Marquette University
e-Publications@Marquette

Dissertations (1934 -)

Dissertations, Theses, and Professional Projects

Predicting Mental Health Crisis in Veterans: Early Warning Signs, Precursors and Protective Factors

Priyanka Annapureddy
Marquette University

Follow this and additional works at: https://epublications.marquette.edu/dissertations_mu

Part of the Computer Sciences Commons

Recommended Citation
Annapureddy, Priyanka, "Predicting Mental Health Crisis in Veterans: Early Warning Signs, Precursors and Protective Factors" (2022). *Dissertations (1934 -)*. 2071.
https://epublications.marquette.edu/dissertations_mu/2071

PREDICTING MENTAL HEALTH CRISIS IN VETERANS: EARLY WARNING SIGNS, PRECURSORS AND PROTECTIVE FACTORS

by

Priyanka Annapureddy

A Dissertation submitted to the Faculty of the Graduate School,
Marquette University,
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy

Milwaukee, Wisconsin

December 2022

ABSTRACT

PREDICTING MENTAL HEALTH CRISIS IN VETERANS: EARLY WARNING SIGNS, PRECURSORS AND PROTECTIVE FACTORS

Priyanka Annapureddy

Marquette University, 2022

Mental Health (MH) conditions have recently increased to a large extent due to socio-demographic changes. Posttraumatic Stress Disorder (PTSD) is one of the most common mental health disorders prevalent in US. PTSD is even more troubling at double the rate in combat veterans leaving their service compared to general population. Severity of PTSD is associated with risk taking behaviors such as substance abuse, non-suicidal self-injury, and sexual risk behaviors. Psychological disorders are often preceded by early warning signs and recognizing the early warning signs of PTSD will help in preventing the returning or worsening of PTSD symptoms. Ecological momentary assessment (EMA) studies are more sophisticated in tracking fluctuations of symptoms real-time, and they are effective in monitoring for crisis events in veterans.

Mobile applications are commonly used means to gather such EMA information from participants. Our research focuses on developing interpretable machine learning (ML) models using socio-demographic data and EMA data from natural settings to predict high PTSD risk in veterans and those who engage in risky behaviors. Findings from these models can be integrated with existing m-health frameworks to generate text alerts to the mentors when the crisis patterns are observed in their mentees. Such an integrated crisis prediction and alerting system would add benefit to peer mentors to plan intervention.

ACKNOWLEDGMENTS

Priyanka Annapureddy, M.S.

First, I would like to express my deepest gratitude to my Ph.D. advisor Dr. Praveen Madiraju for his constant mentorship and incredible guidance during my journey in this program. Dr. Praveen was also my inspiration for pursuing Ph.D. at Marquette. He always helped me through his guidance and support in a way that I could focus and give my best in research. I am grateful for all his mentorship and kindness.

Next, I would like to thank Dr. Franco Zeno for providing me constant instruction and feedback that ensured my research is enriched. His suggestions helped me to develop methodology in my work. Dr. Zeno was also generous in providing one-one meetings which helped me to work in the right direction. I am grateful to Dr. Zeno for all his time and support.

I am thankful to Dr. Sheikh Iqbal Ahamed for being on my dissertation committee and for his continuous encouragement and support in the project. I also would like to thank QRF team members Mark Flower, Dr. Niharika Jain, Dr. Sabirat Rubya for the informative discussions and providing me their guidance during the research. I am also thankful to Dr. Nasim Yahyasoltani for her teaching during the course work which helped me to advance my knowledge. I specially want to thank my teachers from high school who inspired me and introduced the love of learning. I want to acknowledge the care and support of all my friends in my life whom I can never forget.

I am grateful to my devoted family members for providing me the unconditional support in every step of my life. Without them, I could not get through the hard times and be the person I am today.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ALGORITHMS

CHAPTER 1: INTRODUCTION AND BACKGROUND

1.1 Mental Health

Issues related to mental health have been on rise according to World Health Organization (WHO) [6] due to changes in social landscape. Nearly one in five US adults every year experience mental illness in their lives [2] which include various disorders with varying severity. Factors such as stress, poverty, rapid social change, violence, environmental, family and relationship problems greatly affect mental health [2]. Psychological disorders are found to be associated with risk taking behaviors like substance abuse, aggression, suicide etc., [1]. They are also associated with medical comorbidities [3], which causes increase in patient suffering, and health care costs. An increase in the occurrence of mental illness and its burden on healthcare has mandated a growing interest in treatment and prevention.

1.2 PTSD, Symptoms and Risk Factors

In this study, we focus on Post-Traumatic Stress Disorder (PTSD) and its associated patterns of onset and persistence. PTSD is a psychological disorder that can affect individuals after exposure to significant trauma. In the aftermath of trauma, people experience symptoms of stress and anxiety. These symptoms start to appear any time after the traumatic event, it can be within days or even years after exposure. For some, the symptoms vanish without any intervention, but for others they go on to develop PTSD [110]. PTSD symptoms are characterized into four symptom clusters in the Diagnostic Statistic Manual (DSM-5). They include intrusive memories or reliving the experience; avoidance of people and trauma related situations; negative alterations in thoughts and mood that can interfere with the daily activities; hyperarousal or easily being startled [111]. These symptoms are presented in Figure 1.1.

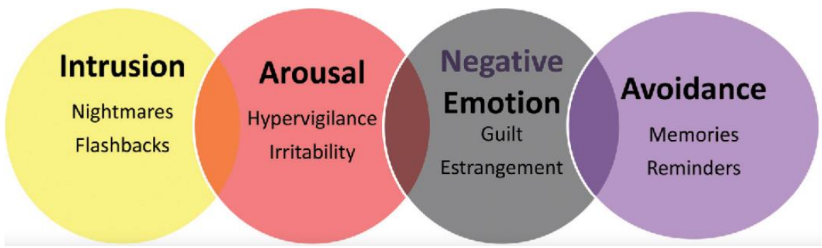

Figure 1.1. PTSD symptom cluster (source *Jorge, R. 2015.*)

Additionally, anxiety and depression are considered as the co-occurring symptoms of PTSD. Exposure to traumatic events is not uncommon in military veterans and the prevalence of PTSD in US veterans is high.

1.2.1 PTSD Vulnerability in Veterans

While in military service veterans can be exposed to a wide range of stressors and trauma events ranging from training accidents to intense combat exposure, to military sexual trauma [112]. PTSD prevalence in veterans is twice the rate of general population in US [113]. It is estimated that of the veterans who returned from OEF/OIF wars, nearly one third have been diagnosed with PTSD or other mental health conditions [114]. Veterans with these mental health conditions face additional challenges while transitioning to civilian life. Veterans returning from service often have difficulties adapting to the change and undergo multiple stressors during this process [115]. Though there are a wide range of estimates provided by different studies, a meta-analysis done by [4], estimated that average PTSD prevalence in OEF/OIF veterans is 23% [4].

PTSD affects the medical comorbidities; a study on 163 male combat related veterans exposed that cell counts of RBC, WBC and platelets are significantly higher in veterans with PTSD group and these were associated with inflammation [10]. Elevated plasma AVP levels were observed with PTSD, and these were significantly correlated with symptoms of avoidance [116]. A number of high-risk behaviors are also associated with military PTSD. For example, rates of non-suicidal self-injury and substance use is high in veterans with PTSD [5]. Suicide rate is also observed to be higher in OEF/OIF veterans compared with previous war cohorts and general population [6]. It is estimated that younger veteran suicide rates increased 26% from 2005 to 2007 [7]. Alcohol use is high in veterans with PTSD compared to non-PTSD veterans and anxiety and depression were reported as the reasons for heavy drinking [9]. Emotions like anger, hostility, and aggression are common in veterans with PTSD [8], and the level of difficulty regulating these emotions correlates with PTSD symptom severity [8] and risk-taking behaviors [117,118].

1.2.2 Risk Factors of PTSD Identified from Literature

Research indicates that personal characteristics play an important role in the development of PTSD [11,12,13]. Although PTSD occurs following a traumatic experience, not everyone who gets exposed to trauma develops PTSD. Demographic, Psycho-social, lifestyle, medical comorbidities and pre-deployment factors influence the impact of traumatic event on veteran’s mental health. Age, gender [12], unemployment [13], physical inactivity [15], lack of family and friends support [14] are some of the socio-demographic risk factors identified from literature. Lifestyle related variables like increased weight [15] is also understood to be a risk factor of PTSD. Psychological factors like self-reported anxiety, depressive symptoms [16], and emotion are also considered as risk factors. Emotion dysregulation is not only a symptom [8,17] of PTSD but is a predictor of the development of more severe PTSD symptoms [18,19]. Medical comorbidities like higher blood cell counts are indicative of PTSD risk in veterans. These immune markers in blood along with C-reactive protein at the time of pre-deployment are also the risk factors of later PTSD development [119]. Mitochondrial metabolites like lactate, citrate, eicosanoids, and glutamine along with plasma cortisol level are also predictive of later PTSD [16]. It was proved that genetic factors also play a role in developing PTSD after trauma and the risk is strongest among women. Risk factors identified by earlier works are presented in Figure 1.2. All these risk factors offer benefit for early prediction and identification of at-risk veterans. It is common that individuals have more than one risk factor, the correlation methods and feature selection methods used in the above techniques cannot detect the behavioral combinations of them. Risk factors alone also do not provide information about any intermittent mental health crisis events. Mental illness can trigger crisis episodes in individuals. Crisis events are common in any domain, whether it is a natural disaster, manufacturing, financial or mental health, they share similarities. There is a generalized definition of crisis that is applicable to mental health as well.

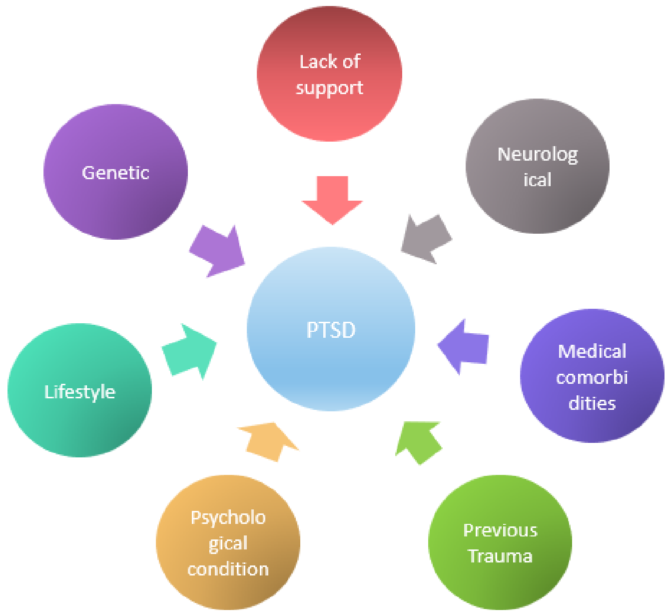

Figure 1.2. Risk Factors of PTSD

1.3 Crisis Theory, Risk Factors, Precursors and Early Warning Signs

Crisis is defined as a state when a previously existing equilibrium is disturbed because of an unexpected event or series of events that create high levels of uncertainty [20,122]. The event can be either related to a human made catastrophes, technological system failure, natural disaster, political, economic issues, social loss of a significant person, biological which is a physical illness, a stressful event which causes emotional upset. etc., [20,21,22]. Depending on the nature of events these crisis events cause property or reputational loss, physical or emotional pain [20,27]. Some characteristics of crisis period specified are "it awakens the unresolved problems" [20], bridges the gap between pre-existing conditions and current events [21]. Whether it is an organizational crisis, or natural disaster, economic crisis or individual crisis, the most universal component associated with it is the perception of threat. Duration of crisis can be acute or extreme (long-term) and calls for action [20,22] and during crisis period, tension mounts to peak and falls [21]. It was argued by [22] that crisis if not attended may develop into a disaster. In this section we explore the definitions of precursors, warning signs, risk factors and their relationship to long-term and acute crisis events.

1.3.1 Precursors

A crisis can be either acute or long-term depending on the length of period of its occurrence. Any major catastrophic event is often preceded by intermediate events called precursors [24,25]. National Academy of Engineering workshop defined precursor as any event or group of events that must occur for an accident to occur [123]. The NASA precursor analysis handbook [124] defines an accident precursor as ''an anomaly that signals the potential for more severe consequences that may occur in the future, due to causes that are discernible from its occurrence today.'' Some definitions of precursors include both conditions and events, it was proposed in [26], to exclude conditions that contribute but do not constitute an accident. The pre-existing conditions can be attributed as risk factors. Precursors are also defined as an event or situation in light of small set of changes in behaviors would have led to a consequential adverse event [23]. These precursor events are understood to be repetitive and are subjected to phase transitions. It was noted that there was increase in the occurrence of these precursors close to the critical point, this was noted in the case of large-scale earthquakes [25] and geo technical failures [24]. A review of literature marked the following characteristics for precursors [26]: 1. These are defined as off-nominal events but not the conditions; 2. Can be real or postulated; 3. Should follow an initiating off-nominal event; 4. Exhibit state transitions; 5. Increase in occurrence of precursors close to critical point. In this work, we consider precursors to be "*A group of events following an initiating event that has the potential for severe consequences in future*".

1.3.2 Early Warning Signs

There are different definitions of early warning signs in literature. These are defined as the immediate antecedents of crisis [28]. In other work [26], these are defined as signals of the events which together constitute an accident sequence. It was stated that pre crisis period is one of the stages of crisis, which consists of warning signs [20]. These warning signs are associated with near time adverse events which reveal the developing of a future event within short periods of time. However, warning signs are not always identified to represent close relation, the time frame within which they appear change with the nature of application. In a work by [103], warning signs were also defined as weak signals that strengthened over time, these were observed as early as years ago. Signals of an impending earthquake were observed in 1 year span [107]. Warning signs are perceived differently depending up on the crisis in study. In mental health, warning signs are defined as small changes in behavior that indicate acute risk and are understood to have proximal relationship and correlation with it [28]. For the purpose of this paper we define the warning sign to a crisis event as "indicators of acute crisis events ". In this work, we are interested in finding the early warning signs which indicate risk taking behaviors within a week.

1.3.3 Risk Factors

Risk factors are the pre-existing conditions (e.g., age, abuse history etc.,) that can contribute or aggravate a crisis and can be regarded as contributing factors. Unlike the warning signs risk factors suggest long-term probabilistic risk possibly a lifetime. In [28], differentiating characteristics between risk factors and warning signs were stated. According to them, risk factors are defined as population dependent and have limited implications for intervention whereas warning signs call for specific intervention. Risk factors are static whereas warning signs are episodic and variable. Risk factors are defined to have more objective quality whereas warning signs are subjective.

Pictorial representation of risk factors, precursors, acute crisis events, early warning signs and long-term crisis and how they are interrelated is shown in Figure 1.3.

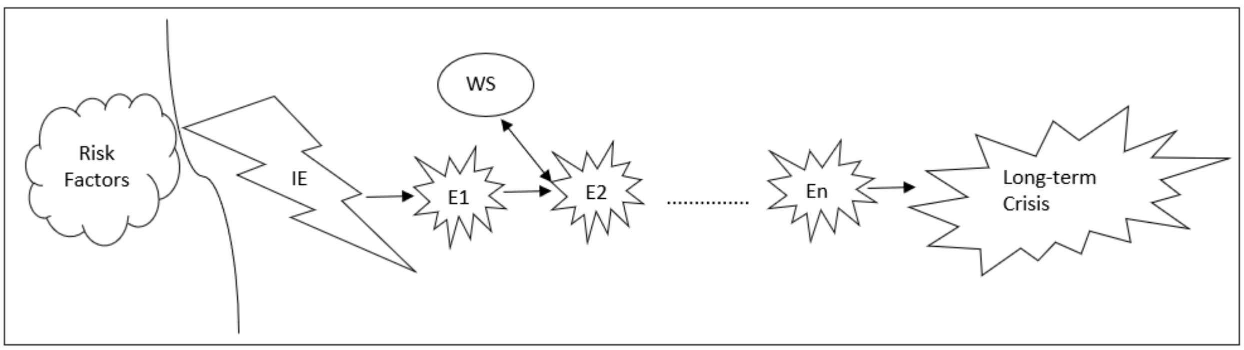

Figure 1.3. Pictorial representation of risk factors, events, warning signs. IE: Initiating event; E1..En: Crisis events; WS: Warning sign; Groups of events E1-En: Precursors

1.4 Mental health Crisis

The concept of crisis is applicable to mental health as well. A mental health crisis is defined as changes in person’s actions, feelings, and behaviors. These marked changes are considered as crisis events. Internal and external stressors in daily life when not perceived properly can trigger crisis episodes in individuals. It was argued that there could be multiple occurrences of these crisis events and their length varies from very short to longer periods [20,21]. These crisis events are also understood to be subjective [20,21]; i.e., the crisis events vary from an individual or group to another group. Inability to cope with the crises can lead to risk taking behaviors, and/or put them at risk of being unable to care for themselves. According to National Alliance on Mental Illness, being prepared in advance to a mental health crisis can help in avoiding the crisis. To prevent hazardous crisis outcome or be prepared for the crisis events, it is essential that these events need to be reliably identified. The definition of crisis theory is applied here to find the risk factors, precursors and early warning signs to long-term and acute crisis events respectively. In this research, unchanged high PTSD symptoms or onset of high PTSD symptoms are considered as long-term crisis [61]. Engaging in risk-taking behaviors are considered as acute crisis events. Therefore, the focus of this research is in identifying the precursors and early warning signs of long-term and acute crisis in veterans who were exposed to combat trauma.

1.5 Ecological Momentary Assessment

Precursors and early warning signs can be identified by monitoring for current symptoms in participants. One of the most sophisticated methods for monitoring ongoing symptoms is Ecological Momentary Assessment (EMA) where real time dynamic experiences are collected from the participants [120]. EMA are effective to track the fluctuations in individual behaviors, emotions, and subjective experiences in participant's usual environment. The advantage of EMA is that data is sampled multiple times in natural environment as against to laboratory setting, thereby providing quality data. The data needs to be sampled is decided by the researchers. There are two main categories of EMA sampling strategies: Time based, and Event based. In the time-based method, the responses are gathered at various times of the day and in the event-based methods, the responses are gathered at the occurrence of events. The choice of the method depends on the interest of the study whether to capture daily events or particular events .

The other focus in EMA studies is that how this data needs to be collected. There have been two methods of collecting this data: active and passive methods. In the passive methods, EMA are synchronized with data collected from mobile devices like geolocation, phone calls and texts. Whereas in active methods participants self-report the information in the form of structured interviews or surveys. With the recent advances in technology, technology offered solutions are being used. The use of mobile phones and portable medical based sensors in healthcare is called m-health [121]. EMA via m-health can assist the participants to report their behavioral and health patterns.

1.6 m-Health and ML for Mental Health

As m-health includes the mobile technology such as smartphones, many m-health apps are being developed to support healthcare professionals in diagnostic procedures and patients to monitor their health. The goal of m-health data is to provide an actionable information, however m-health data is raw that is not understandable by the users. Machine Learning (ML) plays a key role in discovering the unrecognized patterns from this data and predicts the clinical targets or disorders. ML has been used to examine most mental health disorders, including depression, anxiety, etc [32,57]. Supervised machine learning algorithms can model the complex relationships between variables and are effective in classification and prediction. ML together with m-health have proven effective in providing effective healthcare including mental health. ML is one of the applications of Artificial Intelligence (AI) where the system learns from the patterns in training data and makes it predictions on the new data without being programmed explicitly. Various Machine Learning (ML) algorithms evolved for building clinical decision support systems. EMA data collected through m-health app when compounded with the analytical capabilities of ML algorithms, can contribute to an effective decision support system. ML in mental health is an emerging research area with many ML enabled systems built for identifying the symptoms, mental illness and disease progression.

1.7 Literature Review

There is increased interest in studies on mental health due to change in social life-styles. Table 1.1 lists the predictive models used in mental health detection.

1.7.1 ML Works in Mental Health

A review of literature demonstrates that ML techniques are found to be robust and scalable in mental health domain. They were applied for diagnostic and prognostic modelling of mental health disorders like depression, anxiety, stress, suicide, substance use, PTSD etc.. [51,57,32]. Diagnostic modelling refers to detecting or identifying the disorder in individuals, whereas prognostic modelling means predicting the onset or progression of mental health conditions. In both types of modelling, supervised classification techniques were the most utilized ML method. Data varied from unstructured data, sensor-based data, voice, clinical data, neuro imaging etc., social media data, MRI and speech based data were also used by [29,30] for identification of depressive symptoms, sensor based data was used to detect psychiatric emergencies [31]. The commonly used algorithms in all these works are SVM, RF, DT, Naïve Bayes, NN and boosting models. Prognostic modelling of mental health involved multiple disorders, and depression was the most analyzed disorder in research. Anxiety, stress, autism and

PTSD are the next most commonly studied disorders. Table 1.1 shows the summary of ML techniques used for the prediction of mental health disorders.

