval per row.

	5	4	3	2	1
จำนวนคำถามในส่วนของการประเมินภาวะซึมเศร้ามีความเหมาะสม					
ไม่มีคำถามส่งผลกระทบทางจิตใจของผู้ตอบ					
ชุดคำถามนี้ให้ความรู้สึกที่ดีต่อผู้ถูกประเมินมากกว่าคำถามที่ใช้ทดสอบสุขภาพจิตโดยตรง					
ภาษาที่ใช้มีความถูกต้องและสามารถสื่อความให้เข้าใจได้ถูกต้อง					
หลักการทำงานของโมเดลนี้มีความน่าเชื่อถือ					

9. ความพึงพอใจในภาพรวม *

Mark only one oval per row.

	1	2	3	4	5
โดยภาพรวมท่านมีความพึงพอใจกับตัวอย่างการบูรณาการข้ามศาสตร์นี้	◯	◯	◯	◯	◯

10. ข้อเสนอแนะ (Suggestion)

This content is neither created nor endorsed by Google.

Google Forms

Appendix C. Conference proceeding

2023 International Conference On Cyber Management And Engineering (CyMaEn)

A Machine Learning Approach for Depression Screening in College Students Based on Non-Clinical Information

Pornrat Nison
The Center of Excellence in Artificial Intelligence and Emerging Technologies
School of Information Technology
Mae Fah Luang University, Thailand
pornrat37@gmail.com

Pattaramon Vuttipittayamongkol*
The Center of Excellence in Artificial Intelligence and Emerging Technologies
School of Information Technology
Mae Fah Luang University, Thailand
*Corresponding: pattaramon.vut@mfu.ac.th

Pimrat Boonyapuk
School of Nursing
Mae Fah Luang University, Thailand
pimrat@mfu.ac.th

Kemachart Kemavuthanon
School of Information Technology
Mae Fah Luang University, Thailand
kemachart.kem@mfu.ac.th

2023 International Conference On Cyber Management And Engineering (CyMaEn) | 978-1-6654-9329-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/CyMaEn57228.2023.10051001

***Abstract*—Depressive disorder is particularly prevalent in college students. Visiting a mental health expert such as a psychologist or a psychiatrist is a common and reliable practice to get depression screening. However, it is often costly and time-taking, and an individual may feel embarrassed to see a psychologist. Although taking a self-administered standard questionnaire is an alternative, existing questionnaires often contain a lot of items and even some aggressive questions. This discourages the taker from completing the questionnaire with precise answers or even leads to denying the assessment. In this paper, we present a machine-learning approach for detecting potential college students suffering from depression based on their general information. The items used for training the model are such as demographic information, physical health problems, relationships with family and friends, and aspects towards the university life while none of the direct questions from the mental health measure was included. Furthermore, the imbalanced distribution of depression and non-depression samples, which hinders the learning process and causes a bias towards the majority class, was addressed. Several resampling methods were employed to alleviate the class imbalance, and hence improve the classification. Results show that the average predicting accuracy in both classes up to 0.66 could be achieved. This paves the way for the research community to progress in machine learning-based depression screening approaches.**

***Index Terms*—depression, mental health screening, college students, machine learning, classification, imbalanced data**

I. Introduction

Major depressive disorder (MDD), also known as clinical depression, is one of the most common mental disorders among college students [1]. Depression in students can be caused by socioeconomic status [2] and declining academic performance [1]. Entering a new society as an incoming freshman as well as high expectations from the parent elevate the level of stress and anxiety in college students, which could eventually lead to depression and college dropout [3]. Physical impacts on depressed students include risky sexual behavior, smoking, alcohol and drug dependency, self-harm behaviors, and in the worst scenario, the risk of suicide.