Table 1.1 Related work on use of ML methods in mental health

Mental Health	Related work	Data type	Algorithm
Depression	Perlis,2013	Clinical	RF, NB, Regression, SVM
	Kessler et al. 2016	Survey	DT, Regression
	van Breda et al.,2016	Survey	RF, SVM
	Wahle et al.,2016	Sensors	RF, SVM
	Ryu et al.,2015	Survey	Gradient Boosting
PTSD	Saxe et al.,2017	Clinical	RF, SVM
	Rosellini et al.,2018	Interview	DT, RF, Regression, SVM, Super learner
	Kessler et al.,2014	WMH surveys	Regression, RF, Super learner
	Wshah et al., 2019	survey	LR, RF, SVM, NB, Ensemble
	Galatzer-Levy et al., 2014	ED	SVM, RF, Adaboost
	Zandvakili et al., 2020	MRI	LARS
	Schultebraucks et al., 2020		RF, SVM
Anxiety	Bermejo et al.,2013	MRI	DT
	Panagiotakopoulos et al.,2010	EHR	ARM
	Hoogendoorn et al.,2016		DT, Regression, RF
Anxiety, Depression and PTSD	Park et al.,2018)	Social Media	Clustering
Suicide/self-harm	Metzger et al.,2017	EHR	ARM, DT, NB, RF, SVM, Regression

1.7.1.1 PTSD in general population:

It was established that biomarkers can be used to distinguish individuals with PTSD, and it subtypes from healthy cohorts using ML methods [33]. However, this is a model for diagnostic prediction of PTSD. Recent years have shown increased interest into the early prediction of PTSD status at the end of follow up period. There are many works done by researchers for predicting PTSD. All these works involved various kinds of data like neuro imaging [48], ED room observations [48], sensor based [49], speech markers [50], survey responses etc.,. All these studies focused on early prognosis of PTSD. [45]

showed that PTSD can be predicted accurately within 10 days of trauma incident through conducting surveys using Metricwire a mobile app. Loss of interest in activities, sleep difficulty are some of the important predictive variables identified in their model. Among the various algorithms used in their work, ensemble model returned better performance. socio demographics of the participants were not included in their work. Kessler et al., 2014 have used world mental health surveys to build predictive model to identify people at high risk of PTSD after trauma exposure. RF, regression and super learner were employed in their models, their evaluation results showed that super learner an ensemble model outperformed all other algorithms. Another study by [47] worked on predicting long-term PTSD. Their model included features from demographics, ED observations, and telephonic interviews, SVM was the best performer among the models they have used. Strobl et al., 2012 demonstrated that multimodal data like neuroimaging and quality of life increased the accuracy of predictions. A recent study conducted by used MRI images to predict four subscales of PCL-5 using functional networks of the brain [46]. In another study Saxe et al., 2017 applied ML models to predict childhood PTSD in acutely traumatized children [51]. Their feature set included variables belonging to childhood development, demographics, parent symptoms, genes, and other child symptoms and functioning. Causal analysis was used to find the risk factors of PTSD in children. It was found that prior PTSD, prior loss, acute stress in parents as some of the risk factors of PTSD.

1.7.1.2 ML works in veterans: All the studies above have worked on general population, and the multi modal risk factors of PTSD were reported. Similar data and methods were used in predicting the probable combat related PTSD in veterans. Probabilistic and regression methods were used for estimating the likelihood of PTSD. Other works involved use of more sophisticated ML methods. [52] examined the pre-deployment related risk factors of PTSD in veterans. They used biological, clinical, neurocognitive variables and self-reported information of anxiety and depression for predicting PTSD within 90-180 days after their return from duty. RF and SVC were used for the classification model, better performance was observed with SVM. In a work on UK military veterans [53], supervised ML algorithms RF, SVC, Artificial Neural Networks (ANN) were applied on a sample data consisting of serving and ex-service military veterans. Their models achieved higher accuracies but the model’s sensitivity of predicting true positive labels is relatively low (0.69-0.70) giving more potential for false negatives.

ARM techniques were also used for discovering the knowledge in the mental health domain. ARM methods have been used in the data mining analysis of psychological problems in college students [54]. FP-growth algorithm was used to generate positive and negative association rules. Wang etc.al 2019, [55] have applied ARM on medical claims data for identifying the medical comorbidities of mental disorders. Their study found a high association between digestive system disease and psychiatric disorders. Apriori algorithm was used for mining the rules based on support – confidence framework. In another study, ARM was used to find the profiles of traumas and life stressors that can predict the presence of anxiety and depression [56] in Srilankan war survivors. In their work, self-report questionnaires of anxiety and depression were used to identify the risk factors. ARM was used by Panagiotakopoulos et al., to find the relationship between context data (location, time, age, gender, physical conditions, symptoms etc.,) and stress level of the patient [57].

1.8 ML Strategies for PTSD Research

Looking at the preliminary works in PTSD space, it is understood that most of the ML models were oriented towards predicting an individual as at risk of PTSD or not. But they fail to predict the severity of PTSD, which is found to be more associated with risk taking behaviors.

In most of the studies discussed in the above section, RF, DT, regression, SVC and ensemble algorithms are the most used methods. In all the works that involved ensemble algorithm, ensemble was clearly the winner, SVM also outperformed other algorithms. Given the potential of SVC to handle the non-linear data, they proved to perform better in both classification and regression tasks. They work well with high dimensional data. Random forests are an ensemble of multiple decision trees built on random subsets of features. Decision tree is made up of multiple decision points forking into decision paths, with each of them terminating in a class label. Decision tree is interpretable, and the decision paths contribute as rules. But these rules of decision trees are very few and are not reliable and can lead to overfitting. In RF, which is an ensemble, random subset of features is used for splitting at each node. By emphasizing on random features, complex feature patterns are captured, and chance of overfitting can be reduced. Since RF constitutes of deep decision trees, understanding why a decision was made from each individual tree is not feasible. However, they explain the individual feature importance, but they cannot be used as an explanatory model. Some studies in mental health show that association rules were used to mine the patterns and relationships between the variables. Identified patterns provide new knowledge about the association between symptoms and presence or absence of a disorder. But when there are multiple rules (especially hundreds and thousands) suggesting the presence and absence of a disorder, arriving at a cohesive conclusion of risk in a patient might be difficult. Combined effect of these mined rules must be evaluated to predict the risk in an individual.

In all the methods discussed, models are evaluated or considered to be predictable based on accuracy, sensitivity and AUC metrics. But these do not explain why a prediction was made. Failure to explain an outcome can result in lack of trust in the predictions. Removing the black-box nature of a model and increasing the transparency ensures that generated predictions are unbiased and reliable. For this, our research takes a different approach of building an interpretable ML model for the prediction of PTSD outcome. To date, this was not implemented in the predictive modelling of mental health disorders as per our knowledge.

This research work proposes to address these gaps by

1. Building a multiclass classification model to stratify PTSD risk in veterans into three levels of low risk, medium risk, and high risk. Further, the work is extended to
2. Mining precursors to long-term crisis in veterans using ARM and build associative classifier to predict the PTSD severity.

3. Identifying early warning signs to acute crisis events that can indicate the risk-taking behaviors in veterans.

CHAPTER 2: DATA AND ANALYSIS

2.1 Community based Peer Mentoring by Dryhootch

It was indicated that veterans after leaving their regular service are at increased risk of mental health disorders including PTSD [58]. The goal of this work is in developing crisis alert system based on behavioral markers collected from m-health app for veterans. Though Department of Defense and Department of Veteran Affairs provide mental health services to veterans, they cannot reach out to all places where veterans reside. Therefore, many outreach programs like “texting campaign”, social media campaigns and community-based guidance programs have emerged to support veterans who recently separated from military service [59]. Community based organizations serving veterans provide more comfortable and welcoming environment to the veterans. Dryhootch of America (DH) is one of the community-based organizations which provides peer mentor support to the OEF/OIF returning veterans trying to reintegrate into the society [60]. They offer a 12-week program where the veteran’s health and progression of symptoms are observed through weekly surveys. The veteran support is provided by qualified veteran mentors who went through similar phase and have overcome the challenges. It is understood that veterans feel comfortable to share their personal experiences with the co-veteran mentors than with anyone outside their community [59].

2.2 m-Health Enabled Community Outreach

The peer-mentor program at DH is a 12-week program, during which veterans are assessed for risk taking behavior, symptom changes and social functioning by mentors. Ecological Momentary Assessment (EMA) techniques are used to monitor the veterans for 12-weeks. The advantage of EMA is that repeated measures can be taken in individuals' natural settings. The weekly EMA surveys at DH capture the crisis events in veterans for the mentors to analyze the risk in their mentees. Initially, paper-based surveys were conducted by mentors, and it was later understood to be ineffective as it became difficult for mentors to assess each mentee condition effectively [59]. As most of the veterans in the program were younger veterans and were comfortable with the use of smartphones and internet technology, a technology mediated solution was then built. The system was called iPeer and has two modules veteran and peer-mentor mobiles apps [59]. The veteran version of the app is for the mentees to provide their feedback through surveys about their wellbeing, and the mentor can view this information in the mentor version of the app.

2.3 QRF App

Quick Reaction Force (QRF) application consists of QRF mentee and mentor versions. Figure 2.1 shows the mentee and mentor QRF app. Mentee version of the app is used by mentee veterans to take weekly EMA survey and these survey responses can be analyzed by the mentor in the mentor version of the app. The surveys are available in the form of a check-in to mentees, on clicking the available check-in button, mentee gets questions for the survey.

QRF Mentor version of the app is used by the mentors. When a mentor logs in, he or she will get a screen with a list of all their mentees. For each mentee, a graphical view representing the survey responses over a broader time period is shown. [61] focused on providing visualization of the mentee self-reports to the mentors. The purpose of this dashboard is to aid mentors in detecting early warning signs of crisis in their mentees and determine the need to reach out their mentees in acute crisis. This research builds upon their work to develop a personalized crisis alert system based on the predictive modelling of PTSD. This work focuses on application of Machine Learning (ML) in enhancing the personalized analysis for peer mentors and to integrate it with existing QRF system to generate text alerts to the mentors for their intervention.

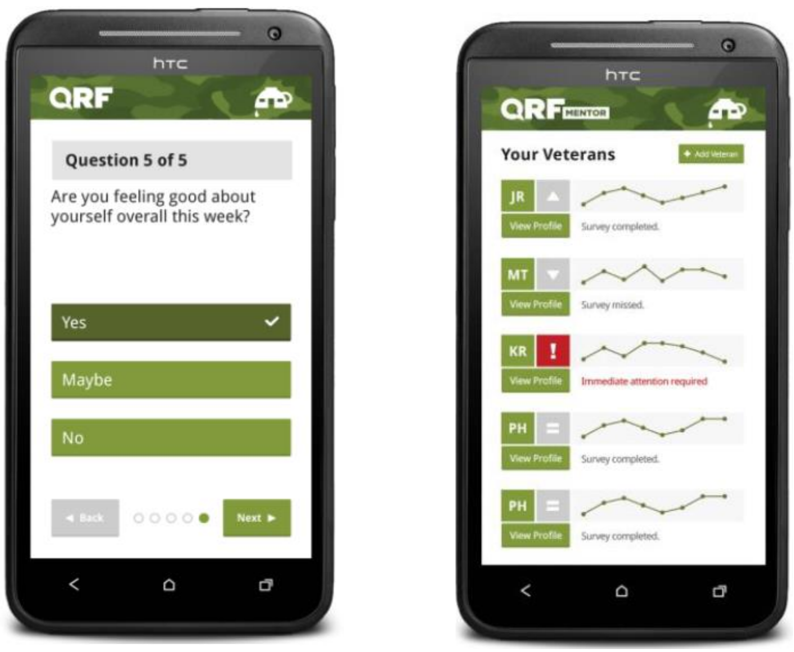

Figure 2.1. a) QRF mentee version of the app (left side) and b) mentor version of the app (right side)

2.4 Data Description

Data for the study comes from QRF, a smartphone application initially developed by [59] in partnership with Dryhootch of America. Initial set of veterans who enrolled for 12-week program at DH consists of 305 participants, their socio demographic characteristics and other measures including PTSD diagnostic score were collected at the beginning of the program (baseline). QRF study used a repeated measures approach, where the baseline survey repeated at six and twelve weeks which are called midpoint and discharge surveys. Along with these, the weekly EMA surveys capture the crisis events in participants during the 12-week program. The details of these surveys are provided in the next subsections.

2.5 Data Variables

2.5.1 Baseline Variables

Many researchers have identified the risk factors of PTSD in their works. The risk factors are understood to increase the risk of PTSD onset when exposed to trauma. They include demographic variables [12], alcohol and substance use [5,6], lack of family and friends support [14], physical inactivity [62], and unemployment [13]. To measure these factors in the participants, various questionnaires involving AUDIT, SAS, DRRI, smoking were conducted at baseline, midpoint and discharge times of the 12-week program. Along with these, participants PTSD diagnostic score (PCL-5) is also measured at those time points. Demographic characteristics and the branch of military service of the participants were also included in the study. All these variables collected at the time of baseline are called as the baseline variables and are considered in the current work. The following subsections describes the baseline variables included in the study.

2.5.1.1 Demographic variables: These include demographic information like age, gender, school enrollment, and the branch of military (Army, AirForce, Navy etc.,) veteran served. Some of the questions used are listed in Table 2.1.

Table 2.1 Demographic questionnaire used in survey

1.	Gender	0	Female
		1	Male
		2	Transgender
		98	Other
		97	Refused
2.	Are you currently enrolled in school?	1	Yes
		0	No
3.	In which branch(es)/component(s) of the military did you serve? (Check all that apply)	1	Army
		2	Navy
		3	Air Force
		4	Marine
		5	Coast
		6	National
		7	Active
		8	Reserve
		9	Other

2.5.1.2 AUDIT: It is a 10-item questionnaire developed by World Health Organization (WHO) to assess alcohol consumption, drinking behaviors and alcohol-related problems. A score of 8 or greater is considered as problematic alcohol use. From this questionnaire, three questions related to heavy alcohol use are used in the baseline survey, these are shown in Table 2.2.

Table 2.2 AUDIT questionnaire used in survey

1.	How often do you have a drink containing alcohol?	0	Never
		1	Monthly or less
		2	Two to four times a month
		3	Two or three times per week
		4	Four or more times per week
2.	When you are drinking, how many drinks do you typically have?	0	1 or 2
		1	3 or 4

		2	5 or 6
		3	7 to 9
		4	10 or more
3.	How often do you have six or more drinks in one day?	0	Never
		1	Less than monthly
		2	Monthly
		3	Two or three times per week
		4	Four or more times per week

2.5.1.3 Smoking: This smoking questionnaire is conducted to measure the smoking patterns in participants. It contains questions to understand the smoking habits in participants and their dependency on smoking. Following question in Table 2.3 is used as one of the baseline variables.

Table 2.3 Smoking questionnaire used in survey

1.	How many cigarettes a day do you smoke?	1	10 or less
		2	11-20
		3	21-30
		4	31 or more

2.5.1.4 SAS: It is composed of questions to assess the participants interest in hobbies, daily activities, and job. It is also measuring social activeness and engaging in community activities by the participants. Table 2.4 lists the SAS questions that are used to form the baseline variables.

Table 2.4 SAS questionnaire used in survey

1.	Are you employed?	1 10 or less 2 11-20 3 21-30 4 31 or more
2.	How interested are you in your job?	1 Very 2 Moderately 3 A little 4 Not at all
3.	Do you pursue your job with?	1 A lot of 2 Some enjoyment 3 Little enjoyment 4 None
4.	Do you pursue these home related activities with?	1 A lot of 2 Some enjoyment 3 Little enjoyment 4 None
5.	Are you interested in hobbies/leisure?	1 Very 2 Moderately 3 A little 4 Not at all
6.	To what extent are you involved in the community (such as clubs, church, etc) ?	1 Fully 2 Moderately 3 Slightly 4 Not at all
7.	How important do you consider your physical appearance?	1 Very 2 Moderately 3 Not very 4 Not at all

2.5.1.5 DRRI-2: Deployment Risk and Resilience Inventory -2 (DRRI-2) is a successor of DRRI, it measures deployment related risk and resilience factors in veterans deployed to overseas military missions. It measures post deployment family functioning, post deployment psychosocial experiences like family stressors and others. Here are some of the drri questions included in the study:

Table 2.5 DRRI questionnaire used in survey

1.	I went through a divorce or have been left by a partner or significant other.	1 Yes 2 No
2.	I had problems getting access to adequate healthcare.	1 Yes 2 No
3.	I have experienced stressful legal problems (for example, being sued, suing someone else, or being in a custody battle).	1 Yes 2 No
4.	I experienced a natural disaster (for example, a hurricane), a fire, or an accident in which I or someone close to me was hurt or had serious property damage.	1 Yes 2 No
5.	I have witnessed someone being seriously assaulted or killed.	1 Yes 2 No
6.	I have seriously physically injured by another person (for example, hit or beaten up).	1 Yes 2 No
7.	My family members and/or friends make me feel better when I am down.	1 Strongly Disagree 2 Somewhat Disagree 3 Neither Agree nor Disagree 4 Somewhat Agree 5 Strongly Agree

2.5.2 PCL-5

All the participants are assessed for their PTSD diagnostic status before entering into the program. There are a wide number of ways for evaluating the diagnostic status of PTSD. PTSD is widely assessed using PTSD Checklist (PCL) which consists of self-report questionnaire outlined by Diagnostic Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV). This was later revised as PCL-5 in accordance with updated DSM-V criteria, which contains a 20-item questionnaire with each item measuring the symptoms of PTSD on a scale of 0-4. The total score ranges from 0-80 and cut-off score is used for provisional diagnosis of PTSD. Studies on PTSD have a mixed use of PCL and PCL-5 and wide range of cut-off scores were proposed for different population. The most used cut-off score is 33 in general population, if the PCL-5 is greater than 33 in an individual, he/she is considered to have PTSD. For veterans, a wide range of cutoff score between 38-50 based on PCL were reported [63]. It was further demonstrated in other studies [5,64] that veterans with PCL score greater than 50 are more likely to engage in risk taking behaviors like alcohol and substance abuse etc., These behaviors have a recurring effect on PTSD and can lead to self-harm. The cut-off score proposed by these works is based on PCL, whereas in the current study PCL-5 checklist is used. To find a corresponding score in PCL-5 for high-risk behaviors, our work used the findings from [65] in which a correspondence between PCL and PCL-5 are reported. According to their findings, PCL score of 50 corresponds to a PCL-5 score of 39. Therefore, a PCL-5 score of 39 is used as the criteria to identify veterans at high-risk in this study.

2.5.3 Weekly EMA data

Risk factors alone do not contribute to the decision-making process in mental health. The current signs and symptoms also play a major role in the judgement process [66]. The weekly EMA surveys were aimed at capturing the current symptoms of the participants. In most other works, EMA questionnaire was either designed by clinicians or trained professionals and they were aimed at general population. The patients perceptive of symptoms is completely ignored in their works. The method adopted by QRF team is different, veterans were involved in the design process, EMA questionnaire was framed after a deep discussion with veterans [60]. For example, the veterans requested a question on health as even minor illness could aggravate psychological symptoms. The final set of survey questionnaire shown in table 2.6 consisted of 6 questions related to how veterans are doing with stress, sleep, risk taking behaviors, health, self-worthiness, and whether peer mentor contacted them during that week. The questions regarding the symptoms have three choices whether they are feeling better, same, or worse compared to last week. The perception of crisis and coping abilities are understood to vary from individual to individual. This subjective nature of crisis is also captured in the EMA survey with the responses of “Better” and “Worse”.

Table 2.6 Weekly EMA questionnaire used in survey

1.	Have you engaged in any risky behavior (as you define it) this week?	1	Less than last week
		2	Same as last week
		3	More than last week
2.	How well did you sleep this week?	1	Better than last week
		2	Same as last week
		3	Worse than last week
3.	Has your health changed this week?	1	Better
		2	Same
		3	Worse
4.	How stressful has this week been?	1	Less than last week
		2	Same as last week
		3	More than last week
5.	Are you feeling good about yourself overall this week?	1	Yes
		2	Maybe
		3	No
6.	Did your Dryhootch peer mentor talk to you this week?	1	Yes
		2	No

2.6 Preprocessing the Input and Output Variables

2.6.1 Input Variables

2.6.1.1 Baseline variable one-hot encoding: All the baseline variables and weekly EMA survey data are considered as input variables for the study. Baseline variables are all categorical whereas EMA variables are ordinal. It is observed that most of these variables have multiple categories. As one-hot encoding of these variables could result in large number of categories, initially these variables categories are grouped to few categories with the help of domain experts. For example, for the audit question, "how often do you have drinks containing alcohol? " one and two responses are considered medium alcoholic use, three and four are considered high alcoholic use. Similar grouping of the categories is applied to other baseline variables.

2.6.1.2 EMA symptoms aggregation: With the case of EMA variables, each survey symptom is evaluated for twelve weeks. And each survey symptom response is encoded to numerical values 1, 2 and 3 for better, same and worse. The total score is evaluated for every week. Along with total score, 12-week responses are aggregated to see whether there were "two symptoms worse in a week", or "three symptoms worse in a week", or "two symptoms worse for two consecutive weeks", or "three symptoms worse for two consecutive weeks". Symptoms are also aggregated independently to see their individual effect on crisis events. Grouping of these symptoms like "sleep worse for two weeks", "sleep worse for three weeks" and "sleep worse for two consecutive weeks" are evaluated and similar summaries are performed on other EMA symptoms stress, health, risk and self-worthiness. Number of contacts made by the peer mentor during the program is also aggregated and used as a predictor variable.