There are several ways to determine whether an individual is suffering from depression. Getting a direct consultation and examination by a psychiatrist is the most reliable practice; however, it is costly and takes up time. Various self-administered psychometric tests for assessing depression are also widely available [4]. However, these tests are generally lengthy. A large number of questions discourages the taker to complete the test with precise answers. As a consequence, the assessment result would not be accurate. The main drawback of these traditional assessments is that the person realizes that he or she is being tested for depression. This may bring negative feelings to the subject and refusal of the assessment.

Emerging methods based on machine learning has been proposed to predict the likelihood of developing depression. Cho et al. used data consisting of sociological factors, physiological factors (e.g., the prevalence of chronic diseases), psychological factors, and laboratory tests(e.g.,blood cell counts) to create a classification model [5]. LASSO, a regression method that also performs variable selection [6], was used to learn the data. The predicting accuracy of over 80% in both classes was obtained with the 13 remaining variables. However, the requirement of laboratory results makes this method not readily accessible to average people.

Kumar et al. presented the use of several standard learning algorithms as well as neural network-based algorithms to predict the severity level of psychological disorders [7]. The

This work was financially supported in part by MFU Learning Innovation Institute, Mae Fah Luang University and the Center of Excellence in Artificial Intelligence and Emerging Technologies, School of Information Technology, Mae Fah Luang University.

Authorized licensed use limited to: Mae Fah Luang University. Downloaded on May 15,2023 at 09:31:40 UTC from IEEE Xplore. Restrictions apply.

focused psychological problems were anxiety, depression, and stress. For each disorder, machine learning was employed to learn and predict the five severity levels based on 14 Likert-scaled items of a standard questionnaire used by psychiatrists that is the Depression, Anxiety, and Stress Scale questionnaire (DASS) [8]. Priya et al. later proposed a similar work [9]. It should be pointed out that the use of machine learning would not be encouraged in this case. The severity levels were directly proportional to the sum of the answer scores; thus, using machine learning would not be a faster and more accurate way to classify the disorder level.

In [10], Zulfiker et al. properly trained their predictive models to classify depressed and non-depressed persons based on socio-demographic and psychosocial information. Several feature selection techniques, namely, SelectKBest, mRMR, and Boruta, were applied resulting in 13 to 15 out of 30 remaining features for model learning. Results showed accuracy and recall of over 90% on most methods. However, some question items on the psychosocial information, for example, "Whether the participant has any suicidal thought recently or not", could make the subject feel intimidated and uncomfortable to answer.

In this paper, we present the use of general information solely to predict depression in college students with machine learning methods and further demonstrate classification improvement carried out using data resampling methods to address the class imbalance issue. The main contributions of this work are as follows. First, we investigated the possibility of using only the general information related to college students in building machine learning models to detect students with depression. Unlike many existing works, we have dropped aggressive and intimidating questions that can be perceived by the subject as direct implications of depression from the required data for training a model. Our hypothesis was that some general information can be associated with depression symptoms and machine learning algorithms could aid such a process. Second, the class imbalance problem, which is common in health-related datasets and exists in many depression datasets, is addressed to improve the classification, especially the detection of students with depression.

II. Methodology

In this section, the methodology of the proposed work is provided in detail. First, the dataset is introduced, and data preprocessing is discussed. Then, the flow of the experiment settings, which consists of two separate experiments, is outlined. A detailed explanation of each experiment will be given. This includes the learning algorithms and the data resampling methods used. Finally, evaluation metrics are listed. The flow of the experimental framework is illustrated in Fig. 1.

A. Dataset and data selection

The dataset used in this work was obtained from the research project "The Situation and Educational Productivity System of Undergraduate Level at Mae Fah Luang University" in 2017. The data was collected in form of a questionnaire, which consists of six parts: general information, mental health measures, emotional intelligence assessment, learning style assessment, assessment of cognitive skills and assessment of student's skills according to the higher education qualifications framework. The questionnaire contains a total of 270 questions, 70 of which are the Thai Mental Health Questionnaire (TMHQ) [11]. The TMHQ includes 20 indicators of depression, and most of them can be easily perceived by non-experts as mental health questions.

In this work, only the first part of the original questionnaire was used. This part consists of 45 items covering demographic information, physical health problems, relationships with family and friends, existence of problems in different aspects, etc. The dataset has 4,642 instances representing the number of students participating in the questionnaire. Among these students, 811 of them had indications of depression-related symptoms. We identified them according to the 20 depression test scores available in the second part of the data based on the rubric of TMHQ. Thus, this results were not diagnoses made by a doctor. However, for the ease of discussion, we will simply refer to them as depression samples, and the others as non-depression samples. The difference in the number of samples in the two categories makes the dataset imbalanced with 3831:811 or 4.72 ratio of the negative (non-depression) to the positive samples (depression).

B. Data preprocessing

Firstly, the items with too much variation in responses were removed. This includes, for example, home address, relaxing activities, and health problems. Moreover, items containing a large number of missing values and excessive outliers were also excluded from the dataset. These were student income and family income, which contain 1,575 and 1,524 missing values, respectively. After the removals, there were 37 remaining attributes. The variables with few missing values that are GPAX and number of close friends, which have 11 and 6 missing values, respectively, were retained for the experiment. The samples with these missing values were instead removed from the dataset with a minimal impact on the dataset size.

The 13 schools were grouped into 3 main categories: 1) Science and Technology, 2) Health and Biomedical Sciences, and 3) Social Sciences and Humanities based on the literature [2]. Furthermore, we employed one-hot encoding to convert categorical data to numerical data. To avoid bias brought by diverse ranges of feature values, the standard scale (z-score) was adopted. Finally, for training and assessment, we split the dataset into 80:20 for training and testing, respectively.

C. Experimental framework

This section explains the experimental setup and evaluation metrics. Two experiments were carried out. In the first experiment, we investigated the performance of different learning algorithms to classify samples with depression and non-depression on the original dataset. Then, in the second experiment, several resampling methods were applied to im-

Authorized licensed use limited to: Mae Fah Luang University provided by UniNet. Downloaded on May 15,2023 at 09:31:40 UTC from IEEE Xplore. Restrictions apply.

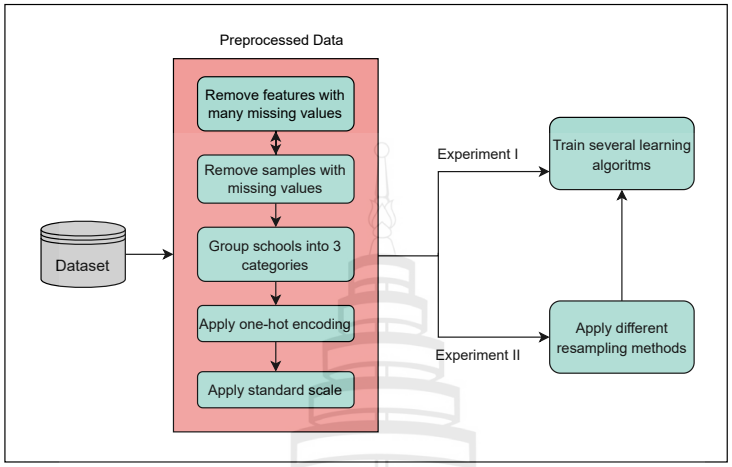

Fig. 1: The experimental framework

prove the classification performance. The detailed setup of each experiment and evaluation metrics used are as follows.

In Experiment I, the learning algorithms used were decision tree (DT), k-nearest neighbor (kNN), logistic regression (LR), random forest (RF), bagging (Bag), and Adaboost (Ada). These algorithms were selected because they are non-deep learning methods that are commonly used for real-world applications with small-to-medium-sized data. The default parameters setting of the algorithms in the *sklearn* package in *python* were used.