2.6.2 Output Variables

2.6.2.1 Long-term crisis (high PTSD): PCl-5 score is considered as one of the output variables, a cut-off score on this value will be used to label the participants as at high PTSD or not. The level of the PTSD symptoms is seen to be related to high-risk behaviors; a study conducted by [5] demonstrated that a higher level of PTSD symptoms was associated with increased risk-taking behaviors. The results of their work based upon 394 veterans showed that total risk frequency to be high for individuals with PCL-M score greater than 50. This association was also asserted by another study [64] which reported that personnel with PCL scores higher than 50 were likely to have increased alcohol use, aggression, and impulsive behaviors. As higher levels of PTSD symptoms can lead to high-risk behaviors in veterans, this work tries to know early in the program if the PCL-5 score at the time of discharge is likely to be in the severe range. There is no study suggesting the levels in PCL-5 severity score but a correspondence between PCL-5 and PCL scores is shown in [65]. Their work on veteran population concluded that PCL-5 scores of 25, 31 and 39 correspond with PCL-S cut scores of 39, 44, and 50. From these observations, a “high” risk level cut off of 39 with PCL-5 as opposed to 50 with PCL will be used to determine high PTSD in participants.

2.6.2.2 Acute crisis (risky behaviors): Participants engaging in risk taking behaviors is considered as acute crisis events. EMA responses by the participants provide this information whether a participant has engaged in risky behaviors. If a participant has at least one response worse for risk taking behavior, he or she is labelled as having an acute crisis event.

2.6.3 Missing Values

Data in the current study contains missing values, not all surveys are taken by all participants; hence an additional feature is generated to track the missed surveys. Some participants miss the discharge surveys too, in such cases their discharge PCL-5 score is not available. Such samples are filtered from processing. Filtering the samples with unavailable discharge PCL-5 left 83 of 305 samples.

CHAPTER 3: RISK STRATIFICATION MODEL FOR LONG-TERM CRISIS

3.1 Introduction

This work aims to monitor for early warning signs of PTSD symptom severity in veterans undergoing a community-based peer veteran support program. PCL-5 is used as a measure for identifying PTSD symptom severity. Based on PCL-5 score and its reported association with risky behaviors (details provided in section 3 of the paper), three categories low, medium, and high PTSD are used in the current work. Differentiating veterans with high PTSD symptoms from those with mild and low symptoms helps us to know who are at high risk of engaging in unsafe behaviors and who are just above the diagnostic cut score so that a peer mentor intervention can be provided accordingly. The machine learning studies done in PTSD [45,67] proposed methods to identify individuals at risk, and not at risk of PTSD, to our knowledge this study is the first of its kind to include the high-risk category to differentiate individuals with high PTSD symptoms. Supervised Machine Learning (ML) models are developed to predict the defined PTSD risk categories of participants using Ecological Momentary Assessment (EMA) based self-reported data. This chapter outlines the computational approach used to categorize a participant into one of these three PCL-5 categories low, medium, and high using supervised machine learning techniques. The objective of this research aim is to answer the following research questions:

1. How effective are the crisis events captured by EMA techniques in discriminating the participants with high discharge time PTSD from medium and low risk levels?

2. How early in the program can a participant's discharge time PTSD status be identified?

Multinomial classifiers which can directly learn all the classes were employed. The major contributions of this research aim are (1) the use of three categories to differentiate participants into different risk levels low, medium, and high instead of just risk and not at risk. (2) Identification of the earliest time point in the 12-week rehabilitation program when the predictions can generate a warning or alert to aid early intervention. (3) This work used an ensemble of under-sampling and oversampling methods to handle class imbalance in data. (4) Weighted soft voting classifier is used to combine the predictions from best performing primary models.

3.2 Preprocessing

A PCL-5 cut-score proposed by many validation studies is often in the range 28 to 38 to determine a provisional diagnosis of PTSD in clinical settings, pending verification by a clinician, or to ascribe status for research [68,69]. A PCL-5 cut off 39 which is associated with risk taking behaviors as discussed in section 2 is used as "high" risk level. Lower cut-off of 28 in the proposed diagnostic cut score range is considered to increase the detection of probable PTSD, therefore, the range 28 to 39 was considered "medium" risk level. Any score below 28 was given the "low" risk level.

The distribution of discharge PCL-5 score shown in Fig. 3.1 indicates class imbalance with 40 for “low”, 14 for “medium” and 26 for “high” classes. Performance of the classifiers is dependent on the distribution of classes. Because many standard classifiers assume balanced class distribution in data, their learning and recognition are

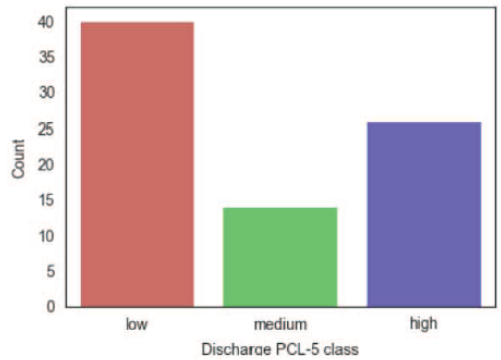

Figure 3.1 Figure showing distribution of class variable

more biased towards the majority class increasing the misclassifications for minority class [39]. Many methods have been proposed to overcome this issue, sampling is one of the class-imbalance learning methods that works by under-sampling or over-sampling the training sets. In the current work both under-sampling and over-sampling methods were applied to the QRF data and an ensemble of them is used for final classification. Among the group of over-sampling methods, Synthetic Minority Oversampling Technique (SMOTE) was used here to balance the distribution of classes. SMOTE avoids overfitting by generating new cases of minority class in the same sample space [32]. The other method used here for class imbalance problem is under-sampling which reduces the majority class instances. The limitation with under-sampling is it ignores most of the majority class information. This is overcome in Easy Ensemble (EE) method, details of it are provided in the following subsection.

3.3 Method

As stated earlier, data preprocessing was done to extract the features from multilevel data. Aggregate functions were used for this purpose. However, all the features identified may not be good indicators of crisis. Feature selection is done to select only the features that are associated with the output variable. Feature selection is a technology for feature dimensional reduction. Thus, by reducing the dimension of the data, it can reduce the complexity of models and thereby avoid overfitting.

3.3.1 Correlation Analysis

A correlation analysis was performed to measure the association between survey scores, discharge PCL-5 and other participant characteristics. Their correlation coefficients are shown in Table 3.1. The weekly survey scores are seen to be positively correlated with the target discharge PCL-5 score with r-value ranging between 0.1 to 0.31. School enrollment is seen to be positively associated with the number of surveys missed by the participant. This implies that veterans who are enrolled in school tend to miss the surveys. Also, school enrollment is negatively correlated with discharge PCL-5 with a r-value -0.20 indicating that veterans enrolled in school are likely to have lower discharge PCL-5 scores.

Table 3.1 Correlation analysis among variables

	disch_pcl5	gender	school_enrollment	Military Component: Army	Navy	Air Force	Marine Corps	National Guard	Active Duty	Reserve
score.week1	0.23	-0.05	-0.47	-0.16	0.09	0.27	-0.14	0.12	-0.23	-0.18
score.week2	0.1	0.12	-0.43	-0.08	0.01	0.21	-0.1	0.15	-0.21	-0.13
score.week3	0.19	0.17	-0.43	-0.04	-0.04	0.23	-0.11	0.12	-0.15	-0.11
score.week4	0.29	0.08	-0.45	-0.08	0.08	0.23	-0.21	0.05	-0.36	-0.3
score.week5	0.31	-0.06	-0.43	-0.17	0.11	0.2	-0.15	0.12	-0.25	-0.19
score.week6	0.23	-0.01	-0.38	-0.16	0.09	0.24	-0.13	0.04	-0.14	-0.15
score.week7	0.18	0.08	-0.49	0.04	0.02	0.08	-0.19	0.14	-0.23	-0.25
score.week8	0.26	0.09	-0.39	0.01	-0.07	0.21	-0.13	0.05	-0.22	-0.19
score.week9	0.26	0.06	-0.36	0.01	-0.06	0.13	-0.12	0.06	-0.25	-0.24
score.week10	0.26	0.12	-0.43	0	-0.01	0.13	-0.2	0.04	-0.25	-0.28
score.week11	0.26	0.08	-0.51	-0.07	0.01	0.24	-0.2	-0.07	-0.26	-0.31
Missed_surveys.week1	-0.2	0.03	0.47	0.13	-0.11	-0.21	0.14	-0.18	0.28	0.19
Missed_surveys.week7	-0.19	-0.1	0.55	0.06	-0.07	-0.19	0.21	-0.19	0.31	0.23
Missed_surveys.week12	-0.2	-0.11	0.55	0.02	-0.05	-0.19	0.25	-0.18	0.3	0.25
disch_pcl5	1	-0.15	-0.2	0.01	0.08	0.04	-0.16	0.1	-0.15	-0.26

3.3.2 SMOTE

The distribution of discharge PCL-5 score shown in Figure 3.1 indicates class imbalance with 40 for “low”, 14 for “medium” and 26 for “high” classes. Performance of the classifiers is dependent on the distribution of classes. Because many standard classifiers assume balanced class distribution in data, their learning and recognition are more biased towards the majority class increasing the misclassifications for minority class [71]. Many methods have been proposed to overcome this issue, sampling is one of the class-imbalance learning methods that works by under-sampling or over-sampling the training sets. In the current work both under-sampling and over-sampling methods were applied to the QRF data and an ensemble of them is used for final classification. Among the group of over-sampling methods, Synthetic Minority Oversampling Technique (SMOTE) was used here to balance the distribution of classes. SMOTE avoids overfitting by generating new cases of minority class in the same sample space [70].

3.3.3 ML Algorithms

3.3.3.1 Logistic Regression: Logistic Regression (LR) is one of the standard methods for analyzing binary outcomes. It assumes that feature sets have a linear relationship with the outcome on log odds scale. Each predictor has a weighted co-efficient which describes the strength and direction of relationship to the outcome. The interpretability of these regression co-efficients has led to its wide acceptance in health care where interpretation is of interest. LR is the first model applied on the data sets in this chapter.

3.3.3.2 Ensemble Models: Ensemble learning methods have emerged recently and are most adequate solution for building powerful classification models. Ensemble learning algorithms combine multiple classifiers either weak or strong called base learners to achieve improve accuracy and robustness over single classifier. In an ensemble method, the output of multiple methods is combined and collectively evaluated to make final predictions. Ensemble models like bagging, boosting and voting classifiers are implemented in this chapter.

XGB

XGB is an implementation of Gradient Boosted Decision Trees (GBDT) built on the idea of boosting to improve speed and performance. A series of trees are involved, and a weight is associated with each monitoring in the dataset. New trees are built on the performance of previously created trees. XGB is an optimal approach to boosting.

Easy Ensemble

Easy Ensemble (EE) is used for imbalanced classification. EE is an under-sampling method with ensemble framework [72]. Through random under sampling, several subsets of majority class instances are created. A learner is trained on each of these subsets.

Voting Classifier

Soft voting classifier considers the class probabilities of each base learner and final predictions are made through averaging process unlike hard voting.

3.3.4 Bayesian optimization

The performance and complexity of these algorithms depend on many tunable configuration parameters. These parameters are often hardcoded, or default values are chosen. Parameter tuning plays an important role in ML. There are many methods to optimize these hyper parameters. One method Bayesian Optimization is a principled technique based on Bayes theorem which attempts to find the global optimum of an objective function in a minimum number of steps. The hyper parameters of LR and XGB are selected using Bayesian Optimization technique.

3.4. Classification results

In this section we compare the predicted class labels of 9 weekly models with the actual discharge PCL-5 categories of the participants. The primary purpose of this work is to assess the usage of weekly survey scores in identifying the PTSD severity at discharge. Comparison of results of the base classifiers and the voting classifier from Figure 3.2 and Table 3.2, showed that voting classifier outperforms other classifiers in terms of sensitivity for high-risk class, macro-averaged f-score, and false positive rate at most of the time points (weeks). The results of the voting classifier are based on weighted probabilities of the base classifiers.

Table 3.2 Classification results of ML models

Week	Classifier	Accuracy	Sensitivity of high risk class	Sensitivity of medium risk class	Sensitivity of low risk class	F-score	False positive rate of low risk class
3	EasyEnsemble	50	0.6	0.29	0.57	0.48	0.33
3	LR	62.5	0.6	0.71	0.57	0.63	0.14
3	XGB	66.67	0.8	0.57	0.57	0.66	0.2
3	**Voting**	**70.83**	**0.8**	**0.57**	**0.71**	**0.7**	**0.2**
4	EasyEnsemble	58.33	0.8	0.14	0.71	0.52	0.33
4	LR	66.67	0.6	0.71	0.71	0.67	0.25
4	XGB	70.83	0.8	0.43	0.86	0.69	0.27
4	**Voting**	**70.83**	**0.8**	**0.43**	**0.86**	**0.69**	**0.2**
5	EasyEnsemble	70.83	0.9	0.29	0.86	0.66	0.31
5	LR	70.83	0.7	0.71	0.71	0.71	0.22
5	XGB	66.67	0.7	0.43	0.86	0.66	0.3
5	**Voting**	**70.83**	**0.8**	**0.43**	**0.86**	**0.69**	**0.2**
6	EasyEnsemble	66.67	0.8	0.29	0.86	0.63	0.27
6	LR	58.33	0.5	0.57	0.71	0.59	0.29
6	XGB	62.5	0.9	0.14	0.71	0.56	0.25
6	**Voting**	**66.67**	**0.8**	**0.43**	**0.71**	**0.64**	**0.2**
7	EasyEnsemble	54.17	0.8	0	0.71	0.43	0.43
7	LR	62.5	0.5	0.71	0.71	0.63	0.17
7	XGB	54.17	0.7	0.14	0.71	0.49	0.36
7	**Voting**	**70.83**	**0.9**	**0.43**	**0.71**	**0.68**	**0.25**
8	EasyEnsemble	58.33	0.8	0.14	0.71	0.51	0.33
8	LR	62.5	0.4	0.57	1	0.61	0.2
8	XGB	62.5	0.9	0.14	0.71	0.55	0.31
8	**Voting**	**66.67**	**0.9**	**0.29**	**0.71**	**0.61**	**0.25**
9	EasyEnsemble	62.5	0.7	0.14	1	0.55	0.3
9	LR	50	0.4	0.43	0.71	0.5	0.33
9	XGB	66.67	1	0.14	0.71	0.58	0.23
9	**Voting**	**70.83**	**0.8**	**0.43**	**0.86**	**0.68**	**0.2**
10	EasyEnsemble	62.5	0.7	0.14	1	0.55	0.3
10	LR	50	0.4	0.43	0.71	0.5	0.33
10	XGB	66.67	1	0.14	0.71	0.57	0.29
10	**Voting**	**70.83**	**0.8**	**0.43**	**0.86**	**0.68**	**0.2**
11	EasyEnsemble	58.33	0.7	0.14	0.86	0.52	0.3
11	LR	58.33	0.5	0.43	0.86	0.57	0.29
11	XGB	66.67	1	0.14	0.71	0.58	0.23
11	**Voting**	**70.83**	**0.8**	**0.43**	**0.86**	**0.68**	**0.2**

The average performance of the models for all the weeks is shown in Figure 3.3f. The boosting classifiers, EE and XGB performed well for high and low risk categories, with an average sensitivity of 0.83 (EE) and 0.73 (XGB) for “high” PTSD and 0.74 (EE) and 0.88 (XGB) for “low” categories. But these classifiers show poor performance with medium risk class. The average sensitivity for medium risk class is 0.17. LR is a good performer of medium risk class but not with high and low classes compared to other classifiers. It has an average recall of 0.58 for class medium and around 0.74 and 0.51 for classes high and low, respectively. FPR of low-risk class appeared to be also less for LR compared to other base classifiers with an average of 0.25. FPR of the class low is the proportion of medium and high-risk samples falsely labelled as class low. The goal of this work is not only to improve the predictions for the severity class, but it is equally important to minimize the number of participants from severity classes “medium” and “high” from being falsely labelled under the “low” risk class as this would reduce the crisis services to veterans actually in need of immediate intervention. When the FPR for class low is compared among all the classifiers, FPR is the least with 0.2 during weeks 3 to 6 for the voting classifier. The other performance metrics of voting classifier are sensitivity for severity class “high” is between 0.71 to 0.86 and for moderate risk class is between 0.43 to 0.57 for the weeks 3 to 6, the macro-averaged f-score is between 0.65 to 0.7 for the same weeks. When all the performance measures are compared, voting classifier alone performs consistently better for all the weeks with any of those metrics, therefore voting classifier is a better choice here.

To identify the earliest time point (week) that has better predictions, the voting classifier results are compared across all the weeks in Table 3.2, it is observed that weeks 4 and 5 during the first half of the 12-week program and weeks 9 to 11 during the second-half have higher sensitivity for high risk class and least FPR for low risk class. The macro averaged f-score of the voting classifier at weeks 4 and 9 are 0.69 and 0.68; while the recall of high, medium and low risk classes is 0.8, 0.43 and 0.86 respectively for both the weeks. The FPR of the “low” class is consistent with 0.2 at these time points.

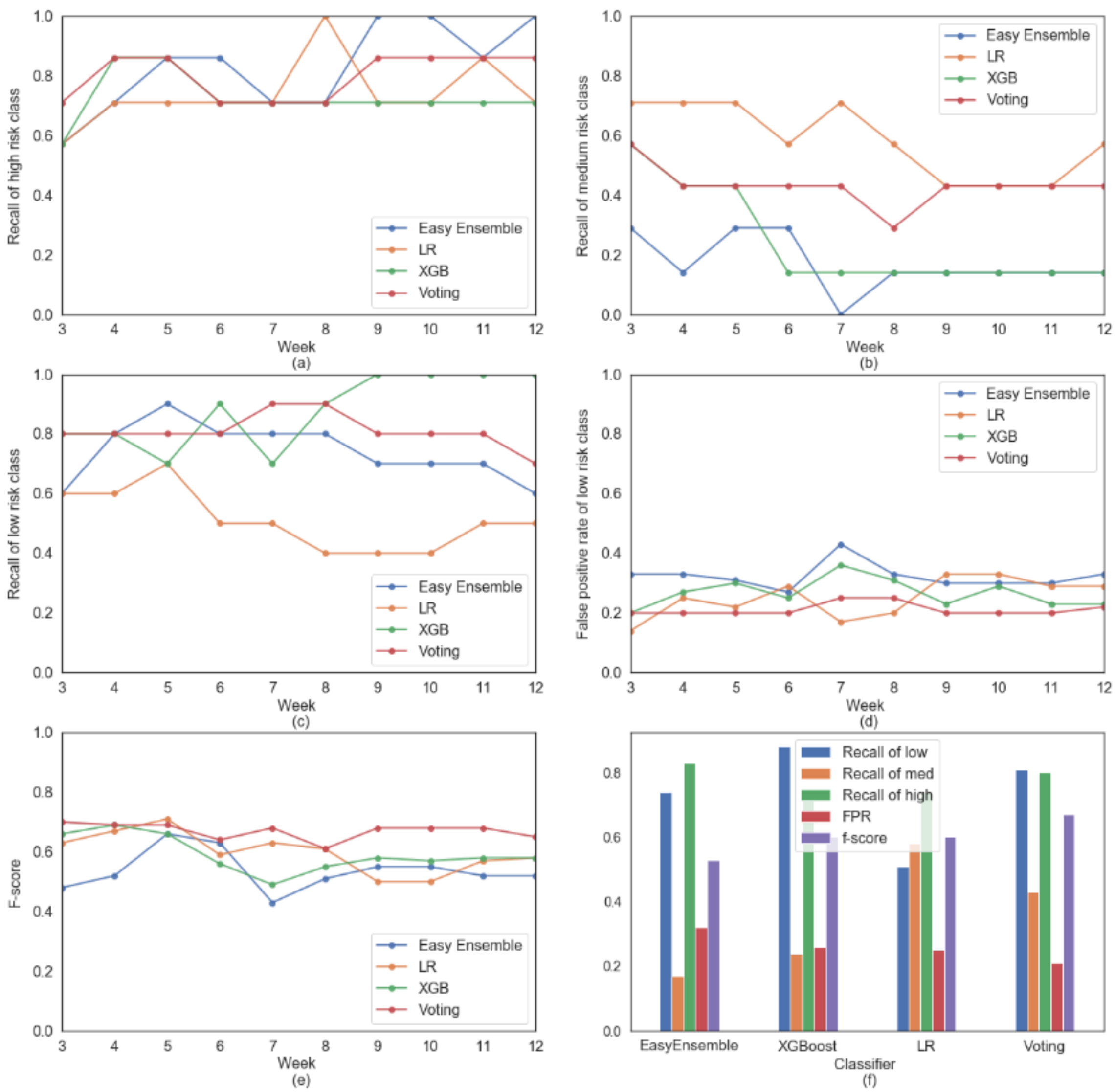

Figure 3.2 a) Sensitivity of high-risk class; b) sensitivity of medium risk class; c) sensitivity of low risk class; d) false positive rate of low risk class; e) Macro average f-score of all classes; f) bar chart showing average recall of the classifiers.