In Experiment II, different well-known data resampling methods that are well-known and easy to apply were used in an attempt to alleviate the data imbalance problem. The methods used were random undersampling (RUS), random oversampling (ROS), SMOTE [12], SMOTE-NC [13], borderline-SMOTE (BL-SMOTE) [14], and ADASYN [15]. We followed the parameter settings for all of these methods as presented in their original works.

To evaluate the classification results, common evaluation metrics for binary-class problems were adopted. The measures include recall (also known as sensitivity or the true positive rate), specificity (the true negative rate), and G-mean. The measures are formulated as shown in Equation 1, 2, and 3, where TP is the true positive, P is the number positive samples, TN is the true negative, and N is the number negative samples.

$$recall = \frac{TP}{P} \tag{1}$$

$$specificity = \frac{TN}{N} \tag{2}$$

TABLE I: Classification performance of different learning algorithms

	DT	kNN	LR	RF	Bag	Ada
Recall	**0.32**	0.1	0.09	0.01	0.09	0.12
Specificity	0.83	0.97	0.97	**0.99**	0.96	0.97
G-mean	**0.52**	0.31	0.3	0.1	0.29	0.34

$$G\text{-}mean = \sqrt{recall * specificity} \tag{3}$$

III. Results and Discussion

A. Experiment I

Table I shows the performance of the learning algorithms across different measures. The highest value in each measure is shown in bold. It can be seen that all learning algorithms provided extremely low recall, where the lowest value was 0.01. Even though the best recall of 0.32 was achieve by decision tree, it is less than random probability indicating failure in detecting depression cases. In contrast, high specificity was achieved across the different algorithms. Specifically, all but DT gave specificity close to 1, where the highest value was 0.99. The major difference in recall and specificity obtained was potentially caused by the imbalance distribution of the sample classes. This is because typical learning algorithms are designed to maximize the overall accuracy and hence bias towards the majority class [16]. DT offered the highest g-mean of 0.52 indicating the optimal trade-off between the accuracy of the positive class and the negative class; however, a higher g-mean would be preferable.

Authorized licensed use limited to: Mae Fah Luang University provided by UniNet. Downloaded on May 15,2023 at 09:31:40 UTC from IEEE Xplore. Restrictions apply.

TABLE II: Recall and specificity across different resampling methods

	DT		kNN		LR		RF		Bag		Ada	
	Recall	Spec	Recall	Spec	Recall	Spec	Recall	Spec	Recall	Spec	Recall	Spec
ROS	0.61	0.73	0.46	0.69	0.59	0.68	0.08	0.98	0.13	0.94	0.61	0.73
RUS	0.57	0.62	**0.63**	0.65	0.59	0.65	**0.66**	0.65	0.47	0.68	**0.65**	0.67
SMOTE	0.29	0.84	**0.62**	0.59	0.59	0.65	0.1	0.98	0.13	0.95	0.3	0.91
SMOTE-NC	0.27	0.83	0.61	0.57	0.6	0.67	0.1	0.97	0.15	0.95	0.31	0.9
BL-SMOTE	0.27	0.83	0.61	0.57	0.6	0.67	0.09	0.97	0.15	0.95	0.31	0.9
ADASYN	0.38	0.84	0.49	0.68	0.47	0.8	0.1	0.97	0.2	0.95	0.3	0.92

TABLE III: G-mean across different resampling methods

	DT	kNN	LR	RF	Bag	Ada
ROS	**0.67**	0.56	0.63	0.28	0.35	**0.67**
RUS	0.59	**0.66**	0.62	**0.65**	0.57	**0.66**
SMOTE	0.49	0.56	0.62	0.31	0.35	0.52
SMOTE-NC	0.47	0.59	0.63	0.31	0.38	0.53
BL-SMOTE	0.47	0.59	0.63	0.3	0.38	0.53
ADASYN	0.56	0.58	0.57	0.31	0.44	0.53

B. Experiment II

Table II and III present results after application of various data resampling techniques. The top four recall and g-mean values are highlighted in bold in Table II and III, respectively. The highest specificity values are not highlighted here because it has been shown in Section III-A that high specificity without a good balance with recall is not meaningful.

As can be seen in Table II, RUS, on average, could provide better recall than other resampling methods and led to the three highest recall values of 0.66, 0.65, and 0.63 with RF, Adaboost, and kNN, respectively. In addition, RUS with these learning algorithms also resulted in top g-mean values of 0.66 and 0.65 as shown in Table III. Although DT and Adaboost with ROS achieved higher g-mean of 0.67, their recall of 0.61 was lower than those of RF and Ada with RUS. Generally, higher recall with decent g-mean is the priority in the medical and health domain; thus, it can be said that RF and Ada with RUS are the two most preferred methods in this work.

The higher performance of RUS over the other resampling methods could be attributed to the removal of some majority class instances from the overlapping region, which enhances the visibility of minority class instances to the learning algorithm [17]. Similarly, ROS and SMOTE-family methods help increase minority class instances in the overlapping region, but all overlapped majority class instances remain. This results in higher recall, yet to some extent. Lastly, Adasyn has a similar objective to remove overlapped majority class instances; however, less removal is performed [16].

IV. Conclusion

This paper proposes a machine learning-based approach for depression detection in college students using none of clinical information or direct questions from mental health measures and presents the use of data resampling methods to handle the under-representation problem of samples with depression. Predictive models were built upon six non-deep learning algorithms across six different resampling techniques applied. Random forest and Adaboost with random undersampling showed the most optimal classification accuracy. The two provided competitive results with the average accuracy of both classes up to 0.66. Results suggest that by using only general information of college students, depression screening using machine learning could be achieved. This could help detect potential students with depression without making them uncomfortable to answer questions in a standard mental health questionnaire, which are often lengthy or intimidating. The findings in this research will pave the way for further improvements on such an approach for the actual usage in the future. A potential attempt is possibly injecting some questions from a standard mental health test that are carefully selected so they would not find them aggressive or intimidating.

References

[1] S. Rudenstine, K. McNeal, T. Schulder, C. K. Ettman, M. Hernandez, K. Gvozdieva, and S. Galea, "Depression and anxiety during the covid-19 pandemic in an urban, low-income public university sample," *Journal of Traumatic Stress*, vol. 34, no. 1, pp. 12–22, 2021.

[2] D. D. Ebert, C. Buntrock, P. Mortier, R. Auerbach, K. K. Weisel, R. C. Kessler, P. Cuijpers, J. G. Green, G. Kiekens, M. K. Nock *et al.*, "Prediction of major depressive disorder onset in college students," *Depression and anxiety*, vol. 36, no. 4, pp. 294–304, 2019.

[3] K. W. Lun, C. Chan, P. K. Ip, S. Y. Ma, W. Tsai, C. Wong, C. H. Wong, T. Wong, and D. Yan, "Depression and anxiety among university students in hong kong," *Hong Kong Med J*, vol. 24, no. 5, pp. 466–472, 2018.

[4] E. Bernaras, J. Jaureguizar, and M. Garaigordobil, "Child and adolescent depression: A review of theories, evaluation instruments, prevention programs, and treatments," *Frontiers in psychology*, vol. 10, p. 543, 2019.

[5] S.-E. Cho, Z. W. Geem, and K.-S. Na, "Predicting depression in community dwellers using a machine learning algorithm," *Diagnostics*, vol. 11, no. 8, p. 1429, 2021.

[6] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[7] P. Kumar, S. Garg, and A. Garg, "Assessment of anxiety, depression and stress using machine learning models," *Procedia Computer Science*, vol. 171, pp. 1989–1998, 2020.

[8] L. Parkitny and J. McAuley, "The depression anxiety stress scale (dass)," *Journal of physiotherapy*, vol. 56, no. 3, p. 204, 2010.

[9] A. Priya, S. Garg, and N. P. Tigga, "Predicting anxiety, depression and stress in modern life using machine learning algorithms," *Procedia Computer Science*, vol. 167, pp. 1258–1267, 2020.