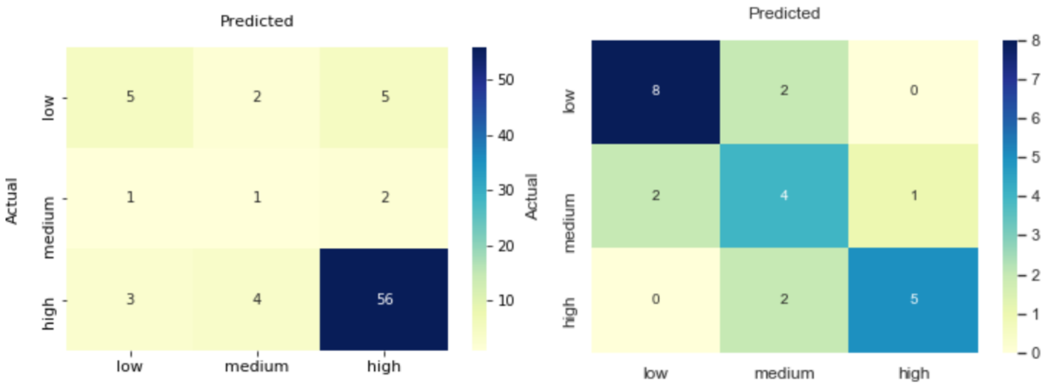

Figure 3.3 Confusion matrix of soft voting classifier (a) considering participants with missing values as high risk (b) filtering the participants with missing values

The confusion matrix of the voting classifier at week 4 is shown in Figure 3.3b, out of seven test samples that actually belong to high-risk class, two of them are falsely labelled under medium risk class and none under low-risk class. Though the false negative rate of high-risk class is 0.75 here, the actual risk of not predicting the class correctly is still less because the false negatives are labelled under medium risk class. Similarly, of three false negatives of medium risk class, two are labelled as low risk and one under the high-risk class.

The overall macro averaged f-score for the voting classifier at week 4 is 0.69. All these results are obtained after filtering the missed discharge PCL-5 scores. ML models were also tested removing the filter considering the participants with missing values under high risk, the highest f-score achieved from this approach was only 0.58. Confusion matrix of these results is shown in Figure 3.3a. Filtering the missed discharge surveys increased the overall f-score by 19 percent. Missed discharge surveys could be due to administrative oversight and the missingness may not be attributed to PTSD risk.

This work evaluated the use of self-report questionnaires in monitoring PTSD symptoms and ability to predict PTSD severity in veterans. The results of the ML algorithms applied to QRF data demonstrated that PTSD severity levels can be predicted early during the 12-week peer support program to a significant degree by using the weekly self-report data, baseline symptoms, and sociodemographic information. The current work reinforces the ability of ML methods in predicting the crisis.

CHAPTER 4: INTERPRETABLE PREDICTIVE MODEL FOR LONG-TERM CRISIS

In this chapter, we focus on second research aim i.e. to build a prediction model for long-term crisis that is interpretable and can explain the factors or precursors that influenced the prediction. Often psychological disorders are preceded by warning signs indicative of upcoming crises [73]. The ability to identify precursors helps in planning interventions to help mitigate long-term crises.

Participants enrolled in the qrf program have their PTSD symptoms evaluated at baseline and discharge. Among the participants enrolled in the 12-week program, some of the participants tend to have same level of high PTSD symptoms from baseline to discharge, some tend to recover from baseline high PTSD symptoms by the time of discharge, and some eventually develop high PTSD symptoms. There are various factors affecting or being associated with these changes in PTSD levels. Knowledge of these factors or patterns helps peer mentors know about the problematic patterns in their mentees and can plan personalized intervention. Therefore, the primary goal of this research is to build a interpretable machine learning model that can identify the precursors to high PTSD symptoms at discharge in veterans; and compare these patterns of persistent PTSD symptoms, recovery and onset of PTSD in the participants. Though there have been many works which focused on developing PTSD symptoms, to our knowledge this is the first work to focus on finding and comparing the patterns of changes in PTSD symptoms. This objective is aimed at answering the following research questions:

1. What are the risk factors and precursors of persistent PTSD, and onset?

2. Do the risks and pathways of persistent PTSD and PTSD onset overlap?
3. What are the protective factors of recovering from high PTSD?

There have been multiple studies to determine pre-characteristics that increase the likelihood of risk of mental illness in individuals [14,15,48,74]. Statistical and Machine Learning (ML) methods were adopted which include correlation methods such as chi-square, Pearson product moment correlation, and other embedded ML algorithms, to name a few [75,66,76]. These methods can identify only the independent risk variables, but the effects of the variable combinations and interactions among them were not explored. Further, the articulation of risk factors alone does not capture the current state of mental health for the individual. It is crucial that risk factors along with current symptoms of possible mental health issues within an individual need to be considered for more comprehensive clinical decision-making efforts [77]. Ecological Momentary Assessment (EMA) techniques are more sophisticated in capturing daily changes in symptoms. Therefore, current study uses EMA symptoms captured during the 12-week program along with the baseline characteristics as inputs for the interpretable ML model for the identification of precursor events to high PTSD at discharge in veterans.

4.1 Interpretable Model and Advantages

ML models have been in use in the data-driven decision making process across various applications including healthcare. As the adoption of ML models for real-life decision making increased, there is increased concern about being able to understand and trust the predictions became more important. This is because ML models are increasingly used for sensitive applications where mistakes can have catastrophic consequences. Therefore, practitioners are looking for models that are interpretable and explainable.

Interpretability and explainability are commonly used terms to explain model predictions. Interpretability is defined as “the ability to explain or to provide the meaning in understandable terms to a human” [79]. On the other hand, explainability is “associated with the notion of explanation as an interface between humans and a decision maker that is, at the same time, both an accurate proxy of the decision maker and comprehensible to humans” [79]. Explainability is an active characteristic for the model, whereas interpretability is a passive and inherent component of the method. ML models are generally categorized into white box and black box depending on their transparency. White box models are transparent, and it is easy to understand the logic that drives the decision. Algorithms like Decision Trees, Logistic Regression and Linear Regression are some examples of white box models. Explainability can be achieved from these models and the explanations these provide do not change across the dataset. Black box models are complex, and the predictions are less interpretable. SVM, neural networks and boosted trees are some of the black box models. Although, algorithms like DT and LR were preferred by some practitioners due to their interpretable nature, these algorithms lose their interpretability under different conditions. Decision tree is a simple and effective rule extraction method where a rule is a decision path that is traceable from node to the leaf. As the depth of the tree and number of nodes increases, model loses its interpretable nature. DT is also known to overfit. In a linear regression model, a linear model is fit to the data and the weights of the variables can be used to interpret predictions. But this does not perform well when there are non-linear relationships with the target variable and when the input variables are highly correlated. There have been many works which offer alternatives for converting non-interpretable model into interpretable one. One of such methods is the local model-agnostic interpretability methods which offer explanations to the black box model predictions. This was done by fitting simpler models to the local neighborhood of the instance to be explained with the assumption that behavior of the instance to be explained is similar to the behavior of its neighborhood. One of the first works using this approach was LIME, in which linear model was fitted in the selected neighborhood of the instance to be explained. Logic-based approaches which hold global explanations have been in place too. These were extended by [80] to the case of boosted trees. SHAP is an example of both local and global explanation. Explanations of these methods are based on feature importance and may not work with unstructured data, disadvantage of these methods is that they are time consuming as they need to run multiple evaluations of the model to provide explanations.

Interpretability and explanation are important in high-risk environment where false predictions have a significant impact especially in medical diagnosis when they can cost the life of patients. The debugging of the false predictions is easy with explanatory models in a high-risk environment. There are different explain methods, the explanation this work considers is relating the feature values of an instance to the predictions in a human understandable way. Understanding the interactions between the features that lead to crisis not only provides information about risky patterns but can also aid mentors to understand the risk factors in their mentees while providing intervention.

In this paper, we propose a flexible framework to build an interpretable model using ARM that can explain the predictions. It was argued in [81] that class association rules are better suited to provide explanations than linear models and decision trees. In their work, k-optimal class association rules were mined in the neighborhood of the instance to be explained.

Although studies discussed in the section 1.2 have identified the risk factors of PTSD, complex interaction among the risk variables and with current symptoms and how they contribute to the progression of PTSD is not yet studied. To mine the combinatorial pattern of risk factors and current symptoms and to build an explainable model from them, this research proposes to use the Association Rule Mining (ARM) based classifier in this work.

4.2 Association Rule Mining

Association Rule Mining (ARM) which is one of the important branches of Data Mining (DM) methods was initially developed for market basket analysis but now is widely used in other domains including medicine [82, 83, 84] and mental health [85]. ARM methods are proven to be an effective way of discovering the patterns for the outcome of interest. ARM is a rule-based machine learning approach to find interesting patterns or associations between input and output variables. These are represented in the form of rules and these associations are based on the co-occurrence of the variables.

An ARM rule is of the form X => Y, meaning if X exists, Y also coexists. The left-hand side(LHS) of the rule X is called the antecedent and right-hand side(RHS) of the rule Y is the consequent. The LHS and RHS of the rule is a boolean condition on feature values. The strength of the association between antecedent and consequent is represented by measures like Support, Confidence, Lift and Coverage. The definitions of these are described below:

4.2.1 Measures

Support of the rule is the proportion of instances that match the rule from the total amount of data in the database (D), it indicates how frequently the rule occurs.

$$Support(X => Y) = \frac{n(X \cap Y)}{n(D)}$$

Confidence is the proportion of the instances that match the rule over the number of instances which contain only the antecedent. It is also the conditional probability of X and Y occurring together given X

$$Confidence(X => Y) = \frac{n(X \cap Y)}{n(X)}$$

Lift is a measure of how interesting a rule is and the value represents the association to be evaluated. It represents a non-trivial correlation between antecedent and consequent. A lift less than one means negative association and a lift greater than one means positive association between the antecedent and consequent. A lift equals to one implies no associations can be found.

$$Lift(X => Y) = \frac{Support(X => Y)}{Support(X)Support(Y)}$$

Coverage is the support of the antecedent of the rule.

$$Coverage(X => Y) = Support(X)$$

Based on these measures, rules can be either predictive or interesting. High confidence rules can be predictive and rules with high lift are interesting which best explain a dataset [81]. Measures are also used to impose rule constraints so as to generate optimal sets of rules. The most commonly used rule constraints are min support and min confidence thresholds. Every rule that satisfies these pre-specified threshold values will be considered optimal.

4.3 Class Association Rules

Association rules provide patterns among all the variables. To better guide peer mentors about their mentees discharge PTSD outcome (high PTSD or low PTSD), this research tries to find out association rules with consequent limited to discharge PTSD only. Sociodemographic characteristics, other baseline characteristics gathered and progressive EMA symptoms are all included as variables for the antecedent. Here we try to investigate the prevalence of these behavioral characteristics in relation to long-term crisis in veterans. Such subset of association rules with the rule consequent limited to class variables only are called Class Association Rules (CARs). The set of rules with positive class label are grouped as positive class rules and those with negative class label are called negative class rules. In medical terms, a rule shows the association between risk factors and presence or absence of a disorder.

4.4 Associative Classifier

CARs generated using ARM will be used in building a classification model. The use of ARM for classification is called associative classifier. Flowchart of the associative classifier used in this work is shown in Figure 4.1. Recent studies have shown that associative classification achieved good accuracies and the performance was better than traditional ML models [86,87]. Various classification methods were used in different studies. The first associative classifier CBA used the rule with highest confidence for classification. In their algorithm, all the class rules will be matched for the data instances, and only the rule having highest confidence will be used to predict the class label for the instance. The use of multiple rules and their combined effect will be more effective to make a prediction in comparison with a single high confidence rule. In the later work, Classification based on multiple association rules CMAR was proposed [86], all the matched rules were grouped based on the class labels and the predicted label for a data instance is based on the class of the group with highest weighted chi-square. In [89], probability score which is the average of weighted confidence is calculated for each instance and the final prediction is based on this probability. ROC and AUC curves were used to determine the cutoff values. This method may not work with class imbalanced data since weighted average is based on the confidence measure. In WCBA, harmonic mean of support and confidence is used as the strength of the rule [88]. The class with highest average HM of the matched rules is selected. Minimum class complement score [90], and overall coverage by [87] were other used measures for classification.

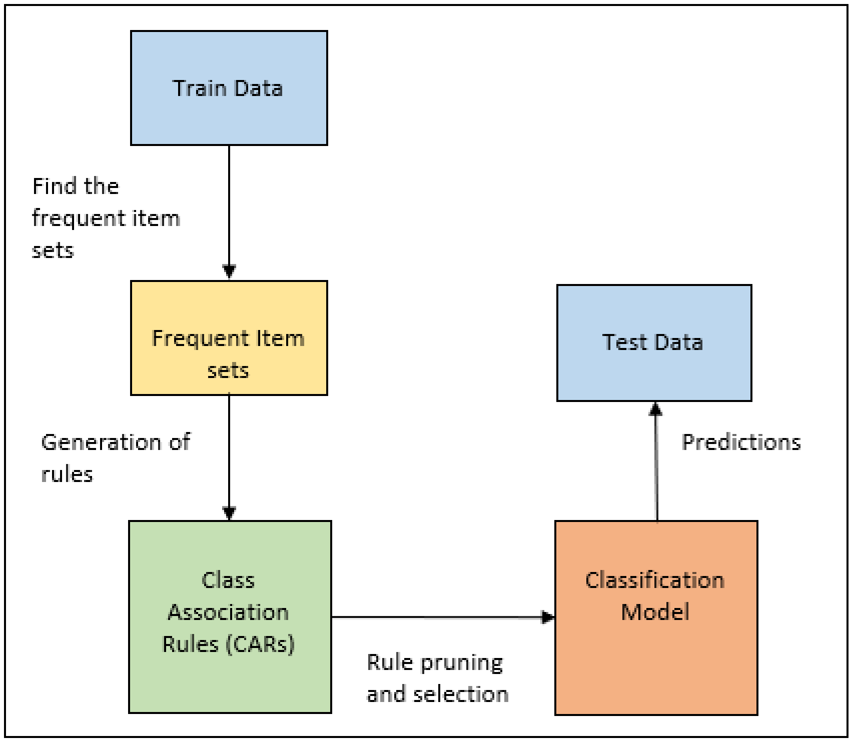

Fig 4.1. Associative classifier flow chart

4.5 Preprocessing

4.5.1 Input Variables

4.5.1.1 Baseline variables

In order to find precursors to long-term crisis in veterans, the current study incorporated variables like gender, marital status, school enrollment, employment status, family and friends support. Lifestyle variables like current smoking, heavy alcohol use, interest in hobbies, community engagement, social support, and interest in physical appearance are also included. These characteristics measured at the time of baseline in the 12-week QRF peer mentor support program are considered as input variables in this study. The variables of ARM are typically binary, they indicate the presence or absence of a risk factor. Therefore, all the baseline variables feature engineered in section 2 are one-hot encoded to convert into binary in this study.

4.5.1.2 EMA Variables

Along with these, the changes in current symptoms/behaviors captured using weekly EMA surveys are also included as predictor variables. The weekly EMA questionnaire, which has multiple measurements, is aggregated. Warning signs are summarized as whether the participant had “at least one sleep symptom worse”, “two sleep symptoms worse for two consecutive weeks”, “at least two sleep symptoms worse” and “at least three sleep symptoms worse”. Similar aggregated measures were generated for other EMA variables related to stress, health, risky behaviors, and self-worthiness. Number of peer mentor contacts is dichotomized to capture if it is “less than or equal to 2” or “3 to 4” or “greater than 4”.

4.5.2 Class Variables

The class label of interest in this study is based on the presence or absence of PTSD severity at the discharge time. Target labels are generated based on the change in high PTSD symptoms from baseline to discharge. If high PTSD symptoms prevail in participants from baseline to discharge, they are given persistent category, and those who have lessened their high PTSD symptoms from baseline to discharge are labeled as recovery. Participants who have low PTSD symptoms at the time of baseline, if they have developed new PTSD symptoms during the program they are categorized as onset, and if they continued to have low PTSD symptoms they are called low risk. The prevalence of high PTSD symptoms is evaluated using PCL-5 score.

4.5.3 Missing Values

Participants tend to miss weekly surveys, failure to take the survey is also considered as a warning sign by veterans [61]. Missed surveys are also tracked by including the variables that represent if the participant "missed one survey", or "missed two surveys" or "missed three surveys". All the variables related to baseline characteristics, summarized EMA responses, missingness of survey are considered as predictors of the study. If a veteran missed the discharge survey, the participant is filtered from the study. For missingness of information in baseline variables, the missingness is also added as a category to each of the baseline variable.

4.5.4 Dependency Analysis

Correlation analysis is done to evaluate the association between baseline feature variables and EMA variables which are associated with persistent, onset and recovery of high PTSD symptoms which are shown in Table 4.1.

Table 4.1 Correlation analysis between variables

	Persistent high PTSD		**Recovery from high PTSD**		**Onset of high PTSD**	
Variable	**corr**	**p-value**	**corr**	**p-value**	**corr**	**p-value**
Consumes 5 or more alcoholic drinks when drinking	0.22	0.000119	-0.1	0.08814	0.22	0.000119
divorced	0.17	0.002329	-0.03	0.642052	0.17	0.002329
Peer mentor contacted for 3 to 4 weeks	0.17	0.002859	-0.04	0.490698	0.17	0.002859
Peer mentor contacted for more than 4 weeks	0.11	0.056999	0.22	0.000079	0.11	0.056999
Any three symptoms worse in a week	0.42	0	0.1	0.085337	0.42	0
Sleep worse for at least two weeks	0.41	0	0.17	0.002205	0.41	0
Stress worse for at least two weeks	0.37	0	0.22	0.000105	0.37	0
Health worse for at least one	0.36	0	0.1	0.092662	0.36	0

week						
Sleep worse for at least three weeks	0.36	0	0.12	0.043409	0.36	0
Stress worse for at least three weeks	0.36	0	0.18	0.00125	0.36	0
Stress worse for two consecutive weeks	0.34	0	0.15	0.007078	0.34	0
Feeling self worthy for two consecutive weeks	0.14	0.011943	0.31	0	0.14	0.011943
Stress better for at least two weeks	0.14	0.01248	0.24	0.000023	0.14	0.01248
Feeling self worthy for at least two weeks	0.13	0.023344	0.29	0	0.13	0.023344
Stress better for at least one week	0.13	0.021051	0.22	0.000089	0.13	0.021051
Stress better for two consecutive weeks	0.12	0.04311	0.2	0.000403	0.12	0.04311

4.6 Method

To alert peer mentors about the likelihood of high discharge PTSD symptoms in their mentees, this research aims at discovering the precursors to persistent and onset of high PTSD symptoms. These precursors provide knowledge of socio-demographic characteristics, progressive current symptoms captured from EMA surveys that are associated with severity. Here we try to investigate the prevalence of these behavioral characteristics in relation to long-term crisis in veterans. Along with these protective factors that are associated with recovery from high PTSD symptoms will also be identified. To discover association rules relevant to the discharge outcome variable, the rule generation process focused on rules with high PTSD and low PTSD only.

To find the precursors to persistent discharge PTSD severity, onset and recovery, this study considered only the participants who took the discharge survey. After discussion with domain experts and peer mentors, the participants are divided into two groups: baseline high risk group and baseline low risk group based on high or low PTSD symptoms at the beginning of the program (baseline). CARs are generated from discharge data from participants in these two groups to find the patterns of persistent high PTSD and recovery (low PTSD) in the high baseline PTSD group and Onset and no-risk in the low baseline PTSD group. Pictorial representation of the categorization of these groups is shown in Figure 4.2.