Authorized licensed use limited to: Mae Fah Luang University provided by UniNet. Downloaded on May 15,2023 at 09:31:40 UTC from IEEE Xplore. Restrictions apply.

[10] M. S. Zulfiker, N. Kabir, A. A. Biswas, T. Nazneen, and M. S. Uddin, "An in-depth analysis of machine learning approaches to predict depression," *Current research in behavioral sciences*, vol. 2, p. 100044, 2021.

[11] S. Phattharayuttawat, T. Ngamthipwattana, and K. Sukhatungkha, "The norm profile for 'the thai mental health questionnaire'," vol. 15, no. 20, p. 212, 1999.

[12] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[13] M. Mukherjee and M. Khushi, "Smote-enc: A novel smote-based method to generate synthetic data for nominal and continuous features," *Applied System Innovation*, vol. 4, no. 1, p. 18, 2021.

[14] H. Han, W.-Y. Wang, and B.-H. Mao, "Borderline-smote: a new over-sampling method in imbalanced data sets learning," in *International conference on intelligent computing*. Springer, 2005, pp. 878–887.

[15] B. Tang and H. He, "Kerneladasyn: Kernel based adaptive synthetic data generation for imbalanced learning," in *2015 IEEE congress on evolutionary computation (CEC)*. IEEE, 2015, pp. 664–671.

[16] P. Vuttipittayamongkol, E. Elyan, and A. Petrovski, "On the class overlap problem in imbalanced data classification," *Knowledge-based systems*, vol. 212, p. 106631, 2021.

[17] P. Vuttipittayamongkol and E. Elyan, "Neighbourhood-based undersampling approach for handling imbalanced and overlapped data," *Information Sciences*, vol. 509, pp. 47–70, 2020.

Authorized licensed use limited to: Mae Fah Luang University provided by UniNet. Downloaded on May 15,2023 at 09:31:40 UTC from IEEE Xplore. Restrictions apply.

Researcher Profiles

1. Asst. Prof. Dr. Pattaramon Vuttipittayamongkol

- **Position** Associate Dean
- **Affiliation** School of Information Technology, Mae Fah Luang University
- **Education**

2017 - 2020	Ph.D. in Data Science, Robert Gordon University, UK
2011 - 2013	M.Sc. in Electrical Engineering, University of Southern California, USA
2007 - 2011	B.Sc. in Electrical Engineering, University of Illinois at Urbana-Champaign, USA

- **Publications**

  - Nison, P., Vuttipittayamongkol, P., Boonyapuk, P., & Kemavuthanon, K. (2023, January). A Machine Learning Approach for Depression Screening in College Students Based on Non-Clinical Information. In 2023 International Conference On Cyber Management And Engineering (CyMaEn) (pp. 413-417). IEEE.
  - Vuttipittayamongkol, P., & Arreeras, T. (2022, December). Data-driven Industrial Machine Failure Detection in Imbalanced Environments. In 2022 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM) (pp. 1224-1227). IEEE.
  - Arreeras, T., Sittivangkul, K., Tiwong, S., & Vuttipittayamongkol, P. (2022, December). Exploration of Risky Riding Behavior on Last Mile Food Delivery using Motorcycle Rider Behavior Questionnaire: Evidence From Chiang Rai. In 2022 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM) (pp. 1088-1091). IEEE.
  - Elyan, E., Vuttipittayamongkol, P., Johnston, P., Martin, K., McPherson, K., Jayne, C., & Sarker, M. K. (2022). Computer vision and machine learning for medical image analysis: recent advances, challenges, and way forward. Artificial Intelligence Surgery, 2.

- Elyan, E., Hussain, A., Sheikh, A., Elmanama, A. A., Vuttipittayamongkol, P., & Hijazi, K. (2022). Antimicrobial Resistance and Machine Learning: Challenges and Opportunities. IEEE Access, 10, 31561-31577.