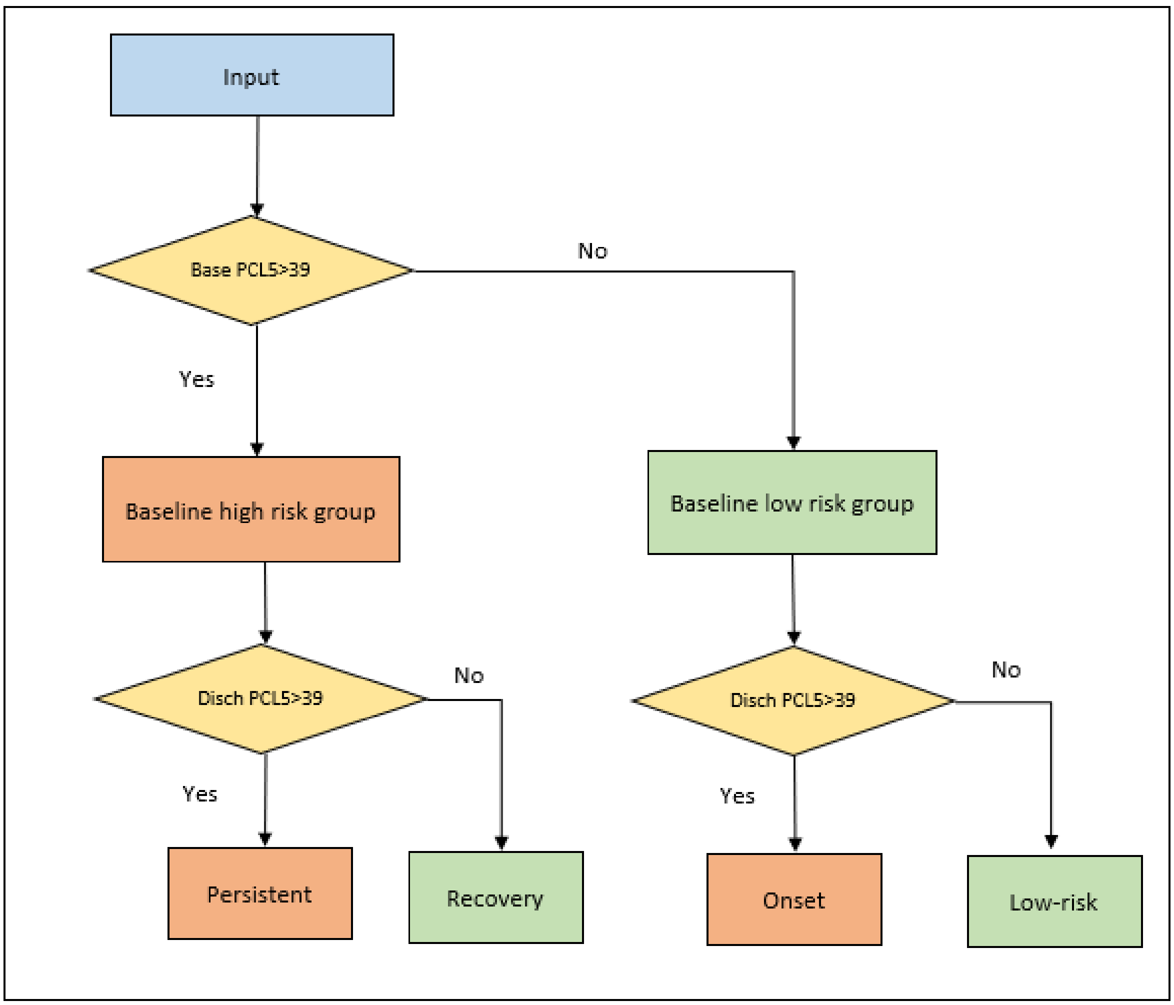

Figure 4.2. Flow diagram showing the grouping of participants.

4.6.1 Apriori Algorithm

ARM is based on finding the frequent rules that define relationship between unrelated frequent items in the data. Apriori algorithm is used in this work to generate the association rules of long-term crisis. Apriori algorithm involves two steps: Initially, all the frequent item sets satisfying minimum support threshold are found. Next all the association rules are mined from these frequent item sets and are verified for minimum confidence. The rules satisfying minimum confidence threshold are passed and the rest are ignored.

4.6.2 Pitfalls of Support Confidence Framework

Traditional ARM uses common minimum support and minimum confidence thresholds for rule mining. Use of minimum support and minimum confidence filters out noisy rules. This approach was followed by many algorithms like CMAR [86], WCBA [88] and others [87]. However, this doesn't work well when the distribution of classes is imbalanced. Setting min support too high can result in loss of relevant minority class rules, whereas a low value of min support can pull irrelevant majority class rules which overfit the data. This problem of common minimum support with imbalanced data was addressed by Liu et al. (2003). In their work, the use of different support thresholds for the rules of different classes was proposed, these thresholds are based on the proportion of class labels. This approach was followed in PCBA algorithm [89] which is a modified form of CBA to overcome the class imbalance.

In this work, we propose a new method of filtering the noisy rules, here instead of support we use Class Support (CS) to meet the minimum threshold. CS is defined as the support of the rule within the class instead of measuring against the complete data [90]. It measures the strength of the rule within the class of interest.

$$Class\ Support(X => Y) = \frac{n(X \cap Y)}{n(Y)}$$

For the confidence threshold we use support of the class as the threshold for the rules of respective classes [89].

4.6.3 Avoiding False Discoveries:

The process of rule mining outputs all the rules that satisfy user defined thresholds for interesting measures, however all these rules are not interesting. All the mined rules may not represent significant correlation and the association between the constituent items might have occurred by chance. Such randomly occurring rules not only increase the size of the rule set, but they can also lead to false discoveries. To avoid those false discoveries and to select an optimal set, different rule pruning methods are endorsed by research works. In CMAR [86], χ2 testing is done to retain only the rules where X and Y are positively correlated. Chi-square can perform erratically with class imbalanced datasets [91]. To deal with class imbalanced data, various measures have been proposed by researchers. One of them is the complement class support (CCS) [90] that captures the strength of the rule in the complement class. Smaller the value of CCS, the stronger the rule is. In a similar work to handle class imbalance [92], Fisher exact test was used for initial running pruning and subsequent rules are tested for Class Correlation Ratio (CCR). CCR was defined as “the measure of positive correlation of the antecedent with the class it predicts relative to the alternative class”.

In the current study, we used Odds Ratio (OR) which is found to be suitable for datasets with unequal class distribution [93]. OR is the ratio of odds of the event (disorder) happening in the presence of an exposure (risk factors) to the odds of the event happening in the absence of the exposure.

$$OR = \frac{\frac{n(exposed\ cases)}{n(unexposed\ cases)}}{\frac{n(exposed\ non-cases)}{n(unexposed\ non-cases)}}$$

If OR > 1, exposure is associated with the odds of the outcome. However, OR does not demonstrate the statistical significance of the association, the use of Confidence Interval (CI) and p-value help to determine the significance [94]. In the current process of generating the CARs, we use a commonly used significance level of 0.05 for p-value to keep the probability of error below 5 percent. The use of p-value for testing the statistical significance of the rules was also reinstated in these works [95,96]. All the positive and negative class rules which are statistically significant with p-value below 0.05 will be retained and the remaining rules will be discarded. This step eliminates any noisy rules. We modify the apriori algorithm to generate the rule as soon as a frequent item set is created, and the statistical significance of the rule will also be tested at the same time. This avoids the separate rule generation step and noisy rules can be filtered when they are created. The steps followed in the rule generation and pruning process is outlined in Algorithm 4.1.

Our proposed algorithm will generate statistically significant class association rules from frequent item sets that satisfy user defined class support and p-value thresholds.

4.6.4 Redundant Rules:

Even after rule pruning methods, ARM generates a large number of rules, typically thousands of rules making it difficult for a human interpreter. And these rules increase proportionally with increase in frequent item sets. Most of these rules may be redundant and they can be removed by using interesting measures. Three kinds of redundancies in rules were observed in this study.

First type of redundancy is the most widely defined form by other researchers [97]. A rule X => Y is considered to be redundant in this work, if there exists another rule X1 => Y such that X1 subset of X (and X1 != X) and support and confidence of rules X1 => Y' and X => Y are equal. For example of the rules shown in Table 4.2, a rule with antecedent "no family support, one health symptom worse" is subset of "no family support, one health symptom worse, little interest in hobbies" and if both the rules have same confidence and support rule 2 is considered redundant. Therefore, Non-redundant association rules are the generalized rules with minimal antecedents.

Second kind of redundancy considered in this study is "A rule X => Y is considered to be redundant in this work, if there exists a superset X2 for X such that confidence (X2 => Y) > confidence (X => Y)". For example of the rules shown in Table 4.3, a rule "Divorced, having 1 or 2 drinks in a day" is subset of "Divorced, having 1 or 2 drinks in a day, No family support"; and rule 1 has lower confidence than rule 2. In this case rule 1 is discarded and rule 2 with highest confidence is considered.

Third kind of redundancy observed is in the weekly EMA attributes. "A rule is considered redundant if there exists another rule with the same confidence but with the same EMA symptom observed for a smaller number of weeks". For example of the rules shown in Table 4.4, if there are two rules "One health symptom worse" and "two health symptoms worse" and both have the same confidence, then the first rule is considered, and the rest are filtered.

Table 4.2. Table showing example redundant rules of first type

Rule no	Antecedent	Consequent	Confidence(%)	Support(%)
1	Baseline PCL5 > 39, No family support , Health symptom worse for at least one week	High risk	100	6.02
2	Baseline PCL5 > 39, No family support , Health symptom worse for at least one week, Little interest in hobbies	High risk	100	6.02
3	Baseline PCL5 > 39, No family support , Health symptom worse for at least one week, Sleep symptom worse for at least one week	High risk	100	6.02

Table 4.3. Table showing redundant rules of second type

Rule no	Antecedent	Consequent	Confidence(%)	Support(%)
1	Baseline PCL5 > 39, Divorced, Drinks_1or2	High risk	70.5	14.45
2	Baseline PCL5 > 39, Divorced, Drinks_1or2, No family support	High risk	80	9.63
3	Baseline PCL5 > 39, Divorced, Drinks_1or2, Female	High risk	100	4.81

Table 4.4. Table showing redundant rules of third type

Rule no	Antecedent	Consequent	Confidence (%)	Support(%)
1	Baseline PCL5 > 39, Divorced yes, Drinks_1or2, No family support, Health symptom worse for at least one week	High risk	100	13.25
2	Baseline PCL5 > 39, Divorced yes, Drinks_1or2, No family support, Health symptom worse for at least two weeks	High risk	100	8.43
3	Baseline PCL5 > 39, Divorced yes, Drinks_1or2, No family support, Health symptom worse for at least three weeks	High risk	100	7.22

4.6.5 Ranking of Rules:

Rules for each of the persistent, recovery, onset and low-risk groups are ordered in the decreasing order of confidence, support and then by the size of the antecedent. Ranking of the rules is important in the process of rule selection for building a classifier. These rules will be used for predicting the class label in participants.

```
Algorithm1: Generating class association rules
Input: n,minsupp
Output: Positive class rules; Negative class rules
{
 POS=empty set; NEG=empty set
 k=1
 Do {
     Sk = Generate all candidate k-itemset which include class variables
     For each item i in Sk
       Generate class association rules r like X=>Y and X=>Y' (Y,Y' are positive and negative class)
       For each r
         minconf = support_count(Y)/n
         if (class_support(r) >= minsupp) and (confidence(r) >= minconf)
            if the dependency between X and Y is statistically significant
              if(Y)
                POS=POS+r
              End if
              if(Y')
                NEG=NEG+r
              End if
            End if
         End if
       End for
       k=k+1
     }
     while (Sk+1 != empty)
     Return POS, NEG
}
```

Algorithm 4.1. Rule generation and pruning process

4.6.6 Majority voting

In our work, we used hard majority voting of the rules for class prediction. This classification method is borrowed from ML powerful ensemble classifiers, which combine different classification models into meta classifiers [98]. Ensemble classifier assigns a class label based on the majority vote. Our classification process follows a similar approach to predict the class label. In this method, for each participant, all the set of positive and negative class rules are scanned and matched. The matching rules for each participant are used for voting. The final predicted output class is the class with the highest majority of votes by the rules. The algorithm for labelling the participants based on positive and negative class rules is outlined in Algorithm 4.2. Based on this method, two classification models will be built, one using only statistically significant rules, and the other using all the generated rules. These two models will be evaluated and compared for their false positive rate. The classification models were built for the baseline high risk group using the positive and negative class rules. The crisis patterns and predictions generated from this aim can be used by mentors for planning intervention.

```
Algorithm2: Prediction
Input: D, POS, NEG
Output:
{
 For each i in D
   For each r in POS
      if r matches i
         p_rules=p_rules+r
      End if
   End For
   For each r in NEG
      if r matches i
         n_rules=n_rules+r
      End if
   End For
   if count(p_rules>=n_rules)
      Assign Positive class
   else
      Assign negative class
   End if
 End For
 Compare the assigned labels with actual labels to find the true postive rate and flase negative rate for the classes
}
```

Algorithm 4.2. Classification process from CARs

4.7 Results

In this section, rules identified among the persistent, onset and recovery groups involves the combination of baseline and EMA characteristics. EMA characteristics identified for persistent and onset groups are considered the precursors of long-term crisis and baseline characteristics of recovery group are called the protective factors.

4.7.1 Rules Identifying Persistent PTSD

Patterns of persistent PTSD severity (high risk) and recovery (low risk) are mined in the baseline high risk group. There are in total 33 participants in this group, 23 of them continued to have PTSD severity (high risk) at the time of discharge and 10 had low risk. Table 4.5 shows rules with persistent and recovery high PTSD along with OR, CI and p-values. Overview of the rules indicate that worse health, sleep and stress are the most occurring patterns in persistent risk veterans.

Baseline personal characteristics like having no family or friends support, unemployment, being divorced, and little interest in hobbies are the other markers of persistent high PTSD. From EMA data, poor health for at least one week along with no family support has nearly 61% class support with high confidence and OR 10. Worse health symptoms are also accompanied by no enjoyment in home related activities. Stress being worse for two consecutive weeks along with unemployment or little interest in hobbies has also marked higher confidence and support for persistent high PTSD. The identified patterns will further improve our understanding of the pathways and precursors of such behavior. These findings are consistent with the literature which shows low family and friends support are associated with higher rates of PTSD [14]. No family and friend support as expected increased the likelihood of persistent high PTSD. Stress can be caused by an individual's perception of an event; they can be internal or external events [100]. Individual's protective factors and coping ability play an important role. Having no family and friends support together with worse sleep symptoms are more positively correlated with persistent PTSD severity in the positive class rules. This is in line with the findings of [101], where lack of social and emotional support was found to be associated with disturbed sleep.

4.7.2 Rules Identifying Recovery from High PTSD

Similarly, the positive aspects of these variables are found as the protective factors of recovery. Most of the baseline characteristics observed in the recovery group are they show moderate interest in hobbies, have family and friends support, they enjoy home related activities and they engage in community activities moderately. Frequency of high alcoholic drinks is also low in this group. These baseline characteristics when accompanied with better sleep symptoms and self-worthiness in participants showed 100% confidence of recovery. Self esteem was understood to be related to various facets of mental health (Rosenberg 1981). Rules in the table show that nearly 57% of participants who recovered have moderate interest in hobbies and sleep (or) self-worthiness better for at least one week. 71% of participants who recovered engage in community activities and they show enjoyment towards home related activities. 43% of the participants have moderate interest in hobbies and better EMA health symptoms. The odds ratio for participants having these baseline and EMA characteristics is between 18-22 (CI: 2-280).

4.7.3 Rules Identifying Onset of High PTSD

The rules identified by these participants discover the patterns in the participants who developed PTSD symptoms during the program. These rules show the baseline characteristics and EMA variables that are indicative of developing high PTSD symptoms. These rules are shown in the Table 4.5. Prior works on veterans identified that participants with specific risk factors are at greater risk of developing PTSD. Medical comorbidities is one of the identified risk factors, findings of our study also supports this. Having prior health problems is one of the baseline characteristics in the onset rules. Nearly 66% of the participants in this group have health problems and are heavy alcoholic users. The OR in this group is 60 (CI: 3- 100). Number of cigarettes smoked is also high in this group, unemployed, divorced, and do not engage in community related activities. The EMA symptoms of interest in these rules are engaging in risky behaviors, worse stress and sleep symptoms. Being unemployed and engaging in risky behaviors for at least two weeks have high likelihood to develop high PTSD symptoms. Another interesting baseline characteristic observed in these rules is, these participants also showed interest in physical appearance. Having interest in physical appearance and had problems getting access to health care, there is 75% likelihood that they develop high PTSD symptoms. It can be observed from the table that frequency of use of alcoholic drinks 6 at a time is high in these participants with 100% confidence and odds ratio of 60 (CI:). EMA patterns observed and that are of interest in this group are: engaging in risky behaviors at least two times during the program and had two symptoms worse for at least two consecutive weeks. Being an alcoholic user and having health problems or engaging in risky behaviors for at least two weeks or having two symptoms worse for two consecutive weeks are some of the important rules identified by CARs. Being unemployed and engaging in risky behaviors for at least two weeks also has 100% confidence and 60 OR. These participants also had peer mentor contacts made for less than or equal to 2 weeks and they had stress worse two consecutive weeks. Confidence of the rule is 50%, it means 50% likelihood that they develop high PTSD symptoms, and the class support of this rule is 66%. Which means of all the participants who had onset of PTSD symptoms, 66% were contacted by peer mentors for at most 2 weeks and were having stressful symptoms. As these participants had low PTSD symptoms at the time of baseline, probably were less contacted by peer mentors.

4.7.4 Rules Identifying Low PTSD:

It is observed that participants who continued to have low PTSD symptoms from baseline to discharge are 96% non-alcoholic users and 80% of them have low alcoholic use.

4.7.5 Classification results

The set of persistent rules and recovery rules were used to vote whether a participant is likely at risk of persistent high PTSD or not during discharge time. The class label of the rules which have majority voting is assigned to the participant. These assigned labels of the participants are compared with the actual discharge outcomes. The true positive rate of predicting persistent high PTSD and low PTSD for training and test data is shown in Table 4.5. Measure used to evaluate the performance of classifier is the True positive rate (TPR) of predicting the discharge label in participants. TPR is defined as the percentage of participants correctly labelled against the true label. TPR of predicting persistent high PTSD correctly is 87% and that of low-risk class is 70% in the baseline high risk group. Similar process was done for participants in low baseline risk group, both the onset rules and low risk rules were used for predicting the discharge label in these participants. 50% of the participants who developed high PTSD symptoms were predicted correctly and in the low risk group the TPR is 58%. Though these predictions are not far better than a random model due to limited data, our model identifies the precursors which could be used by peer mentors. Unique characteristics of this study is that, this study lists the baseline characteristics along with the current symptoms that are indicators/precursors to long-term crisis.

The predictions of the associative classifier can be used to notify peer mentors about the likelihood of risk in their mentees. The advantage of this associative classifier is that peer mentors can also be presented with the decision rules that predict the risk in their mentees. This not only increases the reliability or trust in the predictions but also provides information about the risk factors and patterns in veterans to plan intervention.

Table 4.5. Rules of persistent, recovery and onset of high PTSD

Rule	Support (%)	Class support (%)	Confidence (%)	Odds Ratio (95% CI)	p-value
Persistent (Baseline PTSD high -> Discharge PTSD high)					
No family and friends support, Health worse for at least one week => High PTSD	12	58.8	100	55 (4,313)	<0.0001
Divorced, Sleep worse for two consecutive weeks, Stress worse for two consecutive weeks => High PTSD	12	58.8	100	47.7 (3.4,260)	<0.0001
Unemployed, One health symptom worse, Stress worse for two consecutive weeks => High PTSD	11	53	100	47.7 (3.4,260)	<0.0001
Divorced, No family and friends support, Sleep worse for at least three weeks => High PTSD	10	47.1	100	30.6 (2,154)	<0.0001
No family and friends support, Little interest in hobbies, Sleep worse for at least one week => High PTSD	8	41.2	100	26.1 (1.6,129)	0.00014
No family support, Served in army, Sleep worse for at least one week => High PTSD	8	41.2	100	26.1 (1.6,129)	0.00014
Unemployed, Sleep worse for at least one week, Stress worse for two consecutive weeks => High PTSD	12	58.8	90.9	54 (4.6, 365)	<0.0001
Divorced, Engaging in risky behaviors for at least two weeks	6	29.4	100	18.3	0.0016
Missed a survey, Sleep worse for at least one week, Stress worse for two consecutive weeks	6	29.4	83.3	14.7 (3,87)	0.005

=> High PTSD					
Served in army, No family and friends support, Stress worse for at least one week => High PTSD	8.4	41.2	77.8	9.6 (1.07, 85)	0.0006
Doesn't enjoy home related activities, Any two symptoms worse in a week => High PTSD	8.3	70.6	92.3	6.7 (1.4,152)	0.025
Doesn't enjoy home related activities, One health symptom worse => High PTSD	50	61	100	11 (1.1,313)	0.044
Doesn't enjoy home related activities, Cigarettes use high=> High PTSD	45.8	64.7	100	12.8 (1.2,130)	0.043
Recovery (Baseline PTSD high -> Discharge PTSD low)					
Moderate interest in hobbies, Feeling self worthy for at least two weeks => Low PTSD	6 (18)	60	83.3	22 (2,232)	0.01
Male, Sleep better for two consecutive weeks => Low PTSD	6 (18)	60	83.3	22 (2,232)	0.01
Moderate interest in hobbies, Health better for at least one week	5 (15)	50	80	14.7 (1.3,157)	0.02
Having family and friends support, Moderate interest in hobbies => Low PTSD	4 (12)	40	80	14.7 (1.3,157)	0.02
Missed surveys for less than two weeks, Sleep better for two consecutive weeks => Low PTSD	6 (18)	60	66.6	7 (1.02, 48)	0.047
Having family and friends support, Feeling self worthy for at least two weeks => Low PTSD	6 (18)	60	66.6	7 (1.02, 48)	0.047
Moderate interest in hobbies => Low PTSD	8 (24)	80	62.5	6.7 (1.18, 38)	0.03
Moderate engagement in community activities, enjoyment in home activities	5 (20.8)	71.4	71.4	18.8 (2.1, 170.7)	0.009
Cigarettes use low, Not exposed to a trauma incident	3 (12.5)	42.9	100	12.8 (1, 157.8)	0.046
Moderate engagement in community activities, Sleep better for two consecutive weeks	3 (12.5)	42.9	100	12.8 (1, 157.8)	0.046
Onset (Baseline PTSD low -> Discharge PTSD high)					
Consumes 6 or more drinks in a day for at least 3 times in a week	6.1	66.7	100	60 (2.7, 1358)	0.01
Takes drink containing alcohol for more than 2	6.1	66.7	100	60 (2.7,	0.01

times a week				1358)	
Alcohol user, have health problems	6.1	66.7	100	60 (2.7, 1358)	0.01
Unemployed, Engaging in risky behaviors for at least two weeks	6.1	66.7	100	60 (2.7, 1358)	0.01
Have family support, Engaging in risky behaviors for at least two weeks	6.1	66.7	100	60 (2.7, 1358)	0.01
Divorced yes, Had problems getting access to healthcare	9.1	100	75	87 (4.3, 1775)	0.004
Interest in physical appearance, , Had problems getting access to healthcare	9.1	100	75	87 (4.3, 1775)	0.004
Community engagement is none, Had problems getting access to healthcare	6.1	66.7	66.7	58 (2.6, 1313.9)	0.011
Cigarettes use low, , Had problems getting access to healthcare	9.1	100	75	87 (4.3, 1775)	0.004

4.7.6 Comparison of Precursors Among Three Groups

In this section, we compare the precursors identified by the rules among three groups persistent, recovery and onset groups. Comparison of these helps us in answering the research questions, if the participants of persistent and onset groups share similar pathways. Tables 4.6 and 4.7 show the shared and diverse characteristics in persistent, onset and recovery groups. One common characteristic in persistent and onset groups is number of cigars use being high and they do not have family support. Contrasting characteristic observed between both the groups is persistent PTSD group doesn't have family support and onset group have family support. Recovery group is also observed to have family support. Therefore, it can be observed that family support played a role in the recovery process but has nothing to do from stopping to develop high PTSD symptoms. Persistent group doesn't enjoy home related activities whereas recovery group show interest in home related activities. High alcoholic use is an indicator of onset of high PTSD symptoms. Participants with high frequency of alcohol use are at great risk of developing high PTSD symptoms whereas in participants with high PTSD symptoms at baseline and low frequency of alcohol predicted the likelihood of recovery in these participants.

Comparison of EMA characteristics showed that, different EMA symptoms were prominent in different groups. Stress and sleep were seen to be observed the most in persistent high PTSD symptoms, whereas engaging in risky behaviors was dominant in the onset group. Although stress and sleep were also worse in the onset group. Participants who had better self worthiness showed signs of recovery from high baseline PTSD symptoms. Having better sleep for two weeks also showed signs of recovery. Sleep is observed to be one good precursor to predict the risk of high PTSD symptoms in participants.

Table 4.6. Comparison of EMA characteristics between three groups

Persistent high PTSD	**Onset of high PTSD**	**Recovery from high PTSD**
Health worse for at least one week	Health worse for at least one week	N/A
Sleep worse for two consecutive weeks	Sleep worse for at least two weeks	Sleep better for at least two weeks
Stress worse for two consecutive weeks	Stress worse for two consecutive weeks	N/A
Any two symptoms worse in a week	Any two symptoms worse for two consecutive weeks	N/A
Engaging in risky behaviors for at least two weeks	Engaging in risky behaviors for at least two weeks	N/A
N/A	N/A	Feeling self worthy for at least two weeks

Table 4.7. Comparison of baseline characteristics between three groups

Persistent high PTSD	**Onset of high PTSD**	**Recovery from high PTSD**
Cigarettes use high	Cigarettes use high	Cigarettes use low
No Family support	Had family support	Has family support
Doesn't enjoy home related activities	N/A	Enjoys home related activities
Doesn't engage in community activities	Doesn't engage in community activities	Moderate engagement in community activities
Little interest in hobbies	N/A	Hobbies interest moderate
N/A	Interest physical appearance	Interest in physical appearance
N/A	Consumes 6 or more drinks in a day for at least 3 times in a week? Yes	Consumes 6 or more drinks in a day for at least 3 times in a week? No
N/A	Takes drink containing alcohol for more than 2 times a week? Yes	When drinking takes less than 1 or 2 drinks
Fatality	N/A	N/A
Unemployed	Unemployed	N/A
N/A	Has Health problems	N/A
Divorced	N/A	N/A
	Peer mentor contacted for less than two weeks	

CHAPTER 5: EARLY WARNING SIGNS OF ACUTE CRISIS EVENTS

Identifying early warning signs is understood to be a potentially useful way to avert mental health crisis. Many research works recognize the importance of responding to early warning signs to prevent relapse [101,102]. Intervention strategies can prevent the escalation of early warning signs into crisis. The goal of this research aim is to focus on acute crisis events. Acute crisis span for short periods of time and are intermittent in nature.

5.1 The Need to Predict Crisis Events

It is suggested that prior information is important for the preparation of a person and their resolution to stress [21]. These crisis patterns not following the expected pattern of recovery is an indication of aid and effective intervention is needed. With proper knowledge, peer mentors can render valuable mental health service to their peer veterans in crisis. Discovering such rules would allow for accumulation of more knowledge of crisis patterns and for the development of effective techniques for intervention. Knowledge of the factors which precipitate into crisis in individuals is proved to be valuable and can be used for tailoring different intervention strategies by peer mentors. Additionally, crisis information allows peer mentors to free up their efforts and to spend their time effectively on more serious cases. “The goal of intervention is to prevent chronic” conditions [21]. It is known that successful crisis solutions will have further implications towards handling future life stressors and in the prevention of mental disorder [125]. This research aims to explore the variables associated with risk taking behaviors which are the warning signs. Warning signs are intended to identify those at imminent risk and facilitate intervention.

5.2 Related Work

Warning signs are apparent in any domain. Earlier works have provided a proof of principle that crises had early indicators of occurrence of hazard. These warning signs were identified in medical [101,102], financial [103], nuclear disaster [104]. In a review on warning signs of suicide in clinical practice [28], author differentiates warning signs from risk factors. It is pointed out that risk factors are studied over a long period of time and have little clinical relevance. Whereas warning signs are associated with near time risk rather than distal relationship. In other work [103], author tried to find out the emerging warning signs of financial crisis in September 2008 from weak signals. Wor2vec was used to find weak signals, and these were evaluated at different time periods using AT function to evaluate their evolving strength. Weak signals that were strengthening over time were considered as early warning signs. All these works present the warning signs, [105] proposes method to evaluate these warning indicators to improve reliability by avoiding false alarms. Statistical approach has been adopted by [101] to detect early warning signs of depression. Momentary changes were captured using smartphone 10 times a day. Statistical summaries (variance, correlation, and autocorrelation) of these repeated measures were analyzed using kernel change point detection to identify the change points. These change points yield as early warning signs to a potential upcoming psychotic disorder. [107] introduced CQUAKE, an earthquake monitoring tool which monitors changes atmospheric parameters to provide early warning information about an impending earth quake.

Within this project, the weekly EMA item for risk taking behavior is considered as acute crisis in this work. To adequately identify early warning signs to risk taking behaviors in veterans, we propose to use AC defined in the previous chapter.

5.3 Risky Behaviors in Veterans

It was observed that OEF/OIF veterans engage in risky behaviors more often than general population especially during mental illness [5]. It was further investigated that risky behaviors were most correlated with high PTSD symptoms. In the current study the association between baseline characteristics, EMA current symptoms and risky behaviors is further investigated. The purpose of this is to evaluate whether any specific characteristics could predict the risk behaviors in veterans.

Several studies have identified the risk factors of PTSD, these risk indices included long-term risk factors which provide very little information about acute crisis events, and these are not evaluated in natural settings. There is a need to identify the warning signs of short-term risk and examine how they vary across subgroups. In this work, we propose a method for the prediction of short term or imminent risk of crisis events or risky behaviors. In this work, we seek to identify the early warning signs or markers of acute crisis in veterans. Being informed about developing symptoms or early warning signs can help reduce the severity of crisis. The following are the research objectives of this research aim:

1. Can the weekly EMA symptoms captured predict the upcoming acute crisis in participants with reasonable accuracy?

2. What are the warning signs that indicate acute crisis?

3. How different are these ML identified warning signs from veterans' implicit theories of warning signs?

5.4 Preprocessing

The goal of this section of the study is in finding the patterns of risky behaviors by veterans during the 12-week rehabilitation program. This information is collected in the weekly EMA surveys to know if there was a risk-taking behavior by the veteran during the past week. The patterns that are identified serve as early warning signs to risky behaviors and can be by peer mentors to tailor targeted intervention.

The data preprocessing involves considering only the veterans who took the discharge survey. Participants who missed the discharge survey are excluded from this study. Whether or not the veteran engaged in risk-taking behavior that week is assessed via EMA question "Have you engaged in any risky behavior (as you define it) this week?" stated in chapter 2. The response of 3 which is worse was considered as engaging in risky behaviors (1) and others as not (0). Total responses reflecting the number of risky behaviors was calculated. A binary variable to capture whether participants engaged in risky behaviors at least once(1) or not (0) is created which is the outcome variable. The other EMA variables are aggregated until the prior week of risky behavior to capture the sleep, stress and health patterns that indicate possible acute crisis. The summarized variables included whether sleep is worse for 2weeks, 3weeks and so on until 11 weeks. Similar aggregations are performed for stress, health, and self-worthiness. Additionally, veterans' peer mentors' perception, or what we viewed as their implicit theories around crisis warning signs were included. These included " any two symptoms worse for at least one week" , "any two symptoms worse for two consecutive weeks", "three symptoms worse for at least one week”, and “three symptoms worse for two consecutive weeks”. These variables are included to verify whether the model identifies these pattens as warning signs. All the baseline variables as discussed in the section 4.2 are also considered in this research aim.

The sample size of this study consists of 108 entries. Whether or not the veteran engaged in risky behaviors is indicated by the class variable “risk flag”. The data distribution of this variable is shown in the Figure 5.1. The data distribution shows the sample is class balanced data, with 60 risk taking behaviors and 48 being the participants who didn’t engage in risk taking behaviors at least once.

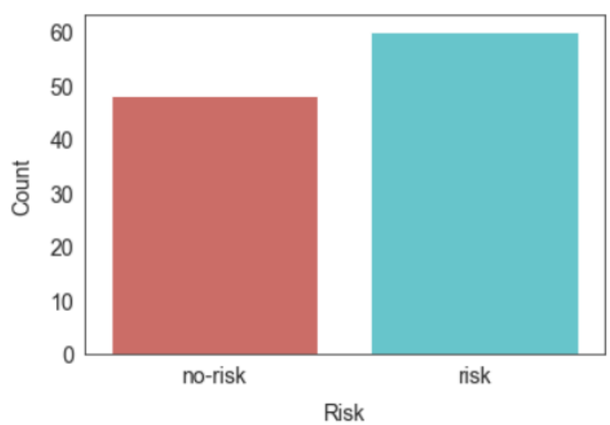

Fig 5.1. Data distribution of class variable

5.5 Analysis of Risky Behaviors by PTSD Levels

The risk behavior characteristics are compared among the participants with persistent, onset and recovery from high PTSD symptoms. Descriptive statistics such as mean, median and mode were used. The average number of times participants without high PTSD symptoms during the 12-week program engaged in risky behaviors is 0.3 (sd = 0.6). But this is different in participants who had high PTSD symptoms either at the time of baseline or discharge. The average and standard deviation of risky behaviors in relevant groups are shown in the Table 5.1. For the participants who joined the program with high PTSD symptoms have a mean value of about 1 (sd: 1-1.4). However, the occurrence of these behaviors is comparatively high in the group who developed high PTSD symptoms by the time of discharge(onset). The average occurrence is approximately 2 with standard deviation of 1.4. The results are consistent with many other studies showing associations between risk behaviors and high PTSD symptoms.

Table 5.1. Average risk taking behavior observed in the sample in three groups

	Average risk-taking behaviors	**Std**
Persistent	1.2	1.1
Onset	1.8	1.4
Recovery	0.8	1.5
Other	0.3	0.6

5.6 Week with Most Occurring Risky Behaviors

12 weeks of the program are compared to find the period when occurrences are most prominent. Knowing the period at which veterans are more likely to engage in risky behaviors is important for mentors, so that they can pay more attention to their mentees. The measure for this analysis is the number of participants engaged in risky behaviors in a week. Both the first occurrences and repetitive occurrences are compared across all the weeks. If a participant engages in risky behaviors in that week for the first time during

the program, then it is considered as “first occurrence” and whether or not it is first or repetitive occurrence it is considered as “all occurrence”. Figure 5.2 shows the percentage of total occurrences for each week. It is observed from the data that on an average 8 participants are likely to engage in risky behaviors in any week, and of them 4 participants are engaged in risky behaviors for the first time. Weeks 6,8,9 ad 10 are observed to have higher percentage of risk-taking behaviors by the participants. At week 6, 9% of the participants are more likely to have reported their first occurrence. Percentages are evaluated here because the raw number of participants remaining in the intervention declined over time.

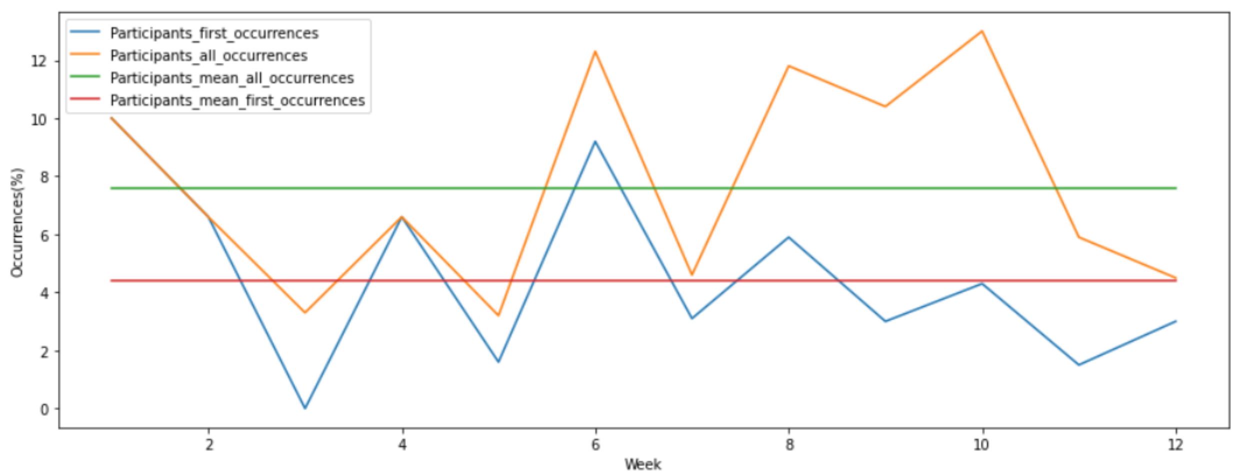

Figure 5.2. Percentage of risky behaviors across twelve weeks

5.7 Correlation Analysis

Correlation analysis is done to evaluate the association between baseline feature variables and EMA variables which are associated with risk taking behaviors. All the variables (high alcoholic use, exposure to trauma incident, marital status, stressful legal problems, worse EMA symptoms) are significantly correlated with risk taking behaviors in the participants. All the input variables are artificially dichotomized and the dependent variable being continuous, Point biserial correlation is used to measure the degree of relationship between independent variables and dependent variable. *P*-value is used to test if this relation is statistically significant or not. The *p*-value shows the probability that this strength may occur by chance. Correlation analysis between the variables revealed that having worse stress and self-worthiness for at least 6 to 8 weeks has a moderate correlation value with risk behaviors with high statistical significance. Having stressful legal problems and exposure to trauma have less correlation between 0.28-0.33. Having any two symptoms worse for two consecutive weeks has strong correlation of 0.7 with very high statistical significance. Similar correlation is observed with having any 3 symptoms worse for atleast one week. Having reported zero worse sleep, stress or self-worthiness have a moderate negative correlation with risk taking behaviors. The results also noted another observed association in the literature, high PTSD symptoms being associated with risk-taking behaviors.

5.8 Method for Mining the Patterns

This research aim uses association rule mining for discovering the patterns of acute crisis events which in this study are the risk-taking behaviors in veterans. Rules are generated from frequent item sets with a min support threshold of 7% and 0.7 confidence. All the redundant rules are filtered by applying the redundancy rules discussed in section 4.6.4. There were nearly 2000 non-redundant rules generated from the algorithm. As all these rules may not represent the true association and some of them could have occurred by chance, odds ratio is calculated to test the statistical significance of these rules. P-value of 0.01 is used as the significance level, which implies probability that association occurred by chance is 1%. To find the optimal set of rules that improve prediction accuracy, rule selection is performed. Support and confidence independently cannot be used for rule selection as they can lead to high bias and variance. For example, if we use a high confidence threshold, we will be missing the rules with low confidence but have good support in the database. On the other hand, having low confidence threshold results in many uninteresting rules. Therefore, to maintain a balance between these two, this work used harmonic mean of support and confidence (weight) as the criteria for rule selection. To find the optimal threshold for weight that has higher accuracy and sensitivity in predicting the positive class, these measures are evaluated for various values. Plot showing these measures for various thresholds is shown in Figure 5.3.

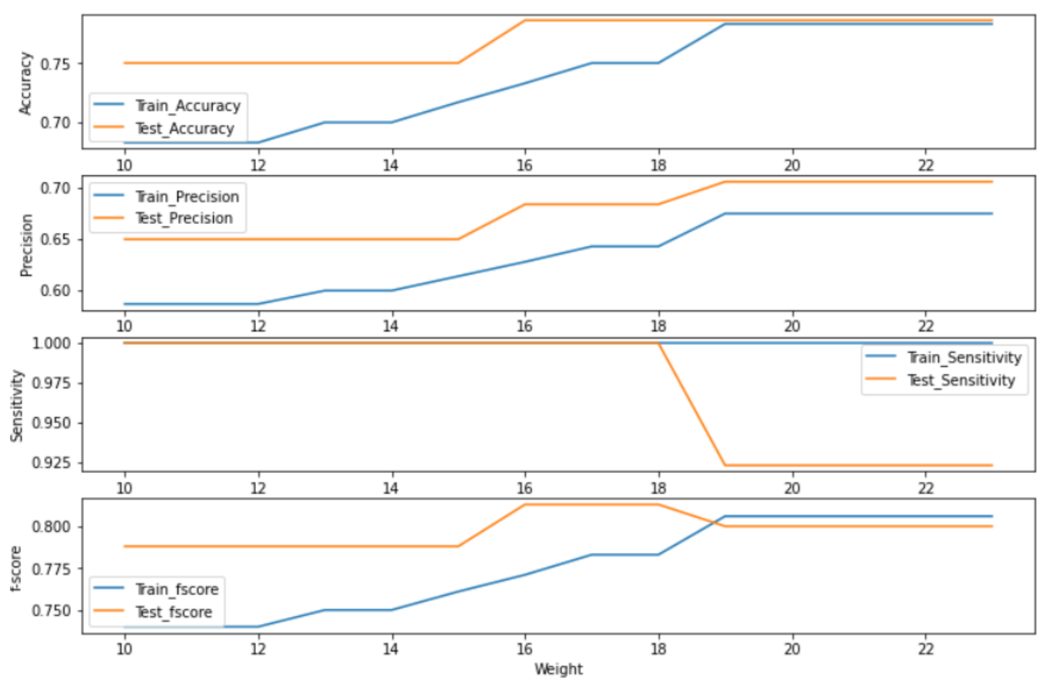

Figure 5.3. Performance metrics for various weight thresholds

From weight 19, precision and f-score of the positive class is high and the accuracy and sensitivity of the positive class are also stable. Therefore, weight 19 is used as the criteria for rule selection and these rules will be used for predicting the risk label in participants. There were in total 35 rules selected and some of these are shown in the Table 5.2

Table 5.2. Identified rules of acute crisis events

Rank	**Rule**	**Support (%)**	**Confidence (%)**	**Odds Ratio (95% CI)**	**p-value**
1	Interest in physical appearance, Any two symptoms worse for two consecutive weeks => Acute crisis	28.3	94.4	79.5 (9.7,646)	0.002
4	Any two symptoms worse for two consecutive weeks, Any three symptoms worse in a week => Acute crisis	31.7	90.5	57.5 (11.6,283)	0.001
5	Fatality, Any two symptoms worse for two consecutive weeks => Acute crisis	28.3	89.5	38.9 (8,187.5)	<0.001
6	Social life limited to few people, Any two symptoms worse for two consecutive weeks => Acute crisis	28.3	89.5	38.9 (8,187.5)	<0.001
8	Fatality, Social life limited to few people, Any two symptoms worse for two consecutive weeks	25	88.2	27.3 (5.7,130)	<0.001
10	Has health problems, Any two symptoms worse for two consecutive weeks => Acute crisis	25	88.2	34.5 (7,165.6)	<0.001
11	Social life limited to few people, Stressful legal problems, Any two symptoms worse in a week	25	88.2	7.4(2.5,21.8)	<0.001
12	Interest in physical appearance, Any three symptoms worse in a week => Acute crisis	28.3	85	16.5 (5.1,52.6)	0.014
14	Fatality, Any three symptoms worse in a week => Acute crisis	26.7	84.2	18.6 (5.4,63.7)	0.016
18	Social life limited to few people, Any three symptoms worse in a week => Acute crisis	31.7	82.6	17.5 (5.6,53.9)	0.002

19	Has health problems, Any three symptoms worse in a week => Acute crisis	26.7	80	11.8 (3.9,35.3)	0.001
20	Stressful legal problems, Any two symptoms worse in a week	26.7	80	5.9 (2,16)	<0.001
21	No family support, Any two symptoms worse in a week => Acute crisis	28.3	73.9	9.9 (3.4,28.4)	0.014
22	Cigarettes use high, Any two symptoms worse in a week => Acute crisis	28.3	68	5.1 (2,13.1)	0.042

5.9 Results

5.9.1 Early Warning Signs Identified

Figure 5.4 shows the pictorial representation of the rules from Table 5.2, all the baseline characteristics are shown in blue and EMA symptoms in brown. From these rules, the most observed EMA symptoms in participants engaging in risky behaviors are “any two symptoms worse in a week”, “any three symptoms worse in a week” and ”any two symptoms worse for two consecutive weeks”. Showing interest in physical appearance with two symptoms worse for two consecutive weeks is seen as the important pattern with 94.4% confidence and OR 79.5. Fatality of dear ones is another important baseline variable which along with any two symptoms worse for two consecutive weeks has nearly 90% confidence and 38.9 OR. The other most important baseline variable observed in the rules is social life limited to few people. OR of this variable is 1 and is not significant to explain the association with the acute crisis. But this variable along with current EMA symptom of experiencing any two symptoms worse for two consecutive weeks has OR of 38.9 and 90% confidence. The other baseline variables observed with this EMA characteristic are having health problems, no family and friends support, and going through stressful legal problems. Participants with all these baseline characteristics and experiencing any three symptoms worse in a week and has a significant association with acute crisis with 80-85% confidence of engaging in risky behaviors. From the results, it can be summarized that these EMA characteristics “any two symptoms worse in a week”, “any three symptoms worse in a week” and ”any two symptoms worse for two consecutive weeks” are some of the early warning signs of acute crisis events in veterans.