- Khoory, B. J., Keuning, M. W., Fledderus, A. C., Cicchelli, R., Fanos, V., Khoory, J., ... & Abu Hilal, M. (2022). Psychosocial Impact of 8 weeks COVID-19 quarantine on Italian parents and their children. Maternal and Child Health Journal, 26(6), 1312-1321.

- Vuttipittayamongkol, P., Tung, A., & Elyan, E. (2021). A Data-Driven Decision Support Tool for Offshore Oil and Gas Decommissioning. IEEE Access, 9, 137063-137082.

- Vuttipittayamongkol, P., Tung, A., & Elyan, E. (2021, July). Towards Machine Learning-Driven Practices for Oil and Gas Decommissioning–Introduction of a New Offshore Pipeline Dataset. In Proceedings of the 9th International Conference on Computer and Communications Management (pp. 111-116).

- Vuttipittayamongkol, P., Elyan, E., & Petrovski, A. (2021). On the class overlap problem in imbalanced data classification. Knowledge-based systems, 212, 106631.

- Vuttipittayamongkol, P., & Elyan, E. (2020). Improved overlap-based undersampling for imbalanced dataset classification with application to epilepsy and parkinson's disease. International journal of neural systems, 30(08), 2050043.

- Vuttipittayamongkol, P., & Elyan, E. (2020). Overlap-based undersampling method for classification of imbalanced medical datasets. In Artificial Intelligence Applications and Innovations: 16th IFIP WG 12.5 International Conference, AIAI 2020, Neos Marmaras, Greece, June 5–7, 2020, Proceedings, Part II 16 (pp. 358-369). Springer International Publishing.

- Vuttipittayamongkol, P., & Elyan, E. (2020). Neighbourhood-based undersampling approach for handling imbalanced and overlapped data. Information Sciences, 509, 47-70.

- Vuttipittayamongkol, P., Elyan, E., Petrovski, A., & Jayne, C. (2018). Overlap-based undersampling for improving imbalanced data classification. In Intelligent Data Engineering and Automated Learning–IDEAL 2018: 19th International Conference, Madrid, Spain, November 21–23, 2018, Proceedings, Part I 19 (pp. 689-697). Springer International Publishing.

- Vuttipittayamongkol, P. (2016, January). Predicting factors of academic performance. In 2016 Second asian conference on defence technology (ACDT) (pp. 161-166). IEEE.

2. Dr. Kemachart Kemavuthanon

- **Position** Assistant Dean
- **Affiliation** School of Information Technology, Mae Fah Luang University
- **Education**

  2021 Ph.D. in Engineering, Information Science and Technology, Tokai University, Japan

  2008 M.Sc. in Information Technology Management, Mahanakorn University of Technology, Thailand

  2006 B.Sc. in Information Technology, Mahanakorn University of Technology, Thailand

- **Publications**

  - Nison, P., Vuttipittayamongkol, P., Boonyapuk, P., & Kemavuthanon, K. (2023, January). A Machine Learning Approach for Depression Screening in College Students Based on Non-Clinical Information. In 2023 International Conference On Cyber Management And Engineering (CyMaEn) (pp. 413-417). IEEE.
  - Kemavuthanon, K., Rueangsirarak, W., Kemavuthanon, S., & Temdee, P. (2023, March). Tracking and Notification Application for Arrhythmia Monitoring System. In 2023 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON) (pp. 554-559). IEEE.
  - Kemavuthanon, K., Manomaivibool, P., & Kemavuthanon, S. (2023, March). Integrate system design and development of Global Sustainable Tourism Criteria system (GSTC). In 2023 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON) (pp. 51-55). IEEE.
  - Kemavuthanon, K., Manomaivibool, P., & Eungwanichayapant, A. (2022, November). System Design of a Toxic waste management system Empirical study for Chiang Rai province. In 2022 6th International Conference on Information Technology (InCIT) (pp. 157-161). IEEE.
  - Kemavuthanon, K., & Uchida, O. (2021, April). System Design and Analysis major factors to developing Mobile-based Apps of natural disasters. In 2021 IEEE 6th International Conference on Computer and Communication Systems (ICCCS) (pp. 1122-1127).