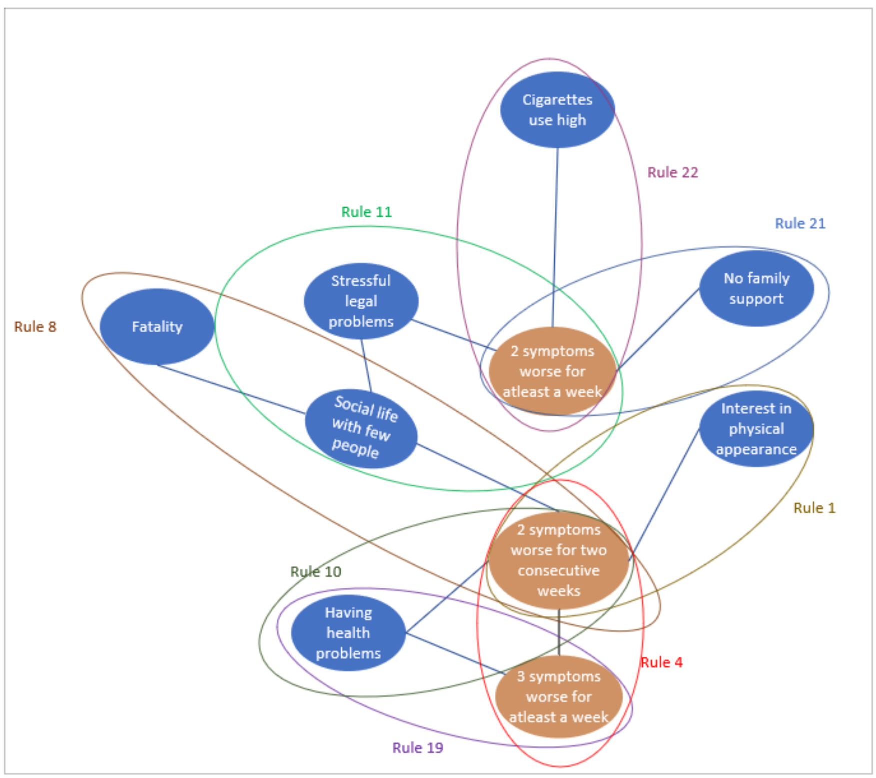

Figure 5.4. Identified rules of acute crisis events

Characteristics in blue represent baseline; In brown represents EMA symptoms; Baseline and EMA symptoms enclosed in a oval represent a rule; Ex: Rule 11 (in green) consists of stressful legal problems, social life with few people, 2 symptoms worse for at least a week

Confidence of these rules is shown in Figure 5.5 For each of the 35 rules, the number of participants the rule predicts correctly as risk are considered true positives, and false predictions are considered as false positives. True positives and false positives of these rules are plotted as shown in Figure 5.6. Numbers in the chart represent rank of the rule. The chart shows that most of these selected rules occupy lower right-hand side of the chart which indicate high true positives and low false positives. Rules are ranked in the decreasing order of confidence and support, rule with highest confidence is assigned rank 1.

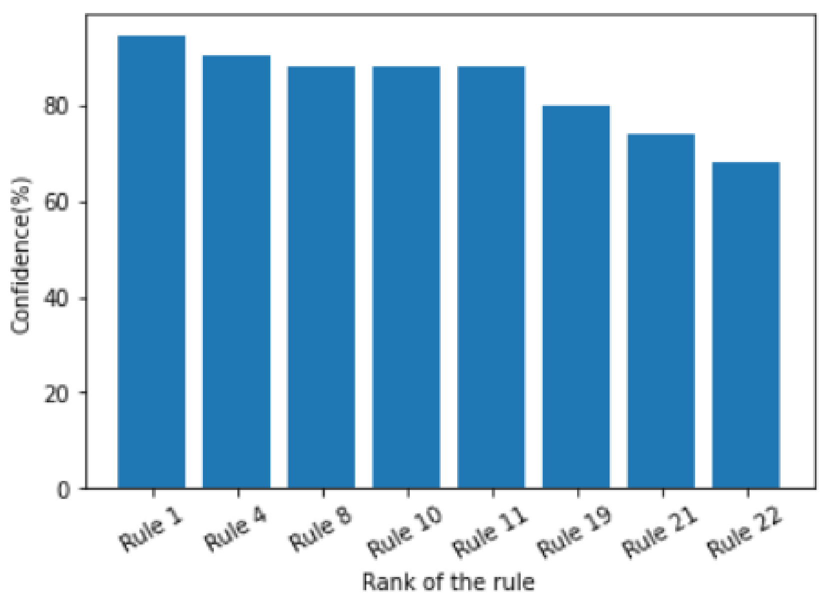

Figure 5.5 Confidence of the rules

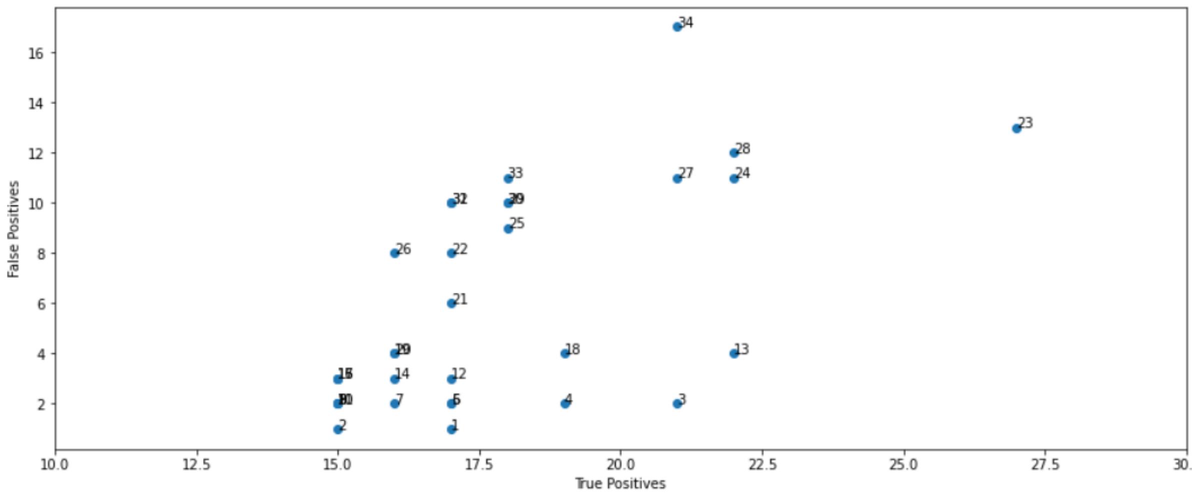

Figure 5.6. True positives vs false positives of the identified rules (Numbers indicate rank of the rules)

5.9.2 Classification Results

The rules identified in the previous section are used to predict if the participant is at risk of engaging in at least one risky behavior. If at least one rule votes a participant as likely to engage in risky behaviors, he/she is considered at risk of acute crisis event or else is considered as not at risk. These rules are evaluated on both train data and unseen test data. The accuracy which is the percentage of correct predictions against the total number of participants, is shown for both train and test data. Along with it the sensitivity of predicting the correct risk label in the risk group is also evaluated. The accuracy in the test data is 82% and the sensitivity of predicting the risk flag which represents the true positives relative to the actual positives is 92.3%. These are in accordance with the accuracy (81.7%) and sensitivity (96.3%) of the train data. The f1 score of both train and test data is 0.82. These results show that the risk rules are reasonably good at predicting the acute crisis events in veterans and the early warning signs identified by the rules can be used by the peer mentors for providing personalized interventions in their mentees.

Hence the associative classifier built in this study not only predicts the true label in these participants but also provides the indicators which are the early warning signs behind the prediction.

CHAPTER 6: CONCLUSION

6.1 Summary

This dissertation addresses several challenges to predict long-term crisis (high PTSD) and intermittent acute crisis events (risk taking behaviors) in veterans in addition to identifying the precursors and early warning signs that are prior indicators to them. Finding of these precursors and early warning signs are based on the definition of crisis theory. The knowledge of these information serve peer mentors in providing intervention to their mentees who are predicted to be at risk of crisis. Some of the unique and major contributions of this work are introducing an interpretable ML model which provides reasoning to the predictions.

The precursors identified to long-term crisis (persistent and onset of high PTSD) and the protective factors in those who recovered from high PTSD observed in chapter 4 are very contrasting. The results showed that having family and friends support, interest in hobbies and engaging in community activities played key role in participants recovery during the rehabilitation program. Absence of these baseline characteristics were observed to have worse sleep symptoms for at least two consecutive weeks and lack of employment with worse stress symptoms. Having health problems at the time of baseline is one of the risk factors for developing high PTSD. Participants who developed high PTSD during the program were also observed to have engaged in risk taking behaviors more frequently than others. The other interesting pattern in these participants is showing interest in physical appearance. This requires further investigation in future work due to insufficiency of data in this group. These precursors identified very much exhibit the characteristics of precursors defined in crisis theory in chapter 1. These are repetitive, there could be multiple occurrences of these events during the 12-week rehabilitation program.

Interest in physical appearance, social life with few people, stressful legal problems, having no family and friends support and fatality of dear ones were most observed in participants who engaged in risky behaviors at least once. The early warning signs observed in these participants are experiencing at least two to three symptoms worse for two consecutive weeks prior to engaging in first risky behavior. These findings are in line with veterans peer mentors' implicit theory of early warning signs.

Though these results were generated from small volume of data, the results are promising as they are in line with the findings from literature. The accuracy and sensitivity of the predictions generated from these patterns were also reasonably good ranging between 80-96 %. Thus, it can be concluded that associative classifier built in this work was not only able to identify patterns that indicate long-term and acute crisis in veterans, it also serves as a predictive model with explainable predictions.

6.2 Limitations

Measures in this study are limited to veteran's self-report, without further testing. Other tools may be needed such as collecting peer mentor feedback to increase data reliability and prevent biases. Though the associative classifier generated predictions with reasonable accuracy and sensitivity, this model has the limitation of being trained on limited set of data. It is advised that future work may consider extending these findings on larger set of participants.

6.3 Future Work

Rules identified in this work can be used to generate alerts to peer mentors to provide intervention programs. Targeted and tailored intervention programs can be implemented for veterans who meet these specific characteristics. The baseline characteristic observed in veterans at high risk was showing interest in physical appearance which needed more data support. From the literature, it was observed that age and health related changes are associated with concerns about physical appearance. It was stated that body image dissatisfaction can have consequences on psychological well-being and self-esteem [107,108]. People with self-presentational concerns may experience lowered self-esteem. It is also understood to result in social isolation in older adults. The future work may further investigate to find if the participants have concerns about their physical appearance. Also the rules identified in this work do not represent causation, this work can be extended to generate causal association rules by future works.

BIBLIOGRAPHY

[1] Whiteford, H.A. et al. 2013. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. The Lancet. 382, 9904 (Nov. 2013), 1575–1586.

[2] https://www.nami.org/mhstats

[3] Janssen, E. M., McGinty, E. E., Azrin, S. T., Juliano-Bult, D., & Daumit, G. L. (2015). Review of the evidence: prevalence of medical conditions in the United States population with serious mental illness. General hospital psychiatry, 37(3), 199-222.

[4] Fulton JJ, Calhoun PS, Wagner HR, Schry AR, Hair LP, Feeling N, Elbogen E, Beckham JC. The prevalence of posttraumatic stress disorder in Operation Enduring Freedom/Operation Iraqi Freedom (OEF/OIF) Veterans: a meta-analysis. J Anxiety Disord. 2015 Apr;31:98-107. doi: 10.1016/j.janxdis.2015.02.003. Epub 2015 Feb 19. PMID: 25768399.

[5] Strom, Thad Q. et al. 2012. "An Exploratory Examination of Risk-Taking Behavior and PTSD Symptom Severity in a Veteran Sample." Military Medicine 177(4): 390–96.

[6] Bullman, T., Schneiderman, A. Risk of suicide among U.S. veterans who deployed as part of Operation Enduring Freedom, Operation Iraqi Freedom, and Operation New Dawn. Inj. Epidemiol. 8, 40 (2021). https://doi.org/10.1186/s40621-021-00332-y

[7] https://www.nbcnews.com/health/health-news/suicide-rate-young-veterans-26-percent-flna1c9441420

[8] S.R. Miles, D.S. Menefee, J. Wanner, A. Teten Tharp, T.A. Kent The relationship between emotion dysregulation and impulsive aggression in veterans with posttraumatic stress disorder symptoms J. Interpers. Violence, 31 (10) (2016), pp. 1795-1816

[9] McDevitt-Murphy, M. E., Fields, J. A., Monahan, C. J., & Bracken, K. L. (2015). Drinking motives among heavy-drinking veterans with and without posttraumatic stress disorder. Addiction research & theory, 23(2), 148-155.

[10] Baker, D. G., Ekhator, N. N., Kasckow, J. W., Dashevsky, B., Horn, P. S., Bednarik, L., & Geracioti Jr, T. D. (2005). Higher levels of basal serial CSF cortisol in combat veterans with posttraumatic stress disorder. American Journal of Psychiatry, 162(5), 992-994

[11] Li Y, Lv Q, Li B, Luo D, Sun X, Xu J. The role of trauma experiences, personality traits, and genotype in maintaining posttraumatic stress disorder symptoms among child survivors of the Wenchuan earthquake. BMC Psychiatry. 2020 Sep 7;20(1):439. doi: 10.1186/s12888-020-02844-1. PMID: 32894097; PMCID: PMC7487586.

[12] Green, B. L., Korol, M., Grace, M. C., Vary, M. G., Leonard, A. C., Gleser, G. C., & Smitson-Cohen, S. (1991). Children and disaster: Age, gender, and parental effects on PTSD symptoms. Journal of the American Academy of Child & Adolescent Psychiatry, 30(6), 945-951.

[13] Prigerson, H. G., Maciejewski, P. K., & Rosenheck, R. A. (2001). Combat trauma: trauma with highest risk of delayed onset and unresolved posttraumatic stress disorder symptoms, unemployment, and abuse among men. The Journal of nervous and mental disease, 189(2), 99-108.

[14] Maoz, H., Goldwin, Y., Lewis, Y. D., & Bloch, Y. (2016). Exploring reliability and validity of the Deployment Risk and Resilience Inventory-2 among a nonclinical sample of discharged soldiers following mandatory military service. Journal of Traumatic Stress, 29(6), 556-562.

[15] Maguen, S., Madden, E., Cohen, B., Bertenthal, D., Neylan, T., Talbot, L., ... & Seal, K. (2013). The relationship between body mass index and mental health among Iraq and Afghanistan veterans. Journal of general internal medicine, 28(2), 563-570.

[16] Schultebraucks, K., Qian, M., Abu-Amara, D. et al. Pre-deployment risk factors for PTSD in active-duty personnel deployed to Afghanistan: a machine-learning approach for analyzing multivariate predictors. Mol Psychiatry (2020). https://doi.org/10.1038/s41380-020-0789-2

[17] M.J. McDermott, M.T. Tull, K.L. Gratz, S.B. Daughters, C. Lejuez The role of anxiety sensitivity and difficulties in emotion regulation in posttraumatic stress disorder among crack/cocaine dependent patients in residential substance abuse treatment J. Anxiety Disord., 23 (5) (2009), pp. 591-599.

[18] Seligowski, A. V., Lee, D. J., Bardeen, J. R., & Orcutt, H. K. (2015). Emotion regulation and posttraumatic stress symptoms: A meta-analysis. Cognitive behaviour therapy, 44(2), 87-102.

[19] Forbes, C. N., Tull, M. T., Rapport, D., Xie, H., Kaminski, B., & Wang, X. (2020). Emotion dysregulation prospectively predicts posttraumatic stress disorder symptom severity 3 months after trauma exposure. Journal of Traumatic Stress, 33(6), 1007-1016.

[20] Eastham, K., Coates, D., & Allodi, F. (1970). The concept of crisis. Canadian Psychiatric Association Journal, 15(5), 463-472.

[21] Miller, K. (1963). The concept of crisis: Current status and mental health implications. Human Organization, 22(3), 195-201.

[22] Shaluf, I. M., & Said, A. M. (2003). A review of disaster and crisis. Disaster Prevention and Management: An International Journal.

[23] Fu, Y. S., Reagan, J. W., & Richart, R. M. (1981). Definition of precursors. Gynecologic oncology, 12(2), S220-S231.

[24] Sornette, Didier. 2002. “Predictability of Catastrophic Events: Material Rupture, Earthquakes, Turbulence, Financial Crashes, and Human Birth.” Proceedings of the National Academy of Sciences of the United States of America 99(SUPPL. 1): 2522–29.

[25] T. Szwedzicki (2004) Warning Signs to Geotechnical Failure of Mining Structures, International Journal of Surface Mining, Reclamation and Environment, 18:2, 150-163, DOI: 10.1080/13895260412331295402

[26] Joseph H. Saleh, Elizabeth A. Saltmarsh, Francesca M. Favarò, Loïc Brevault, Accident precursors, near misses, and warning signs: Critical review and formal definitions within the framework of Discrete Event Systems, Reliability Engineering & System Safety, Volume 114, 2013, Pages 148-154, ISSN 0951-8320, https://doi.org/10.1016/j.ress.2013.01.006.

[27] DiRaddo, J. Douglas, and Stephen E Brock. 2012. “Is It a Crisis?” Principal Leadership 12: 12–16.

[28] Rudd, M. David et al. 2006. “Warning Signs for Suicide: Theory, Research, and Clinical Applications.” Suicide and Life-Threatening Behavior 36(3): 255–62.

[29] Wu, J-L, Yu, L-C and Chang, P-C (2012) Detecting causality from online psychiatric texts using inter-sentential language patterns. BMC Medical Informatics and Decision Making 12, 72.

[30] Kliper, R, Portuguese, S and Weinshall, D (2016) Prosodic analysis of speech and the underlying mental state. In Serino, S, Matic, A, Giakoumis, D, Lopez, G and Cipresso, P (eds), Pervasive Computing Paradigms for Mental Health. MindCare 2015. Communications in Computer and Information Science, vol 604. Cham: Springer, pp. 52–62.

[31] Alam, M. G. R., Haw, R., Kim, S. S., Azad, M. A. K., Abedin, S. F., & Hong, C. S. (2016). EM-Psychiatry: an ambient intelligent system for psychiatric emergency. IEEE Transactions on Industrial Informatics, 12(6), 2321-2330.

[32] Perlis, R. H. (2013). A clinical risk stratification tool for predicting treatment resistance in major depressive disorder. Biological psychiatry, 74(1), 7-14.

[33] Siegel, C. E., Laska, E. M., Lin, Z., Xu, M., Abu-Amara, D., Jeffers, M. K., ... & Marmar, C. R. (2021). Utilization of machine learning for identifying symptom severity military-related PTSD subtypes and their biological correlates. Translational psychiatry, 11(1), 1-12.

[34] Kessler, R. C., van Loo, H. M., Wardenaar, K. J., Bossarte, R. M., Brenner, L. A., Cai, T., ... & Zaslavsky, A. M. (2016). Testing a machine-learning algorithm to predict the persistence and severity of major depressive disorder from baseline self-reports. Molecular psychiatry, 21(10), 1366-1371.

[35] van Breda, W., Pastor, J., Hoogendoorn, M., Ruwaard, J., Asselbergs, J., & Riper, H. (2016, June). Exploring and comparing machine learning approaches for predicting mood over time. In International Conference on Innovation in Medicine and Healthcare (pp. 37-47). Springer, Cham.

[36] Wahle, F., Kowatsch, T., Fleisch, E., Rufer, M., & Weidt, S. (2016). Mobile sensing and support for people with depression: a pilot trial in the wild. JMIR mHealth and uHealth, 4(3), e5960.

[37] Ryu, E., Takahashi, P. Y., Olson, J. E., Hathcock, M. A., Novotny, P. J., Pathak, J., ... & Sloan, J. A. (2015). Quantifying the importance of disease burden on perceived general health and depressive symptoms in patients within the Mayo Clinic Biobank. Health and quality of life outcomes, 13(1), 1-10.

[38] Rosellini, A. J., Dussaillant, F., Zubizarreta, J. R., Kessler, R. C., & Rose, S. (2018). Predicting posttraumatic stress disorder following a natural disaster. Journal of psychiatric research, 96, 15-22.

[39] Kessler, R. C., Rose, S., Koenen, K. C., Karam, E. G., Stang, P. E., Stein, D. J., ... & Carmen Viana, M. (2014). How well can post-traumatic stress disorder be predicted from pre-trauma risk factors? An exploratory study in the WHO World Mental Health Surveys. World Psychiatry, 13(3), 265-274.

[40] Bermejo, P., Lucas, M., Rodríguez-Montes, J. A., Tárraga, P. J., Lucas, J., Gámez, J. A., & Puerta, J. M. (2013, May). Single-and multi-label prediction of burden on families of schizophrenia patients. In Conference on Artificial Intelligence in Medicine in Europe (pp. 115-124). Springer, Berlin, Heidelberg.

[41] Panagiotakopoulos, T. C., Lyras, D. P., Livaditis, M., Sgarbas, K. N., Anastassopoulos, G. C., & Lymberopoulos, D. K. (2010). A contextual data mining approach toward assisting the treatment of anxiety disorders. IEEE transactions on information technology in biomedicine, 14(3), 567-581.

[42] Hoogendoorn, M., Berger, T., Schulz, A., Stolz, T., & Szolovits, P. (2016). Predicting social anxiety treatment outcome based on therapeutic email conversations. IEEE journal of biomedical and health informatics, 21(5), 1449-1459.

[43] Park, A., Conway, M., & Chen, A. T. (2018). Examining thematic similarity, difference, and membership in three online mental health communities from Reddit: a text mining and visualization approach. Computers in human behavior, 78, 98-112.

[44] Metzger, M. H., Tvardik, N., Gicquel, Q., Bouvry, C., Poulet, E., & Potinet-Pagliaroli, V. (2017). Use of emergency department electronic medical records for automated epidemiological surveillance of suicide attempts: a French pilot study. International journal of methods in psychiatric research, 26(2), e1522.

[45] Wshah, S., Skalka, C., & Price, M. (2019). Predicting Posttraumatic Stress Disorder Risk: A Machine Learning Approach. JMIR mental health, 6(7), e13946. https://doi.org/10.2196/13946

[46] Zandvakili, A., Barredo, J., Swearingen, H.R. et al. Mapping PTSD symptoms to brain networks: a machine learning study. Transl Psychiatry 10, 195 (2020). https://doi.org/10.1038/s41398-020-00879-2

[47] Galatzer-Levy I.R., Karstoft K.-I., Statnikov A., Shalev A.Y. Quantitative forecasting of PTSD from early trauma responses: a machine learning application. J. Psychiatr. Res. 2014;59:68–76.

[48] Bryant, R. A. (2003). Early predictors of posttraumatic stress disorder.Biological Psychiatry,53, 789–795. doi:10.1016/S0006-3223(02)01895-4.

[49] Wahle, F., Kowatsch, T., Fleisch, E., Rufer, M., & Weidt, S. (2016). Mobile sensing and support for people with depression: a pilot trial in the wild. JMIR mHealth and uHealth, 4(3), e5960.

[50] Marmar, C. R., Brown, A. D., Qian, M., Laska, E., Siegel, C., Li, M., ... & Vergyri, D. (2019). Speech-based markers for posttraumatic stress disorder in US veterans. Depression and anxiety, 36(7), 607-616.

[51] Saxe, G.N., Ma, S., Ren, J. et al. Machine learning methods to predict child posttraumatic stress: a proof of concept study. BMC Psychiatry 17, 223 (2017). https://doi.org/10.1186/s12888-017-1384-1

[52] Sah, R., Ekhator, N. N., Jefferson-Wilson, L., Horn, P. S., & Geracioti Jr, T. D. (2014). Cerebrospinal fluid neuropeptide Y in combat veterans with and without posttraumatic stress disorder. Psychoneuroendocrinology, 40, 277-283.

[53] Rona, R. J., Jones, M., Sundin, J., Goodwin, L., Hull, L., Wessely, S., &Fear, N. T. (2012). Predicting persistent posttraumatic stress dis-order (PTSD) in UK military personnel who served in Iraq: A longi-tudinal study. Journal of Psychiatric Research,46, 1191–1198. doi:10.1016/j.jpsychires.2012.05.009.UK

[54] Zhao, L., Hao, F., Xu, T., & Dong, X. (2017). Positive and negative association rules mining for mental health analysis of college students. EURASIA Journal of Mathematics, Science and Technology Education, 13(8), 5577-5587.

[55] Wang, C. H., Lee, T. Y., Hui, K. C., & Chung, M. H. (2019). Mental disorders and medical comorbidities: Association rule mining approach. Perspectives in psychiatric care, 55(3), 517-526.

[56] Jayawickreme, N., Atefi, E., Jayawickreme, E., Qin, J., & Gandomi, A. H. (2020). Association rule learning is an easy and efficient method for identifying profiles of traumas and stressors that predict psychopathology in disaster survivors: the example of

Sri Lanka. International journal of environmental research and public health, 17(8), 2850.

[57] Panagiotakopoulos, T. C., Lyras, D. P., Livaditis, M., Sgarbas, K. N., Anastassopoulos, G. C., & Lymberopoulos, D. K. (2010). A contextual data mining approach toward assisting the treatment of anxiety disorders. IEEE transactions on information technology in biomedicine, 14(3), 567-581.

[58] Brown, W. B. (2011). From war zones to jail: Veteran reintegration problems. Justice Policy Journal, 8(1), 1-48.

[59] Rizia, Rizwana et al. 2015. “Collaborative Design with Veterans: Identifying Challenges of Designing Mhealth Solution for Veterans.” 2015 17th International Conference on E-Health Networking, Application and Services, HealthCom 2015: 358–62.

[60] Franco, Zeno et al. 2016. “Crisis Warning Signs in MHealth for Military Veterans : A Collaborative Design Approach.” (May).

[61] George, O., Rizia, R., Hossain, M. F., Johnson, N., Echeveste, C., Mazaba, J. L., ... & Rein, L. (2019). Visualizing Early Warning Signs of Behavioral Crisis in Military Veterans: Empowering Peer Decision Support. In ISCRAM.

[62] Weiss, N. H., Tull, M. T., & Gratz, K. L. (2014). A preliminary experimental examination of the effect of emotion dysregulation and impulsivity on risky behaviors among women with sexual assault–related posttraumatic stress disorder. Behavior modification, 38(6), 914-939.

[63] Tsai, J., Pietrzak, R. H., Hoff, R. A., & Harpaz-Rotem, I. (2016). Accuracy of screening for posttraumatic stress disorder in specialty mental health clinics in the US Veterans Affairs Healthcare System. Psychiatry research, 240, 157-162.

[64] J. M. Brown, J. Williams, R. M. Bray, and L. Hourani, “Postdeployment Alcohol Use, Aggression, and Post-Traumatic Stress Disorder,” Mil. Med., vol. 177, no. 10, pp.1184–1190, 2012.

[65] Wortmann, Jennifer H. et al. 2016. “Psychometric Analysis of the PTSD Checklist-5 (PCL-5) among Treatment-Seeking Military Service Members.” Psychological Assessment 28(11):1392–1403.

[66] Stein, M. B., & Kennedy, C. (2001). Major depressive and post-traumatic stress disorder comorbidity in female victims of intimate partner violence. Journal of affective disorders, 66(2-3), 133-138.

[67] S. I. Omurca and E. Ekinci, “An alternative evaluation of post traumatic stress disorder with machine learning methods,” INISTA 2015 - 2015 Int. Symp. Innov. Intell. Syst. Appl. Proc., no. Ml, 2015.

[68] Ashbaugh, S. Houle-Johnson, C. Herbert, W. El-Hage and A. Brunet, "Psychometric Validation of the English and French Versions of the Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5)", PLOS ONE, vol. 11, no. 10, p. e0161645, 2016. Available: 10.1371/journal.pone.0161645.

[69] J. Sveen, K. Bondjers, and M. Willebrand, “Psychometric properties of the PTSD checklist for dsm-5: A pilot study,” Eur. J. Psychotraumatol., vol. 7, 2016.

[70] N. Chawla, K. Bowyer, L. Hall and W. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique", Journal of Artificial Intelligence Research, vol. 16, pp. 321-357, 2002.

[71] M. Kubat, R. C. Holte, and S. Matwin, “Machine learning for the detection of oil spills in satellite radar images,” Mach. Learn., vol. 30, no. 2–3, pp. 195–215, 1998.

[72] Liu, X. Y., Wu, J., & Zhou, Z. H. (2008). Exploratory undersampling for class-imbalance learning. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 39(2), 539-550.

[73] Franco, Z., Hooyer, K., Roushan, T., O'Brien, C., Johnson, N., Watson, B., ... & Ahamed, S. I. (2018, January). Detecting & Visualizing Crisis Events in Human Systems: an mHealth Approach with High Risk Veterans. In ISCRAM.

[74] Cottler, L. B., Compton, W. M., Mager, D., Spitznagel, E. L., & Janca, A. (1992). Posttraumatic stress disorder among substance users from the general population. American journal of Psychiatry, 149(5), 664-670.

[75] Green, M. A., & Berlin, M. A. (1987). Five psychosocial variables related to the existence of post-traumatic stress disorder symptoms. Journal of Clinical Psychology, 43(6), 643-649.

[76] Zhang, J., Richardson, J. D., & Dunkley, B. T. (2020). Classifying post-traumatic stress disorder using the magnetoencephalographic connectome and machine learning. Scientific reports, 10(1), 1-10.

[77] Tucker, Raymond P., Kevin J. Crowley, Collin L. Davidson, and Peter M. Gutierrez. 2015. “Risk Factors, Warning Signs, and Drivers of Suicide: What Are They, How Do They Differ, and Why Does It Matter?” Suicide and Life-Threatening Behavior 45(6): 679–89.

[78] Waring, J., Lindvall, C., & Umeton, R. (2020). Automated machine learning: Review of the state-of-the-art and opportunities for healthcare. Artificial intelligence in medicine, 104, 101822.

[79] Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Bar-
bado, Salvador Garcia, Sergio Gil-Lopez, Daniel Molina, Richard Benjamins, Raja Chatila, and Francisco Herrera.

Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible
ai. Information Fusion, 58:82–115, 2020. ISSN 1566-2535.
doi:https://doi.org/10.1016/j.inffus.2019.12.012.

[80] Ignatiev, A., Narodytska, N., & Marques-Silva, J. (2019). On validating, repairing and refining heuristic ML explanations. arXiv preprint arXiv:1907.02509.

[81] Rajapaksha, D., Bergmeir, C., & Buntine, W. (2020). LoRMIkA: Local rule-based model interpretability with K-optimal associations. Information Sciences, 540, 221-241.

[82] Li, Dingcheng, Gyorgy Simon, Christopher G Chute, and Jyotishman Pathak. 2013. "Using Association Rule Mining for Phenotype Extraction from Electronic Health Records." AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science 2013: 142–46.
http://www.ncbi.nlm.nih.gov/pubmed/24303254%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3845788.

[83] Khedr, Ahmed M., Zaher Al Aghbari, Amal Al Ali, and Mariam Eljamil. 2021. "An Efficient Association Rule Mining from Distributed Medical Databases for Predicting Heart Diseases." IEEE Access 9: 15320–33.

[84] Orphanou, Kalia et al. 2018. "Incorporating Repeating Temporal Association Rules in Naïve Bayes Classifiers for Coronary Heart Disease Diagnosis." Journal of Biomedical Informatics 81(February): 74–82. https://doi.org/10.1016/j.jbi.2018.03.002.

[85] Park, So Hyun, Shin Yi Jang, Ho Kim, and Seung Wook Lee. 2014. "An Association Rule Mining-Based Framework for Understanding Lifestyle Risk Behaviors." 9(2): 1–9.

[86] Li, Wenmin, Jiawei Han, and Jian Pei. 2001. "CMAR: Accurate and Efficient Classification Based on Multiple Class-Association Rules." Proceedings - IEEE International Conference on Data Mining, ICDM: 369–76.

[87] Mattiev, Jamolbek, and Branko Kavšek. 2020. "A Compact and Understandable Associative Classifier Based on Overall Coverage." Procedia Computer Science 170(2019): 1161–67. https://doi.org/10.1016/j.procs.2020.03.050.

[88] Alwidian, Jaber, Bassam H. Hammo, and Nadim Obeid. 2018. "WCBA: Weighted Classification Based on Association Rules Algorithm for Breast Cancer Disease." Applied Soft Computing Journal 62: 536–49.
http://dx.doi.org/10.1016/j.asoc.2017.11.013.

[89] Chen, Wen Chin, Chiun Chieh Hsu, and Yu Chun Chu. 2012. "Increasing the Effectiveness of Associative Classification in Terms of Class Imbalance by Using a Novel Pruning Algorithm." Expert Systems with Applications 39(17): 12841–50.

[90] Imbalanced, Classifier F O R, and Class Distribution. 2006. "School of IT Technical Report CCCS : A Top-Down Associative Classifier For Imbalanced Class Distribution

Technical Report No 584 Bavani Arunaslam And Sanjay Chawla School Of Information Technologies.” (584).

[91] Forman, G. (2003). An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res., 3(Mar), 1289-1305.

[92] Verhein, Florian, and Sanjay Chawla. 2007. “Using Significant, Positively Associated and Relatively Class Correlated Rules for Associative Classification of Imbalanced Datasets.” Proceedings - IEEE International Conference on Data Mining, ICDM: 679–84.

[93] Jamali, Ilnaz, Mohammad Bazmara, and Shahram Jafari. 2012. “Feature Selection in Imbalance Data Sets.” International Journal of Computer Science Issues 9(3): 42–45.

[94] Szumilas, M. (2010). Explaining odds ratios. Journal of the Canadian academy of child and adolescent psychiatry, 19(3), 227.

[95] Liu, Guimei, Haojun Zhang, and Limsoon Wong. 2011. “Controlling False Positives in Association Rule Mining.” Proceedings of the VLDB Endowment 5(2): 145–56.

[96] Hämäläinen, Wilhelmiina, and Matti Nykänen. 2008. “Efficient Discovery of Statistically Significant Association Rules.” Proceedings - IEEE International Conference on Data Mining, ICDM: 203–12.

[97] Zaki, M. J. (2004). Mining non-redundant association rules. Data mining and knowledge discovery, 9(3), 223-248.

[98] Dietterich, T. G. (2000, June). Ensemble methods in machine learning. In International workshop on multiple classifier systems (pp. 1-15). Springer, Berlin, Heidelberg.

[99] Mason, J. W. (1975). A historical view of the stress field. Journal of human stress, 1(2), 22-36.

[100] Ailshire, J. A., & Burgard, S. A. (2012). Family relationships and troubled sleep among US adults: examining the influences of contact frequency and relationship quality. Journal of health and social behavior, 53(2), 248-262.

[101] Cabrieto, J., Adolf, J., Tuerlinckx, F., Kuppens, P., & Ceulemans, E. (2019). An objective, comprehensive and flexible statistical framework for detecting early warning signs of mental health problems. Psychotherapy and psychosomatics, 88(3), 184-186.

[102] Allan, S., Bradstreet, S., McLeod, H. J., Gleeson, J., Farhall, J., Lambrou, M., ... & Gumley, A. I. (2020). Perspectives of patients, carers and mental health staff on early warning signs of relapse in psychosis: a qualitative investigation. BJPsych open, 6(1).

[103] El Akrouchi, Manal, Houda Benbrahim, and Ismail Kassou. 2020. “Monitoring Early Warning Signs Evolution through Time.” ACM International Conference Proceeding Series.

[104] Lipscy, P. Y., Kushida, K. E., & Incerti, T. (2013). The Fukushima disaster and Japan's nuclear plant vulnerability in comparative perspective. Environmental science & technology, 47(12), 6082-6088.

[105] Boettiger, Carl, and Alan Hastings. 2012. "Quantifying Limits to Detection of Early Warning for Critical Transitions." Journal of the Royal Society Interface 9(75): 2527–39.

[106] Cabrieto, Jedelyn et al. 2019. "An Objective, Comprehensive and Flexible Statistical Framework for Detecting Early Warning Signs of Mental Health Problems." Psychotherapy and Psychosomatics 88(3): 184–86.

[107] Cervone, G., Kafatos, M., Napoletani, D., & Singh, R. P. (2006). An early warning system for coastal earthquakes. Advances in Space Research, 37(4), 636-642.

[108] Yazdani, N., Hosseini, S. V., Amini, M., Sobhani, Z., Sharif, F., & Khazraei, H. (2018). Relationship between body image and psychological well-being in patients with morbid obesity. International journal of community based nursing and midwifery, 6(2), 175.

[109] Satghare, P., Mahesh, M. V., Abdin, E., Chong, S. A., & Subramaniam, M. (2019). The relative associations of body image dissatisfaction among psychiatric out-patients in Singapore. International Journal of environmental research and Public health, 16(24), 5162.

[110] "What Is Posttraumatic Stress Disorder (PTSD)?" Psychiatry.org - What Is Posttraumatic Stress Disorder (PTSD)?, https://psychiatry.org/patients-families/ptsd/what-is-ptsd.

[111] Jorge, R. E. (2015). Posttraumatic stress disorder. Continuum: lifelong learning in neurology, 21(3), 789-805.

[112] "Va.gov: Veterans Affairs." Combat Exposure, 19 Sept. 2018, https://www.ptsd.va.gov/understand/types/combat_exposure.asp.

[113] Gates, M. A., Holowka, D. W., Vasterling, J. J., Keane, T. M., Marx, B. P., & Rosen, R. C. (2012). Posttraumatic stress disorder in veterans and military personnel: epidemiology, screening, and case recognition. Psychological services, 9(4), 361.

[114] Matthew Tull, PhD. "Veterans of the OEF/OIF Conflicts Are at High Risk for PTSD." Verywell Mind, Verywell Mind, 27 June 2022, https://www.verywellmind.com/oefoif-veterans-and-posttraumatic-stress-ptsd-symtpoms-2797314.

[115] Morin, R. (2011). The difficult transition from military to civilian life. Washington, DC: Pew Research Center.

[116] De Kloet, C. S., Vermetten, E., Geuze, E., Wiegant, V. M., & Westenberg, H. G. M. (2008). Elevated plasma arginine vasopressin levels in veterans with posttraumatic stress disorder. Journal of psychiatric research, 42(3), 192-198.

[117] Morawetz C, Mohr PNC, Heekeren HR, Bode S. The effect of emotion regulation on risk-taking and decision-related activity in prefrontal cortex. Soc Cogn Affect Neurosci. 2019 Oct 1;14(10):1109-1118. doi: 10.1093/scan/nsz078. PMID: 31680150; PMCID: PMC6970147.

[118] Weiss NH, Sullivan TP, Tull MT. Explicating the role of emotion dysregulation in risky behaviors: A review and synthesis of the literature with directions for future research and clinical practice. Curr Opin Psychol. 2015 Jun 1;3:22-29. doi: 10.1016/j.copsyc.2015.01.013. PMID: 25705711; PMCID: PMC4332392.

[119] Friend SF, Nachnani R, Powell SB, Risbrough VB. C-Reactive Protein: Marker of risk for post-traumatic stress disorder and its potential for a mechanistic role in trauma response and recovery. Eur J Neurosci. 2022 May;55(9-10):2297-2310. doi: 10.1111/ejn.15031. Epub 2020 Nov 23. PMID: 33131159; PMCID: PMC8087722.

[120] Burke LE, Shiffman S, Music E, Styn MA, Kriska A, Smailagic A, Siewiorek D, Ewing LJ, Chasens E, French B, Mancino J, Mendez D, Strollo P, Rathbun SL. Ecological Momentary Assessment in Behavioral Research: Addressing Technological and Human Participant Challenges. J Med Internet Res. 2017 Mar 15;19(3):e77. doi: 10.2196/jmir.7138. PMID: 28298264; PMCID: PMC5371716.

[121] Price, M., Yuen, E. K., Goetter, E. M., Herbert, J. D., Forman, E. M., Acierno, R., & Ruggiero, K. J. (2014). mHealth: a mechanism to deliver more accessible, more effective mental health care. Clinical psychology & psychotherapy, 21(5), 427-436.

[122] Roux-Dufort, C. (2007). Is crisis management (only) a management of exceptions?. Journal of contingencies and crisis management, 15(2), 105-114.

[123] Kunreuther, H. C., Bier, V. M., & Phimister, J. R. (Eds.). (2004).” Accident precursor analysis and management: reducing technological risk through diligence”. National Academies Press.

[124] NASA. Nasa accident precursor analysis handbook. Available from: https://ntrs.nasa.gov/api/citations/20120003292/downloads/20120003292.pdf. [accessed 2022]

[125] Caplan, Gerald: (Ed.) Prevention of Mental Disorders in Children. New York, Basic Books, Inc. 1961.