- Kemavuthanon, K., & Uchida, O. (2020). Classification of social media messages posted at the time of disaster. In Information Technology in Disaster Risk Reduction: 4th IFIP TC 5 DCITDRR International Conference, ITDRR 2019, Kyiv, Ukraine, October 9–10, 2019, Revised Selected Papers 4 (pp. 212-226). Springer International Publishing.

- Kemavuthanon, K., & Uchida, O. (2020). Integrated question-answering system for natural disaster domains based on social media messages posted at the time of disaster. Information, 11(9), 456.

- Kemavuthanon, K., & Uchida, O. (2019, August). Category Classification of Tweet Messages at the Time of Disaster. In IEICE Conferences Archives. The Institute of Electronics, Information and Communication Engineers.

- Kemavuthanon, K., & Uchida, O. (2019, March). Social media messages during disasters in Japan: an empirical study of 2018 Osaka North Earthquake in Japan. In 2019 IEEE 2nd International Conference on Information and Computer Technologies (ICICT) (pp. 199-203). IEEE.

- Kemavuthanon, K. (2017, October). Integrated e-project collaborative management system: Empirical study for problem-based learning project. In 2017 9th International Conference on Information Technology and Electrical Engineering (ICITEE) (pp. 1-5). IEEE.

- Panwong, P., & Kemavuthanon, K. (2014). Problem-based learning framework for junior software developer: Empirical study for computer programming students. Wireless personal communications, 76, 603-613.

3. Dr. Pimrat Boonyapuk

- **Position** Associate Dean
- **Affiliation** School of Nursing, Mae Fah Luang University
- **Education**

  2015 Ph.D. in Nursing, Case Western Reserve University, USA

  2002 M.A. in Educational Psychology and Guidance, Kasetsart University, Thailand

  1995 B.Sc. in Boromarajonani College of Nursing, Thailand

- **Publications**

  - Nison, P., Vuttipittayamongkol, P., Boonyapuk, P., & Kemavuthanon, K. (2023, January). A Machine Learning Approach for Depression Screening in College Students Based on Non-Clinical Information. In 2023 International Conference On Cyber Management And Engineering (CyMaEn) (pp. 413-417). IEEE.
  - Boonyapuk P., Mekloi C., Sundern R., & Singkhorn O. (2021). The development of Mental Health Caring System for Students in Higher Education Institutions. Journal of Nursing and Education 2021; 14(4): 33-49.
  - Thohinugn U., Boonyapuk P., & Kodyee S. (2022). Predictive Factors of Parental Behaviors on Promote Development of Executive Functions (EF) Skills among Early Childhood in Chaing Rai Province. Journal of The Royal Thai Army Nurses 2022; 23(1): 344-352.
  - Santiprasitkul S., Junlapeey P., Boonyapuk P., & Srimanthayamas, V. (2019). Patient Safety Culture: Perception of State Hospital Nurses. Journal of The Royal Thai Army Nurses 2019; 20(2): 327-339.
  - Potjanamart C., Kasemkitwatana S., Moampanane K., Suthipong, J, Boonyapuk P., & Wiangosot w. (2010). Way of life, behavior and health impact of alcohol consumption among agriculture workers: A case study in Mae Khum Sub District, Mae Chan District, Chiang Rai Province. Journal of Public Health and Development. 2010; 8(3): 319-333.
  - Santiprasitkul S.,Wongnat, S., Mitranan, S., Srimantayamas V. & BoonyapukP. (2009). Health status and health risk behaviors among students at Mae Fah Luang University. Journal of Public Health. 2009; 39(3): 296-307.